

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Dynamic analysis in productivity, oil shock, and recession

Permalink

<https://escholarship.org/uc/item/0gj2t1pq>

Author

Katayama, Munechika

Publication Date

2008

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Dynamic Analysis in Productivity, Oil Shock, and Recession

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

in

Economics

by

Munechika Katayama

Committee in charge:

Professor James D. Hamilton, Chair
Professor Marjorie Flavin
Professor Takeo Hoshi
Professor Valerie A. Ramey
Professor Krislert Samphantharak

2008

Copyright
Munechika Katayama, 2008
All rights reserved.

The dissertation of Munechika Katayama is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2008

DEDICATION

To my parents.

TABLE OF CONTENTS

	Signature Page	iii
	Dedication	iv
	Table of Contents	v
	List of Figures	vii
	List of Tables	viii
	Acknowledgements	ix
	Vita and Publications	x
	Abstract	xi
Chapter 1	The Predictive Power of the Interest Rate for Industry-level TFP: A State-space Approach	1
	1.1 Introduction	1
	1.2 Method	3
	1.2.1 Measuring the Industry-level Productivity Growth	4
	1.2.2 State space Model	6
	1.2.3 Data	12
	1.3 Empirical Results	13
	1.4 Extension	20
	1.5 Conclusion	23
Chapter 2	Declining Effects of Oil-price Shocks	25
	2.1 Introduction	25
	2.2 Transportation Industry Background	31
	2.2.1 Petroleum Consumption for Transportation	31
	2.2.2 Regulation in the Transportation Industry	31
	2.3 The Model	35
	2.3.1 Aggregators	36
	2.3.2 Households	36
	2.3.3 Intermediate-good Firms	37
	2.3.4 Transportation Firms	40
	2.3.5 The Monetary Authority	42
	2.3.6 Real Price of Oil	42
	2.3.7 Symmetric Equilibrium	42
	2.4 Results	43
	2.4.1 Calibration	43
	2.4.2 Baseline Results	46
	2.4.3 Changes in the Economy Over Time	48
	2.5 Conclusion	57

Chapter 3	Improving Recession Probability Forecasts in the U.S. Economy . . .	58
3.1	Introduction	58
3.2	Basic Models and Results	60
3.2.1	Recession Probability Forecasting Model	60
3.2.2	Data	62
3.2.3	Evaluating Forecasting Performance	64
3.2.4	Results	67
3.3	Allowing More Flexible Skewness and Excess Kurtosis	81
3.3.1	Generalized Edgeworth Expansion	81
3.3.2	The GEE Results	83
3.3.3	Current Predictions by the GEE	85
3.4	Conclusion	86
Appendix A	Appendix to Chapter 1	91
A.1	The Kalman filter and maximum likelihood estimation	91
A.2	Gradient and Information Matrix	92
Appendix B	Appendix to Chapter 2	93
B.1	Calculating the Ratio of $T/\tilde{Y}$	93
B.2	Characterizing Equilibrium	94
Appendix C	Appendix to Chapter 3	96
C.1	Detailed Expression for (3.14)	96
C.2	Parameter Estimates based on the Recent Data	96
References		96

LIST OF FIGURES

Figure 1.1:	Decomposition of variance in the Solow residual	14
Figure 1.2:	Comparison of the Impulse Response Functions	15
Figure 2.1:	Different Responses to the Oil-price Shock: Before and After 1984 . .	27
Figure 2.2:	Petroleum Consumption by Sector 1949–2005	29
Figure 2.3:	Real Price of Oil	30
Figure 2.4:	Petroleum Consumption in the Transportation Industry by Mode . .	32
Figure 2.5:	Impulse Responses to a 10% Increases in the Oil Price: Baseline Results	47
Figure 2.6:	Comparison of IRFs to a 10% Increase in the Oil Price: Case 1 and Case 2	50
Figure 2.7:	Comparison of IRFs to a 10% Increase in the Oil Price: Case 3 and Case 4	54
Figure 2.8:	Comparison of IRFs to a 10% Increase in the Oil Price: Before and After 1984	55
Figure 3.1:	Predictions with Univariate Probit Models	70
Figure 3.2:	Comparison of Out-of-Sample Predictions	79
Figure 3.3:	Recursive Out-of-Sample vs. Leaving 2-years Out	87
Figure 3.4:	Comparison of Forecasting Accuracy: Recursive vs. Leaving 2-years Out	88
Figure 3.5:	Admissible Set of Skewness and Excess Kurtosis with the Logistic ($\mu = 0, \sigma^2 = 1.2$)	89
Figure 3.6:	Comparison of Functional Forms	89
Figure 3.7:	Out-of-Sample Predictions by the GEE	90
Figure 3.8:	Current Predictions by the GEE	90

LIST OF TABLES

Table 1.1:	Granger-causality Test for 1960:I - 1996:IV	13
Table 1.2:	Likelihood Ratio Test Results for 2- and 3-industry Model	16
Table 1.3:	Subcategories of Manufacturing Industries	17
Table 1.4:	Wald Test Results for Subcategories in Manufacturing	18
Table 1.5:	Wald Test Results for Detailed Manufacturing Industries	19
Table 1.6:	Estimation Results with Capital Utilization	22
Table 2.1:	Summary of Parameter Values in the Baseline Specification	44
Table 2.2:	Summary of Changes in the Peak Response of Aggregate Value Added	49
Table 3.1:	List of the Cumulative Distribution Functions Considered	63
Table 3.2:	List of Variables	65
Table 3.3:	Variable Rankings with Univariate Probit Models	68
Table 3.4:	Top 20 Univariate Models	71
Table 3.5:	Top 20 Bivariate Models [QPS]	73
Table 3.6:	Top 20 Bivariate Models [LPS]	73
Table 3.7:	Top 20 Bivariate Models [KS]	74
Table 3.8:	Top 20 Models [QPS]	75
Table 3.9:	Top 20 Models [LPS]	75
Table 3.10:	Top 20 Models [KS]	76
Table 3.11:	Importance of the Two Margins	77
Table 3.12:	Out-of-Sample Forecasting Performance with the Generalized Edge- worth Expansion	84
Table C.1:	Parameter Estimates based on the Recent Data	97

ACKNOWLEDGEMENTS

I would like to express sincere appreciation to my advisor, James D. Hamilton. I am truly grateful and indebted for his encouragement and guidance at every stage of my research and this dissertation. I would not be able to accomplish this work and fulfill my dream of becoming a professor without his stimulating suggestion and support. His thought, attitude, enthusiasm, and sincerity have great influence on my life. His econometrics lecture was one of the best lectures I have taken and I will not forget what I have learned from him as a researcher and as an individual. I hope that one day, I will become an insightful researcher, a influential teacher, and a great person like him.

I would also like to thank my committee members, Marjorie Flavin, Takeo Hoshi, Valerie A. Ramey, and Krislert Samphantharak. Their comments and suggestions have contributed to improve this work. Lessons I have learned from them will influence on my future research as well. In addition, I have learned many practical issues through being a research assistant for Takeo Hoshi. This experience will be my invaluable asset.

I am grateful to my friends for being there. They have helped me go through the ups and downs in my life. My deepest gratitude goes to many of my colleagues at the department of economics for stimulating discussions and conversations about research, life, and everything else. In particular, I would like to thank Yoichi Arai, Shigeru Fujita, Hibiki Ichiue, Munir Jalil, Yong-Gook Jung, Hiroaki Kaido, Young-Do Kim, Kwan Hwan Kim, Julie Lee, Tatsuyoshi Okimoto, Young Joon Park, Karmen Suen, and Tomoki Tanemura, among many others. I am grateful for everything that we have shared together.

My years in San Diego as a graduate student have been exciting and challenging. There were unforgettable and pleasant experiences. At the same time, there were tough moments that I have never encountered in my life. Without the support from all the above people, I have never overcome such obstacles.

Above all, I would like to thank my family, especially my parents, Masato and Reiko Katayama, for their continuous support and encouragement. Without them, I would not be able to achieve anything. I would like to dedicate this dissertation to them.

VITA

1977	Born, Tokyo, Japan
1999	B.A. in Economics, Keio University, Tokyo, Japan
2001	M.A. in Economics, State University of New York at Buffalo, Buffalo
2002–2008	Teaching and Research Assistant, University of California, San Diego
2008	Ph.D. in Economics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Dynamic Analysis in Productivity, Oil Shock, and Recession

by

Munehika Katayama

Doctor of Philosophy in Economics

University of California San Diego, 2008

Professor James D. Hamilton, Chair

The first chapter analyzes the predictive power of the interest rate for various industry-level measures of productivity growth. Although industry-level data are only available at an annual frequency, by using the state space model and the Kalman filter, it is possible to perform the Granger-causality test at a quarterly frequency. The results highlight the heterogeneous nature of predictive power and suggest that the nonexogeneity of the Solow residual reported by Evans (1992) is due to manufacturing industries. In addition, two case studies on industries in which we can obtain an appropriate measure of capital utilization rate show that the forecasting ability of the interest rate diminishes after taking account of variable capital utilization in TFP growth.

The second chapter studies macroeconomic consequences of oil-price shocks. Output responses to oil-price shocks not only tend to be weaker, but also to peak earlier recently. This chapter builds a model that incorporates a realistic structure of US petroleum consumption and explores three possible explanations for the changes. The first is based on deregulation in the transportation sector, which has brought more competition and improved efficiency in the industry. The second is overall improvements in use of energy. The third is less persistence of the oil-price shock. Under realistic parameter values, it is demonstrated that all three factors play an important role quantitatively. These three factors together could account for a large portion of the changes over time.

The third chapter examines forecasting models of recession probabilities. There are two margins to improve forecasting accuracy: including additional variables and using different functional form. Using out-of-sample and cross validation methods, I systematically compare the performance of various forecasting models that differ in terms of

variables included and functional forms used. I find substantial gains from including additional variables, such as the S&P 500 and non-farm employment growth, together with the term spread. Additionally, we can further improve forecasting accuracy by utilizing a non-Normal distribution. I also explore this possibility by using the generalized Edgeworth expansion. Resulting predictions outperform ones from a probit model in all three measures of forecasting accuracy considered.

Chapter 1

The Predictive Power of the Interest Rate for Industry-level TFP: A State-space Approach

1.1 Introduction

Although productivity shocks play an important role in modern macroeconomic models, we have limited knowledge about their usual empirical measures, the Solow residual. Although the Solow residual plays an important role in the RBC literature as an exogenous driving force, Evans (1992) casts doubt on its exogeneity by showing that the observed Solow residual is Granger-caused by money, interest rates, and government spending.

There have been many studies that try to understand the nature of the observed Solow residual and they have offered several possibilities why the conventionally measured Solow residual might fail to be the true measure of technological progress, such as variable factor utilization, variable markup rate, increasing returns to scale, and sectoral reallocation. Partly because of data availability, however, most studies are based on aggregate data. We argue that we also need to understand the nature of technological progress at the disaggregated level. It is true that the aggregate one-sector model is handy to analyze, but at the same time, our economy consists of heterogeneous industries. So we will obtain better insights from more disaggregated level of analysis. For instance, it is possible that some industries are better characterized by increasing returns

to scale, but other industries are not.

There are several studies analyzing industry-level productivity. Many of them focus on manufacturing industries, such as Basu and Fernald (1995), Basu (1996), Basu and Kimball (1997), and Burnside, Eichenbaum, and Rebelo (1995, 1996), among others. A few exceptions are Basu and Fernald (2001), Basu, Fernald, and Shapiro (2001), and Basu, Fernald, and Kimball (2004), which cover the entire industries in the economy and try to establish the link between the industry-level productivity and the aggregate productivity. Most of the above studies take account of possibilities that contaminate the observed measure of technological progress. However, the fundamental question raised by Evans (1992) still remains unanswered at the disaggregated level.

This paper focuses on the following question: What are the sources of the Granger-causality found in the aggregate Solow residual? More specifically, we decompose the aggregate Solow residual into industry-level TFP growth and investigate sources of the Granger-causality within highly aggregated groups of industries and more disaggregated industries systematically.

Although Evans (1992) uses quarterly data, industry-level productivity series are only available at an annual frequency. It is often the case that macroeconomic variables of interest are available at different frequencies. In such circumstances, we usually need either to convert high frequency variables into lower frequency or to interpolate low frequency variables with appropriate signals. Both approaches have some problems. If we convert high frequency data into a lower frequency, we may lose some important information. On the other hand, there is no guarantee that the variable used to interpolate low frequency series is a good proxy for the true time series behavior and it is likely to create an errors-in-variable problem.

In this paper, we investigate an alternative way to handle observations at different frequencies. Our approach is based on the state space model. Together with the Kalman filter, the state space model provides a highly flexible framework that can deal with unobserved component in the estimation model. For example, Hamilton (1985) and Burmeister, Wall, and Hamilton (1986) use the state space model to uncover the market's expected rate of inflation based on observed interest rates and inflation.

Recently it has also been popular to use the state-space representation and the Kalman filter to take DSGE models to the data. In order to cope with the stochastic singularity problem in DSGE models, people include a variety of additional structural

disturbances so that the number of shocks equals to the number of observables in the estimation, or add measurement errors.¹ In such situations, regardless of whether one uses maximum likelihood estimation or a Bayesian approach, it is necessary to rely on the Kalman filter. There are many other advantages of the state space model, for example, its capabilities of handling missing observations, temporal and contemporaneous aggregation, and data irregularities such as those arising from data revisions. However, most applications of the state space model have been limited to simple introduction of unobserved components as described above.

It is not new to use the state space model in order to cope with missing observations (see, for example, Harvey and Pierse, 1984). However, such applications mainly focus on uncovering missing observations in time series. By combining missing observations and temporal aggregation, we will be able to analyze time series data at different frequencies in one-step without converting data frequencies beforehand. In such a unified framework, we can perform statistical tests based on estimates, which are entirely consistent with observed realizations even at different frequencies. Not only do we decompose the aggregate Solow residual, but also we perform Granger-causality test at a quarterly frequency by using the state space model in order to explore the source of the Granger-causality.

The rest of the paper is organized as follows. Section 1.2 describes how to measure the industry-level TFP based on gross outputs and explains how we incorporate annual data into the state space model with a quarterly frequency. Section 1.3 updates findings in Evans (1992) and analyzes the Granger-causality at the aggregate level using the disaggregated data. Section 1.4 extends our framework to control for variable capital utilization. Finally Section 1.5 concludes.

1.2 Method

In this section, we describe our procedure. Section 1.2.1 explains how to measure the industry-level productivity and establish the link between the industry-level productivity and the aggregate productivity measures. Details about the state space model and the Kalman filter are given in Section 1.2.2. Description of the data is in Section 1.2.3.

¹See, for example, Smets and Wouters (2003), Ireland (2004), and references therein.

1.2.1 Measuring the Industry-level Productivity Growth

When we analyze the relation between the aggregate economy and underlying industry activities, we will encounter a dilemma. Although GDP is value added, an appropriate measure of output at the industry or firm level is gross output if we are interested in the industry-level productivity. Value added is defined as the difference between the value of gross output and the cost of intermediate inputs. Goods produced by industries or firms are either consumed as final products or used as intermediate inputs for other production processes. To avoid double counting, an aggregate measure of output is value added, which is net of payments to intermediate inputs. If we use aggregate value added, all flow of intermediate inputs will be canceled out and hence it does not cause any problem of ignoring intermediate inputs. However, interindustry flows of intermediate goods do not cancel out even at the one-digit industry level.

Not only does value added have no physical counterparts in terms of tangible goods and services as industry-level outputs, but also we need an additional restrictive assumption that the gross-output production function takes the following form:²

$$Q_i = F(A_i G(K_i, L_i), E_i, M_i). \quad (1.1)$$

where Q_i is gross output in industry i , $G(\cdot)$ is a value-added function, K_i and L_i are primary inputs of capital and labor, A_i is a technology index, E_i energy usage in production process, and M_i is material inputs. For notational simplicity, time-subscripts are suppressed. This separability condition requires that the marginal rate of substitution between capital and labor is independent of energy and material inputs and implies that technological innovation only improves the productivity of capital and labor, but not intermediate inputs. A more general specification would be

$$Q_i = Z_i F(K_i, L_i, E_i, M_i). \quad (1.2)$$

Once difference between (1.1) and (1.2) is that Z_i is gross-output-augmenting technology, while A_i is value-added-augmenting technology. In order to distinguish A_i and Z_i , hereafter, we call $\frac{dA}{A}$ the Solow residual and TFP (or KLEM TFP) refers to $\frac{dZ}{Z}$.

While we want to establish the link between the aggregate Solow residual and the industry-level TFP, simple aggregation of the industry-level TFP does not work

²This weak separability condition has been well-known in the production function literature. Sato (1975, Ch.2) provides a comprehensive summary. In the context of productivity analysis, see Jorgenson (1990) and Hulten (2001).

because there exist conceptual differences as described above. The dilemma is coming from the fact that appropriate measures of outputs and specifications of the production function are different.

To overcome the dilemma, we will take the following strategy: (i) Calculate the industry-level technological progress; (ii) Aggregate over industries to construct the aggregate measure of technological progress based on gross output; (iii) Using an accounting identity, convert the gross-output based measure of productivity into the one based on the value-added function concept.³

Under perfect competition and constant returns to scale, a standard growth accounting approach based on (1.2) gives us the industry-level TFP

$$\frac{dZ_i}{Z_i} = \frac{dQ_i}{Q_i} - \left\{ c_i^K \frac{dK_i}{K_i} + c_i^L \frac{dL_i}{L_i} + c_i^E \frac{dE_i}{E_i} + c_i^M \frac{dM_i}{M_i} \right\}, \quad (1.3)$$

where $c^X = \frac{P_i^X X_i}{P_i^Q Q_i}$ for any variable X_i and P^X denotes a price of any good X_i .

Aggregation over industries can be done by defining the Divisia index of aggregate KLEM TFP as follows:

$$\frac{dZ}{Z} = \sum_{i=1}^N \omega_i^Q \frac{dZ_i}{Z_i}, \quad (1.4)$$

where $\omega_i^Q = \frac{P_i^Q Q_i}{\sum_{i=1}^N P_i^Q Q_i}$. Notice that (1.4) holds at any level of aggregation, regardless of whether we are aggregating over the entire economy or a subset of the economy.

In order to convert the aggregate measure of KLEM TFP into the aggregate Solow residual, we will start from an accounting identity. Let Y be aggregate real value added. Since nominal value added is the difference between the value of gross output and the expenditure for energy and material inputs,

$$P^Y Y = P^Q Q - P^E E - P^M M,$$

where variables without subscripts are aggregate measures. Following Sato (1976), we define the Divisia quantity index for aggregate real value added

$$\frac{dY}{Y} = \frac{P^Q Q}{P^Y Y} \frac{dQ}{Q} - \frac{P^E E}{P^Y Y} \frac{dE}{E} - \frac{P^M M}{P^Y Y} \frac{dM}{M} = (1 + s^E + s^M) \frac{dQ}{Q} - s^E \frac{dE}{E} - s^M \frac{dM}{M}, \quad (1.5)$$

where $s^X = \frac{P^X X}{P^Y Y}$ for any variable X .

³This strategy is similar to the one used in Basu and Fernald (2001), Basu, Fernald, and Shapiro (2001), and Basu, Fernald, and Kimball (2004).

Using $\frac{dQ}{Q} = \frac{dZ}{Z} + c^K \frac{dK}{K} + c^L \frac{dL}{L} + c^E \frac{dE}{E} + c^M \frac{dM}{M}$ and $(1 + s^E + s^M)c^X = \frac{P^Q Q}{P^Y Y} \frac{P^X X}{P^Q Q} = \frac{P^X X}{P^Y Y} = s^X$ for any X , we can rewrite (1.5) to obtain

$$\frac{dY}{Y} = s^K \frac{dK}{K} + s^L \frac{dL}{L} + s^Q \frac{dZ}{Z}.$$

Since the Solow residual $\frac{dA}{A}$ is computed as $\frac{dA}{A} = \frac{dY}{Y} - s^K \frac{dK}{K} - s^L \frac{dL}{L}$, we obtain the relation between the aggregate Solow residual and the industry-level KLEM TFP as

$$\frac{dA}{A} = s^Q \frac{dZ}{Z}.$$

Since the value of gross output is greater than the value of value added by construction, this tells us that the value-added based measure of productivity is greater in magnitude than the TFP based on all inputs. Using (1.4), we can achieve the relation between the aggregate Solow residual and the industry-level TFPs:

$$\frac{dA}{A} = \sum_{i=1}^N d_i \frac{dZ_i}{Z_i}$$

where $d_i = \frac{P_i^Q Q_i}{\sum_{i=1}^N P_i^Y Y_i}$ is called the Domar weight (Domar, 1961).

Notice that any change in assumptions, such as imperfect competition, increasing returns to scale, or variable factor utilization, will alter the form of (1.3). However, the remaining steps are unchanged.

So far, technological progress has been characterized as a continuous-time growth rate (i.e., the Divisia index). In our empirical study, we employ a commonly used discrete approximation to the Divisia index, namely the Törnqvist index. The Törnqvist index is defined as the differences in logs and the weights are equal to the average of current and previous periods.

1.2.2 State space Model

A state space model consists of two sets of equations:

$$\text{State equation} \quad \boldsymbol{\xi}_{t+1} = \boldsymbol{\Phi} \boldsymbol{\xi}_t + \boldsymbol{\varepsilon}_{t+1} \quad (1.6)$$

$$\text{Observation equation} \quad \mathbf{y}_t = \boldsymbol{\Psi}_t \boldsymbol{\xi}_t + \mathbf{b}_t + \boldsymbol{\nu}_t \quad (1.7)$$

where $\boldsymbol{\xi}_t$ is the state vector containing unobserved variables and $\mathbf{y}_t$ is a vector of observables. $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}_t$, and $\mathbf{b}_t$ are matrices of parameters. The $\boldsymbol{\varepsilon}_{t+1}$ and $\boldsymbol{\nu}_{t+1}$ are vectors of white noise:

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s'] = \begin{cases} \mathbf{Q} & \text{for } t = s \\ \mathbf{0} & \text{otherwise} \end{cases} \quad E[\boldsymbol{\nu}_t \boldsymbol{\nu}_s'] = \begin{cases} \mathbf{R} & \text{for } t = s \\ \mathbf{0} & \text{otherwise} \end{cases}$$

$$E[\boldsymbol{\varepsilon}_t \boldsymbol{\nu}_s'] = \mathbf{0} \quad \text{for all } t \text{ and } s.$$

The state equation describes the dynamics of unobserved variables and the observation equation relates observed data with unobserved variables contained in the state vector.

To carry out an empirical estimation, we need to make use of the Kalman filter due to the existence of unobservables. The Kalman filter is an algorithm to recursively estimate the state vector at time t given an initial condition and observations up to time t . Intuitively, the Kalman filter delivers the optimal estimator of unobserved components as each new piece of information becomes available.⁴ When all observations have been processed, it yields the optimal estimator of the state vector that is consistent with observed realizations, given a model of interactions between observed and unobserved variables. This estimator contains all the information necessary to make optimal predictions of future values of the state and the observations.

Given parameter values and an initial condition, the Kalman filter gives us the sequence of optimal estimates of the state vector, $\{\hat{\boldsymbol{\xi}}_{t+1|t}\}_{t=1}^T$, and the associated sequence of MSEs, $\{\mathbf{P}_{t+1|t}\}_{t=1}^T$, where

$$\begin{aligned} \hat{\boldsymbol{\xi}}_{t+1|t} &= E[\boldsymbol{\xi}_{t+1} | \Omega_t], \\ \mathbf{P}_{t+1|t} &= E\left[\left(\boldsymbol{\xi}_{t+1|t} - \hat{\boldsymbol{\xi}}_{t+1|t}\right)\left(\boldsymbol{\xi}_{t+1|t} - \hat{\boldsymbol{\xi}}_{t+1|t}\right)'\right], \\ \Omega_t &\equiv [\mathbf{y}_1', \mathbf{y}_2', \dots, \mathbf{y}_t']'. \end{aligned}$$

We will maximize the log-likelihood function. At each iteration of numerical maximization, the Kalman filter is invoked and the log-likelihood function is evaluated based on $\{\hat{\boldsymbol{\xi}}_{t+1|t}\}_{t=1}^T$ and $\{\mathbf{P}_{t+1|t}\}_{t=1}^T$, given parameter values at each iteration.

If there are missing observation, we just need to either exclude those missing observations from a vector of observables (i.e., the dimension of $\mathbf{y}_t$ varies over time) or assign zeros for them. Even when we introduce missing observations in the Kalman filter, they do not affect the derivation and validity of the Kalman filter. The procedure of computing the gradient of the log-likelihood function and the information matrix remains unaffected as well.⁵

⁴Derivation of the Kalman filter is well described in Hamilton (1994, Ch.13). Key equations are reproduced in Appendix.

⁵See Harvey (1989) for more details. The procedure and formulas are presented in Appendix.

Specifications in the State Equation

Let us define

$$\mathbf{k}_t = [\Delta w_{1,t} \quad \cdots \quad \Delta w_{N,t} \quad \Delta \mathbf{x}'_t]', \quad (1.8)$$

where $\Delta w_{i,t}$ is the unobserved quarterly TFP for the industry i at time t , and $\mathbf{x}_t$ is a vector of macroeconomic variables that will be specified below. As will be discussed later, it is necessary to include the current value and six lags of the unobserved technological growth at a quarterly frequency into the state vector $\boldsymbol{\xi}_t$. Then the state vector can be represented as

$$\boldsymbol{\xi}_{t+1} = [\mathbf{k}'_{t+1} \quad \mathbf{k}'_t \quad \cdots \quad \mathbf{k}'_{t-5} \quad 1]'. \quad (1.9)$$

We could include a constant as an exogenous variable in the state equation. Instead, we include 1 in the last element of the state vector because it is more convenient in order to incorporate a constant term in the dynamics. We assume that $\mathbf{k}_{t+1}$ follows VAR(4):

$$\mathbf{k}_{t+1} = \mathbf{c} + \Phi_1 \mathbf{k}_t + \Phi_2 \mathbf{k}_{t-1} + \Phi_3 \mathbf{k}_{t-2} + \Phi_4 \mathbf{k}_{t-3} + \tilde{\boldsymbol{\varepsilon}}_{t+1}, \quad (1.10)$$

where

$$\Phi_j = \begin{bmatrix} \rho_{11}^{(j)} & \cdots & \rho_{1N}^{(j)} & \mathbf{g}_1^{(j)'} \\ \vdots & \ddots & \vdots & \vdots \\ \rho_{N1}^{(j)} & \cdots & \rho_{NN}^{(j)} & \mathbf{g}_N^{(j)'} \\ \mathbf{h}_1^{(j)} & \cdots & \mathbf{h}_N^{(j)} & \mathbf{f}^{(j)} \end{bmatrix} \quad \text{for } j = 1, \dots, 4,$$

$E[\tilde{\boldsymbol{\varepsilon}}_{t+1} \tilde{\boldsymbol{\varepsilon}}'_{t+1}] = \mathbf{Q}$, and $\mathbf{Q}$ is assumed to be diagonal.

We are interested in whether macroeconomic variables help predict the behavior of the unobserved quarterly technological progress. That is, we will test the null hypothesis that $\mathbf{g}_i^{(1)} = \cdots = \mathbf{g}_i^{(4)} = \mathbf{0}$ for the industry i .

The remaining components of the state equation are just identities, such as

$$\mathbf{k}_t = \mathbf{k}_t, \quad \cdots, \quad \mathbf{k}_{t-5} = \mathbf{k}_{t-5}, \quad 1 = 1. \quad (1.11)$$

So the corresponding parts of disturbances are set to be zero.

Specifications in the Observation Equation

Industry-level TFP In order to carry out the Granger-causality test for disaggregated industries at a quarterly frequency, we relate the observed annual industry-level TFP growth with the unobserved quarterly TFP growth rates. It is worthwhile to consider

the relation between the annual technological growth and the quarterly technological growth because it will determine the specification of the observation equation.

There are at least two possible ways of measuring the annual technology level using quarterly series. One is to observe the annual technology level at a particular point in time (e.g., in the fourth quarter). The other one is to measure the annual technology level as the average of the quarterly technology levels.

We assume that annual observations of the technology level are available in the fourth quarter of each year. Let $\{w_{i,t}\}_{t=1}^T$ and $\{z_{i,t/4}\}_{t/4=1}^{T/4}$ be sequences of log of quarterly and annual technology levels in industry i , respectively. The time subscript t is a quarterly time index. If the annual technology level is measured, say, in the fourth quarter, we have $z_{i,t/4} = w_{i,t}$ and thus the annual TFP growth is expressed as $\Delta z_{i,t/4} \equiv \Delta w_{i,t}^* = w_{i,t} - w_{i,t-4}$. On the other hand, if we assume that the annual technology level is the average of quarterly technology levels, then

$$z_{i,t/4} = \frac{1}{4} \{w_{i,t} + w_{i,t-1} + w_{i,t-2} + w_{i,t-3}\}$$

and hence

$$\begin{aligned} \Delta z_{i,t/4} &= \frac{1}{4} \{ \Delta w_{i,t}^* + \Delta w_{i,t-1}^* + \Delta w_{i,t-2}^* + \Delta w_{i,t-3}^* \} \\ &= \frac{1}{4} \{ \Delta w_{i,t} + 2\Delta w_{i,t-1} + 3\Delta w_{i,t-2} + \\ &\quad 4\Delta w_{i,t-3} + 3\Delta w_{i,t-4} + 2\Delta w_{i,t-5} + \Delta w_{i,t-6} \}. \end{aligned} \quad (1.12)$$

Obviously the former specification may be subject to various factors. For example, if there is the Christmas effect on the industry-level TFP, as reported in Barsky and Miron (1989), observing the technology level in the fourth quarter does not reflect true level of technology.

Notice that since (1.12) is an identity, we do not include either a constant term or an error term. Furthermore, since $\Delta z_{i,t/4}$ is the annual TFP growth observed in the fourth quarter, (1.12) holds only in the fourth quarter. The vector of observables $\mathbf{y}_t$ does not contain $\Delta z_{i,t/4}$ in other quarters, and hence the dimension of $\mathbf{y}_t$, $\mathbf{\Psi}_t$, $\mathbf{b}_t$, and $\mathbf{\nu}_t$ depends on time. Let N be the number of industries and let n and m denote the number of quarterly observable series and the size of the state vector. Then in the fourth quarter, $\mathbf{y}_t$, $\mathbf{b}_t$, and $\mathbf{\nu}_t$ are $((n + N) \times 1)$ vectors and $\mathbf{\Psi}_t$ is an $((n + N) \times m)$ matrix. In other quarters, the dimension of $\mathbf{y}_t$, $\mathbf{b}_t$, and $\mathbf{\nu}_t$ is $(n \times 1)$ and $\mathbf{\Psi}_t$ becomes an $(n \times m)$ matrix.

Aggregate Solow residual As described above, the aggregate Solow residual is the weighted sum of the industry-level TFP with the Domar weight. It is likely that the conventionally derived aggregate Solow residual differs from the measure of aggregate technological progress based on the TFP at the industry level. It is because the conventionally derived aggregate Solow residual is based on the assumption that the aggregate value-added production function exists and there will be aggregation errors as well. We still believe that the conventionally derived Solow residual will give us some useful signals for unobserved quarterly TFP growth. Hence, we model the relation between the aggregate Solow residual and the industry-level TFPs by including a constant term and an aggregation error term as follows:

$$\Delta a_t = \sum_{i=1}^N d_i \Delta w_{i,t} + b + \nu_t, \quad (1.13)$$

where Δa_t is the growth rate of the conventionally derived Solow residual, and $d_i = \frac{P_i^Q Q_i}{P^Q Y}$ is the Domar weight.

Macroeconomic variables We observe macroeconomic variables every quarter. Thus, corresponding observations are nothing but identities.

$$\Delta \mathbf{x}_t = \Delta \mathbf{x}_t \quad (1.14)$$

This means that a part of components in the state vector is perfectly observable at a quarterly frequency.

Matrix Form

Equations (1.12)-(1.14) form our observation equation, $\mathbf{y}_t = \boldsymbol{\Psi}_t \boldsymbol{\xi}_t + \mathbf{b}_t + \boldsymbol{\nu}_t$. To summarize, in the matrix form:

$$\mathbf{y}_t = \begin{cases} [\Delta a_t \quad \Delta \mathbf{x}'_t]' & \text{if } t = \text{Q1, Q2, or Q3} \\ [\Delta a_t \quad \Delta \mathbf{x}'_t \quad \Delta z_{1,t/4} \quad \cdots \quad \Delta z_{N,t/4}]' & \text{if } t = \text{Q4} \end{cases} \quad (1.15)$$

$$\Psi_t = \begin{cases} \begin{bmatrix} \psi'_1 & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & 0 \\ \Psi_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 \end{bmatrix} & \text{if } t = \text{Q1, Q2, or Q3} \\ \begin{bmatrix} \psi'_1 & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & \mathbf{0}' & 0 \\ \Psi_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 0 \\ \frac{1}{4}\Psi_3 & \frac{1}{2}\Psi_3 & \frac{3}{4}\Psi_3 & \Psi_3 & \frac{3}{4}\Psi_3 & \frac{1}{2}\Psi_3 & \frac{1}{4}\Psi_3 & 0 \end{bmatrix} & \text{if } t = \text{Q4} \end{cases} \quad (1.16)$$

$$\mathbf{b}_t = \begin{cases} \begin{bmatrix} b & \mathbf{0}' \end{bmatrix}' & \text{if } t = \text{Q1, Q2, or Q3} \\ \begin{bmatrix} b & \mathbf{0}' & 0 & \dots & 0 \end{bmatrix}' & \text{if } t = \text{Q4} \end{cases} \quad (1.17)$$

$$\boldsymbol{\nu}_t = \begin{cases} \begin{bmatrix} \nu_t & \mathbf{0} \end{bmatrix}' & \text{if } t = \text{Q1, Q2, or Q3} \\ \begin{bmatrix} \nu_t & \mathbf{0} & 0 & \dots & 0 \end{bmatrix}' & \text{if } t = \text{Q4} \end{cases} \quad (1.18)$$

where

$$\boldsymbol{\psi}'_1 = [d_1 \quad \dots \quad d_N \quad \mathbf{0}] \quad (1.19)$$

$$\Psi_2 = [\mathbf{0}' \quad \dots \quad \mathbf{0}' \quad \mathbf{I}] \quad (1.20)$$

$$\Psi_3 = \begin{bmatrix} 1 & & 0 & 0 \\ & \ddots & & 0 \\ 0 & & 1 & 0 \end{bmatrix} \quad (1.21)$$

$$E[\boldsymbol{\nu}_t \boldsymbol{\nu}_t'] = \mathbf{R}_t = \begin{cases} \begin{bmatrix} r & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \text{if } t = \text{Q1, Q2, or Q3} \\ \begin{bmatrix} r & \mathbf{0}' & \mathbf{0}' \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} & \text{if } t = \text{Q4} \end{cases} \quad (1.22)$$

For the state equation, equations (1.9)-(1.11) characterize the dynamic behavior of the state variables. In the matrix form, the state equation (1.6) can be expressed as:

$$\begin{bmatrix} k_{t+1} \\ k_t \\ k_{t-1} \\ k_{t-2} \\ k_{t-3} \\ k_{t-4} \\ k_{t-5} \\ 1 \end{bmatrix} = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \Phi_4 & 0 & 0 & 0 & c \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ 0' & 0' & 0' & 0' & 0' & 0' & 0' & 1 \end{bmatrix} \begin{bmatrix} k_t \\ k_{t-1} \\ k_{t-2} \\ k_{t-3} \\ k_{t-4} \\ k_{t-5} \\ k_{t-6} \\ 1 \end{bmatrix} + \begin{bmatrix} \tilde{\varepsilon}_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (1.23)$$

where

$$E[\varepsilon_{t+1}\varepsilon'_{t+1}] = \begin{bmatrix} Q & 0 & \dots & 0 & 0 \\ 0 & 0 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & & & 0 & 0 \\ 0' & \dots & \dots & 0' & 0 \end{bmatrix} \quad (1.24)$$

Q is assumed to be diagonal.

1.2.3 Data

For the industry-level data, we use the data set compiled by Dale Jorgenson and his colleagues. The data set consists of the industry-level production inputs and outputs as well as corresponding price indices.⁶ It covers 35 industries of the U.S. private economy, which corresponds to one-digit non-manufacturing industries and two-digit manufacturing industries, from 1958 to 1996.

One advantage of Jorgenson's data set is that variables in the data set are constructed such that they are suitable for economic analysis. For example, labor inputs are constructed from the *Census of Population Survey*, the *Current Population Survey*, and the National Income and Product Account (NIPA). Using data from those primary sources, they construct labor inputs (hours worked) by taking account of different types of workers, such as sex, class of workers (employee, self-employed, or unpaid), age, and education. Capital inputs are also carefully constructed in a similar fashion, by classifying capital inputs into very detailed categories.⁷

⁶ Available at <http://post.economics.harvard.edu/faculty/jorgenson/data/35klem.html>

⁷ More detailed description of the data set can be found in Jorgenson, Gollap, and Fraumeni (1987), Jorgenson (1990), and Jorgenson and Stiroh (2000).

Table 1.1: Granger-causality Test for 1960:I - 1996:IV

Variables	W	df	p -value
M1	7.799	4	0.099
T-Bill	18.474	4	0.001
Gov	2.236	4	0.693
Oil	1.195	4	0.879

Although there are a couple of alternative data sets, there are some shortcomings for our purpose. For example, although it has longer sample periods, the Multifactor Productivity database from the BLS does not cover the entire U.S. economy at the disaggregated level.⁸ Another industry-level data set covering the entire economy is available from the BEA. The BEA data set contains gross output, intermediate inputs, nominal factor shares, and so on. Necessary series are available only from 1977, however. Furthermore, there is a discontinuity in terms of industry classification due to the revision in the SIC code. For these reasons, we will employ Jorgenson's data set.

For macroeconomic variables, all variables used here follow the description in Evans (1992).

1.3 Empirical Results

We start off our empirical analysis by updating the Granger-causality test in Evans (1992). Table 1.1 reports the Granger-causality test results for the sample period 1960:I - 1996:IV. With an extended sample period, the predictive power of macroeconomic variables declines. In his study, only oil price fails to Granger-cause the aggregate Solow residual. On the contrary, in this sample period, money and government expenditure no longer Granger-cause the Solow residual at the 5% level. The predictive power of the interest rate still remains statistically significant at the 1% level. Variance decomposition results are shown in Figure 1.1. While Evans (1992) reports that about 70% of the variance for 16-quarter-ahead forecasting error for the Solow residual can be attributable to its own innovations, in this extended sample, own innovations to the Solow residual account for about 75% of the variance. If we look at the percentage

⁸Although NBER Industry Database can be another candidate, it covers only manufacturing industries as well.

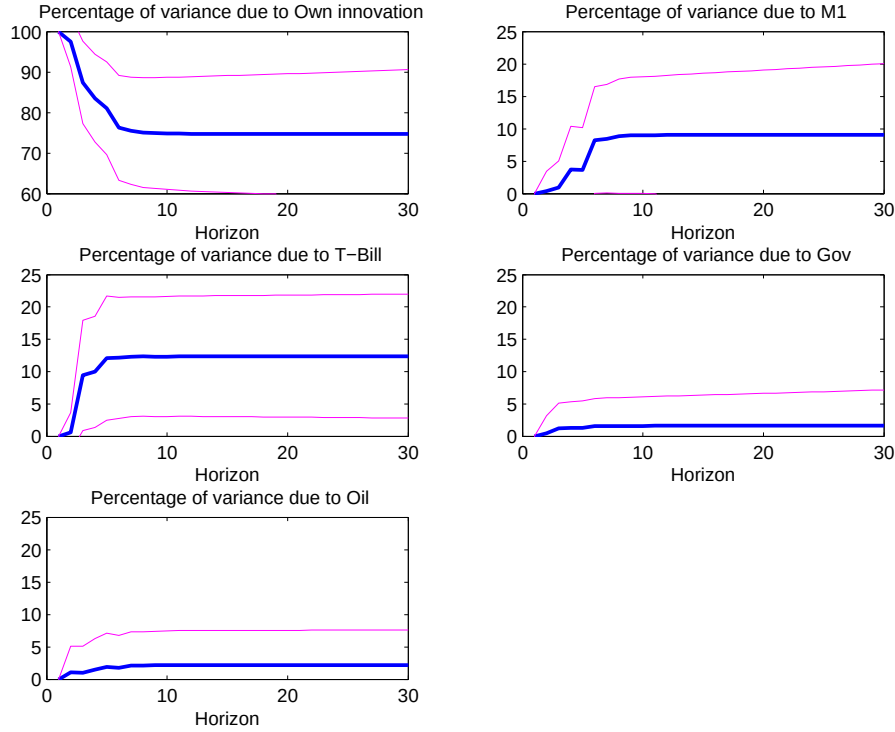


Figure 1.1: Decomposition of variance in the Solow residual

Note: Following Evans (1992), the order of orthogonalization is as follows: Solow residual, M1, T-Bill, Oil, and Gov. Thick lines represent point estimates and thin lines are corresponding two-standard-error band computed by Monte Carlo with 1000 repetitions.

of forecasting variance explained by individual variables, the effect of the interest rate appears to be quantitatively significant as well.

For the reason described above, we will focus on the predictability of the interest rate on the aggregate Solow residual and try to see the source of the predictability at the disaggregated level. Thus, a vector of macroeconomic variables $\Delta \mathbf{x}_t$ in (1.8) becomes just a scalar, changes in 3-month T-Bill rates. Focusing on the interest rate also helps handle disaggregated data. As the level of disaggregation increases, it will be difficult to include all variables used in Evans (1992) since the number of parameters and the dimension of the state vector increase dramatically.

We will first decompose the aggregate Solow residual into two industries (i.e., $N = 2$), namely non-manufacturing and manufacturing. Figure 1.2 compares the impulse response function from the 2-industry model with the one from a bivariate VAR(5) of the aggregate Solow residual and interest rate. The lag length of VAR is chosen based on

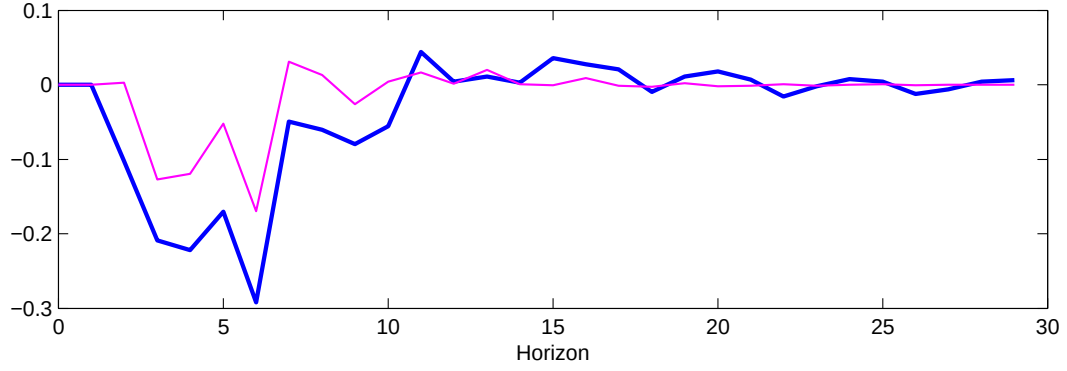


Figure 1.2: Comparison of the Impulse Response Functions

Note: The thick line is the impulse response of the aggregate Solow residual to one unit interest rate innovation obtained from the 2-industry model. The thin line represents the impulse response function based on bivariate VAR(5) of the aggregate Solow residual and interest rate. The lag length is supported by LR criterion, AIC, and BIC.

the LR criterion, AIC, and BIC. They are responses to one unit interest rate innovation. As you can see, the impulse response from the state space model successfully mimics the one obtained from VAR. This suggests that our framework that deals with missing observations and contemporaneous and temporal aggregation using the state space model may help identify the source of the interest rate predictive power.

The upper panel of Table 1.2 shows the maximum log-likelihood values and the likelihood ratio test results for the 2-industry model. The null hypothesis is that $g_i^{(1)} = \dots = g_i^{(4)} = 0$, for the industry i . The LR test indicates that the interest rate does not have any predictive power for non-manufacturing industry, which is a majority of the economy. Surprisingly, including the interest rate variable into the non-manufacturing equation does not help improve the log-likelihood value at all. The null hypothesis is accepted at the 5% level. On the other hand, the unobserved quarterly TFP for manufacturing industry is Granger-caused by the interest rate. The null hypothesis is rejected even at the 1% level of significance.

Given this result, we will further explore details in manufacturing industries. For this reason, we split the manufacturing industry into two groups, non-durable and durable manufacturing industries. Now a vector of unobserved quarterly TFP, $\Delta \mathbf{w}_t$, becomes a (3×1) vector. Coefficients matrices in the model should also be expanded accordingly.

Table 1.2: Likelihood Ratio Test Results for 2- and 3-industry Model

	LL Value	d.f.	LR Stat
2-industry Model			
Unrestricted	-454.27		
Restricted Non-manufacturing	-454.96	4	1.38
Restricted Manufacturing	-463.80	4	19.06**
3-industry Model			
Unrestricted	-497.78		
Restricted Non-durable	-508.54	4	21.52**
Restricted Durable	-508.47	4	21.38**

Note: The 5% level critical value for $\chi^2(4) = 9.488$. The 1% level critical value for $\chi^2(4) = 13.277$. ** indicates rejecting the null hypothesis of no Granger-causality at the 1% level.

Let $\Delta w_{2,t}$ denote the unobserved quarterly TFP for manufacturing industry, and also let $\Delta w_{(2,1),t}$ and $\Delta w_{(2,2),t}$ be the unobserved quarterly TFP for non-durable and durable industries, respectively. According to the discrete version of (1.4), we have

$$\Delta w_{2,t} = \gamma_1 \Delta w_{(2,1),t} + \gamma_2 \Delta w_{(2,2),t}, \quad (1.25)$$

where γ_1 and γ_2 are gross output share of non-durable and durable in manufacturing industry, respectively. Given the estimated parameter values in the 2-industry model, we can infer elements in Φ_i that correspond to non-manufacturing and interest rate equations by using the relation in (1.25). Given the estimates for Φ_j in the 2-industry model,

$$\hat{\Phi}_{j,2\text{ind}} = \begin{bmatrix} \hat{\rho}_{11}^{(j)} & \hat{\rho}_{12}^{(j)} & \hat{g}_1^{(j)} \\ \hat{\rho}_{21}^{(j)} & \hat{\rho}_{22}^{(j)} & \hat{g}_2^{(j)} \\ \hat{h}_1^{(j)} & \hat{h}_2^{(j)} & \hat{f}^{(j)} \end{bmatrix},$$

we will get coefficient matrices Φ_j for the 3-industry model as

$$\Phi_{j,3\text{ind}} = \begin{bmatrix} \hat{\rho}_{11}^{(j)} & \gamma_1 \hat{\rho}_{12}^{(j)} & \gamma_2 \hat{\rho}_{12}^{(j)} & \hat{g}_1^{(j)} \\ \rho_{1,(2,1)}^{(j)} & \rho_{(2,1),(2,1)}^{(j)} & \rho_{(2,1),(2,2)}^{(j)} & g_{(2,1)}^{(j)} \\ \rho_{1,(2,2)}^{(j)} & \rho_{(2,2),(2,1)}^{(j)} & \rho_{(2,2),(2,2)}^{(j)} & g_{(2,2)}^{(j)} \\ \hat{h}_1^{(j)} & \gamma_1 \hat{h}_2^{(j)} & \gamma_2 \hat{h}_2^{(j)} & \hat{f}^{(j)} \end{bmatrix}. \quad (1.26)$$

Parameters in the shaded area in (1.26) are unknown. Since our purpose is to investigate the source of what happens at the higher level of aggregation with more disaggregated

Table 1.3: Subcategories of Manufacturing Industries

Non-durable Manufacturing		Durable Manufacturing	
Consumer Products	Food and kindred products	Consumer Products	Furniture and fixtures
	Tobacco		Misc. manufacturing
	Textile mill products	Materials	Lumber and wood
	Apparel		Stone, clay, glass
	Paper and allied		Primary metal
	Printing, publishing and allied		Fabricated metal
Materials	Leather	Processing and Assembly	Machinery, non-electrical
	Chemicals		Electrical machinery
	Petroleum and coal products		Motor vehicles
	Rubber and misc plastics		Transp. equip. & ordnance
			Instruments

data, for comparability of results we just estimate those parameters in the shaded region by fixing other parameters at the estimates from the 2-industry model. This also helps us deal with the increasing number of coefficients with more disaggregated data.

The lower panel of Table 1.2 presents the maximum log-likelihood values and the associated LR test statistics for the 3-industry model. The LR test results suggest that the interest rate does Granger-cause the industry-level quarterly TFP in both non-durable and durable manufacturing industries. The null hypothesis of no Granger-causality is rejected at the 1% level for both industries.

Now we will take a detailed look at both non-durable and durable manufacturing industries. To do so, we will categorize them into three subgroups based on characteristics of those individual industries. This step is necessary because there are many individual industries in both non-durable and durable industries (10 and 11 subindustries, respectively). If we move from non-durable or durable to individual industries, we will face the large number of parameters to be estimated at once. In order to manage such circumstances, we need to introduce an intermediate step.

Three subcategories in manufacturing industry are consumer products manufacturing, materials producing manufacturing, and processing and assembly manufacturing. The detailed classification of individual industries is shown in Table 1.3.

Based on the subcategories described above, we assume additional restrictions

Table 1.4: Wald Test Results for Subcategories in Manufacturing

Non-durable Manufacturing		Durable Manufacturing	
Consumer Products	8.536	Consumer Products	17.086**
Materials	34.179**	Materials	10.883*
		Assembly & Processing	25.433**

Note: Numbers reported are the Wald test statistics. The 5% level critical value for $\chi^2(4) = 9.488$. The 1% level critical value for $\chi^2(4) = 13.277$. ** indicates rejecting the null hypothesis of no Granger-causality at the 1% level.

on coefficients to be estimated. These restrictions are about feedbacks from other groups of industries. At this level of disaggregation, some groups of industries are irrelevant to other groups of industries in terms of production structure. If that is the case, TFP for other industries should not help predict TFP for a particular industry. For instance, none of the manufacturing industries uses tobacco products as material inputs, and thus technological progress in the tobacco industry should not help forecast TFPs in other industries. Such a priori restrictions are derived from the sparsity of the *use* table of the input-output accounts. If usage of outputs from a particular industry as intermediate inputs is less than 5% of total material inputs, we set the corresponding coefficient of feedback term to be zero.⁹ These a priori restrictions on feedback terms help reduce the number of parameters to be estimated, especially at the very disaggregated level.

We will test non-durable and durable industries one by one. As before, we estimate parameters related to subcategories of interest, but leave other parameters fixed at the estimates from the less disaggregated model. Suppose, for example, we take subcategories of non-durable manufacturing. Then coefficients in the equations describing the behavior of TFP in non-manufacturing and durable manufacturing are fixed at the estimates from the 3-industry model, and vice versa.

Table 1.4 summarizes the Wald test results for subcategories of non-durable and durable industries. The Wald test for the null hypothesis of no Granger-causality suggests that we fail to reject the null only for non-durable consumer products manufacturing. Other subcategories of manufacturing industries continue to show the evidence of the Granger-causality. Except for materials producing durable industry, the null hypothesis is rejected at the 1% level.

⁹These restrictions are similar to the idea of Shea (1993). He uses the sparsity of the IO table in order to search for instrument variables at disaggregated industries. We use 1977 benchmark IO table to determine the sparsity.

Table 1.5: Wald Test Results for Detailed Manufacturing Industries

Non-durable - Materials	
Chemicals	25.024**
Petroleum & Coal Products	85.801**
Rubber & Misc Plastics	20.451**
Durable - Consumer Products	
Furniture & Fixture	22.477**
Misc. Manufacturing	21.773**
Durable - Materials	
Lumber & Wood	25.853**
Stone, Clay, Glass	17.213**
Primary Metal	33.351**
Fabricated Metal	32.061**
Durable - Processing & Assembly	
Non-electrical Machinery	8.107
Electrical Machinery	4.320
Motor vehicles	13.830**
Transportation Equipment	20.837**
Instruments	13.587**

Note: Numbers reported are the Wald test statistics. The 5% level critical value for $\chi^2(4) = 9.488$. The 1% level critical value for $\chi^2(4) = 13.277$. ** indicates rejecting the null hypothesis of no Granger-causality at the 5% and 1% level.

Finally, we will turn to analyzing the Granger-causality found at the highly aggregated level with very disaggregated data. We take non-durable materials producing and three subcategories of durable manufacturing industries, and decompose each subindustry one by one. We will look at the model in which one of those subcategories is broken down into individual industries. For instance, if we focus on the non-durable material producing industries, industries included in the model will be non-manufacturing, non-durable consumer products producing, durable manufacturing, and those individual industries. As described above, we impose a priori restrictions on coefficients.

A summary of the Wald test results appears in Table 1.5. As we can see, the interest rate Granger-causes the unobserved quarterly TFP for most of individual industries, except for both non-electric and electric machinery. The null hypothesis is

rejected at the 1% level significance.

1.4 Extension

In this section, we will explore one of possibilities that interest rates can Granger-cause the aggregate Solow residual. Although several explanations have been offered, we will restrict our attention to a possibility of variable capital utilization. In Jorgenson's data set, the capital utilization rate is assumed to be constant. In the presence of variable capital utilization rate, instead of (1.3), the true technological progress can be expressed as:

$$\frac{dZ_i}{Z_i} = \frac{dQ_i}{Q_i} - \left\{ c^K \left(\frac{dK_i}{K_i} + \frac{dU_i}{U_i} \right) + c^L \frac{dL_i}{L_i} + c^E \frac{dE_i}{E_i} + c^M \frac{dM_i}{M_i} \right\},$$

where K_i represents the physical stock of capital, U_i is the rate of capital utilization, and $U_i K_i$ is assumed to provide service flow of capital. However, according to the growth accounting, we calculate technological progress as $\frac{dQ_i}{Q_i} - \left\{ c^K \frac{dK_i}{K_i} + c^L \frac{dL_i}{L_i} + c^E \frac{dE_i}{E_i} + c^M \frac{dM_i}{M_i} \right\}$. This means that the observed measure of technological progress is contaminated by growth rate of capital utilization.

Although the hypothesis of variable capital utilization is quite realistic, it is difficult to implement it into an empirical estimation. It is because we do not have an appropriate series for capital utilization. In general, the Federal Reserve's capacity utilization series are conceptually misleading as a proxy for capital utilization as argued by Shapiro (1989, 1996). So several researchers suggest to employ the use of materials (Basu, 1996) or the use of energy (Burnside, Eichenbaum, and Rebelo, 1995) as a proxy for capital utilization, by assuming low substitution between value added and those variables used as proxies. In stead, Shapiro (1996) advocates using the workweek of capital in order to proxy for capital utilization. Based on the Census Bureau's *Survey of Plant Capacity*, Beaulieu and Matthey (1998) construct the workweek of capital series that is publicly available. However, their data set covers only from 1974 to 1992 at an annual frequency.

Beaulieu and Matthey (1998) convincingly argue that for a group of industries that can be characterized as continuous processing industries, the FRB's measure of capacity utilization is likely to be better than the workweek of capital as a proxy for capital utilization.¹⁰ The reason is that for those continuous processing industries, almost

¹⁰Their argument is based on the classification done by Matthey and Strongin (1997). The US four-

all of the variation in capital utilization arises from variation in the speed or intensity of capital use, not the duration of operations, which is captured by the workweek of capital.

Among those individual industries that appear to show the evidence of the Granger-causality in Section 1.3, chemicals and petroleum and coal products fall into the category of the continuous-processing industries and corresponding the FRB's capacity utilization series are available from 1948.¹¹ Thus, we will take these two industries and see whether taking account of unobserved capital utilization based on the FRB's capacity utilization eliminates the sign of the Granger-causality. Since we are only interested in either chemicals or petroleum and coal products, we combine other industries together. Thus the economy consists of chemicals (or petroleum products) and the rest of industries.

Although the capacity utilization series can serve as a proxy for capital utilization, there is another issue that may affect our estimation. As Shapiro (1989) argues, there might be a substantial error in the capacity utilization data that comes from interpolating the annual data on capacity. To cope with this issue, we assume that the capacity utilization series is subject to measurement error in log-level. That is,

$$u_{i,t}^* = u_{i,t} + m_{i,t}$$

where $u_{i,t}^*$ is the observed log-level of utilization of the industry i , $u_{i,t}$ is the true log-level of utilization, and $m_{i,t}$ represents a white noise measurement error. This leads to an additional observation equation

$$\Delta u_{i,t}^* = \Delta u_{i,t} + m_{i,t} - m_{i,t-1}.$$

As stated earlier, the observed TFP is a linear combination of the true technological progress and the growth rate of unobserved capital utilization. Thus the obser-

digit manufacturing industries are categorized into three groups. *Variable work period group*, where plants such as automakers tend to vary output through adjusting the workweek; *Continuous-processors*, where manufactures vary their output by adjusting its speed of throughput; *Other manufactures*, where producers change their output rate through adjusting other inputs such as momentary labor input.

¹¹Percentages of four-digit continuous-processing industries in chemicals and petroleum products are 80.6% and 93.7%, respectively.

Table 1.6: Estimation Results with Capital Utilization

		LL Value	d.f.	LR Stat
Chemicals				
Unrestricted		-640.759		
Restricted	$\alpha = \beta = \mathbf{0}$	-691.296	7	101.074**
	$g_2 = \mathbf{0}$	-644.426	4	7.334
Petroleum and coal products				
Unrestricted		-604.059		
Restricted	$\alpha = \beta = \mathbf{0}$	-661.780	7	115.442**
	$g_2 = \mathbf{0}$	-606.503	4	4.888

Note: α is a vector of feedback coefficients from $\Delta u_{2,t}$ and β represents a vector of feedback coefficients from other variables to the capital utilization equation. g is a vector that collects coefficients on the interest rate for the chemical (or petroleum) equation. The 5% level critical values for $\chi^2(7) = 14.067$ and for $\chi^2(4) = 9.488$. The 1% level critical value for $\chi^2(7) = 18.475$ and $\chi^2(4) = 13.277$.

** indicates the rejection of the null hypothesis at the 1% level.

vation equation for the industry-level TFP should be modified as follows:

$$\begin{aligned} \Delta z_{i,t/4} = & \frac{1}{4} \{ \Delta w_{i,t} + 2\Delta w_{i,t-1} + 3\Delta w_{i,t-2} + \\ & 4\Delta w_{i,t-3} + 3\Delta w_{i,t-4} + 2\Delta w_{i,t-5} + \Delta w_{i,t-6} \} + \\ & \frac{c^K}{4} \{ \Delta u_{i,t} + 2\Delta u_{i,t-1} + 3\Delta u_{i,t-2} + \\ & 4\Delta u_{i,t-3} + 3\Delta u_{i,t-4} + 2\Delta u_{i,t-5} + \Delta u_{i,t-6} \}. \end{aligned}$$

Now, instead of (1.13), the aggregate Solow residual can be represented as:

$$\Delta a_t = d_1 \Delta w_{1,t} + d_2 (\Delta w_{2,t} + c^K \Delta u_{2,t}) + b + \nu_t.$$

where the subscript $i = 1$ represents the rest of industries and the subscript $i = 2$ corresponds to either chemicals or petroleum and coal products. Thus $\mathbf{y}_t$ contains Δa_t , Δx_t , $\Delta u_{2,t}^*$ and $\Delta \mathbf{z}_{t/4}$ in the fourth quarter and, as before, $\Delta \mathbf{z}_{t/4}$ does not enter $\mathbf{y}_t$ in other quarters. $\mathbf{k}_t$ now includes two additional components, $\Delta u_{2,t}$ and m_t . The remaining parts of the state space model are unchanged.

Table 1.6 presents the estimated results with the variable capital utilization for chemicals and petroleum and coal products.

Since we have an additional observed series, including the capital utilization into the system does not nest the previous specifications shown in Section 1.3. In order

to see whether incorporating the capital utilization rate matters, first we compare the unrestricted model with the restricted model where we set all feedbacks from the capital utilization to the equations for other variables and all feedbacks from other variables to the equation for the capital utilization rate. That is, now the state-space model consists of two independent sets of equations, equations that inherit the previous specification and an equation of an AR process for the capital utilization rate. The first LR statistic in each panel of Table 1.6 shows strong evidences that the capital utilization is important to explain behaviors of unobserved TFP growth and the interest rate. For both industries, the null hypothesis of no effects is rejected at the 1% significance level.

Now we will look at whether the predictive power of the interest rate vanishes when we control for the unobserved capital utilization. The LR test results (the second LR statistic in each panel of Table 1.6) reveal that once we account for variable capital utilization with an appropriate measure, the quarterly unobserved TFP in chemicals and petroleum and coal products industries are no longer Granger-caused by the interest rate variable. We fail to reject the null hypothesis of no Granger-causality at the 5% level of significance. This is consistent with the findings from Paquet and Robidoux (2001). Although they work with aggregate data in the Canadian economy, they report that once capital utilization is taken into consideration, the capital-utilization-adjusted Solow residual is no longer Granger-cause by a set of monetary variables.¹²

It may sound obvious, but using an appropriate measure of capital utilization is important. We also experiment with growth rate of capacity utilization for manufacturing industry, whose correlation with the growth rate of workweek of capital is 0.49, according to Beaulieu and Matthey (1998). Incorporating inappropriate measure of capital utilization results in rejection of the null hypothesis of no Granger causality from the interest rate to TPF.

1.5 Conclusion

We have been analyzing the predictive power of the interest rate for the aggregate Solow residual. Although the industry-level data is only available at an annual frequency, by using the state space model and the Kalman filter, we are able to perform

¹²Paquet and Robidoux (2001) use the capacity utilization series published by the Statistics Canada to capture the effect of variable capital utilization. They claim that it is conceptually superior to counterparts in other countries.

the Granger-causality test on the disaggregated industry TFP at a quarterly frequency. Our approach will avoid problems arising from either converting high frequency data into low frequency data or interpolating low frequency observations when variables of our interest are available at different frequencies.

In terms of the forecasting power of the interest rate on the industry-level TFP growth, heterogeneity exists among industries. Furthermore, in most cases, the predictive power of the interest rate, if any, is statistically very significant at the 1% level.

Our results suggest that manufacturing industries are responsible for the non-exogenous nature of the aggregate Solow residual reported by Evans (1992). From the 2-industry model, we have found that while the interest rate helps predict TFP growth in manufacturing industry, productivity growth in non-manufacturing industry is not Granger-caused by the interest rate. Although TFP growth in non-durable manufacturing industry as a whole appears to be Granger-caused by the interest rate, exploring the details reveals that the interest rate plays an important role only in three industries that characterized as materials producing industries, such as chemicals, petroleum and coal products, and rubber and misc plastics. Contrary to the findings in non-durable manufacturing, vast majority of durable manufacturing industries shows the evidence of the Granger causality, except for non-electrical machinery and electrical machinery.

We also extend our analysis by taking account of variable capital utilization. Two case studies on chemical and petroleum and coal products, for which we have an appropriate measure of capital utilization, show that once we take account of variable capital utilization, the TFPs in those industries are no longer Granger-caused by the interest rate. This result suggests that the predictive power of the interest rate will disappear if we control for variable capital utilization with an appropriate measure.

Chapter 2

Declining Effects of Oil-price Shocks

2.1 Introduction

Macroeconomic consequences of large increases in the price of oil have been of great concern among economists and policy makers, as well as the general public since two major oil-price shocks hit the economy in the 1970's. In recent years, however, it seems that the effect of oil-price shocks has been decreasing. For example, the Congressional Budget Office reports that:

Contrary to general expectations, the large and persistent rise in energy prices that has occurred over the past two and a half years has not caused substantial problems for the overall U.S. economy. Although many households have had trouble adjusting to the higher prices, the effects on the nation's gross domestic product (GDP), employment, and inflation have thus far been moderate. (Congressional Budget Office, 2006, p.VII)

It is well known that, after World War II, nine out of ten recessions in the US economy were preceded by large increases in the oil price. A general perception is that oil-price increases may cause a prolonged and deep recession. In fact, a number of studies have tested the relation between economic activity and oil prices and have confirmed that it is not a statistical coincidence.¹ There have been substantial hikes in the oil price in

¹Those studies include Hamilton (1983), Burbidge and Harrison (1984), Gisser and Goodwin (1986), Daniel (1997), Carruth, Hooker, and Oswald (1998), and Hamilton (2003). However, Bernanke, Gertler, and Watson (1997) argue that oil-price shocks were not a major cause of economic downturns. See also subsequent discussions in Hamilton and Herrera (2004) and Bernanke, Gertler, and Watson (2004).

the early 2000's. A couple of events, such as the civil unrest in Venezuela in 2002, the Iraq war of 2003, and Hurricanes Katrina and Rita in 2005, triggered increases in the oil price. During 2006, the nominal oil price has reached value more than three times as high as the average price in the 1990's. However, we have not observed any evidence of economic downturn yet, as of the writing of the paper.²

In order to see the possible changes in the US economy over time, Figure 2.1 plots different behavior of macroeconomics variables after a 10% increase in the net oil price increase (NOPI) of Hamilton (1996, 2003). These impulse response functions describe average behavior following a 10% increase in oil prices above their previous 3-year high. Sample periods are split into two sub-samples at 1984, following McConnell and Perez-Quiros (2000). As can be seen clearly, we can observe different patterns in response to the oil-price shock, measured by an increase in the NOPI, in two sub-samples.³ Not only output but also other variables display weaker responses to the oil-price shock in the post-1984 period. This is very robust to different specifications. Furthermore, the peak of the output responses appears earlier in the post-1984 sub-sample than the pre-1984 sub-sample. On average, the output responses peak at 5 quarters after the shock with a 1.76% decline compared to the unshocked path in the pre-1984 sub-sample, whereas, in the post-1984 periods, the responses peak at 2 quarters after the shock only with a 0.35% drop from the unshocked path. The peak response of output declines by 80% and the peak is shifted by 3 quarters. This suggests that there have been some changes in the economy.

This paper takes the evidence presented in Figure 2.1 seriously and explores possible explanations why we expect to observe the weaker responses of macroeconomic variables to the same size of the oil-price shock. Most of existing studies have focused mainly on recessionary consequences of oil-price shocks (e.g., Kim and Loungani, 1991; Rotemberg and Woodford, 1996; Finn, 2000; Leduc and Sill, 2004; Carlstrom and Fuerst, 2006; Aguiar-Conraria and Wen, 2007; Cavallo and Wu, 2007) and have not addressed the issue of weaker effects of oil-price shocks. The only exception is Blanchard and Galí (2007), who ask the same question as this paper does. In order to answer the question,

Barsky and Kilian (2002, 2004) also raise a skeptical view about the role of the oil-price shocks in recessions in the 1970's.

²The GDP-based recession probability index of Chauvet and Hamilton (2006) indicates 7.7% probability of recession as of the third quarter of 2007.

³Early studies, such as Mork (1989) and Hooker (1996), also document that, as the sample period is extended, the relation between oil prices and macroeconomy is weakened.

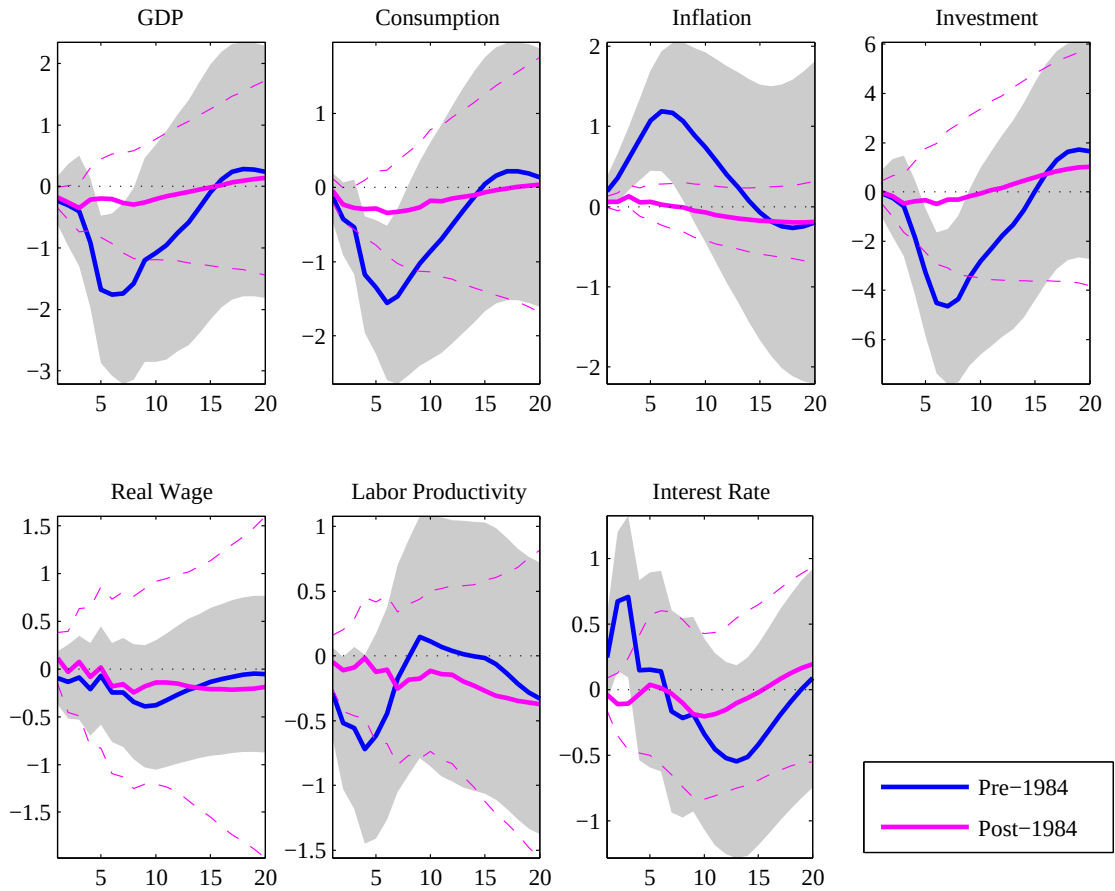


Figure 2.1: Different Responses to the Oil-price Shock: Before and After 1984

Note: Each panel plots responses to a 10% increase in NOPI. These impulse responses are obtained from 8-variable VAR with 4 lags, which is similar to the specification in Christiano, Eichenbaum, and Evans (2005). Variables included are NOPI, real GDP, real consumption, GDP deflator, real investment, real wage, labor productivity, and interest rate. A constant term is included as well. Data are quarterly. Except for NOPI, all variables are logged and retrieved from the FRED database of the FRB St. Louis. Horizontal axes indicate quarters and the vertical axes take percentage deviations from the unshocked path. Shaded areas represent 95% confidence bands for the Pre-1984 IRFs, while dashed lines show those for the Post-1984 IRFs.

they investigate four different hypotheses; (a) good luck (i.e., lack of concurrent adverse shocks), (b) smaller share of oil in production, (c) more flexible labor markets, and (d) improvements in monetary policy. They conclude that all four factors have played an important role in accounting for the mild effects on inflation and economic activity.

Besides the hypotheses Blanchard and Galí (2007) have considered, there are also other possibilities for the reason why we might expect the effect of the oil-price shock on economic activity has become weaker. In this paper, I will analyze three possible explanations for the declining effect of oil-price shocks in detail. The first factor investigated is the outcome of deregulation, especially in the US transportation industry. When we consider the macroeconomic consequences of an oil-price shock, we usually have the industrial/commercial sector in mind in terms of petroleum consumption. It is always helpful to look at data first to see how petroleum is used across sectors. Figure 2.2 shows petroleum consumption in the US economy from 1949 to 2005 obtained from the Energy Information Administration (EIA). It reveals that contrary to common treatment in the existing models, more than a half of the US petroleum consumption is coming from the transportation sector. Thus, it suggests that it is important to pay more attention to the role of petroleum consumption in transportation. Commercial transportation had been one of the most regulated industries in the US economy. This regulatory environment had created inefficiencies and economic rents, and deregulation resulted in lower costs and a more competitive environment. In this paper, in addition to the typical industrial sector, I explicitly model the role of the transportation industry in order to account for the effect of the deregulation in the transportation industry that began in 1980.

Another possibility I will examine is the effect of improvements in energy efficiency and hence less dependence of the US economy on petroleum as a source of energy. After the two oil-price shocks in the 1970's, technological advances enabled us to utilize petroleum more efficiently. For example, vehicles' miles per gallon improved dramatically. According to the Energy Information Administration (2006), from 1973 to 1991, miles per gallon (all motor vehicles) improved by 42%. This more efficient use of oil is apparent in terms of the oil expenditure share in value added. While average oil expenditure share relative to GDP was 3.65% in the pre-1984 periods, the average share declines in the post-1984 periods to 2.75%. This point is similar to one of factors considered in Blanchard and Galí (2007). Since there exist two types of oil usage in the model, industrial sector and transportation sector, it is possible to analyze the consequence of

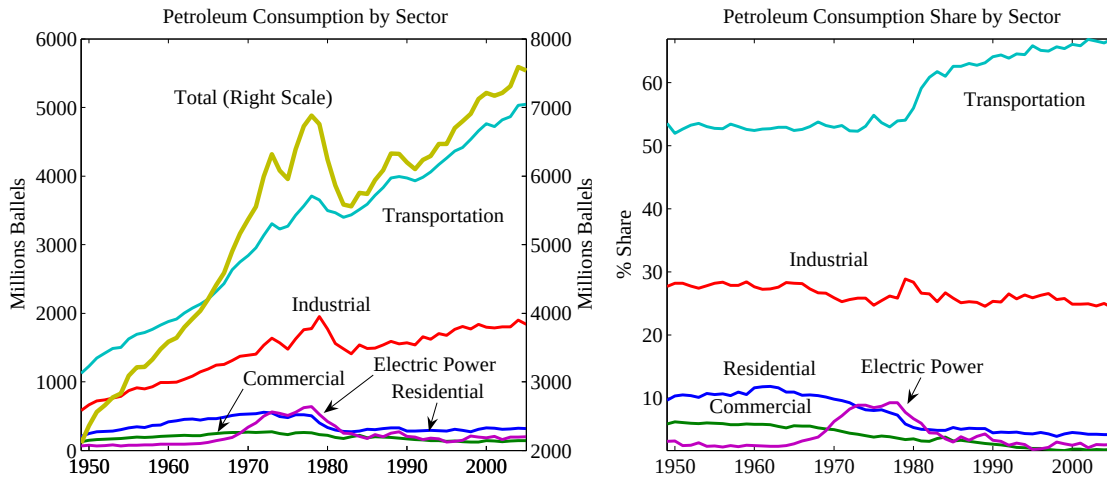


Figure 2.2: Petroleum Consumption by Sector 1949–2005

Sources: Department of Energy, Energy Information Administration (2006), *Annual Energy Review 2005*, Table 5.13a–5.13d.

improved energy efficiency in more detail. In this sense, this paper is complimentary to Blanchard and Galí (2007).

Lastly, I also consider a possible effect of changes in persistence of the oil-price shock. Small changes in the underlying shocks process could alter the responses of macroeconomic variables. Figure 2.3 plots real price of oil measures by the PPI of crude petroleum deflated by GDP deflator. Judging from data, it is clear that the oil-price behavior shows different patterns after the mid 1980's. Particularly, it tends to be less persistent compared to the near-unit-root nature of the oil price we have observed up until the mid 1980's. The goal of the paper is to investigate the role of these three factors in accounting for the two changes observed in the time series data, that is, weaker responses of output and the earlier timing of the peak.

The model is based on a standard model with adjustment costs. I extend the standard sticky-price model to include the role of oil prices into the economy in two ways. As in Finn (2000), variable capital utilization tied to petroleum consumption is introduced in the intermediate-good production process. Another extension is to include a transportation sector as an additional sector in the economy. It is assumed that all produced intermediate products must be delivered to customers by using transportation services. Including transportation firms amplifies the effects of the oil-price shock, because, as will be described in Section 2.2, transportation firms had significant market power before the deregulation. It should be noted that, although there are empirical

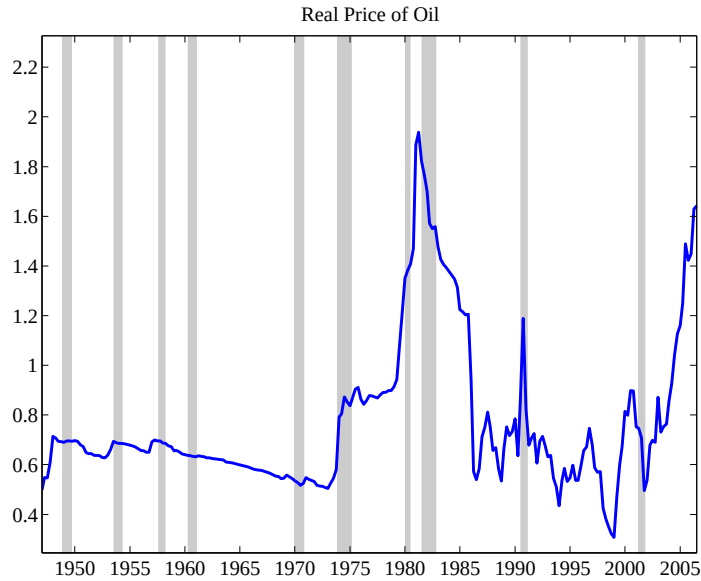


Figure 2.3: Real Price of Oil

studies documenting possible asymmetric effects on the economic activity (see, for example, Davis and Haltiwanger, 2001; Hamilton, 2003)⁴, like other existing models, the structure of the model considered here implies that increases and decreases in the price of oil would have symmetric effects. However, the focus of this paper is on the recessionary consequences following a large increase in the price of oil and on accounting for the changes in the economy over time.

The baseline specification of the model, which corresponds to the pre-1984 state of the economy, well captures important aspects of the economy's response to the oil-price shock, under reasonable parameter values. In terms of accounting for the changes in the economy overtime, each factor individually could account for about a 20-25% decline in the peak response of aggregate value added. Combined together, the three factors result in a 51% reduction in the peak response and shift the timing of the peak by 2 quarters, compared with the baseline specification.

The rest of the paper is organized as follows. Section 2.2 provides some background information on regulation in the US transportation industry, as well as petroleum consumption of the transportation sector. Section 2.3 presents a model that we can use for assessing the three possible factors. Results are in Section 2.4 and Section 2.5 con-

⁴Although the asymmetric effects of oil-price shocks are widely believed, Kilian (2007) summarizes supporting evidence for the symmetric responses of monthly real consumption expenditure based on formal statistical tests and suggests that nonresidential investment could be a source of the asymmetry.

cludes.

2.2 Transportation Industry Background

2.2.1 Petroleum Consumption for Transportation

As mentioned in the Introduction, most of US petroleum consumption comes from the transportation sector. One caveat is that the EIA's definition of the transportation sector includes the supply-side petroleum consumption as well as petroleum use of households for driving automobiles. The EIA does not have data on how much petroleum can be attributed to fuel of household vehicles except for selected years. Based on the information in the EIA's "*Household Vehicles Energy Use: Latest Data & Trends*"⁵, about 52% of petroleum is used by household vehicles in the transportation sector after the late 1980's. Therefore, even when we take account of household vehicle use, the transportation service industry itself consumes as much oil as the whole industrial sector does.

Figure 2.4 displays the share of oil consumption in the transportation sector by mode. As can be seen clearly, within the transportation industry, the trucking industry is a major source of oil consumption. Any investigation of the recessionary consequences of oil-price shocks needs to pay particular attention to the trucking industry.

2.2.2 Regulation in the Transportation Industry

The U.S. transportation industry in the 1970's was heavily regulated by the Interstate Commerce Commission (ICC), which was established by the Interstate Commerce Act of 1887. The ICC regulated prices and competition in interstate transportation. Starting from railroads, the control of the ICC was expanded to a variety of transportation modes, such as trucking, water transportation, oil pipeline, and freight forwarders. The trucking industry is likely to be a good example of a perfectly competitive industry, if there is no regulation in the industry (Olson, 1972). Such a regulatory environment in the transportation industry imposed by the ICC had reduced competition and created significant inefficiency. However, the regulatory environment came to an end in 1980. The Motor Carriers Act of 1980 and the Staggers Rail Act of 1980 trig-

⁵Available at http://www.eia.doe.gov/emeu/rtecs/nhts_survey/2001/.

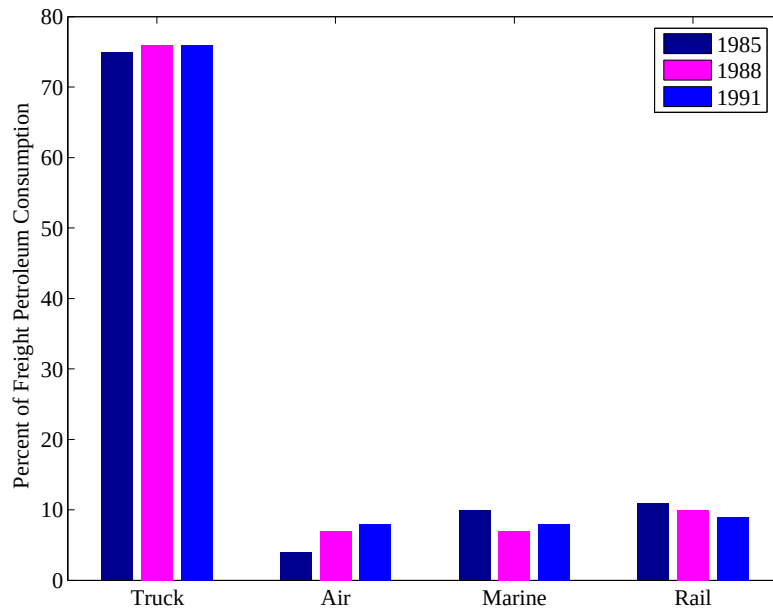


Figure 2.4: Petroleum Consumption in the Transportation Industry by Mode

Source: Department of Energy, Energy Information Administration (1995), *Measuring Energy Efficiency In The United States' Economy: A Beginning*, Figure 5.17

gered deregulation in the transportation industry. Throughout the 1980's, the control of the ICC had been diminishing. Ultimately, all of regulations were removed in 1995 when the ICC was abolished.

Regulations by the ICC primarily consisted of restrictions on operating rights and a requirement of filing price changes. The ICC controlled new entries by granting operating certificates, which specified a scope of operation, such as a list of routes, commodities, and areas that carriers may serve. A new certificate was issued if it can be proven that the new service contributed to the public convenience and necessity and that incumbent carriers were unable to provide the service. New applications were frequently protested by incumbent carriers. This procedure was also applicable to incumbents expanding their scope of operations, such as more variety of commodities and operation routes. Such a difficulty led many firms to purchase the operating authorities of other carriers. The market value of those operating rights could be a measure of the monopolistic rents. In fact, those operating rights were listed as assets on regulated firms' books. Snow and Sobotka (1977) report that the market value of the operating rights amounted to 15-20% of gross annual revenues. The ICC's control of operating rights created not only market power, but also inefficiencies. Restrictions on routes and products often

resulted in empty backhauls. For example, based on the 1976 ICC study, Wright (1983) argues that about 38% of interstate and 71% of intrastate backhauls were empty.

In addition to the entry control, regulated carriers were required to file all changes in their rates with the ICC, thirty days prior to the new rate becoming effective. Within thirty days, the proposal of a new price was subject to objections from anyone, including the ICC itself and competitors (both those in the same transportation mode and from different types of carriers). Once the proposal was protested, the ICC would investigate its reasonableness. Ratemaking and filing to the ICC were usually done in a collective manner, via regionally distributed rate bureaus. Those rate bureaus approved by the ICC had been operated under antitrust immunity since 1948 (the Reed-Bulwinkle Act). Although the ICC required rate bureaus to be open and democratic, investigations done by the U.S. congress and the ICC revealed that “these rate bureaus are, for most part, dominated by small groups of large carriers” (U.S. Congress, Senate, 1980, p. 60). Due to limited resources of the ICC reviewing proposed tariffs,⁶ virtually it was the case that “[t]he carriers set their own rate” (U.S. Congress, Senate, 1980, p. 80).

The rate bureau system ensured that protestants could impose substantial costs on applicants trying to underprice other carriers, and hence created an incentive not to lower prices. An individual carrier could propose selective rate changes. Such proposals were usually forwarded to the standing rate committee, which set up a public hearing on the proposal. Any interested parties could participate, regardless of whether they were directly affected by the proposal. Frequently, those proposals were opposed by other carriers, sometimes causing the ICC to investigate the proposed rate, and making rate changes extremely time-consuming and difficult. The ICC might then suspend the rate and call for a formal investigation of its legality. It was possible for an individual carrier to file an independent action to the ICC without going through the review at the rate bureau. However, most of rate bureaus required the independent action to be informed before taking effect and other carriers had an option of whether to join in the new rate or not. Thus, the rate bureau system under the ICC regime created an incentive not to lower prices and discouraged competitive pricing in the industry.

These regulations and controls by the ICC had created oligopolistic environment in the transportation industry. For example, Winston (1981) estimates the markup over

⁶For example, according to the report from U.S. Congress, Senate (1980), 278,039 tariff publications were filed to the ICC in 1978. On average, 5,000 pages of tariffs are filed daily. However, there were only 60 employees at the ICC who were in charge of checking those tariffs.

marginal cost for various commodity categories. Among regulated commodities, the estimated markup rates range from 29% to 64% (with the mean of 52.2%) for railroads and those for trucking range from 26% to 49% (with the mean of 37.1%).

The deregulation resulted in freer entries and more flexible expansions of the scope of business. The Motor Carriers Act of 1980 eliminated entry barriers. It required the ICC to grant operating rights to any requesting firm that is able to provide the service. The burden of proof was reversed. Although, prior to the Motor Carriers Act of 1980, new entrants had been required to prove the necessity of new service, opponents were asked to provide evidence why the new service was not beneficial. Furthermore, route restrictions were eliminated as well. By 1986, the number of motor carriers holding ICC operating certificates was more than doubled compared with 1980 (Winston, Corsi, Grimm, and Evans, 1990, p.12).

The Motor Carriers Act of 1980 also increased price flexibility. It guaranteed individual motor carriers and freight forwarders complete freedom to change rates as long as price changes are within the $\pm 10\%$ range compared with the level one year prior to a specified point of reference. The ICC could increase the range up to 15% for rate increases in any year that it finds that there is sufficient competition to regulate rates, and that the carriers, shippers and the public would derive a benefit from the increased flexibility. The law prohibited the use of the rate bureau for single-line rate making. Although the rate bureaus themselves were not abolished, expanded independent actions increased pricing flexibility and reduced importance of the rate bureau system.

Overall, the net effect of the deregulation can be summarized as higher efficiency and more competition, which led to lower average costs. Ying and Keeler (1991) estimate the effect of deregulation on rates charged by motor carriers, controlling for various factors, such as changes in primary factor prices and fuel costs. They find that, on average, prices decreased by 15-20% within three years after the deregulation and that prices further dropped by 25-35% by 1985. Nebesky, McMullen, and Lee (1995) estimate and compare the markup rate of the less-than-truckload segment of the trucking industry before and after the deregulation. They find that the estimated markup has declined dramatically after the deregulation. While the estimated markup rate using a 1977 sample is 35.5%, the estimate based on 1988 data is 1.9%.

In terms of the changes in the economy over time, I will interpret the consequences of the deregulation as (i) lower markup due to the more competitive environment

and (ii) improved energy efficiency in the transportation industry. The following section will present the model, which incorporates the role of transportation in the economy.

2.3 The Model

The economy consists of continuums of households, intermediate-good firms, and transportation firms, which are indexed by $h \in [0, 1]$, $i \in [0, 1]$, and $n \in [0, 1]$, respectively, as well as the monetary authority and the aggregator of final-good and labor services.

Each household supplies different types of labor services, which are imperfect substitutes, in a monopolistically competitive market. There exist nominal wage rigidities in the form of quadratic wage adjustment costs.

While intermediate-good firms produce distinct products and sell their output in monopolistically competitive markets, they act as a price-taker in factor markets. Production of intermediate goods requires two additional inputs, other than capital and labor. First, I assume that all produced goods must be delivered to consumers and intermediate-good producers incur the cost of transportation services. Second, the firms also use raw materials as one of factor inputs. The raw materials are assumed to be composed of all intermediate goods. That is, produced intermediate goods will be used to assemble the final good for households' consumption as well as used as materials inputs for the production process.

Transportation firms supply distinct transportation services to intermediate-good firms in monopolistically competitive markets. Their services are distinct in the sense that there exist segmentations in the US transportation industry, such as operating region, products carried, and transportation modes.

There are two direct avenues whereby an increase in the price of oil could have impact on the economic activity. One is variable capital utilization tied with use of energy, as in Finn (2000), Leduc and Sill (2004), and Cavallo and Wu (2007). This is intended here to capture the role of the industrial sector in US petroleum consumption mentioned earlier. The other channel through which the oil-price shock may have an effect in the economy is direct impact on the cost of transportation. In the subsections below, we shall see the details on each sector.

2.3.1 Aggregators

There are two types of aggregators in the economy. The final-good aggregator purchases differentiated goods from intermediate-good firms and assembles the final good, which is sold to households for consumption in a competitive market. The labor-service aggregator hires differentiated labor services from households and transforms them into the composite labor input, which is demanded by both intermediate-good firms and transportation firms in a competitive market.

Both aggregation technologies are based on a typical constant returns to scale CES technology. The final good and the composite labor are produced by

$$Y_t = \left[\int_0^1 Y_t(i)^{(\theta-1)/\theta} di \right]^{\theta/(\theta-1)},$$

$$N_t = \left[\int_0^1 N_t(h)^{(\theta_N-1)/\theta_N} dh \right]^{\theta_N/(\theta_N-1)},$$

respectively, where $Y_t(i)$ represents the gross output of intermediate-good firm i , $N_t(h)$ is the labor service of household h , and $\theta > 1$ and $\theta_N > 1$ are the elasticity of substitution between the different inputs.

Taking the price of inputs as given, the aggregators minimize total costs. The first-order conditions give us constant-elasticity inverse demand functions for intermediate products and labor services

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\theta} Y_t,$$

$$N_t(h) = \left(\frac{w_t(h)}{w_t} \right)^{-\theta_N} N_t,$$

respectively, where $P_t(i)$ is the price of intermediate good i , $w_t(h)$ is real wage for labor services supplied by household h , $P_t = \left[\int_0^1 P_t(i)^{1-\theta} di \right]^{1/(1-\theta)}$ is the aggregate price index, and $w_t = \left[\int_0^1 w_t(h)^{1-\theta_N} dh \right]^{1/(1-\theta_N)}$ is the aggregate wage index.

2.3.2 Households

There is a continuum of households indexed by $h \in [0, 1]$. Households are identical except for the heterogeneity of labor. Each household maximizes the expected utility

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \log(C_t(h)) + \frac{(M_t(h)/P_t)^{1-\sigma}}{1-\sigma} - \frac{N_t(h)^{1+\varphi}}{1+\varphi} \right\},$$

where $0 < \beta < 1$ is the subjective discount factor, $C_t(h)$ is consumption of final goods, $M_t(h)/P_t$ is real money balances, and $N_t(h)$ is hours worked.

Each household carries $M_{t-1}(h)$ units of money and $B_{t-1}(h)$ bonds into period t . In addition to the carry-over, the household receives a lump-sum transfer $\mathcal{T}_t(h)$ from the monetary authority at the beginning of each period. After the maturity of bonds, the household also receives $B_{t-1}(h)$ additional units of money. During the period t , the household provides labor in exchange for labor income. The household also receives nominal dividends from intermediate-good producers and transportation firms, denoted by Π_t and Π_t^T , respectively.

In addition, each household is a monopolistic supplier of differentiated labor services $N_t(h)$ and faces a downward-sloping labor demand curve. Each household sets its own nominal wage subject to quadratic wage adjustment costs, as in Kim (2000). The wage adjustment cost function takes the form

$$AC_t^w(h) = \frac{\phi_w}{2} \left(\frac{W_t(h)}{W_{t-1}(h)} - \pi_s \right)^2 \frac{W_t(h)}{P_t},$$

where $\phi_w > 0$ is a wage adjustment cost parameter, $W_t(h)$ is household h 's nominal wage, π_s is the steady-state rate of inflation, and P_t is the general price index. Thus, household h faces the following budget constraint:

$$\begin{aligned} \frac{M_{t-1}(h) + \mathcal{T}_t(h) + B_{t-1}(h) + W_t(h)N_t(h) + \Pi_t(h) + \Pi_t^T(h)}{P_t} \\ = C_t(h) + \frac{B_t(h)/r_t + M_t(h)}{P_t} + AC_t^w(h), \end{aligned} \quad (2.1)$$

where r_t is the gross nominal interest rate.

2.3.3 Intermediate-good Firms

There is a continuum of intermediate-good producers indexed by $i \in [0, 1]$. Each intermediate firm produces a distinct product $\tilde{Y}_t(i)$ and sells it in a monopolistically competitive market. Unlike typical models, I assume that all goods produced must be delivered to customers and each firm incurs transportation costs. Specifically, each intermediate-good producer must purchase transportation services $T_t(i)$ in order to finalize its production process. We assume that the whole production process is described by

$$\tilde{Y}_t(i) = \min \left[\frac{Q_t(i)}{1 - \chi_Y}, \frac{T_t(i)}{\chi_Y} \right], \quad (2.2)$$

where $\chi_Y \in (0, 1)$ is a technology parameter, which determines complementarity of transportation services, $Q_t(i)$ is firm i 's gross output, and $T_t(i)$ is firm i 's demand for transportation services, which will be defined below.

Following Rotemberg and Woodford (1995) and Basu (1996), we assume that production of gross output involves use of material inputs through the fixed-coefficient technology between value added $V_t(i)$ and material inputs $X_t(i)$, given by

$$Q_t(i) = \min \left[\frac{V_t(i)}{1 - \chi_Q}, \frac{X_t(i)}{\chi_Q} \right], \quad (2.3)$$

where $\chi_Q \in (0, 1)$ is a share of materials cost in the value of gross output. Each firm produces its value added $V_t(i)$ by using capital services $u_t(i)K_t(i)$ and labor input $N_t^I(i)$ with a typical Cobb-Douglas production technology

$$V_t(i) = [u_t(i)K_t(i)]^\alpha [N_t^I(i)]^{1-\alpha}, \quad (2.4)$$

where $u_t(i)$ is capital utilization and $K_t(i)$ is firm i 's beginning-of-period stock of capital.

The material inputs $X_t(i)$ are the CES aggregate of output produced by all intermediate firms, including itself. Here, we assume that $X_t(i)$ is defined by

$$X_t(i) = \left[\int_0^1 X_t(i, j)^{(\theta-1)/\theta} dj \right]^{\theta/(\theta-1)},$$

where $X_t(i, j)$ is firm i 's demand for the product j . Notice that the CES technology is exactly the same as the final-good aggregator uses. Thus, it is equivalent to use the final good as material inputs. The first-order condition for the cost-minimization problem of total material costs implies that firm i 's inverse conditional demand function for intermediate product j is given by

$$X_t(i, j) = \left(\frac{P_t^X(j)}{P_t^X} \right)^{-\theta} X_t(i),$$

where $P_t^X(j)$ is the price of intermediate product j and $P_t^X = \left[\int_0^1 P_t(j)^{1-\theta} dj \right]^{1/(1-\theta)}$. The aggregate demand for firm j 's product as materials inputs is then given by

$$X_t(j) = \int_0^1 X_t(i, j) di = \left(\frac{P_t^X(j)}{P_t^X} \right)^{-\theta} X_t,$$

where $X_t = \int_0^1 X_t(i) di$.

Similarly, $T_t(i)$ is the CES aggregate of transportation services supplied by all transportation firms, that is,

$$T_t(i) = \left[\int_0^1 T_t(i, j)^{(\theta_T-1)/\theta_T} dj \right]^{\theta_T/(\theta_T-1)},$$

where $T_t(i, j)$ is firm i 's demand for type- j transportation services and $\theta_T > 1$.

Each intermediate-good firm accumulates firm-specific capital stock according to the law of motion

$$K_{t+1}(i) = (1 - \delta_t(i))K_t(i) + I_t(i), \quad (2.5)$$

where $I_t(i)$ is firm i 's investment at time t and $\delta_t(i)$ is time-varying depreciation rate. As standard in the variable capital utilization literature (for example, Greenwood, Hercowitz, and Huffman, 1988; Burnside and Eichenbaum, 1996), the capital depreciation rate $\delta_t(i)$ depends on the rate of capital utilization through

$$\delta_t(i) = \frac{(u_t(i))^\kappa}{\kappa}, \quad (2.6)$$

with $\kappa > 1$. As in Finn (2000), Leduc and Sill (2004), and Cavallo and Wu (2007), we assume that capital utilization is tied with use of energy. Higher capital utilization must be accompanied by more petroleum consumption per unit of capital, at an increasing rate. The relationship is specified as

$$\frac{E_t^I(i)}{K_t(i)} = \frac{(u_t(i))^\nu}{\nu}, \quad (2.7)$$

where $E_t^I(i)$ represents firm i 's oil consumption for capital utilization and $\nu > 1$. In order for capital stock to be productive, each firm must purchase the necessary amount of petroleum from the oil producer at the nominal price of P_t^E .

With the production technology (2.2) and (2.3), the optimality condition implies that

$$\tilde{Y}_t(i) = \frac{Q_t(i)}{1 - \chi_Y} = \frac{T_t(i)}{\chi_Y} \quad \text{and} \quad Q_t(i) = \frac{V_t(i)}{1 - \chi_Q} = \frac{X_t(i)}{\chi_Q}. \quad (2.8)$$

Since each intermediate firm sells its distinct product to the final-good aggregator and other intermediate-good firms, it faces its own downward-sloping demand curve. The total demand firm i faces is given by

$$Y_t(i) + X_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\theta} Y_t + \left(\frac{P_t^X(i)}{P_t^X}\right)^{-\theta} X_t = \left(\frac{P_t(i)}{P_t}\right)^{-\theta} (Y_t + X_t) \quad (2.9)$$

The last equality is due to the fact that there are no distinctions between outputs sold to the final-good aggregator and those to intermediate-good firms as materials inputs, so that $P_t(i) = P_t^X(i)$ for all $i \in [0, 1]$ and hence $P_t = P_t^X$. Each firm will make a decision such that it can meet the total demand from the final-good aggregator and other intermediate-good firms.

As in Ireland (2001), each intermediate-good firm is subject to two types of quadratic adjustment costs for price and capital, which are respectively given by

$$AC_t^Y(i) = \frac{\phi_Y}{2} \left(\frac{P_t(i)}{\pi_s P_{t-1}(i)} - 1 \right)^2 Y_t, \quad (2.10)$$

$$AC_t^K(i) = \frac{\phi_K}{2} \left(\frac{K_{t+1}(i)}{K_t(i)} - 1 \right)^2 K_t(i), \quad (2.11)$$

where $\phi_Y > 0$ and $\phi_K > 0$ are adjustment cost parameters and π_s is the steady-state rate of inflation.

Given (2.4)–(2.11), each intermediate firm chooses $P_t(i)$, $u_t(i)$, $K_{t+1}(i)$ and $N_t^I(i)$ in order to maximize its discounted market value defined as

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t}{P_t} \Pi_t(i),$$

where Λ_t is the Lagrange multiplier associated with the household budget constraint (2.1),

$$\frac{\Pi_t}{P_t} = \frac{P_t(i)}{P_t} \tilde{Y}_t(i) - w_t N_t^I(i) - \frac{P_t^E}{P_t} E_t^I(i) - X_t(i) - \frac{P_t^T}{P_t} T_t(i) - I_t(i) - AC_t^K(i) - AC_t^Y(i),$$

P_t^E is the nominal price of oil, and P_t^T is the nominal price of transportation services.

2.3.4 Transportation Firms

There is a continuum of transportation firms indexed by $n \in [0, 1]$. Each transportation firm provides distinct transportation services $T_t(n)$ in a monopolistically competitive market. If there were no regulations, the transportation industry could be considered as a good example of competitive industry. However, due to the regulations described in Section 2.2, transportation firms are monopolistically competitive firms supplying distinctive services. For example, we can think of regional barriers in the industry as one of the sources for creating the distinctiveness.

Transportation services are produced by combining value added, which requires capital and labor, and oil with a fixed-coefficient technology

$$T_t(n) = \min [V_t^T(n), \omega E_t^T(n)], \quad (2.12)$$

where $V_t^T(n)$ is transportation firm n 's value added, which will be defined below, $E_t^T(n)$ is petroleum used for transportation, and $\omega > 0$ is a transportation energy efficiency parameter. Given (2.12), the optimality condition implies that

$$T_t(n) = V_t^T(n) = \omega E_t^T(n) \quad (2.13)$$

The value-added production function of transportation firm n is assumed to be

$$V_t^T(n) = \left[aA_t(n)^{(\gamma-1)/\gamma} + (1-a)N_t^T(n)^{(\gamma-1)/\gamma} \right]^{\gamma/(\gamma-1)}, \quad (2.14)$$

where $A_t(n)$ is transportation firm n 's beginning-of-period stock of transportation equipment, $N_t^T(n)$ is labor inputs, $a \in (0, 1)$, and γ is the elasticity of substitution between capital and labor. Unlike the value-added production function of intermediate-good firms, that of transportation firms takes the form of the CES production function. It is intended to take account of the possibility that transportation equipment and labor are more complements in value-added production of transportation with $\gamma < 1$, compared with that of intermediate-good firms. Such complementarity can arise due to the fact that operating transportation equipment requires certain amount of labor. I have also experimented with a typical Cobb-Douglas value-added production, but the quantitative results on accounting for the changes in the economy overtime are not significantly different.

Each transportation firm accumulates stock of transportation equipment following the law of motion

$$A_{t+1}(n) = (1 - \delta_A)A_t + I_t^A(n), \quad (2.15)$$

where δ_A is the time-invariant depreciation rate of $A_t(n)$ and $I_t^A(n)$ is transportation firm n 's investment. Like intermediate-good producers, each transportation firm is subject to capital adjustment costs, given by

$$AC_t^A(n) = \frac{\phi_A}{2} \left(\frac{A_{t+1}(n)}{A_t(n)} - 1 \right)^2 A_t(n), \quad (2.16)$$

with $\phi_A > 0$.

Transportation firm n faces its own downward-sloping demand curve, which is given by

$$T_t(n) = \left(\frac{P_t^T(n)}{P_t^T} \right)^{-\theta_T} T_t, \quad (2.17)$$

where $P_t^T(n)$ is the price of $T_t(n)$, $P_t^T = \left[\int_0^1 P_t^T(n)^{1-\theta_T} dn \right]^{1/(1-\theta_T)}$ and $T_t = \int_0^1 T_t(n) dn$.

Subject to (2.14) – (2.17), each transportation firm chooses $P_t^T(n)$, $N_t^T(n)$, and $A_{t+1}(n)$ to maximize its discounted market value

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t}{P_t} \Pi_t^T(n),$$

where

$$\frac{\Pi_t^T(n)}{P_t} = \frac{P_t^T(n)}{P_t} T_t(n) - w_t N_t^T(n) - \frac{P_t^E}{P_t} E_t^T(n) - I_t^A(n) - AC_t^A(n).$$

2.3.5 The Monetary Authority

Following Ireland (2001), it is assumed that the monetary authority will set short-term interest rate r_t following a Taylor-type policy rule

$$\log(r_t/r_s) = \rho_y \log(Y_t/Y_s) + \rho_\pi \log(\pi_t/\pi_s) + \rho_\mu \log(\mu_t/\mu_s) + \varepsilon_{r,t},$$

where $\mu_t = \frac{M_t}{M_{t-1}}$ and those variables with s -subscripts indicate the associated steady-state values. Unlike the original policy rule proposed by Taylor (1993), this policy rule allows monetary policy to respond to changes in money growth as well. Although it is well known that the variant of the original Taylor rule estimated using pre-1979 data results in indeterminacy (e.g., Clarida, Galí, and Gertler, 2000), Ireland (2001) notes that this specification makes model's equilibrium unique, even under the pre-1979 estimates.

2.3.6 Real Price of Oil

The log real price of oil follows an exogenous AR(1) process:

$$\log(p_t^E) = (1 - \rho_{p^E}) \log(p_s^E) + \rho_{p^E} \log(p_{t-1}^E) + \varepsilon_{p^E,t}, \quad (2.18)$$

where $p_t^E = P_t^E/P_t$, p_s^E represents the steady-state value the real price of oil, and $|\rho_{p^E}| < 1$.

2.3.7 Symmetric Equilibrium

In a symmetric equilibrium, all agents in each sector make the same decision. The corresponding values are denoted by variables without indices h , i , and n . For example, $N_t^I(n) = N_t^I$ and $N_t^T(n) = N_t^T$ for all $i \in [0, 1]$, $n \in [0, 1]$, and t . The market-clearing condition of the labor market implies that $N_t = N_t^I + N_t^T$. The market-clearing conditions for money and bonds, $M_t = M_{t-1} + \mathcal{T}_t$ and $B_t = B_{t-1} = 0$, have to hold as well.

In the symmetric equilibrium, the household budget constraint (2.1) can be expressed as

$$Y_t - p_t^E(E_t^I + E_t^T) = C_t + I_t + I_t^A + AC_t^W + AC_t^K + AC_t^Y + AC_t^A. \quad (2.19)$$

The equation (2.19) states that the sum of consumption, total investment, and total adjustment costs equal to real value added, which is the value of output produced net of all costs (including energy costs), except for primary-factor payments. Thus, an

appropriate measure of output in the model economy that corresponds to real GDP is real value added $Y_t - p_t^E(E_t^I + E_t^T)$, rather than Y_t .

The system of difference equations that characterize the symmetric equilibrium consists of the first-order conditions for the households, the intermediate-good firms, and the transportation firms, the laws of motion for capital, the monetary policy rule, and the oil-price process. A complete description of the system is summarized in the Appendix.

2.4 Results

This section presents the quantitative results and discusses how the model presented in Section 2.3 responds to an oil-price shock. In order to obtain the solution to the system, I log-linearize the equilibrium conditions of the model around its steady state and use the generalized Schur decomposition described by Klein (2000) and Sims (2002).⁷ Before tuning to showing the quantitative results, I describe the calibration procedure in the following subsection.

2.4.1 Calibration

In order to analyze changes in the economy over time, I need to set the baseline specification such that the model economy corresponds approximately to the state of the pre-1984 period. In order to do so, I use pre-1984 values for some of share parameters, such as the value-added oil expenditure shares. Furthermore, if available, I will utilize some of pre-1979 estimates for parameter values from Ireland (2001) as an approximation. Parameter values assigned in the baseline specification are summarized in Table 2.1.

Those parameter values drawn from the pre-1979 estimates of Ireland (2001) are a discount factor β , the steady-state rate of inflation π_s , the price adjustment cost parameter ϕ_Y and coefficients on the monetary policy rule, ρ_y , ρ_π , and ρ_μ . β is set equal to 0.9974. With the steady-state quarterly inflation rate $\pi_s = 1.0129$, the value of β implies that annualized interest rate is 6.3%.

Frisch labor supply elasticity is set to be unity. The value of σ is assumed to be 10, which is broadly consistent with various estimates presented in Christiano,

⁷This is implemented in DYNARE.

Table 2.1: Summary of Parameter Values in the Baseline Specification

Parameter	Values
β Discount factor	0.9974*
χ_Q Materials cost share	0.47324
χ_V Transportation share	0.0386
σ Money utility parameter	10
φ Frisch labor supply elasticity	1
α Capital share (intermediate-good firms)	0.36
a Capital share (transportation firms)	0.36
γ Transportation firms' elasticity of substitution between capital and labor	0.1
δ_A Depreciation rate for transportation firms	0.025
ϕ_A Capital adjustment cost parameter (transportation firms)	10
ϕ_K Capital adjustment cost parameter (intermediate-good firms)	10
ϕ_W Wage adjustment cost parameter	58
ϕ_Y Price adjustment cost parameter	72.01*
θ Demand elasticity for consumption goods	6
θ_N Demand elasticity for labor	6
θ_T Demand elasticity for transportation services	3.8169
ρ_{p^E} Persistence parameter of the oil price process	0.99
ρ_π Monetary policy coefficient on inflation rate	0.8617*
ρ_y Monetary policy coefficient on output	0.0499*
ρ_μ Monetary policy coefficient on money growth	0.7351*
p_s^E Steady-state real price of oil	1
π_s Steady-state inflation rate	1.0129*
δ_s Steady-state depreciation rate for intermediate-good firms	0.025
s_1 Steady-state value-added expenditure share of oil used for capital utilization	0.0123
s_2 Steady-state value-added expenditure share of oil used for transportation	0.0242
ω Energy efficiency parameter for transportation	3.0334
ν Energy efficiency parameter for capital utilization	1.7373
κ Elasticity of depreciation	1.06

Note: Parameter values with asterisks are those based on the pre-1979 estimates of Ireland (2001).

Eichenbaum, and Evans (2005) with different specification.⁸ Capital share parameters α and a are set equal to 0.36. The steady-state depreciation rate for intermediate-good firms δ_s and the time-invariant depreciation rate for transportation firms δ_A are both equal to 0.025. Capital adjustment cost parameter is assumed to be 10 for both intermediate-good firms and transportation firms. The wage adjustment cost parameter is 58.⁹ Demand elasticity parameters for the final good and for the index of labor services, θ and θ_N , respectively, equal to 6, which implies in the steady-state, price and wage markup rates are 20%.

For parameter values related to transportation firms, I set the demand elasticity parameter θ_T to be 3.8169 for the baseline specification, such that the implied markup is 35.5% and for the value after the deregulation in the transportation industry, I set it to 53.632, so that the implied markup becomes 1.9%. The markup rate is consistent with the estimates of Nebesky, McMullen, and Lee (1995).

Based on the industry-level data compiled by Dale Jorgenson and his colleagues,¹⁰ the average share of materials cost (excluding energy cost) in the value of gross output from 1958 to 1983 is 47.3%, and hence, the materials cost share χ_Q is set equal to 0.473.¹¹ The value of χ_Y can be pinned down by the steady-state share of the transportation industry in the total aggregate gross output from the optimality condition (2.8). The pre-1984 average share of the transportation industry in Jorgenson's data is 3.86%.¹² Thus, I set $\chi_Y = 0.0386$.

We set the steady-state value of the relative price of oil $p_s^E = 1$ and the AR(1) coefficient for the oil price process $\rho_{p^E} = 0.99$. This is obtained from estimating (2.18) using the pre-1984 data on log-real price of oil, which is measured as log of PPI for crude petroleum deflated by GDP deflator.

In terms of percentage deviations of the aggregate value added from the steady-

⁸Levin, Onatski, Williams, and Williams (2006) also have a similar estimate with Bayesian approach. Although their estimate is about 11. This difference in the value of σ does not alter the results presented below.

⁹Wage rigidities introduced by the quadratic wage adjustment cost will have the same log-linearized wage Phillips curve as Calvo-type wage stickiness. This value of wage adjustment cost parameter roughly corresponds to about 20% of households being able to reoptimize their wages in each period, given other parameter values used in the baseline specification.

¹⁰Data were downloaded from <http://post.economics.harvard.edu/faculty/jorgenson/data/35klem.html>. Description of the data set is found in Jorgenson, Gollap, and Fraumeni (1987), Jorgenson and Stiroh (2000), and Jorgenson (1990).

¹¹This number is close to the materials share used in Dotsey and King (2006). Although other authors such as Rotemberg and Woodford (1996), and Huang and Liu (2004), use larger value for the share of materials, this is primarily because they take account of the energy cost as well.

¹²Details on this calculation are given in Appendix.

state value, the oil expenditure shares in aggregate value added play an important role. According to the petroleum consumption data of the EIA and time series data on the price of West-Texas Intermediate from FRED of the FRB St. Louis, the average share of oil consumption in value added from 1949 to 1983 is 3.65%, and the average share after 1983 is 2.75%. According to the EIA's data, the average value of the ratio E^I/E^T is 0.51 in the pre-1984 period and 0.39 after 1984. Therefore, I set $s_1 = p_s^E E_s^I / \{Y_s - p_s^E (E_s^I + E_s^T)\} = 0.0123$ and $s_2 = p_s^E E_s^T / \{Y_s - p_s^E (E_s^I + E_s^T)\} = 0.0242$ for the baseline specification.

Together with other parameter values, these calibrating targets on oil expenditure shares are used to determine the remaining parameter values, which are related to energy efficiency in the model economy, ω , ν , and κ . From the steady-state relation, the value of ω is given by

$$\omega = \frac{(1 + s_1 + s_2) \left(1 + \frac{s_1}{s_2}\right) \chi_Y p_s^E}{(s_1 + s_2) \{1 - \chi_Q(1 - \chi_Y)\}}. \quad (2.20)$$

The values of ν and κ are jointly determined by solving the following non-linear system of equations:

$$\frac{s_1}{s_2} = \frac{(u_s)^\nu}{\nu} \left\{ \frac{1}{\beta} - (1 - \delta_s) + p_s^E \frac{(u_s)^\nu}{\nu} \right\}^{-1} \frac{\Xi_s^Y}{\Lambda_s} (1 - \chi_Q(1 - \chi_Y)) \frac{\alpha \chi_V \omega}{\chi_Y}, \quad (2.21)$$

$$\frac{1}{\beta} - 1 = (u_s)^\kappa \left(1 - \frac{1}{\kappa}\right) + p_s^E (u_s)^\nu \left(1 - \frac{1}{\nu}\right), \quad (2.22)$$

where $u_s = (\kappa \delta_s)^{1/\kappa}$, $\frac{\Xi_s^Y}{\Lambda_s}$ is the real marginal cost of producing an additional unit of V , and $\chi_V = \frac{(1 - \chi_Q)(1 - \chi_Y)}{1 - \chi_Q(1 - \chi_Y)}$. Under the baseline specification, $\omega = 3.0334$, $\nu = 1.7373$, and $\kappa = 1.06$.

2.4.2 Baseline Results

First, we will look at the baseline specification, which corresponds to the pre-1984 state of the economy. Figure 2.5 shows dynamic responses of variables to a 10% increase in the real price of oil.

Aggregate value added, which corresponds to real GDP in data, shows a hump-shaped response and peaks at 5 quarters after the shock. The near-unit-root nature of the oil shock results in strong persistence in IRFs. Other variables, except for inflation, show hump-shaped response as well, due to the price stickiness. Among macroeconomic variables, investment is most severely affected, as pointed out by Aguiar-Conraria and

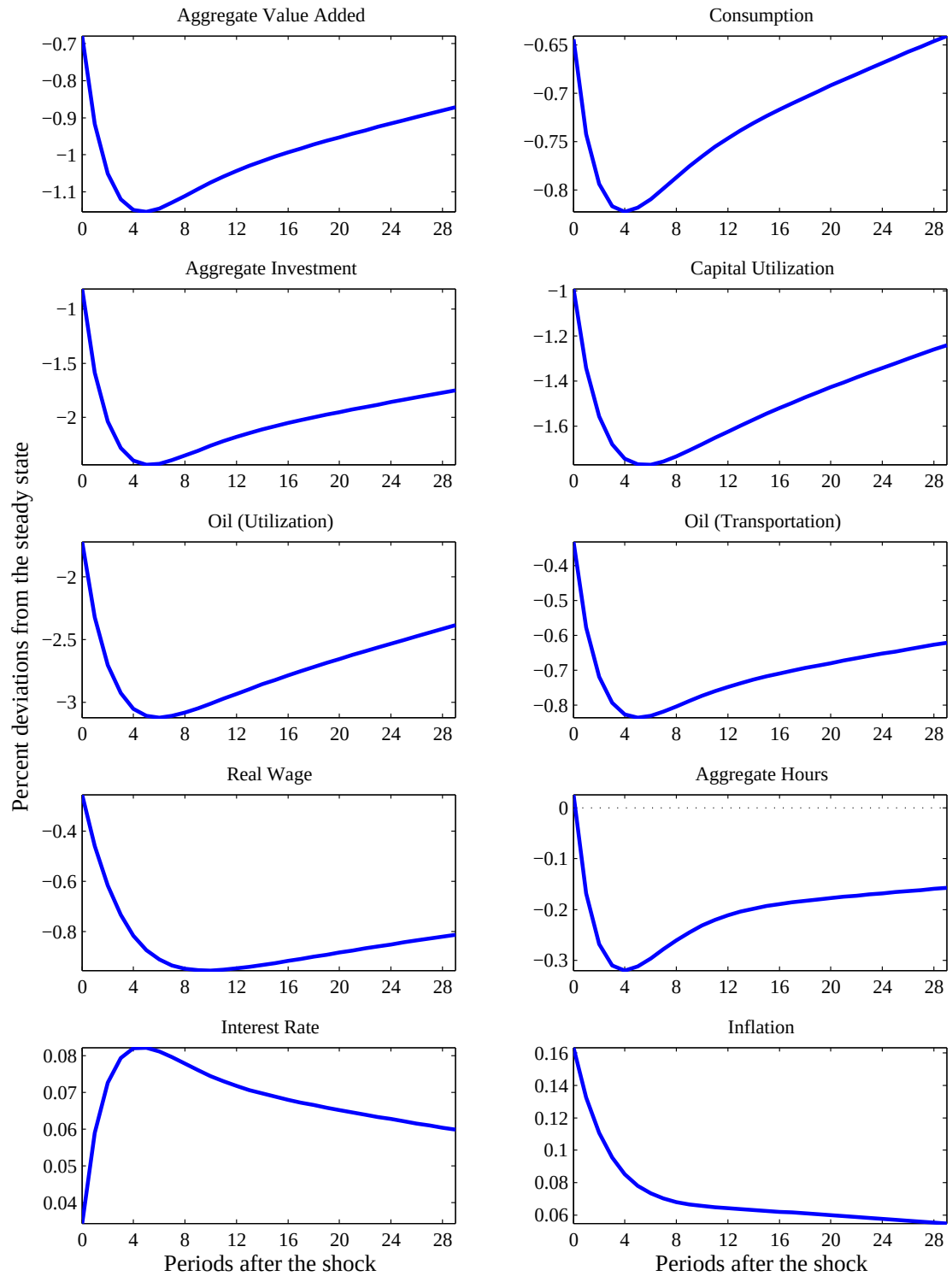


Figure 2.5: Impulse Responses to a 10% Increases in the Oil Price: Baseline Results

Wen (2007). Two types of oil consumption show different patterns in terms of magnitude. Oil consumption for capital utilization is more price elastic compared with oil used for transportation. Aggregate hours worked have a positive response in the first period due to the fact that the income effect slightly dominates the substitution effect, and then decline. Under the monetary policy rule assumed, the Fed increases interest rate in order to fight against inflation triggered by increases in the oil price. Inflation dynamics look like a mirror image of the oil-price process because there are no mechanisms that generate the hump-shaped response of inflation, such as backward indexation or price adjustment costs for inflation.

There are two channels through which the oil-price shocks can affect the economy. The first one is via capital utilization tied with use of oil. The oil-price shock results in a higher cost of utilizing capital stock, and hence a lower rate of capital utilization. This in turn reduces the level of output produced. The second channel is a direct impact on the cost of delivering goods produced.

Intermediate firms' real marginal cost of producing an additional unit of output in terms of log-deviations from the steady state is given by

$$\widehat{mc}_t = \frac{1}{mc_s} \{ (1 - \chi_Q)(1 - \chi_Y)mc_s^V \widehat{mc}_t^V + \chi_Y p_s^T \widehat{p}_t^T \}, \quad (2.23)$$

where $mc_s^V = \frac{\Xi_s^Y}{\Lambda_s}$ represents the steady-state real marginal cost of producing an additional unit of their value added, p_s^T is the real price of transportation services, and hat-variables represent log-deviations from the steady state. Holding everything else constant, an increase in the real price of oil affects both $\widehat{mc}_t^V$ via capital utilization and $\widehat{p}_t^T$ through transportation costs. Transportation firms' markup plays a role in terms of amplifying the responses of macroeconomic variables. The higher the markup, the more expensive the steady-state price of transportation services. This is one of reasons why we expect the lower markup on transportation services can contribute to the weaker response to the same size of the oil-price shock.

Now, we turn to analyze factors that contribute to the observed weaker responses of macroeconomic variables to the oil-price shock.

2.4.3 Changes in the Economy Over Time

In order to understand the declining effects of the oil-price shocks, we will consider three different factors individually and some combinations of those. First, we will

Table 2.2: Summary of Changes in the Peak Response of Aggregate Value Added

		% Change in Peak Responses	Changes in the Timing
Case 1	Lower Markup	-11.41%	± 0
Case 2	Transportation Industry	-19.73%	± 0
Case 3	Improved Energy Efficiency	-24.72%	± 0
Case 4	Less Persistent Shock	-26.52%	-2
Case 5	Post-1984	-51.38%	-2

look at the consequence of more competitive environment in the transportation industry triggered by the deregulation after 1980. Second, we will see the effect of improved energy efficiency in the economy. Lastly, we will look at the consequence of the less persistent oil-price shock. Table 2.2 reports summary of changes in the peak response of aggregate value added under each scenario and associated IRFs are shown in Figure 2.6 – Figure 2.8.

Case 1: Increased Competition among Transportation Firms

In order to see the outcome of the increased competition in the transportation industry, we will only change the value of θ_T from 3.8169 to 53.632 such that the implied steady-state markup of transportation firms changes from 35.5% to 1.9%, which is based on Nebesky, McMullen, and Lee (1995). Other parameter values are left unchanged.

As shown in Figure 2.6 with the label of *Lower Markup*, all variables show weaker responses to the same size of the oil-price shock. The peak response of aggregate value added is reduced by 11.41%. However, there are no changes in the peak response.

Holding other things constant, the immediate effect of the lower markup is a reduction in the steady-state price of transportation services p_s^T . This induces a downward shift in the marginal cost curve of intermediate-good firms. As a result, the economy moves to the new steady state with the higher level of output and the lower level of the general price index. To meet with the higher level of output, all factor inputs will be utilized more than before.

Although the lower markup does not change the dynamics of $\hat{p}_t^T$, the log-deviations of the real marginal cost (2.23) tells us that a decline in p_s^T induces in the deviations in the real marginal cost, by putting less weight on deviations on the price of

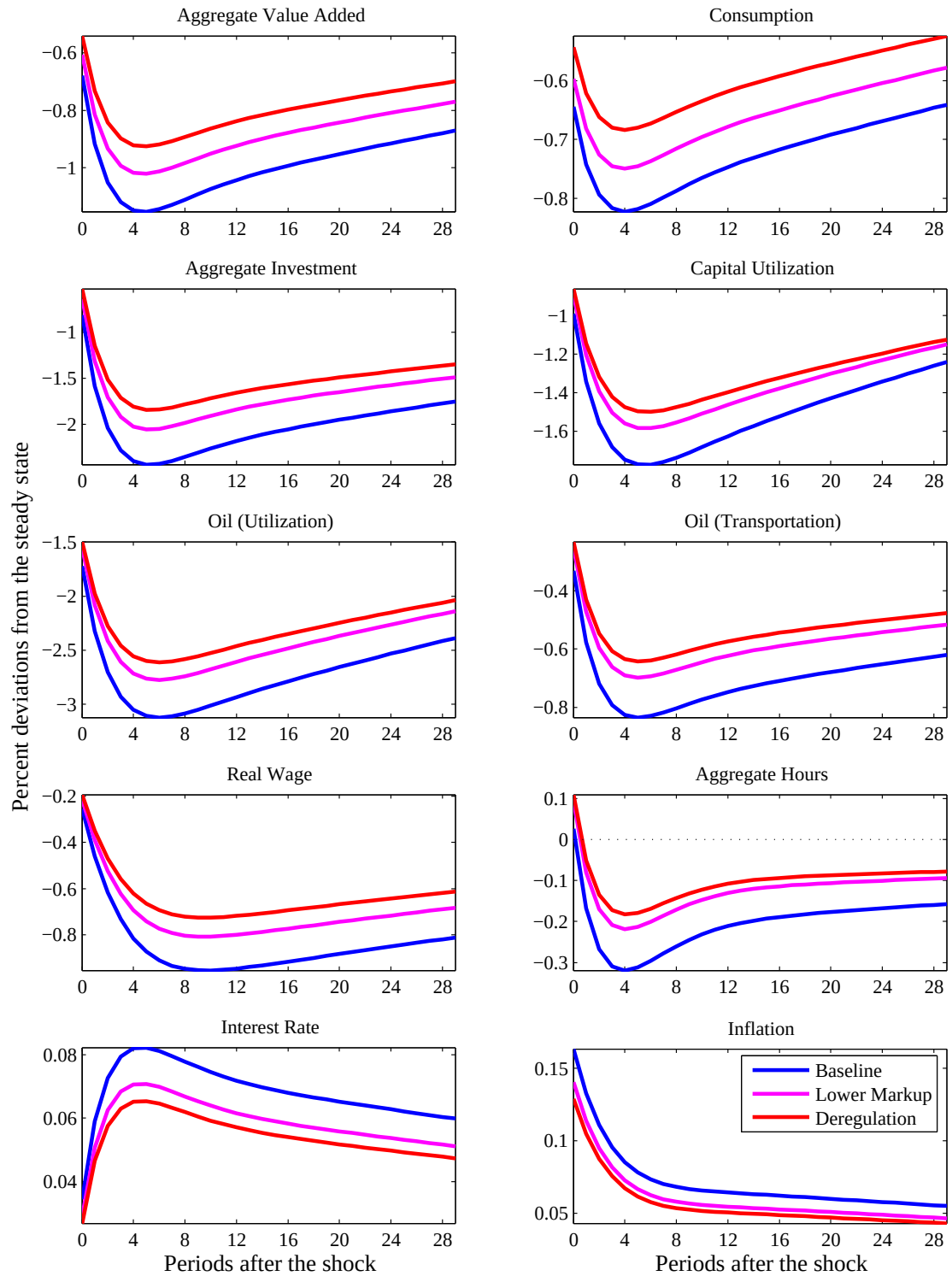


Figure 2.6: Comparison of IRFs to a 10% Increase in the Oil Price: Case 1 and Case 2

transportation firms. In turn, it results in smaller deviations in other real variables.

We can observe another effect of lowering the markup on the inflation dynamics. The intermediate-firms' first-order condition with respect to their prices in the symmetric equilibrium is given by

$$0 = (1 - \theta)Y_t + \theta \{ \chi_Q(1 - \chi_Y) + \chi_Y p_t^T \} Y_t - \phi_Y \left(\frac{\pi_t}{\pi_s} - 1 \right) \frac{\pi_t}{\pi_s} (1 - \chi_Q(1 - \chi_Y)) Y_t \\ + \theta \frac{\Xi_t^Y}{\Lambda_t} (1 - \chi_Q)(1 - \chi_Y) Y_t + \beta \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{\Lambda_t} \phi_Y \left(\frac{\pi_{t+1}}{\pi_s} - 1 \right) \frac{\pi_{t+1}}{\pi_s} (1 - \chi_Q(1 - \chi_Y)) Y_{t+1} \right] \quad (2.24)$$

Log-linearizing the above equation gives us the New-Keynesian-Phillips-curve type relation:

$$\hat{\pi}_t = \frac{\theta}{\phi_Y} \left\{ \frac{(1 - \chi_Q)(1 - \chi_Y) m c_s^V}{1 - \chi_Q(1 - \chi_Y)} \widehat{m c_t^V} + \frac{\chi_Y p_s^T}{1 - \chi_Q(1 - \chi_Y)} \hat{p}_t^T \right\} + \beta \mathbb{E}_t[\hat{\pi}_{t+1}]. \quad (2.25)$$

Thus, immediately from the above equation, a reduction in p_s^T induces smaller deviations in inflation.

Case 2: Transportation Industry (Deregulation)

As described in Section 2.2, another consequence of the deregulation is that it has contributed to improve efficiency in the transportation industry. It is important to consider this contribution of the transportation industry as a whole because it accounts for the majority of US petroleum consumption. Here we will consider the contribution of the transportation industry, by taking account of the lower markup and the improved energy efficiency in the industry.

Unfortunately, we do not have reliable data on how energy efficiency has been improved in the transportation industry. One thing we know is that the average oil expenditure share has dropped from 3.65% in the pre-1984 period to 2.75% in the post-1984 period. Another thing we can observe is changes in petroleum consumption in industrial and transportation sectors. Particularly, the ratio s_1/s_2 has decreased from 0.51 to 0.39 in the post-1984 period. Holding other parameter values fixed and using (2.20), I can infer a new value of ω , which is consistent with the post-1984 characteristic of the US petroleum consumption, by using the above information. As a result, ω increases from 3.0334 to 3.6752.

In order to see the consequence of the deregulation, I will change both θ_T and ω from the values used in the baseline specification. Since improved energy efficiency

in the industry could be because of the deregulation and/or advanced transportation technology, such as higher miles per gallon, this specification does not represent the exact outcome of the deregulation. I still believe that it is a good proxy for the consequence of the deregulation, however. If there still existed restrictions on entry/exit in the industry, then there would be no entrants that are more efficient and there would be less incentive for incumbents to improve efficiency by adopting more fuel efficient technology.

IRFs presented in Figure 2.6 labeled *Deregulation* report the contribution originated from the transportation sector of the economy. All variables show weaker responses to the same size of the shock than the baseline specification. In terms of changes in the aggregate value-added peak responses, the consequence of the deregulation accounts for a 19.73% reduction, compared to the baseline specification.

Improved energy efficiency in the transportation sector, that is, an increase in ω has two effects. First, it reduces the steady-state price of transportation services, which has the same effects as the lower markup. Under this specification, the steady-state price of transportation services decreases by 31%, which is consistent with the empirical finding of Ying and Keeler (1991). Second, it makes deviations of the transportation price smaller.

$$\hat{p}_t^T = \frac{1}{p_s^T} \left\{ mc_s^T \widehat{mc}_t^T + \frac{p_s^E}{\omega} \hat{p}_t^E \right\} \quad (2.26)$$

It is clear that an increase in ω makes $\hat{p}_t^T$ smaller than before. In turn, it makes deviations in the marginal cost smaller.

A 20% increase in ω used in this specification is much smaller than the 42% improvements in the average MPG of all vehicles mentioned in the Introduction. This is particularly because of the fact that the US economy is more dependent on transportation, which is apparent in a large drop in the ratio s_1/s_2 . If the ratio stayed the same, which is equivalent to less reliance on transportation, we would expect a bigger contribution from the transportation industry.

Case 3: Improved Energy Efficiency in the Economy

Another factor that could contribute to the weaker responses of macroeconomic variables is improvements in efficiency of petroleum use. It is quite intuitive that improved energy efficiency might be a candidate for the declining effects of the oil-price shock. As energy efficiency improves, the economy needs to rely less on petroleum and becomes less vulnerable to an increase in the oil price. This is apparent in the smaller

oil expenditure share.

More efficient use of oil could originate from capital utilization and/or transportation. Thus, here I will change both ω and ν , which govern energy efficiency in both transportation and industrial sectors, respectively. Given the value of ω used in Case 2 and other parameter values fixed, ν is changed such that the resulting steady-state oil expenditure share equals to 2.75%. As a result, ν changes from 1.7373 to 1.8864.

Given the same size of shock, IRFs in Figure 2.7 labeled *Energy Efficiency* also show weaker responses of all variables. As a result, the peak response of aggregate value added becomes smaller by 24.72%, compared with the baseline specification.

An increase in ν has an effect similar to an increase in ω . It results in the smaller value of steady-state marginal cost of producing intermediate-good firms' value added mc_s^V and also $\widehat{mc}_t^V$ becomes smaller given the same size of shock in the oil price.

Case 4: Less Persistence of the Shock

It is not hard to imagine that the shape of IRFs would be affected by a small change in the underlying shock process, and thus one might suspect that it could be the whole story why we observe the weaker response of macroeconomic variables to the oil-price shock. Less persistence in the oil price is easily observed. In the post-1984 period, the estimated AR(1) coefficient on the log-real price, measured as log of PPI for crude petroleum deflated by GDP deflator, suggests that it is dropped from 0.99 to 0.92. Thus, I change the persistence parameter ρ_{pE} from 0.99 to 0.92.

IRFs shown in Figure 2.7 labeled *Less Persistence* are the resulting responses. In terms of the peak response, as expected, a small change in the persistence parameter contributes to a large reduction in the peak response, 26.52%. Although there are no changes in the timing of the peak under other factors considered, less persistence in the oil-price process shifts the timing of the peak by 2 quarters.

Case 5: Post-1984

Finally, I combine all factors considered in this paper together, in order to see how much the model could explain changes in the economy over time in response to the oil-price shock. Figure 2.8 compares resulting IRFs with those from the baseline specification. As can be seen clearly, all factors together result in much smaller responses

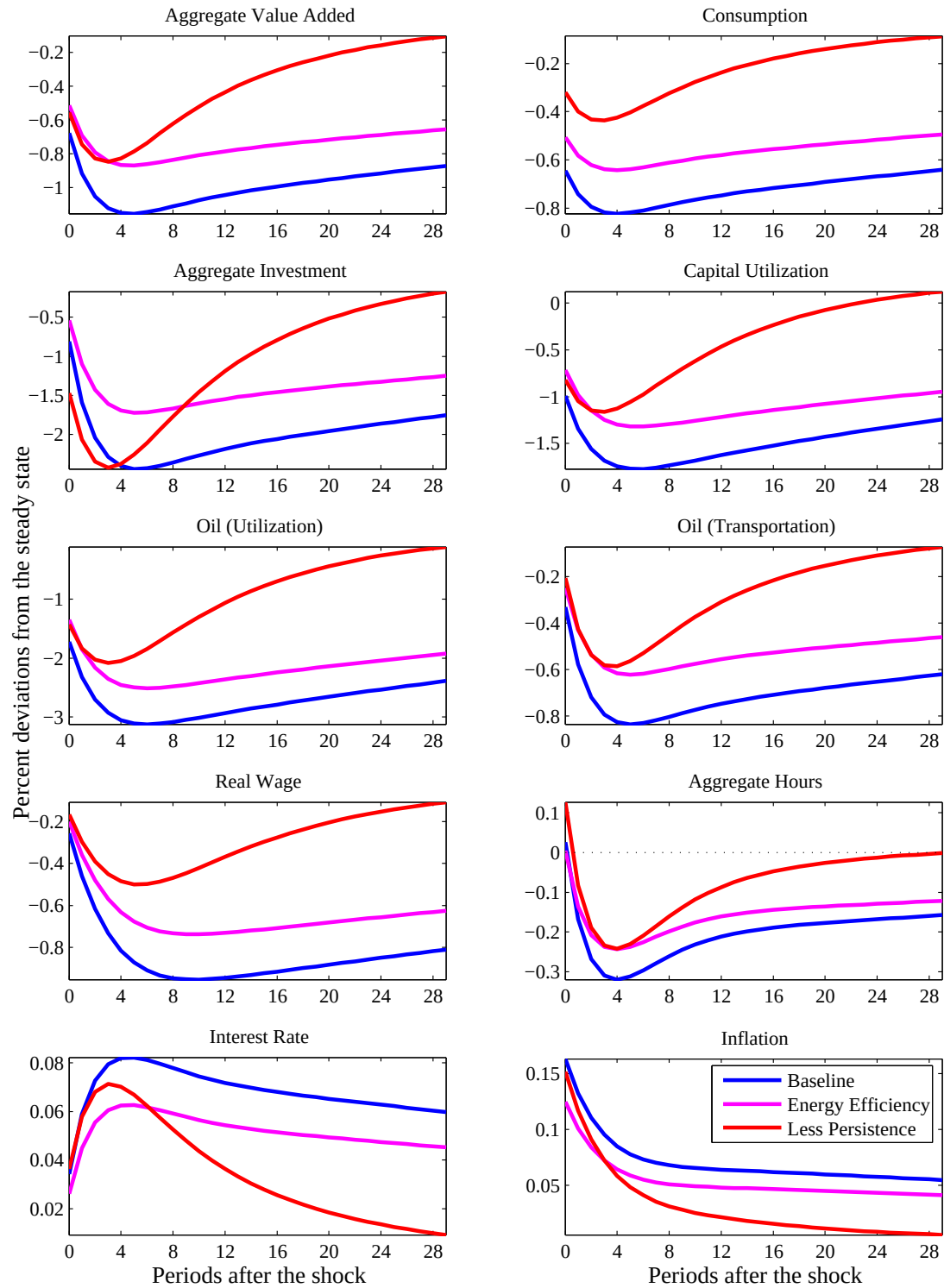


Figure 2.7: Comparison of IRFs to a 10% Increase in the Oil Price: Case 3 and Case 4

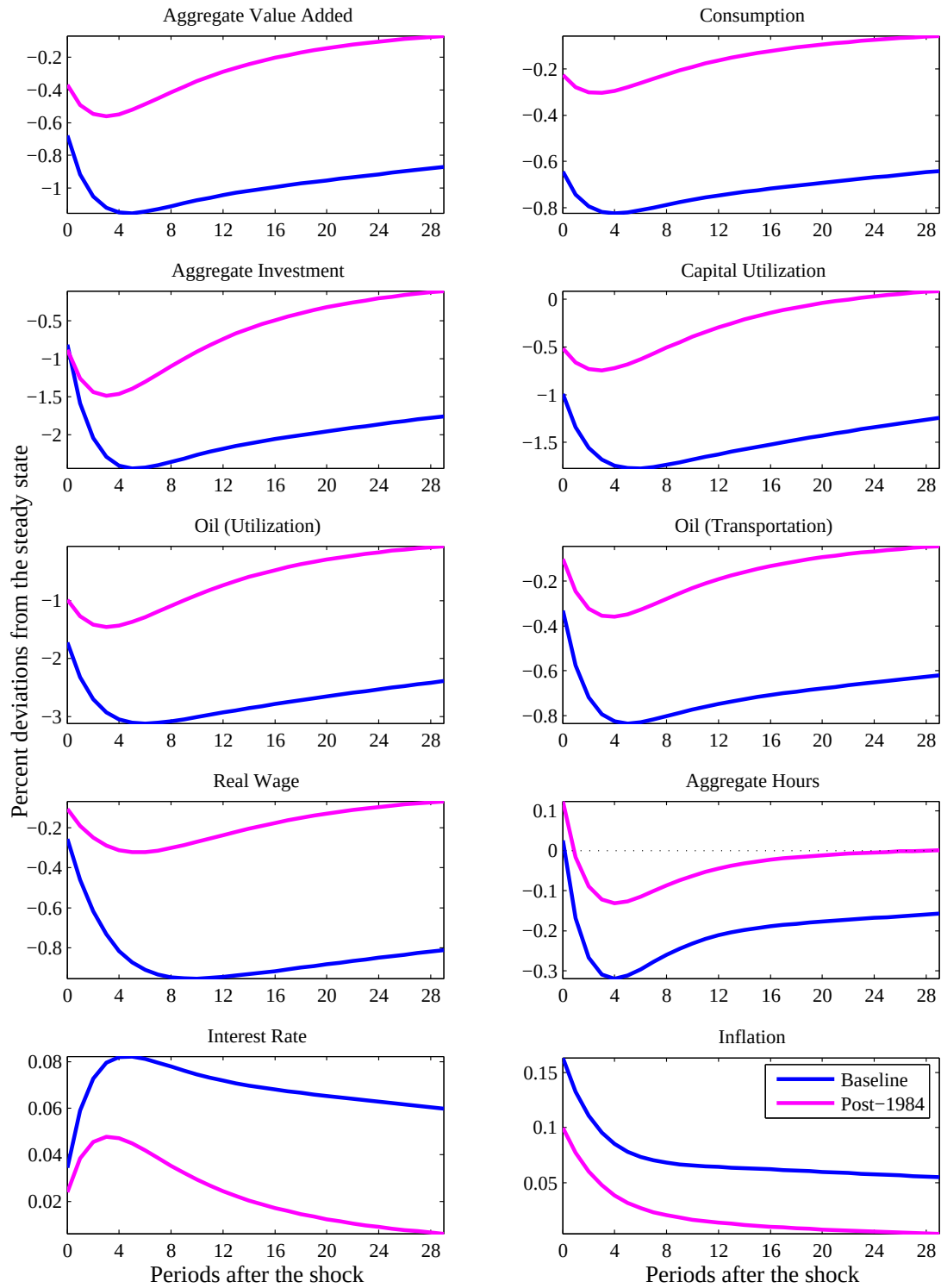


Figure 2.8: Comparison of IRFs to a 10% Increase in the Oil Price: Before and After 1984

to the same size of the oil-price shock. In terms of the changes in the peak response of aggregate value added, it suggests that the peak response is reduced by 51.38% as a whole and the timing of the peak is shifted by 2 quarters due to the less persistent oil-price process.

If one believes the VAR results presented in the Introduction are a good representation of changes in the economy over time, then these factors considered in this paper could account for a majority of the declining effects of the oil-price shock. It should be noted that, however, these results do not suggest that the factors considered here are the only source of changes in the economy over time, or that all changes have made the effects smaller. For example, in the presence of the trade-off between output and inflation stabilization, a strong anti-inflation stance of the monetary authority, which is widely documented, then it would even contribute to much stronger responses of output. However, increased credibility of the Fed could stabilize the output response as demonstrated in Blanchard and Galí (2007).

Alternatively, one might argue that the economy is shifting toward more flexible prices.¹³ Regarding reduced price stickiness, it would contribute to shifting the timing of the peak. The more flexible the price is, the earlier peak we would expect. Since price stickiness has dampening effects on dynamic response of variables, more flexible price would result in stronger responses of aggregate value added, however. Mechanically, less habit persistence, if any, could work in the same way as flexible prices. As argued in Blanchard and Galí (2007), it could be also true that wage rigidities have become smaller over time.

To sum up, by comparing Case 1 and Case 2, we can see that two factors affecting the contribution of the transportation industry, namely more competition and more efficient use of oil in the industry, are equally important. The effect of the deregulation in the transportation industry results in a 20% reduction in the peak response of aggregate value added. It should be noted that in spite of its small share in the economy, the contribution of the transportation industry is non-negligible. The overall effect of more efficient use of oil in the US economy weakens the peak responses of value added by 25%. The results suggest that the industrial sector (intermediate-good firms) has a

¹³It has been widely believed that the average frequency of price change is about one year (for example, Taylor, 1999). But Bils and Klenow (2004) document that in the mid-1990's the average frequency of price change is much more frequent than what typically believed.

slightly larger impact on the weaker response than the transportation sector does, since improved energy efficiency in the transportation industry roughly accounts for about a 10% reduction in the peak response. A less persistent oil-price shock accounts for a 27% reduction in the peak response. Overall, all factors play an important role in accounting for the declining effects of the oil-price shock and are equally important quantitatively.

2.5 Conclusion

We have developed a model that is more realistic in terms of US petroleum consumption, in order to explore possible explanations for the weaker responses of macroeconomic variables to oil-price shocks. Under the baseline specification, which corresponds to the pre-1984 state of the economy, the model well captures important aspects of the economy's response to the oil-price shock. Price stickiness accounts for hump-shaped responses of output measures, which are consistent with what we see in data.

We have explored three developments that may have affected the response of output to an oil-price shock, namely the effect of deregulation in the transportation industry, improved energy efficiency, and less persistent oil-price shocks. All factors are equally important quantitatively. It is worth noting that in spite of its small share, the contribution of the transportation industry triggered by the deregulation is nontrivial. If the US economy becomes less dependent on transportation, we would expect further contribution of the transportation sector. Combining all factors together, the model predicts a 51% reduction in the peak response of aggregate value added and the peak comes 2-quarter earlier.

We can draw one implication from this study. Deregulation in the transportation industry has led to more competitive environment and less reliance on petroleum. Also, technological advancement enables us more efficient use of petroleum overall. Both factors should make the economy less vulnerable to oil-price shocks in the future. From past experience, the general public as well as policy makers might still hold a view that large increases in the price of oil trigger a deep recession. However, we expect that large recessionary consequences of oil-price shocks we observed in the 1970's will not be seen again.

Chapter 3

Improving Recession Probability Forecasts in the U.S. Economy

3.1 Introduction

There are many attempts to forecast recessions in the U.S. economy. Information contained in forward-looking variables can be used for making predictions about the future state of the economy. Stock and Watson (1989) report that the slope of the Treasury yield curve provides useful information in constructing their leading indicator. Estrella and Mishkin (1998) examine the usefulness of financial variables in making predictions about future recession probabilities at various forecasting horizons. They find that the slope of yield curve is the best single predictor of future recession probabilities.¹ Other studies also utilize the predictive content of the term spread in order to forecast recession probabilities.² Although developments in the literature achieve some success in predicting recession, there is a room to improve forecasting accuracy.

To obtain further improvements, there are two margins that we can exploit. One is the question of selecting a set of predictor variables. The other is a choice of functional form. Typical studies on forecasting recession probabilities utilize a simple binary response model, such as a probit or logit model. In this paper, I will explore the importance of the two margins in improving forecasting accuracy.

¹Their finding is consistent with a simple rule of thumb that the Treasury yield curve inversions or negative term spreads are likely followed by recessions in subsequent periods and the term spread has been cited as a leading economic indicator.

²Other studies include Dotsey (1998), Birchenhall, Jessen, Osborn, and Simpson (1999), Estrella, Rodrigues, and Schich (2003), Stock and Watson (2003) and Clements and Galvão (2006).

Within the class of probit/logit models, there is favorable evidence for benefits from including additional information, other than the term spread. For example, Estrella and Mishkin (1998) find some evidence on the usefulness of including an additional variable, such as a New York Stock Exchange index or GDP growth. Wright (2006) reports that including the level of the federal funds rate together with the term spread results in better forecasting performance than the term spread alone.³ King, Levin, and Perli (2007) recently find superior predictive power of the 5-year corporate credit spread on AA-rated firms compared with the 10-year-3-month Treasury term spread over the Great Moderation period (since the mid 1980's). Furthermore, they also report that once the corporate spread is augmented by the term spread, the forecasting performance dramatically improves by reducing "false positive" predictions of recession.

There are also some attempts to extend a simple probit model in the recent literature. For example, Chauvet and Potter (2005) extend a probit model in a way that it allows for business-cycle dependent coefficients (i.e., multiple breaks) and/or autocorrelated errors. Kauppi and Saikkonen (2007) develop dynamic probit models that include lagged explanatory variables and lagged recessionary dummies. Dueker (2005) presents a framework (a Qual VAR) that enables us to treat a qualitative variable as endogenous in a typical VAR framework, so that we can obtain dynamic forecasts of the qualitative variable.

Going beyond the typical univariate probit model with the term spread, such as allowing additional features in a forecasting model and/or including additional predictor variables most likely would improve in-sample forecasting performance. However, at the same time, we also have to worry about a potential over-fitting problem. Hansen (2008) examines the pitfalls of relying on in-sample fit. He shows that too good in-sample fit (over-fitting) tends to be associated with poor out-of-sample fit and that model selection based on in-sample fit is not reliable. In addition, he demonstrates that tendencies of yielding spurious results are much higher if we use in-sample evaluation and are more pronounced when we compare performance of a large number of alternative models. In this paper, therefore, I will use out-of-sample and cross validation methods to examine the importance of the two margins that can contribute to improve forecasting performance.

First, I systematically compare the forecasting performance of 6-month-ahead

³He also finds some evidence favorable for controlling for a term premium proxy.

recession probability predictions in the U.S. economy. I look at all possible combinations of 33 variables up to trivariate models with 6 different functional forms. Results highlight the importance of variable selection. Finding the best combination of predictor variables greatly improve forecasting performance. Especially, among variables considered, the combination of the term spread, changes in the S&P 500 stock price index, and growth rate of non-farm employment is found to achieve the best forecasting accuracy. Furthermore, additional gains can be obtained by using non-Normal cumulative distribution functions. However, there is some mixed evidence on what features of functional form are necessary, in order to improve forecasting performance. Then, I proceed to allow for more flexibility in the functional form by utilizing the generalized Edgeworth expansion of Jarrow and Rudd (1982).

The rest of the paper is organized as follows. Section 3.2 presents basic binary response models and associated results. Section 3.3 introduces the generalized Edgeworth expansion and applies it to making predictions of the U.S. recession probability. Finally, Section 3.4 concludes.

3.2 Basic Models and Results

3.2.1 Recession Probability Forecasting Model

Let y_t represent an NBER recession binary variable, which equals 1 when the economy is in recession in month t and equals 0 in expansion. Typical models of forecasting h -period-ahead recession probabilities using the information available at time t assume that

$$\text{Prob}(y_{t+h} = 1 | \mathbf{x}_t) = F(\boldsymbol{\beta}' \mathbf{x}_t), \quad (3.1)$$

where $F(\cdot)$ is a monotonically increasing function, whose range is the unit interval, $\boldsymbol{\beta}$ is a vector of coefficients associated with a vector of predictors $\mathbf{x}'_t = [1, x_{1,t}, \dots, x_{k,t}]$, and k is the number of variables included.

It is commonly assumed that y_{t+h} is a conditionally independent Bernoulli random variable, so that the likelihood function is given by:

$$L = \prod_{t=1}^T [F(\boldsymbol{\beta}' \mathbf{x}_t)]^{y_{t+h}} [1 - F(\boldsymbol{\beta}' \mathbf{x}_t)]^{1-y_{t+h}}. \quad (3.2)$$

In an empirical analysis, I will set $h = 6$ and focus on 6-month-ahead predictions.

In this formulation, predicting recession probability involves two issues that are possibly related to each other. The first one is to choose a set of predictor variables, so that we can obtain useful signals from data. A workhorse predictor in the literature is the Treasury term spread between 10-year and 3-month bonds, which is due to the finding in Estrella and Mishkin (1998).

Given a choice of $F(\cdot)$, finding a better combination of predictor variables obviously helps improve forecasting accuracy if those variables contain different information and jointly provide useful signals. Since there are no *a priori* variable selection procedures available, our approach is to try all possible combinations of variables, in order to find a better combination of variables that are useful in forecasting recession probabilities. Although knowing the best single predictor is helpful, it is unlikely that we can obtain some insights about a better combination of variables by just looking at forecasting performance of single predictors because those variables that show relatively good forecasting performance tend to contain similar information.

The second issue is how to translate signals into a probability measure between 0 and 1, which is related to a shape of $F(\cdot)$. In order to guarantee that $F(\cdot)$ is monotonically increasing and takes values between 0 and 1, we typically use a known cumulative distribution function (CDF). A popular choice is to use either the Standard Normal CDF (a probit model) or the Logistic CDF (a logit model), or some variants of those (e.g., Chauvet and Potter, 2005; Kauppi and Saikkonen, 2007). However, the shape of $F(\cdot)$ is not necessarily restricted to typical ones. In principle, the CDF of any continuous random variable will suffice.

Differences in the shape of CDFs can be characterized in terms of “skewness” and “excess kurtosis”. It should be noted that it is not appropriate to use terms skewness and excess kurtosis here because we are not talking about characteristics of underlying statistical distributions, but just utilizing functional forms. However, for expositional simplicity, I will use those terms in describing the shape of a CDF. The consequence of allowing skewness and excess kurtosis is as follows. Higher excess kurtosis makes a CDF steeper around the median of $\beta'x$. On the other hand, allowing non-zero skewness makes a CDF asymmetric around $F(\beta'x) = 0.5$. With zero skewness, $F(x) = 1 - F(-x)$ for any $x \in \mathbb{R}$. However, non-zero skewness implies $F(x) \neq 1 - F(-x)$.⁴

⁴In general, changing location and scale parameters does not affect forecasting results and they are fixed to avoid an identification problem. Changing the location parameter just changes the estimate of the constant term and changing the variance just results in different scaling of coefficients β .

In order to understand what features of a CDF are helpful in improving forecasting accuracy, I will consider 6 different CDFs. Table 3.1 summarizes characteristics of the CDFs considered. In addition to the Standard Normal CDF and the Logistic CDF, I consider Student- t , Laplace, Gumbel, and Type III Generalized Extreme Value (GEV3). All location parameters and scale parameters, if applicable, are set to 0 and 1, respectively. These CDFs are chosen and configured to incorporate non-zero skewness and/or higher excess kurtosis that the Standard Normal CDF does not have. The first four CDFs have zero skewness and differ in terms of degree of excess kurtosis. The Logistic CDF has excess kurtosis of 1.2. The degrees of freedom parameter for Student- t is set to be 6.5, such that its excess kurtosis equals 2.4. The excess kurtosis of Laplace is 3. The last two CDFs also have positive skewness. For the Gumbel CDF, skewness is equal to 1.1395 and excess kurtosis is 2.4. The shape parameter of the Type III GEV, s , is set such that its excess kurtosis equals 3 (i.e., $s = -0.1732$). The associated skewness becomes 0.3492. Thus, for a given set of predictor variables, I will be able to infer the importance of allowing for excess kurtosis in $F(\cdot)$ by looking at results based on the first four CDFs. Furthermore, comparing the Gumbel with the Student- t or the Type III GEV with the Laplace enables us to obtain some insight on whether allowing positive skewness is helpful in improving forecasting performance.⁵

3.2.2 Data

Table 3.2 lists all 33 monthly variables considered in this paper. The sample period starts from April 1960, which corresponds to the beginning of the 1960-61 recession. It is chosen to maximize data availability and include as many recession episodes as possible. Since it is often the case that a decision of the NBER Business Cycle Dating Committee has a time lag,⁶ the sample period is ended at the end of December, 2005. The data set contains the term spreads, the credit spreads, various interest rates, employment data, stock price indices, monetary aggregates, and other macroeconomic variables. Most of those are investigated in the earlier studies or used in constructing a composite leading indicator. It also includes variables, to which the NBER Business

⁵It is also possible to introduce negative skewness by using the Type III GEV. However, preliminary estimation indicated that forecasting performance based on the Type III GEV with negative skewness is inferior to others in general. So, I do not report their results.

⁶A substantial time lag of the announcement might be one disadvantage of using the NBER recessions to characterize the state of the economy. This motivates Chauvet and Hamilton (2006) to study algorithms that enable us to construct a real-time business cycle turning point.

Table 3.1: List of the Cumulative Distribution Functions Considered

Type of CDF	$F(x)$	Skewness	Excess Kurtosis
Standard Normal	$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(\frac{-u^2}{2}\right) du$	0	0
Logistic	$\frac{\exp(x)}{1 + \exp(x)}$	0	1.2
Student- t with $\nu = 6.5$	$\int_{-\infty}^x \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{u^2}{\nu}\right)^{-\frac{(\nu+1)}{2}} du$	0	2.4
Laplace	$\frac{1}{2} [1 + \operatorname{sgn}(x) \{1 - \exp(- x)\}]$	0	3
Gumbel	$1 - \exp\{-\exp(-x)\}$	1.1395	2.4
Type III GEV with $s = -0.1732$	$\exp\left\{-(1 + sx)^{-1/s}\right\}$	0.3492	3

Note: All location parameters are set to be 0 and all scale parameters are set to be 1. ν represents degrees of freedom parameter for the Student- t distribution. $\Gamma(\cdot)$ is a gamma function. Excess kurtosis of Student- t is given by $6/(\nu - 4)$ for $\nu > 4$. Skewness of the Gumbel distribution is $\frac{12\sqrt{6}\zeta(3)}{\pi^3}$, where $\zeta(\cdot)$ is a zeta function. s denotes a shape parameter of the generalized extreme value distribution. For the Type III GEV, skewness is given by $\frac{-\Gamma(1-3s)+3\Gamma(1-s)\Gamma(1-2s)-2(\Gamma(1-s))^3}{(\Gamma(1-2s)-(\Gamma(1-s))^2)^{3/2}}$ and excess kurtosis is given by $\frac{\Gamma(1-4s)-4\Gamma(1-s)\Gamma(1-3s)+6\Gamma(1-2s)(\Gamma(1-s))^2-3(\Gamma(1-s))^4}{(\Gamma(1-2s)-(\Gamma(1-s))^2)^2}$.

Cycle Dating Committee pays particular attention in deciding business cycle peaks and troughs. Since Kane (2008) documents the usefulness of employment data in predicting occurrence of recessions, the data set includes employment-related variables as well.

Recently, King, Levin, and Perli (2007) report that, in the period of the Great Moderation, the credit spread on AA-rated firms has particularly good forecasting performance. However, the corporate bond yields used in their study (maturity of 5-year and 10-year) do not cover the entire sample. Since there are not many recessionary periods, especially after the 1980's, using shorter sample periods may have considerable effects on evaluating forecasting performance. Thus, I have decided not to include the credit spreads that King, Levin, and Perli (2007) have used. Instead, as a crude proxy, the data set include spreads between Moody's AAA-, AA-, or A- rated corporate bond

yield (20 years or longer) and the 10-year Treasury bond yield.⁷

When we focus on financial variables as predictors, we do not need to consider a gap between when an observation is made and when it is available for forecasting. However, as shown in Table 3.2, some series are not reported immediately and we need to take account of the information lag, in order to accurately assess forecasting models. In this paper, any variable z_t represents the latest data on z available at month t , instead of an observation at month t . For example, industrial production (IP) has 1 month of the information lag. So, $IP_{2000:01}$ refers to the industrial production data on December 1999.

Because the total number of models increases exponentially as the number of variables included (k) increases, I will restrict my attention to $k \leq 3$. This will result in examining a total of $\sum_{k=1}^3 \frac{33!}{k!(33-k)!} \times 6 = 36102$ forecasting models.

3.2.3 Evaluating Forecasting Performance

In order to evaluate various forecasting models, I will primarily focus on out-of-sample results. Hansen (2008) shows that in-sample and out-of-sample fits are negatively correlated, which implies that good in-sample performance is not a useful indicator of out-of-sample accuracy and that relying on in-sample fit is highly misleading. This over-fitting problem is particularly important, since the likelihood of obtaining spurious results is more pronounced when we search a large number of alternative models. For this reason, I will focus on recursive out-of-sample forecasting evaluation and I will use cross validation as a robustness check.

For recursive (pseudo) out-of-sample forecasting exercises, the out-of-sample prediction starts from January 1989 and ends at the end of the full sample. The out-of-sample period covers the last two recessions in the U.S. economy. By adding observations one by one, I estimate a forecasting model again to produce a forecast for the next month. In reality, when we forecast future recession probability, we are not certain about the true state of the economy in recent months. It is because the decision of the NBER Business Cycle Dating Committee typically involves substantial time lag. To be realistic, I assume that forecasters do not know the true state of the economy for a year and assume that

⁷Since the 20-year Treasury bond yield data has discontinuity between January 1987 and September 1993, I use the 10-year T-bond rate, instead. Thus, precisely speaking, this “credit spread” is a combination of the “true” credit spread and the term spread.

Table 3.2: List of Variables

Predictor	Description	Info. Lag
<i>Interest Rates</i>		0
FF	Federal Funds rate	0
3M	3-month Treasury Bill rate	0
5Y	5-year Treasury Bond rate	0
10Y	10-year Treasury Bond rate	0
AAA	Moody's corporate bond yield, AAA 20 years or longer	0
AA	Moody's corporate bond yield, AA 20 years or longer	0
A	Moody's corporate bond yield, A 20 years or longer	0
<i>Term Spreads</i>		
TS10YFF	10Y-FF Treasury term spread	0
TS10Y3M	10Y-3M Treasury term spread	0
TS10Y5Y	10Y-5Y Treasury term spread	0
<i>Credit Spreads</i>		
CSAAA	AAA - 10Y spread	0
CSAA	AA - 10Y spread	0
CSA	A - 10Y spread	0
<i>Employment Data</i>		
EMP	Non-agricultural employment (log-differenced)	0
CEMP	Civilian employment (log-differenced)	0
UICLAIM	Initial unemployment insurance claims (log-differenced)	1
UNEMP	Unemployment rate	0
UNEMPD	Changes in unemployment rate	0
HOURS	Average weekly hours in manufacturing (log-differenced)	0
<i>Stock Price Indices</i>		
DJ30	Dow Jones 30 average (% changes over 3 months)	0
SP500	S&P 500 stock price index (% changes over 3 months)	0
<i>Monetary Aggregates</i>		
M0	Monetary base (log-differenced)	1
M1	M1 (log-differenced)	1
M2	M2 (log-differenced)	2
<i>Other Macroeconomic Variables</i>		
CLI11	Composite leading indicators (11 series, 1987=100, log-differenced)	1
CPI	CPI, all urban, all items (log-differenced)	1
EXP	Consumer expectation (1966.1 = 100)	0
EXPD	Changes in consumer expectation	0
HOUSE	New private housing units authorized (log-differenced)	1
VENDOR	Vendor performance (slower deliveries diffusion index, %)	0
INCOME	Personal income less transfer payments (log-differenced)	2
IP	Industrial production (log-differenced)	1
SALES	Manufacturing & trade sales (log-differenced)	1

Note: Information lag is measured at the end of month. Strictly speaking, those employment data with zero information lag and vendor performance are not available at the end of the month. However, they will be available at the very beginning of the next month and there are virtually no considerable lags. Thus, they are categorized in the zero information lag variable.

$y_t = \dots = y_{t-11} = y_{t-12}$, that is, the economy is in the same state as a year ago.

Following Clements and Galvão (2006), I will use three different measures for evaluating accuracy of out-of-sample recession probability predictions. The first measure is the probability-analogue of mean squared error, the quadratic probability score (QPS), which is commonly used in evaluating probability forecasts. The QPS is defined as

$$QPS = \frac{2}{T} \sum_{t=1}^T (\hat{p}_t - y_t)^2, \quad (3.3)$$

where $\hat{p}_t$ is the recession probability forecast month t . The QPS takes values between 0 and 2 and smaller value indicates more accurate forecasts.

The second measure is the log probability score (LPS), which is defined as

$$LPS = -\frac{1}{T} \sum_{t=1}^T \{y_t \log(\hat{p}_t) + (1 - y_t) \log(1 - \hat{p}_t)\}. \quad (3.4)$$

The LPS ranges from 0 to $+\infty$ and a smaller value corresponds to more accurate predictions and penalizes larger mistakes more heavily than the QPS.

The last measure is the Kuipers Score (KS), which is given by

$$\begin{aligned} KS &= \frac{\sum_{t=1}^T y_t 1_{(\hat{p}_t > 0.5)}}{\sum_{t=1}^T y_t} - \frac{\sum_{t=1}^T (1 - y_t) 1_{(\hat{p}_t > 0.5)}}{\sum_{t=1}^T (1 - y_t)}, \\ &= \text{hit rate} - \text{false rate}, \end{aligned} \quad (3.5)$$

where $1_{(\cdot)}$ is an indicator function that equals 1 if its argument is true and 0 otherwise. The KS calculates the difference between the hit rate and the rate of false signals by using 50% probability of recession as a cutoff. The KS takes values between -1 and 1 . A score of 1 corresponds to making perfect predictions. The KS evaluates the recession predictions from a slightly different aspect, compared with other two measures. Even when recession predictions never show “strong” indications (say, higher than 50% probability), it is possible to have seemingly good results based on the QPS and LPS. The KS discounts such “weak” predictions. In this sense, the KS captures the strength and accuracy of predictions by using the 50% probability cutoff.

There is a potential problem of just relying on the recursive out-of-sample exercises described above, especially in the context of recession probability forecasting. Since there are not many recessions in the out-of-sample period, it is possible to select a forecasting model that has particularly good performance for the last two recessions, but not for other recessions or a next recession.

In order to robustify the results, I will also carry out a cross-validation type exercise, called leaving 2-years out. The detailed procedures of the leaving 2-years out exercises are as follows. Let $S = \{(y_t, \mathbf{x}_{t-h}) : t = 1, \dots, T\}$ denote a full sample and $L_\tau = \{(y_t, \mathbf{x}_{t-h}) : t = \tau - 12, \dots, \tau + 12\}$ represent a set of excluding observations. For each $\tau = 13, \dots, T - 12$,

- (i) Take $E_\tau = S \setminus L_\tau$ as a training sample.
- (ii) Estimate parameter values β_τ based on E_τ .
- (iii) Make a prediction for y_τ by using β_τ and $\mathbf{x}_{\tau-h}$ and store it.
- (iv) Repeat steps (i) – (iii).

Then calculate QPS, LPS, and KS based on $\{\hat{y}_\tau\}_{\tau=13}^{T-12}$.

3.2.4 Results

Univariate Probit Models

First, we will start off by looking at forecasting performance within a class of univariate probit models of predicting 6-month-ahead recession probabilities ($h = 6$) as a benchmark. Table 3.3 summarizes the out-of-sample forecasting performance of univariate probit models.

The rankings of out-of-sample forecasting accuracy based on the QPS and the LPS have similar patterns, with a few exceptions. The term spread between 10-year and 3-month Treasury bonds (TS10Y3M), which is commonly used, and the term spread between 10-year bond and the Federal Funds rate (TS10YFF) are the best predictors based on the QPS and LPS, respectively. However, the term spread between 10-year and 5-year bonds (TS10Y5Y) has poorer out-of-sample performance in the both measures. In contrast to the term spread, the credit spreads (CSA, CSAA, and CSAAA) have poor forecasting performance. Among other variables, the levels of interest rates (FF and 3M) and the composite leading indicators (CLI11) might be useful predictors as well. However, employment related variables, those variables that the NBER Business Cycle Dating Committee is paying attention to, and monetary aggregates have considerably poorer results.

It is interesting to point out that, in terms of quadratic loss, the conventional term spread (TS10Y3M) outperforms the one based on the Federal Funds rate. However,

Table 3.3: Variable Rankings with Univariate Probit Models

QPS Ranking		LPS Ranking		KS Ranking	
TS10Y3M	0.1399	TS10YFF	0.2396	M2	0.0179
TS10YFF	0.1444	TS10Y3M	0.2401	TS10YFF	0.0000
3M	0.1488	FF	0.2607	TS10Y3M	0.0000
FF	0.1503	3M	0.2637	FF	0.0000
CLI11	0.1541	CLI11	0.2733	3M	0.0000
CPI	0.1579	SP500	0.2877	CPI	0.0000
5Y	0.1610	CPI	0.2964	5Y	0.0000
SP500	0.1627	5Y	0.2993	10Y	0.0000
10Y	0.1650	10Y	0.3081	A	0.0000
A	0.1663	A	0.3094	AAA	0.0000
AAA	0.1665	AAA	0.3098	AA	0.0000
AA	0.1669	AA	0.3108	UICLAIM	0.0000
UICLAIM	0.1695	UICLAIM	0.3131	EXPD	0.0000
CSAAA	0.1712	DJ30	0.3146	IP	0.0000
TS10Y5Y	0.1725	EXPD	0.3216	UNEMPD	0.0000
EXPD	0.1733	IP	0.3227	SALES	0.0000
IP	0.1736	UNEMPD	0.3237	UNEMP	0.0000
CSAA	0.1740	TS10Y5Y	0.3241	EMP	0.0000
UNEMPD	0.1742	SALES	0.3277	HOURS	0.0000
UNEMP	0.1759	UNEMP	0.3284	M0	0.0000
HOURS	0.1765	EMP	0.3284	CEMP	0.0000
M0	0.1767	HOURS	0.3292	HOUSE	0.0000
DJ30	0.1768	M0	0.3302	VENDOR	0.0000
CEMP	0.1769	CEMP	0.3303	M1	0.0000
SALES	0.1775	HOUSE	0.3325	CSA	0.0000
EMP	0.1778	VENDOR	0.3349	CSAA	0.0000
HOUSE	0.1782	EXP	0.3362	CSAAA	0.0000
VENDOR	0.1791	INCOME	0.3480	CLI11	-0.0054
CSA	0.1808	M1	0.3494	EXP	-0.0054
EXP	0.1854	CSA	0.3511	INCOME	-0.0108
INCOME	0.1923	CSAA	0.3694	DJ30	-0.0161
M1	0.1927	M2	0.4214	TS10Y5Y	-0.0215
M2	0.2498	CSAAA	0.4261	SP500	-0.0323

if we penalize larger mistakes more heavily, then the out-of-sample forecasting performance improves by using TS10YFF. That is, TS10Y3M minimizes the quadratic loss, but it results some larger mistakes. The same thing is true for the levels of FF and 3M.

The out-of-sample forecasting performance evaluated by using the KS gives us a completely different picture. Most predictor variables have non-positive values for the KS. In other words, they usually do not give us strong predictions about the occurrence of future recessions or correct predictions are largely offset by false predictions. In the worst case, we will get more false signals, which indicate more than 50% probability of a recession during an expansion, than correct ones. The only variable that has a positive value for the KS is the growth rate of M2. However, other measures of forecasting performance suggest that the performance of M2 is one of the poorest, among predictor variables considered. The opposite can also happen. Although CLI11 has relatively good performance based on other measures, it has a negative value for the KS.

Figure 3.1 presents in-sample (i.e., the entire sample) and out-of-sample 6-month-ahead predictions from univariate probit models with TS10YFF and TS10Y3M. The shaded areas indicate the recession periods. The top panel shows in-sample fit. As can be seen, the term spreads well capture recessions and expansions in early periods, especially those in the 1970's and 1980's. However, it seems that they do not provide strong predictions for the last two recessions.⁸ Particularly, for the 1990-91 recession, the predictions made by the term spreads miss the timing of the recession and they do not have strong signals. This difficulty is also noted in other studies.⁹

The same is true for the out-of-sample forecasts, which are shown at the bottom panel of Figure 3.1. Although both the QPS and LPS select TS10YFF and TS10Y3M as the first two univariate probit models, actual predictions are relatively poor and never exceed 50%. Furthermore, as in the in-sample case, out-of-sample predictions also miss the timing of the 1990-91 recession. This illustrates importance of paying attention to the KS, according to which both models score zero.

⁸In conjunction with the Great Moderation, some researchers find instability of the predictive power of the term spread (for example, Chauvet and Potter, 2005). However, Estrella, Rodrigues, and Schich (2003) do not find evidence of a structural break in the binary response model.

⁹See Stock and Watson (2003) for a more detailed discussion.

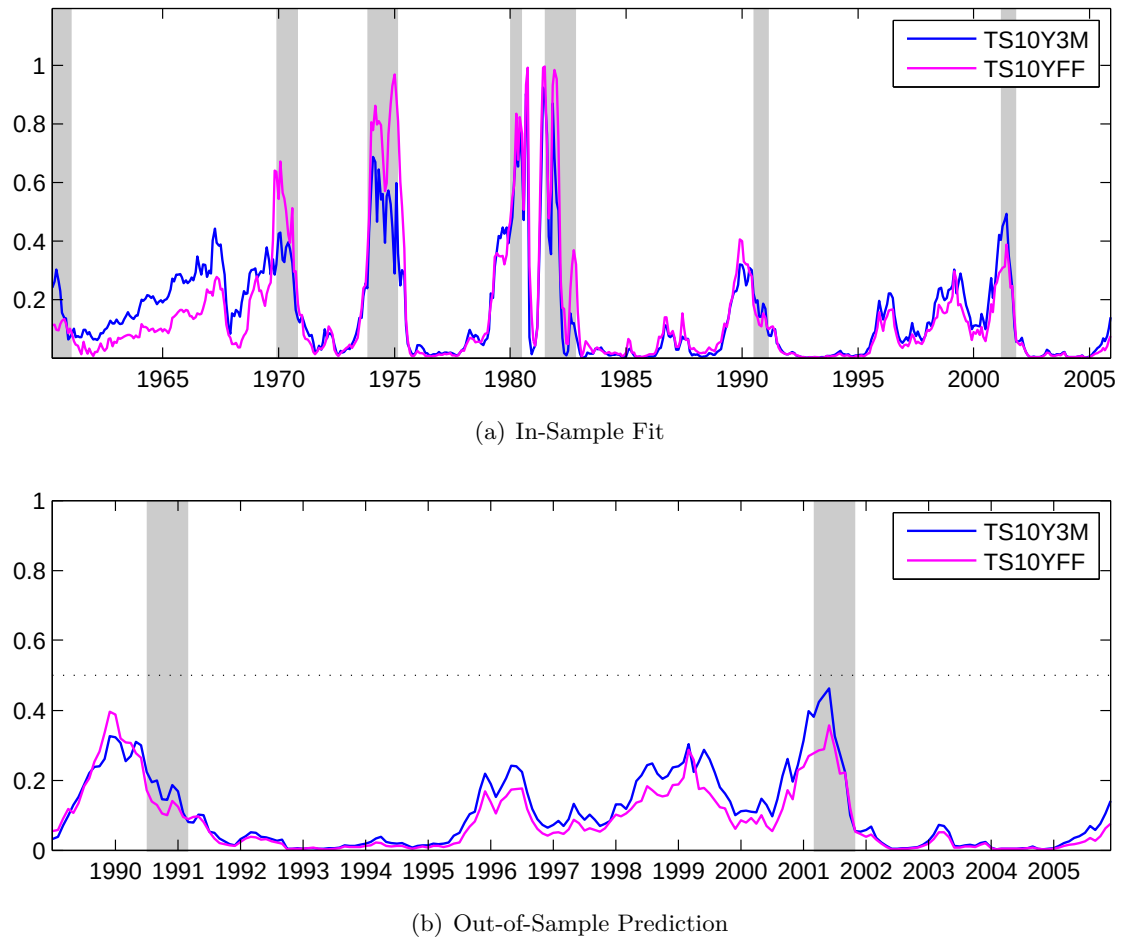


Figure 3.1: Predictions with Univariate Probit Models

Note: The shaded areas represent the NBER recessions. Horizontal axes measures recession probabilities.

Univariate Models

Now we will look at the performance of univariate models based on alternative CDFs, which allow positive skewness and/or excess kurtosis. Table 3.4 summarizes the ranking of univariate models based on out-of-sample fit.

In terms of the QPS, as in the probit model, models with TS10Y3M provide better out-of-sample forecasting accuracy than those with TS10YFF. For TS10Y3M, there is a clear indication of a better prediction for CDFs with excess kurtosis and the Laplace CDF achieves the best QPS. On the other hand, allowing for skewness deteriorates forecasting accuracy, such as GEV3 and Gumbel. However, such a pattern

Table 3.4: Top 20 Univariate Models

QPS Ranking			LPS Ranking			KS Ranking		
CDF	x_1	QPS	CDF	x_1	LPS	CDF	x_1	KS
Laplace	TS10Y3M	0.1360	Gumbel	TS10YFF	0.2362	Laplace	TS10Y3M	0.1111
Student-t	TS10Y3M	0.1388	GEV3	TS10YFF	0.2370	Laplace	CLI11	0.0448
Logistic	TS10Y3M	0.1388	Laplace	TS10Y3M	0.2392	Laplace	SP500	0.0233
Normal	TS10Y3M	0.1399	Normal	TS10YFF	0.2396	GEV3	SP500	0.0233
GEV3	TS10Y3M	0.1410	Normal	TS10Y3M	0.2401	Gumbel	SP500	0.0233
Gumbel	TS10Y3M	0.1421	GEV3	TS10Y3M	0.2401	Normal	M2	0.0179
Normal	TS10YFF	0.1444	Logistic	TS10Y3M	0.2406	Gumbel	TS10YFF	0.0000
GEV3	TS10YFF	0.1446	Gumbel	TS10Y3M	0.2408	GEV3	TS10YFF	0.0000
Logistic	TS10YFF	0.1453	Student-t	TS10Y3M	0.2410	Normal	TS10YFF	0.0000
Student-t	TS10YFF	0.1455	Logistic	TS10YFF	0.2430	Normal	TS10Y3M	0.0000
Gumbel	TS10YFF	0.1459	Student-t	TS10YFF	0.2436	GEV3	TS10Y3M	0.0000
Laplace	TS10YFF	0.1459	Laplace	TS10YFF	0.2463	Logistic	TS10Y3M	0.0000
Gumbel	3M	0.1482	Gumbel	FF	0.2564	Gumbel	TS10Y3M	0.0000
GEV3	3M	0.1484	GEV3	FF	0.2582	Student-t	TS10Y3M	0.0000
Normal	3M	0.1488	Normal	FF	0.2607	Logistic	TS10YFF	0.0000
Gumbel	FF	0.1491	Gumbel	3M	0.2615	Student-t	TS10YFF	0.0000
Logistic	3M	0.1494	GEV3	3M	0.2623	Laplace	TS10YFF	0.0000
Student-t	3M	0.1495	Normal	3M	0.2637	Gumbel	FF	0.0000
GEV3	FF	0.1496	Logistic	FF	0.2640	GEV3	FF	0.0000
Normal	FF	0.1503	Student-t	FF	0.2644	Normal	FF	0.0000

does not hold for TS10YFF and, in fact, the Normal CDF outperforms other CDFs among univariate models with TS10YFF.

If we penalize larger mistakes more heavily than the QPS, the variable ranking based on the LPS suggests some roles played by allowing skewness and excess kurtosis. For TS10YFF, allowing positive skewness and excess kurtosis contributes to reducing large mistakes, compared with the Normal. However, just incorporating excess kurtosis, but not skewness, appears to perform worse than the Normal CDF and also than univariate models with TS10Y3M. Although, within the probit models, TS10Y3M is inferior to TS10YFF, using the Laplace CDF improves the LPS and achieves a better result than TS10YFF with the Normal. It is possible to obtain better forecasting performance by employing non-Normal CDFs. However, it is not clear what additional features of the functional form are helpful and it seems to be variable specific.

The rankings based on the QPS and LPS suggest that information contained in 3M helps reduce quadratic loss, whereas information contained in FF tends to reduce larger mistakes.

In general, most of predictors do not make correct and strong predictions, in terms of the 50% probability cutoff. Only 6 models have positive values for the KS. Especially, the Laplace with TS10Y3M has superior out-of-sample forecasting performance, compared with other univariate models. Although based on other measures, including skewness and/or excess kurtosis improves out-of-sample fit slightly, using the Laplace CDF for TS10Y3M makes significant difference in the KS (increasing from 0 to 0.1111). It is also important to point out that other predictor variables that do not appear in the other rankings are ranked relatively higher in this KS ranking, namely CLI11, SP500 and M2.

Bivariate Models

Now we look at forecasting results based on bivariate models, in order to see the importance of additional information. Tables 3.5 – 3.7 show the top 20 bivariate forecasting models based on the QPS, LPS, and KS, respectively.

The combination of TS10Y3M and SP500 is the best predictor variables based on the QPS and LPS. All three measures of out-of-sample accuracy suggest that it is important to include a measure of term spread (either TS10Y3M or TS10YFF). Interestingly, there are some variables that have relatively poor performance in univariate models and produce good out-of-sample forecasting accuracy, *together with* the term spread measure. According to the rankings based on the univariate models, SP500 and EMP are not a useful single predictor. Especially, EMP has relatively poor out-of-sample forecasting performance. However, they become an important companion variable to the term spread. In fact, those variables with the term spread are better than a combination of two term spreads (TS10Y3M and TS10YFF), which are in the best two variables among univariate models. This suggests that those variables contain some useful information that the term spread does not have and also indicates that the univariate ranking is not a helpful guide for choosing multiple predictors.

The importance of SP500 together with TS10Y3M is consistent with the finding of Estrella and Mishkin (1998). In our bivariate models, even without being combined

Table 3.5: Top 20 Bivariate Models [QPS] Table 3.6: Top 20 Bivariate Models [LPS]

CDF	x_1	x_2	QPS	CDF	x_1	x_2	LPS
Laplace	TS10Y3M	SP500	0.1144	GEV3	TS10Y3M	SP500	0.1984
Student-t	TS10Y3M	SP500	0.1144	Normal	TS10Y3M	SP500	0.1985
Logistic	TS10Y3M	SP500	0.1144	Gumbel	TS10Y3M	SP500	0.1993
Normal	TS10Y3M	SP500	0.1151	Logistic	TS10Y3M	SP500	0.1994
GEV3	TS10Y3M	SP500	0.1160	Student-t	TS10Y3M	SP500	0.1997
Gumbel	TS10Y3M	SP500	0.1173	Laplace	TS10Y3M	SP500	0.2020
GEV3	TS10YFF	SP500	0.1226	Gumbel	TS10YFF	SP500	0.2027
Student-t	TS10YFF	SP500	0.1228	GEV3	TS10YFF	SP500	0.2037
Logistic	TS10YFF	SP500	0.1228	Normal	TS10YFF	SP500	0.2064
Gumbel	TS10YFF	SP500	0.1230	Logistic	TS10YFF	SP500	0.2084
Normal	TS10YFF	SP500	0.1230	Student-t	TS10YFF	SP500	0.2087
Laplace	TS10YFF	SP500	0.1236	Laplace	TS10YFF	SP500	0.2112
Laplace	TS10Y3M	EMP	0.1254	Gumbel	FF	SP500	0.2113
Logistic	TS10Y3M	EMP	0.1268	Gumbel	3M	SP500	0.2129
Student-t	TS10Y3M	EMP	0.1268	GEV3	FF	SP500	0.2146
Normal	TS10Y3M	EMP	0.1277	GEV3	3M	SP500	0.2149
GEV3	TS10Y3M	EMP	0.1283	GEV3	TS10Y3M	EMP	0.2152
Gumbel	3M	SP500	0.1290	Gumbel	TS10Y3M	EMP	0.2154
Gumbel	TS10Y3M	EMP	0.1295	Normal	TS10Y3M	EMP	0.2166
Gumbel	FF	SP500	0.1298	Logistic	TS10Y3M	EMP	0.2172

with the term spread, some bivariate models that contain SP500 perform relatively well. However, it should be mentioned that King, Levin, and Perli (2007) do not find superiority of SP500 in conjunction with the term spread. Rather, they report that a combination of variables, which have better performance in univariate models, also perform better in bivariate models.¹⁰

Unlike the univariate models, combining SP500 with TS10Y3M always achieves the better performance both in terms of the QPS and LPS than those models with TS10YFF. For the combination of TS10Y3M and SP500, allowing only excess kurtosis

¹⁰This could be because of a couple of reasons. First, it could be attributed to the difference in periods used for the out-of-sample forecasting exercises and forecasting horizon. Second, it might be because of the fact that my data set does not include the credit spread measures that they use and perform quite well in their univariate models. Finally, it could be due to the difference in evaluating out-of-sample forecasting performance. They look at average out-of-sample predictions over two test periods, the 2001 recession and the post-2001 expansion.

Table 3.7: Top 20 Bivariate Models [KS]

CDF	x_1	x_2	KS
Laplace	TS10Y3M	EMP	0.2724
GEV3	TS10Y3M	SP500	0.2115
Normal	TS10Y3M	SP500	0.2115
Gumbel	TS10Y3M	SP500	0.2115
Logistic	TS10Y3M	SP500	0.2115
Student-t	TS10Y3M	SP500	0.2115
Laplace	TS10Y3M	SP500	0.2115
Laplace	UICLAIM	SP500	0.1900
Laplace	TS10Y3M	EXPD	0.1667
Laplace	TS10Y3M	VENDOR	0.1667
Laplace	TS10Y3M	CSA	0.1613
Laplace	TS10Y3M	M1	0.1559
Gumbel	UICLAIM	SP500	0.1344
GEV3	UICLAIM	SP500	0.1344
Normal	UICLAIM	SP500	0.1344
Student-t	UICLAIM	SP500	0.1344
Logistic	UICLAIM	SP500	0.1344
Laplace	TS10Y3M	M2	0.1344
Gumbel	EMP	SP500	0.1290
Laplace	IP	SP500	0.1290

helps improve the QPS, while including both skewness and excess kurtosis contributes to improving the LPS and KS. Allowing both skewness and excess kurtosis, in general, improves the LPS by reducing large false predictions. However, there is some mixed evidence for the QPS in terms of the role of skewness and excess kurtosis. At least, for the models with TS10Y3M, it seems better to allow only excess kurtosis, but not positive skewness.

Although forecasting performance generally improves by adding one more predictor, probably the biggest gain appears in the KS. In univariate models, only 6 models out of 198 have positive scores. In bivariate models, 348 models out of 3168 have at least positive values for the KS. Furthermore, the best model (TS10Y3M and EMP with the Laplace) scores more than twice as high as the best univariate model.

Table 3.8: Top 20 Models [QPS]

CDF	x_1	x_2	x_3	QPS
Laplace	TS10Y3M	SP500	EMP	0.1051
Logistic	TS10Y3M	SP500	EMP	0.1074
Student-t	TS10Y3M	SP500	EMP	0.1074
Student-t	TS10Y3M	SP500	A	0.1079
Laplace	TS10Y3M	SP500	A	0.1080
Logistic	TS10Y3M	SP500	A	0.1081
Student-t	TS10Y3M	SP500	AAA	0.1084
Logistic	TS10Y3M	SP500	AAA	0.1086
Student-t	TS10Y3M	SP500	AA	0.1087
Logistic	TS10Y3M	SP500	AA	0.1088
Laplace	TS10Y3M	SP500	AAA	0.1088
Normal	TS10Y3M	SP500	EMP	0.1089
Laplace	TS10Y3M	SP500	AA	0.1090
GEV3	TS10Y3M	SP500	A	0.1092
Normal	TS10Y3M	SP500	A	0.1093
Laplace	TS10Y3M	SP500	IP	0.1093
GEV3	TS10Y3M	SP500	AAA	0.1098
Normal	TS10Y3M	SP500	AAA	0.1098
GEV3	TS10Y3M	SP500	AA	0.1100
GEV3	TS10Y3M	SP500	EMP	0.1100

Table 3.9: Top 20 Models [LPS]

CDF	x_1	x_2	x_3	LPS
GEV3	TS10Y3M	SP500	A	0.1832
Gumbel	TS10Y3M	SP500	A	0.1834
GEV3	TS10Y3M	SP500	AAA	0.1843
GEV3	TS10Y3M	SP500	AA	0.1845
Gumbel	TS10Y3M	SP500	AAA	0.1847
Gumbel	TS10Y3M	SP500	AA	0.1848
Gumbel	5Y	SP500	3M	0.1848
Student-t	TS10Y3M	SP500	A	0.1852
Logistic	TS10Y3M	SP500	A	0.1852
Normal	TS10Y3M	SP500	A	0.1853
GEV3	TS10Y3M	SP500	EMP	0.1857
GEV3	5Y	SP500	3M	0.1857
Gumbel	TS10Y3M	SP500	10Y	0.1858
Gumbel	TS10Y3M	SP500	3M	0.1858
Gumbel	10Y	SP500	3M	0.1858
Gumbel	TS10Y3M	SP500	EMP	0.1859
Student-t	TS10Y3M	SP500	AAA	0.1861
Logistic	TS10Y3M	SP500	AAA	0.1861
Normal	TS10Y3M	SP500	AAA	0.1861
GEV3	TS10Y3M	SP500	10Y	0.1862

Trivariate Models and Overall Rankings

Now we turn our attention to the overall forecasting performance, including all of univariate, bivariate, and trivariate models with 6 different CDFs. Tables 3.8 – 3.10 list the top 20 forecasting models out of 36102 forecasting models, based on different criteria.

The Laplace with TS10Y3M, SP500, and EMP is the best forecasting model based on the QPS and KS. Regardless of the functional form used, this variable combination records relatively good out-of-sample performance. This is consistent with the bivariate results that SP500 and EMP perform relatively better in conjunction with the term spread measure.

However, the combination of TS10Y3M, SP500, and EMP tend to make larger

Table 3.10: Top 20 Models [KS]

CDF	x_1	x_2	x_3	KS
Laplace	TS10Y3M	SP500	EMP	0.3835
Normal	TS10Y3M	SP500	EMP	0.3333
Logistic	TS10Y3M	SP500	EMP	0.3280
Student-t	TS10Y3M	SP500	EMP	0.3280
Logistic	TS10Y3M	SP500	IP	0.3280
Student-t	TS10Y3M	SP500	IP	0.3280
Laplace	TS10Y3M	SP500	IP	0.3280
GEV3	TS10Y3M	SP500	EMP	0.3172
Laplace	UNEMP	SP500	A	0.3172
Laplace	UNEMP	SP500	AA	0.3172
Gumbel	TS10Y3M	SP500	EMP	0.3118
GEV3	TS10Y3M	M2	CSA	0.2957
Gumbel	TS10Y3M	M2	CSA	0.2957
Laplace	TS10Y3M	EMP		0.2724
GEV3	TS10Y3M	SP500	UNEMPD	0.2724
Normal	TS10Y3M	SP500	UNEMPD	0.2724
Logistic	TS10Y3M	SP500	UNEMPD	0.2724
Gumbel	TS10Y3M	SP500	UNEMPD	0.2724
Student-t	TS10Y3M	SP500	UNEMPD	0.2724
Laplace	TS10Y3M	SP500	UNEMPD	0.2724

prediction errors, compared to the combination that includes corporate bond yields on A-rated firms (A), instead of EMP. The combination of TS10Y3M, SP500, and A is the best in terms of the LPS. It is also the second best in the QPS ranking. However, this variable combination does not result in good performance in terms of the KS.

It is worthwhile to point out the importance of TS10Y3M and SP500, regardless of the forecasting accuracy measures. Thus, improving out-of-sample forecasting accuracy could be a task of choosing a good companion variable to this combination. In this sense, other corporate bond yields (AA and AAA) are also good companion variables. Although in the bivariate models, the term spread measure TS10YFF was one of the best predictors together with SP500, augmenting the combination of TS10Y3M and SP500 with other variables outperforms the TS10YFF counterparts.

Except for the combination of TS10Y3M, SP500, and EMP, the variable combinations that are ranked relatively high in the KS ranking do not have good out-of-sample forecasting performance in terms of the QPS or LPS. In other words, making

Table 3.11: Importance of the Two Margins

Models	QPS			LPS			KS		
	Normal	Laplace	GEV3	Normal	Laplace	GEV3	Normal	Laplace	GEV3
Univariate	0.1399	0.1360	0.1410	0.2401	0.2392	0.2401	0.0000	0.1111	0.0000
Bivariate	0.1151	0.1144	0.1160	0.1985	0.2020	0.1984	0.2115	0.2115	0.2115
Trivariate	0.1089	0.1051	0.1100	0.1868	0.1870	0.1857	0.3333	0.3835	0.3172

Note: The univariate model includes TS10Y3M. The bivariate model refers to models with TS10Y3M and SP500 and the trivariate models add EMP to the bivariate models.

strong predictions during recessions involves risks of making positive false signals during expansions, which worsens the QPS and LPS. In out-of-sample forecasting exercises, there is some favorable evidence for allowing excess kurtosis to improve the QPS and KS, whereas adding positive skewness could improve forecasting accuracy in terms of the LPS.

Discussion

Forecasting performance based on all of models indicates that the term spread measure (TS10Y3M) is one of the most important variables, as widely documented in other studies. However, there are two margins to improve forecasting performance. One is to include additional predictors. By combining different information, we can increase accuracy of recession probability predictions. As demonstrated, a particular combination of variables greatly outperforms a typical univariate probit model with TS10Y3M. Based on the forecasting performance of bivariate models, the importance of the stock price index (SP500) is emerged as a companion variable to the term spread measure. In addition, among the all combinations of variables considered in this paper, augmenting the term spread measure by SP500 and EMP results in the best out-of-sample forecasting performance. Another margin is to change a functional form of $F(\cdot)$, in order to improve how to translate signals to a recession probability measure.

It is important to understand how these two margins contribute to better forecasting performance. Table 3.11 summarizes how forecasting performance improves by changing the two margins and illustrates different effects of the margins. We compare probit models with the Laplace and GEV3 counterparts. In general, including additional predictors has greater impact on improving all three out-of-sample forecasting measures. However, it is possible to obtain non-negligible gains from changing the functional form.

For example, starting from the univariate probit model with TS10Y3M, provided that we know SP500 as the best companion variable to the term spread, adding SP500 drastically reduces the QPS and LPS by 17.7% and 17.3%, respectively. Furthermore, the KS increases from 0 to 0.2115. Using the Laplace in the univariate model moderately improves the QPS and results in a large gain in the KS (from 0 to 0.1111), whereas it deteriorates the LPS.

Although marginal gains become smaller, adding the third variable (EMP) further improves the out-of-sample forecasting performance of the probit model. Using the Laplace improves the QPS and KS of the trivariate model. While only allowing for the excess kurtosis (Laplace) cannot improve the LPS, allowing both skewness and excess kurtosis (GEV3) can improve the LPS, compared with the probit model, however. In sum, the QPS and KS can be improved by using the Laplace, whereas using the GEV3 can improve the LPS.

The top panel of Figure 3.2 compares out-of-sample predictions from the univariate probit model with the term spread with the trivariate probit model (TS10Y3M, SP500, and EMP). As can be seen, including additional variables reduces forecasting errors during expansions. Especially, predictions from the univariate model suggest moderate possibilities of recessions during the last half of the 1990's, when the economy was in fact in the longest expansion. Those positive predictions are dampened by including additional variables. At the same time, it contributes to making stronger predictions during recessions, especially for the last recession. However, in exchange for such improvements, the trivariate model also contains a couple of relatively strong false positive signals during expansions, which are not present in the univariate model.

The middle panel of Figure 3.2 plots differences in predictions between the trivariate Laplace model and the trivariate probit model. This suggests some benefits from changing functional form. Using the Laplace, instead of the Normal, will amplify positive signals during the 2001 recession and also will dampen strong false positive signals that happen right after the 1990-91 recession and before and after the 2001 recession. In general, using the Laplace will contribute to amplifying strong signals and to dampening weak signals, so that on average there will be less forecasting errors. However, such a feature works in an unfavorable way during the 1990-91 recession, since the underlying predictions are somewhat weak and miss the correct timing.

The bottom panel of Figure 3.2 plots differences in predictions between the

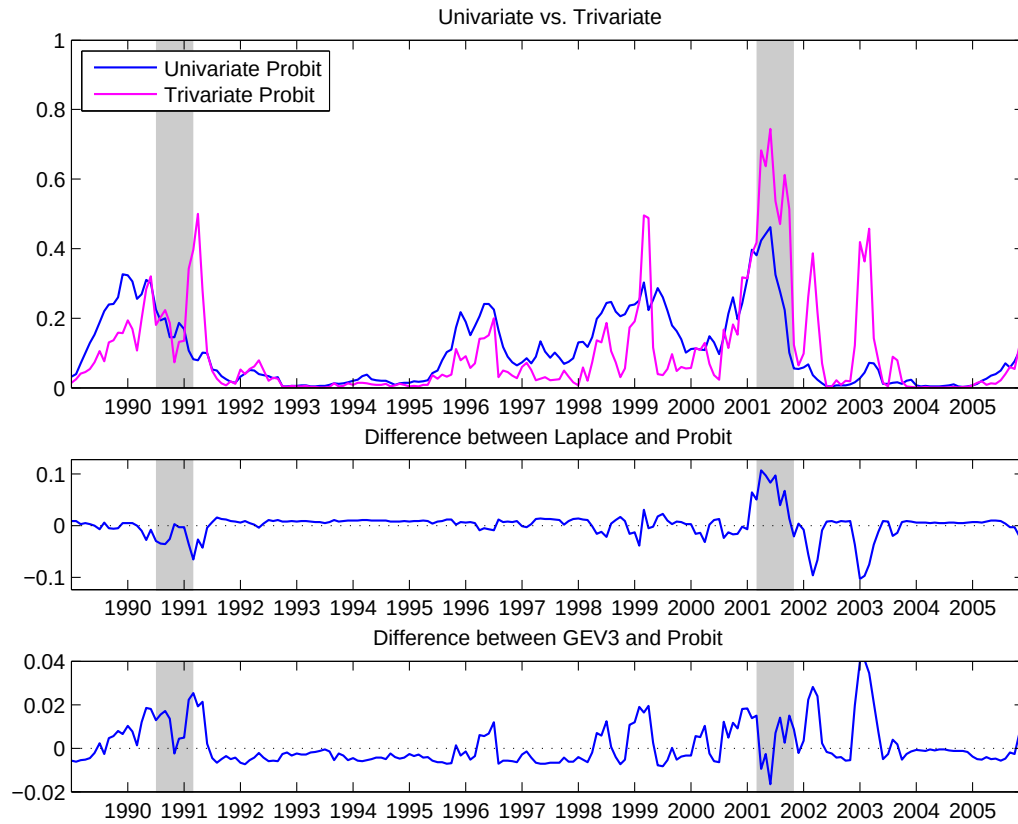


Figure 3.2: Comparison of Out-of-Sample Predictions

Note: Shaded areas indicate the NBER recession episodes. The top panel compares predictions from the univariate probit model with TS10Y3M with those from the trivariate models with TS10Y3M, SP500, and EMP. The middle and bottom panels show differences in predictions by changing the functional form. The middle panel plots the difference between the trivariate Laplace and the trivariate probit and the bottom panel depicts the difference between the trivariate GEV3 and the trivariate probit.

trivariate GEV3 model and the trivariate probit model. Overall, differences between the GEV3 model and the probit model are somewhat small. Unlike the Laplace model, using the GEV3 improves predictions for the 1990-91 recession. However, early periods of the 2001 recession have weaker predictions than the probit model. Although using the GEV3 will reduce false positive predictions during expansions, strong false positive signals are, in fact, amplified.

Although there are some trade-offs in using different functional forms, as indicated in Table 3.11, additional features in a forecasting model can improve out-of-sample forecasts.

Robustness Check

As a robustness check, I also evaluate forecasting performance of all models using the leaving 2-years out exercises. Differences in the results that the leaving 2-years out typically selects TS10YFF as a measure of a term spread, instead of TS10Y3M. Figure 3.3 compares forecasting performance of models including TS10Y3M and TS10YFF in the both exercises.

In general, those models that include TS10Y3M work relatively better in the out-of-sample forecasting exercises, whereas those with TS10YFF have better performance in the leaving 2-years out exercises. A primary reason is that models with TS10YFF make smaller prediction errors in the earlier periods, which are included in the leaving 2-years out exercises and are not covered in the recursive out-of-sample forecasting exercises. As shown in the top panel of Figure 3.1, after the 1980's, both term spreads show similar predictions. However, in earlier periods, predictions based on TS10Y3M are associated with larger forecasting errors. Although it is clear that including a measure of a term spread is important, relative performance could depend on which term spread is used.

Regardless of the choice of the term spread measure, the importance of SP500 and EMP is quite robust in both exercises. Figure 3.4 graphically compares forecasting accuracy based on the recursive out-of-sample predictions with the one based on the leaving 2-years out. The left panels of Figure 3.4 summarize forecasting performance based on the recursive out-of-sample predictions and the right panels plot the one based on the leaving 2-years out. Within each exercise, the top three panels categorize three measures of forecasting accuracy by variable groups and the bottom three panels are based on functional forms. Each panel of Figure 3.4 covers roughly the top 10% forecasting models. For the top three panels, markers associated with a particular value of the out-of-sample accuracy measure (horizontal axis) indicate that the corresponding forecasting model contains variables listed on the vertical axis. For the bottom three panels, the vertical axis indicates which functional form is used.

With the leaving 2-years out exercises, the variable combination of TS10YFF, SP500, and EMP achieves the best results in all three measures of forecasting accuracy. Although the out-of-sample forecasting results suggest that it is helpful to use the corporate bond yield on A-rated firms (A) to improve the LPS, the results based on the

leaving 2-years out do not support the importance of A. Despite its good performance in the out-of-sample predictions that covers the last two recessions, the leaving 2-years out results suggest the predictability of A together with the term spread and SP500 might be somewhat fragile.

In both exercises, there are not many trivariate models that outperform the best bivariate model. However, there are non-trivial gains from including the third predictor variable. In terms of the functional form, it appears that the Laplace CDF contributes to improving the QPS and the KS in both out-of-sample and leaving 2-years out exercises, regardless of the number of variables included. On the other hand, there is favorable evidence for introducing both skewness and excess, in order to improve the LPS.

3.3 Allowing More Flexible Skewness and Excess Kurtosis

3.3.1 Generalized Edgeworth Expansion

Results in the previous section suggest some benefits of allowing for positive skewness or excess kurtosis, in order to improve recession probability forecasts. However, there is some mixed evidence on what features in the functional form are important, compared with a typical probit model. It is possibly because of the fact that I just examine a relatively restrictive set of combinations of skewness and excess kurtosis. In this section, I will examine the effect of allowing for more flexibility in the functional form.

In particular, I will allow for more flexibility in skewness and excess kurtosis. In order to do so, I will utilize the generalized Edgeworth series expansion, which is introduced by Jarrow and Rudd (1982).¹¹ The generalized Edgeworth expansion (GEE) will approximate the “true” distribution $F(s)$ by using a series expansion involving higher moments of the approximating distribution $A(s)$. I assume that both $F(s)$ and $A(s)$ are continuous, that is,

$$\frac{dF(s)}{ds} = f(s) \quad \text{and} \quad \frac{dA(s)}{ds} = a(s) \quad (3.6)$$

exist. Furthermore, I assume that the moments of $F(s)$, $\alpha_j(F)$, exist for $j \leq n$. Given $\alpha_n(F)$ exists, the first $n - 1$ cumulants $\kappa_j(F)$ from $j = 1, \dots, n - 1$ also exist. Jarrow

¹¹They use the generalized Edgeworth expansion to study the problem of option valuation, where the underlying security distribution is unknown.

and Rudd (1982) show that

$$f(s) \approx a(s) + \sum_{j=1}^N E_j \frac{(-1)^j}{j!} \frac{d^j a(s)}{ds^j}, \quad (3.7)$$

where E_j is a function of the cumulants of $F(s)$ and $A(s)$ up to the j th cumulant, which will be given below. By taking integral of (3.7), the “true” $F(x)$ can be approximated by:

$$F(x) \approx \int_{-\infty}^x a(s) ds + \int_{-\infty}^x \left[\sum_{j=1}^N E_j \frac{(-1)^j}{j!} \frac{d^j a(s)}{ds^j} \right] ds. \quad (3.8)$$

Since I am interested in having more flexibility in skewness and excess kurtosis, it is sufficient to set $N = 4$. The third and fourth cumulants are related to skewness and excess kurtosis, respectively, as

$$\kappa_3 = \gamma_1(\kappa_2)^{3/2} \quad \text{and} \quad \kappa_4 = \gamma_2(\kappa_2)^2, \quad (3.9)$$

where γ_1 and γ_2 denote skewness and excess kurtosis, respectively, and κ_2 denotes the 2nd cumulant which corresponds to the variance. Then, E_j for $j = 1, \dots, 4$ are given by:

$$E_1 = \kappa_1(F) - \kappa_1(A), \quad (3.10)$$

$$E_2 = \kappa_2(F) - \kappa_2(A) + E_1^2, \quad (3.11)$$

$$E_3 = \kappa_3(F) - \kappa_3(A) + 3E_1(\kappa_2(F) - \kappa_2(A)) + E_1^3, \quad (3.12)$$

$$E_4 = \kappa_4(F) - \kappa_4(A) + 4(\kappa_3(F) - \kappa_3(A))E_1 + 3(\kappa_2(F) - \kappa_2(A))^2 + 6E_1^2(\kappa_2(F) - \kappa_2(A)) + E_1^4. \quad (3.13)$$

Given a choice of the approximating distribution $A(x)$ (i.e., $\kappa_j(A)$ for $j = 1, \dots, 4$), E_j is a function of the cumulants of $F(x)$ from $\kappa_1(F)$ to $\kappa_j(F)$. For my purpose of using this approximation in the binary response model, location and scale parameters are needed to be fixed for normalization. Otherwise, the coefficient vector β cannot be uniquely estimated. Without loss of generality, I assume that $F(x)$ have zero mean and unit variance and also assume that $A(x)$ has zero mean, so that $\kappa_1(F) = \kappa_1(A)$ and $E_1 = 0$. Then the approximation equation becomes:

$$\tilde{F}(x) \equiv A(x) + \frac{E_2}{2!} \int_{-\infty}^x \frac{d^2 a(s)}{ds^2} ds - \frac{E_3}{3!} \int_{-\infty}^x \frac{d^3 a(s)}{ds^3} ds + \frac{E_4}{4!} \int_{-\infty}^x \frac{d^4 a(s)}{ds^4} ds. \quad (3.14)$$

Although it is easily shown that $\lim_{x \rightarrow \infty} \tilde{F}(x) = 1$ and $\lim_{x \rightarrow -\infty} \tilde{F}(x) = 0$, there is no guarantee that $\tilde{F}(x)$ is monotonically increasing. In order for (3.14) being monotonically

increasing, I need to impose a condition

$$a(x) + \sum_{j=2}^4 E_j \frac{(-1)^j}{j!} \frac{d^j a(s)}{ds^j} \geq 0 \quad \text{for all } x \in \mathbb{R}. \quad (3.15)$$

This condition (3.15) determines admissible combinations of skewness and excess kurtosis of $\tilde{F}(x)$. This admissible set depends on a choice of $A(x)$. In order to cover all CDFs considered in the previous section, I will use the Logistic distribution whose mean equals zero and variance is 1.2 as the approximating distribution $A(x)$. Figure 3.5 illustrates the admissible set of (γ_1, γ_2) with the Logistic distribution, together with skewness and excess kurtosis of other CDFs.¹² Since I cannot estimate the variance of $A(x)$ due to an identification problem, I will restrict my search of skewness and excess kurtosis to combinations within this admissible set in Figure 3.5.

3.3.2 The GEE Results

I will use (3.14) for $F(\cdot)$ in (3.1). Since E_3 and E_4 are functions of γ_1 and γ_2 , respectively, there are 2 additional parameters to be estimated. The exact expression of (3.14) is presented in the Appendix. This is a parsimonious way to cover a relatively large space of skewness and excess kurtosis.¹³

By using the GEE, I will estimate two trivariate models that are found to be the best forecasting models in the out-of-sample forecasting exercises. These two trivariate models include TS10Y3M and SP500. As the third variable, one includes EMP, which resulted in the best QPS and KS, and the other includes A, which resulted in the best LPS.

Figure 3.6 graphically compares functional forms estimated by the GEE with the Normal CDF. In-sample estimates of γ_1 and γ_2 based on the variable combination of TS10Y3M, SP500, and EMP are 1.2486 and 3.5795, respectively. Those estimates using

¹²Basically, the bottom corner of the admissible set depicted in Figure 3.5 moves up and the angle of the bottom corner becomes smaller as the variance increases. For example, with unit variance the area is slightly below the skewness and excess kurtosis of the Gumbel and with larger variance the admissible set passes by the Gumbel's combination.

¹³Instead of using the GEE, we could use more complicated CDFs that can cover a much wider set of skewness and excess kurtosis, such as the one based on the skewed generalized t distribution of Theodossiou (1998). See Hansen, McDonald, and Theodossiou (2007) for other flexible distributions. However, there are basically two problems. First, those functions have more parameters and moments are complicated functions of those parameters, so that it is relatively difficult to control the identification issue in the binary response model. Second, it is often the case that there are no closed form CDFs available, which could increase computational burden.

Table 3.12: Out-of-Sample Forecasting Performance with the Generalized Edgeworth Expansion

Variables	CDF	QPS	LPS	KS
TS10Y3M + SP500 + EMP	GEE	0.1067 [0.1401]	0.1852 [0.2358]	0.4229 [0.5898]
	Normal	0.1089 [0.1486]	0.1868 [0.2560]	0.3333 [0.4925]
	Laplace	0.1051 [0.1420]	0.1870 [0.2488]	0.3835 [0.5629]
	GEV3	0.1100 [0.1486]	0.1857 [0.2510]	0.3172 [0.5000]
TS10Y3M + SP500 + A	GEE	0.1059 [0.1412]	0.1818 [0.2357]	0.1667 [0.5075]
	Normal	0.1093 [0.1516]	0.1853 [0.2772]	0.1667 [0.4599]
	Laplace	0.1080 [0.1498]	0.1872 [0.2656]	0.1667 [0.4815]
	GEV3	0.1092 [0.1487]	0.1832 [0.2603]	0.1667 [0.4555]

Note: Results from the Normal, Laplace, and GEV3 counterparts from the previous section are reproduced for comparison. Values reported in brackets are the corresponding values obtained from the leaving 2-years out exercises.

TS10Y3M, SP500, and A, are 1.5588 and 3.0864, respectively. The GEE estimates from the both trivariate models suggest to include both positive skewness and excess kurtosis. The GEE makes smaller predictions around $x = -1$. While the GEE predictions will be more aggressive than the probit ones at somewhere between $x = -0.5$ and 0.5 , its prediction becomes modest around $x > 1$, given the same information.

Table 3.12 reports performance of predictions made by using the GEE. Regarding the out-of-sample forecasting performance, the GEE models always outperform the probit counterparts.¹⁴ Furthermore, the GEE models almost always have better out-of-sample forecasting performance than other models, except for the QPS with the model that includes EMP. In the previous section, there are no single non-Normal models that outperform the Normal counterpart for the variable combination of TS10Y3M, SP500, and EMP. For example, with TS10Y3M, SP500, and EMP, the Laplace model is better than the probit model in the QPS and KS, but not in the LPS. The GEV3 model is better in the LPS, but not in other two. However, the GEE model in fact beats the probit model in all three measures of out-of-sample forecasting performance.

In addition to the recursive out-of-sample forecasting exercises, I also perform the leaving 2-years out exercises as a robustness check. Table 3.12 also reports results from the leaving 2-years out exercises. As in the recursive out-of-sample exercises, results

¹⁴Although the value of the KS for the model with A, improvements in both the QPS and LPS conditional on the same value of the KS imply that the GEE model is superior than the probit counterpart.

based on the leaving 2-years out exercises also confirm the superiority of the GEE model over other models.

Figure 3.7 compares predictions made by the GEE models with those from the probit models. Predictions from the GEE models amplify the positive signals for the 2001 recession. Furthermore, false positive signals that arise from the trivariate model in 1999 and 2003 are effectively dampened by the GEE. However, for the 1990-91 recession, as in other models, the GEE models do not indicate occurrence of the recession correctly. This suggests that allowing flexibility in the functional form will effectively improve recession probability forecasts. Although there is a possibility to make false predictions worse, there are net gains in forecasting accuracy from using the GEE. Using more flexible functional form can be combined with other extensions to probit models, in order to achieve further improvements in forecasting accuracy.

3.3.3 Current Predictions by the GEE

Results based on the known CDFs suggest to use the forecasting model that includes TS10Y3M, SP500, and EMP, together with the Laplace CDF, which is robust both in out-of-sample and leaving 2-years out forecasting exercises. As we have seen in this section, we can obtain further gains in forecasting accuracy, by using the GEE. It may be of interest to use the GEE model to infer the recent state of the U.S. economy.

Figure 3.8 shows current predictions of the U.S. recession probabilities that are beyond the sample period used in this paper.¹⁵ Predictions made by the GEE suggest moderate slowdown in the U.S. economy starting from the late 2006. However, they do not indicate strong signals of recession. Although we are still uncertain about the state of the economy in this time period, it is likely that there was no recession during 2006 and 2007 if we apply a conventional rule of thumb that more than two consecutive quarters of negative GDP growth are regarded as a recession. Thus, the predictions from the GEE models are perhaps consistent with the state of the economy in the last two years.

¹⁵Parameter estimates based on the data up to 2008:03 are reported in the Appendix.

3.4 Conclusion

Systematically comparing forecasting models of recession probability in the U.S. economy, I have found substantial gains in out-of-sample forecasting accuracy from including additional variables in a typical univariate probit model with the term spread. Furthermore, there is a room to further improve forecasting performance by utilizing non-Normal CDFs, which allow for positive skewness and/or excess kurtosis. Among 36102 forecasting models considered, the combination of the term spread between 10-year and 3-month Treasury bonds, S&P 500, and non-farm employment growth with the Laplace CDF records the best out-of-sample result, in terms of the QPS and KS. On the other hand, the combination of the term spread, S&P 500, and the corporate bond yield on A-rated firms achieves the best out-of-sample LPS. Especially, its superior forecasting performance of S&P 500 and employment growth, together with a term spread measure is robust in the out-of-sample and leaving 2-years out exercises. Although it is always better to use non-Normal CDF, there is some mixed evidence on a better functional form. These models do not always beat the probit counterparts in all three measures of out-of-sample accuracy.

In order to allow for more flexibility in the functional form, I have applied the generalized Edgeworth expansion and I have found that predictions from the GEE are always superior to the probit counterparts. In principle, we can amplify correct predictions and dampen false signals by allowing for more flexible functional form.

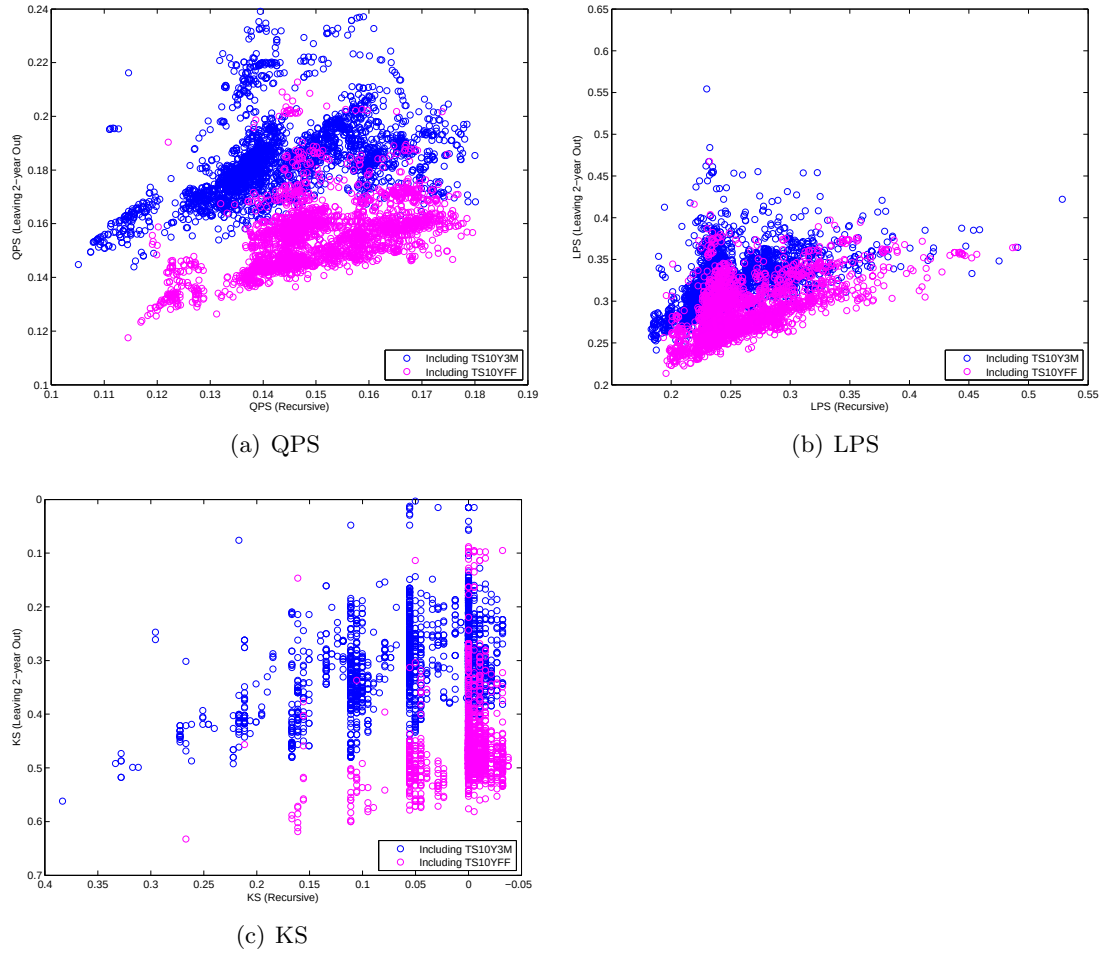


Figure 3.3: Recursive Out-of-Sample vs. Leaving 2-years Out

Note: Horizontal axes take forecasting accuracy measures based on the recursive out-of-sample forecasting. Vertical axes represent those based on the leaving 2-years out exercises. Other models are not plotted here since their performance is inferior to those include the term spread measures.

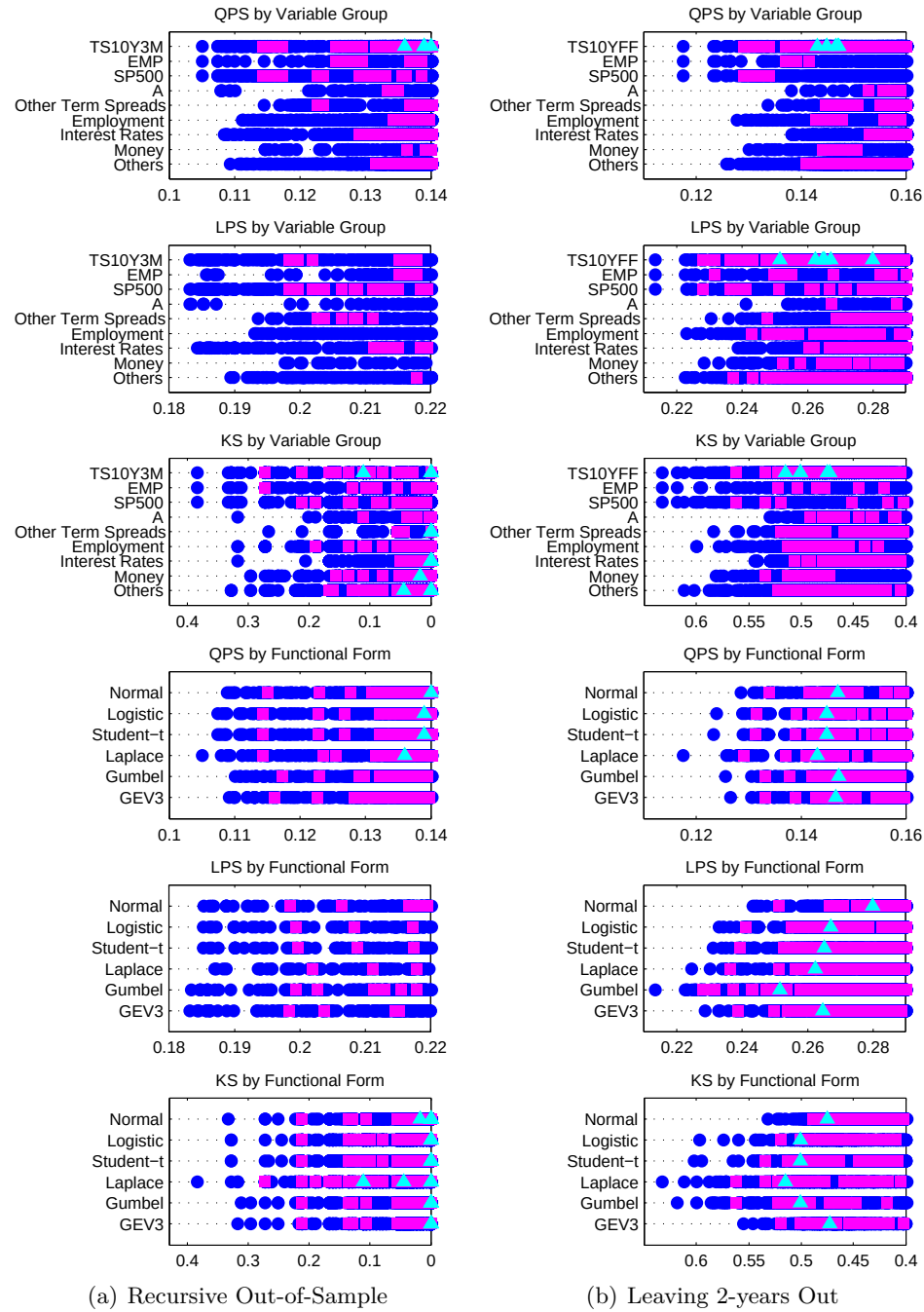


Figure 3.4: Comparison of Forecasting Accuracy: Recursive vs. Leaving 2-years Out

Note: The left panels show forecasting accuracy based on the recursive out-of-sample predictions. The right panels display the based on the leaving 2-years out exercises. Within each exercise, the top three panels and bottom three panels summarize results by variable groups and by functional forms, respectively. Circles represent trivariate models. Squares indicate bivariate models and triangles are univariate models. Horizontal axes for the KS is in the reverse order, such that better models appear on the left.

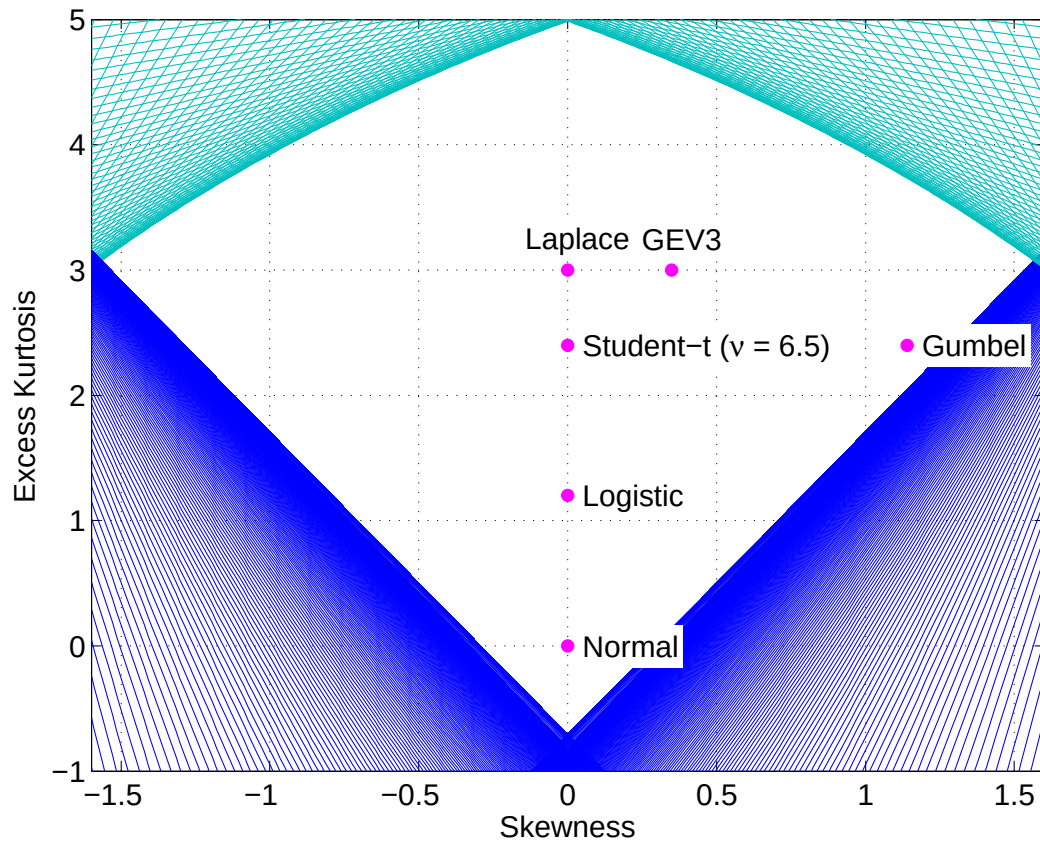


Figure 3.5: Admissible Set of Skewness and Excess Kurtosis with the Logistic ($\mu = 0, \sigma^2 = 1.2$)

Note: The white area bounded by lines approximates the admissible combinations of skewness and excess kurtosis that satisfy (3.15) when the Logistic with mean $=0$ and $\sigma^2 = 1.2$ is used as the approximating distribution. This is drawn by discretizing x .

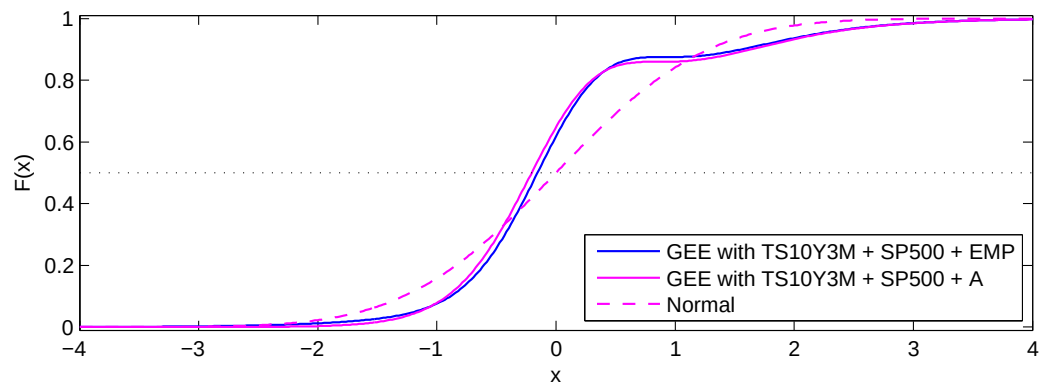
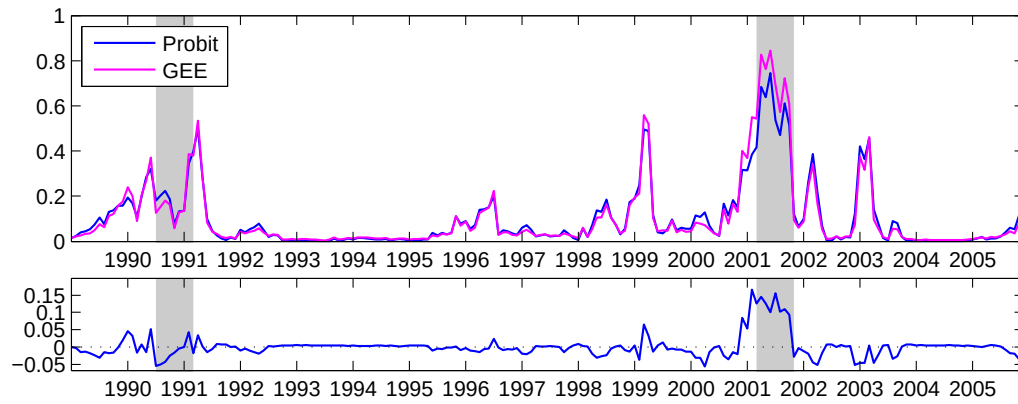
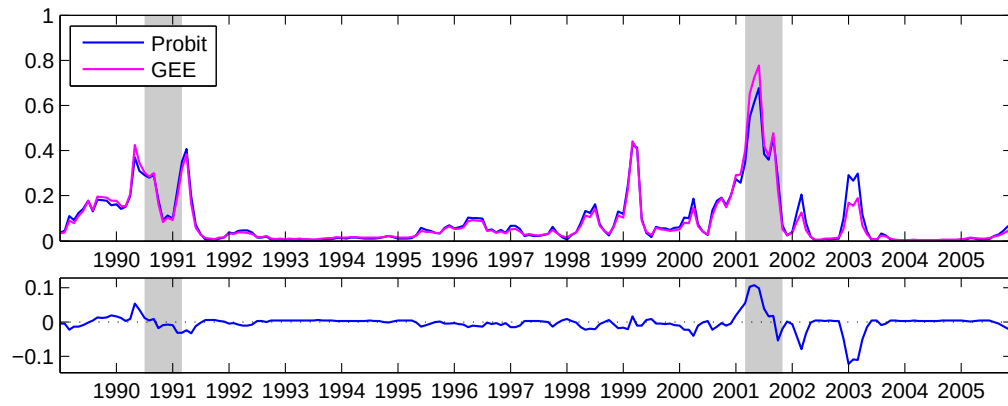


Figure 3.6: Comparison of Functional Forms



(a) Trivariate Model with TS10Y3M, SP500, and EMP



(b) Trivariate Model with TS10Y3M, SP500, and A

Figure 3.7: Out-of-Sample Predictions by the GEE

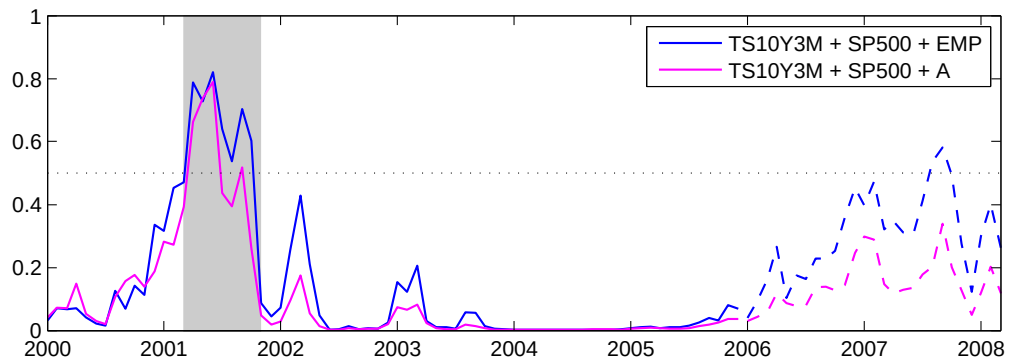


Figure 3.8: Current Predictions by the GEE

Note: Solid lines represent in-sample probabilities and dashed lines indicate recursively calculated current predictions.

Appendix A

Appendix to Chapter 1

A.1 The Kalman filter and maximum likelihood estimation

Given $\hat{\boldsymbol{\xi}}_{1|0}$, $\mathbf{P}_{1|0}$, and parameter values, the Kalman filter recursively calculate $\{\hat{\boldsymbol{\xi}}_{t+1|t}\}_{t=1}^T$ and $\{\mathbf{P}_{t+1|t}\}_{t=1}^T$ as:

$$\begin{aligned}\hat{\boldsymbol{\xi}}_{t+1|t} &= \boldsymbol{\Phi}\hat{\boldsymbol{\xi}}_{t|t-1} + \boldsymbol{\Phi}\mathbf{P}_{t|t-1}\boldsymbol{\Psi}'_t(\boldsymbol{\Psi}_t\mathbf{P}_{t|t-1}\boldsymbol{\Psi}'_t + \mathbf{R}_t)^{-1}(\mathbf{y}_t - \boldsymbol{\Psi}_t\hat{\boldsymbol{\xi}}_{t|t-1} - \mathbf{b}_t), \\ \mathbf{P}_{t+1|t} &= \boldsymbol{\Phi}\mathbf{P}_{t|t-1}\boldsymbol{\Phi}' - \boldsymbol{\Phi}\mathbf{P}_{t|t-1}\boldsymbol{\Psi}'_t(\boldsymbol{\Psi}_t\mathbf{P}_{t|t-1}\boldsymbol{\Psi}'_t + \mathbf{R}_t)^{-1}\boldsymbol{\Psi}_t\boldsymbol{\Phi}\mathbf{P}_{t|t-1}\boldsymbol{\Phi}' + \mathbf{Q}.\end{aligned}$$

The log likelihood function is given by

$$L = \sum_{t=1}^T \left\{ -\frac{n_t}{2} \log(2\pi) - \frac{1}{2} \log(|\boldsymbol{\Omega}_t|) + \mathbf{e}'_t \boldsymbol{\Omega}_t^{-1} \mathbf{e}_t \right\},$$

where n_t is dimension of $\mathbf{y}_t$,

$$\begin{aligned}\mathbf{e}_t &= \mathbf{y}_t - \boldsymbol{\Psi}_t\hat{\boldsymbol{\xi}}_{t|t-1} - \mathbf{b}_t, \text{ and} \\ \boldsymbol{\Omega}_t &= \boldsymbol{\Psi}_t\mathbf{P}_{t|t-1}\boldsymbol{\Psi}'_t + \mathbf{R}_t.\end{aligned}$$

For the estimation, we set $\hat{\boldsymbol{\xi}}_{1|0}$ and $\mathbf{P}_{1|0}$ as follows:

$$\hat{\boldsymbol{\xi}}_{1|0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{P}_{1|0} = 20 \times \begin{bmatrix} \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0}' & \cdots & \cdots & \mathbf{0}' & 0 \end{bmatrix}$$

A.2 Gradient and Information Matrix

The derivative of the log-likelihood function at t with respect to the i^{th} parameter is given by

$$\frac{\partial \log L_t}{\partial \theta_i} = -\frac{1}{2} \text{tr} \left\{ \left(\Omega_t^{-1} \frac{\partial \Omega_t}{\partial \theta_i} \right) \left(\mathbf{I} - \Omega_t^{-1} \mathbf{e}_t \mathbf{e}_t' \right) \right\} - \left(\frac{\partial \mathbf{e}_t}{\partial \theta_i} \right)' \Omega_t^{-1} \mathbf{e}_t.$$

Derivatives of $\mathbf{e}_t$ and Ω_t are

$$\frac{\partial \mathbf{e}_t}{\partial \theta_i} = -\Psi_t \frac{\partial \hat{\xi}_{t|t-1}}{\partial \theta_i} - \frac{\partial \mathbf{b}_t}{\partial \theta_i}, \quad (\text{A.1})$$

$$\frac{\partial \Omega_t}{\partial \theta_i} = \Psi_t \frac{\partial \mathbf{P}_{t|t-1}}{\partial \theta_i} \Psi_t' + \frac{\partial \mathbf{R}_t}{\partial \theta_i}, \quad (\text{A.2})$$

respectively.

Let us define $\mathbf{M}_t = \Phi \mathbf{P}_{t|t-1} \Psi_t'$. Then the Kalman recursion can be expressed as:

$$\begin{aligned} \hat{\xi}_{t+1|t} &= \Phi \hat{\xi}_{t|t-1} + \mathbf{M}_t \Omega_t^{-1} \mathbf{e}_t, \\ \mathbf{P}_{t+1|t} &= \Phi \mathbf{P}_{t|t-1} \Phi' - \mathbf{M}_t \Omega_t^{-1} \mathbf{M}_t' + \mathbf{Q}. \end{aligned}$$

Given (A.1), (A.2), and

$$\frac{\partial \mathbf{M}}{\partial \theta_i} = \frac{\partial \Phi}{\partial \theta_i} \mathbf{P}_{t|t-1} \Psi_t' + \Phi \frac{\partial \mathbf{P}_{t|t-1}}{\partial \theta_i} \Psi_t',$$

we can calculate

$$\begin{aligned} \frac{\partial \hat{\xi}_{t+1|t}}{\partial \theta_i} &= \frac{\partial \Phi}{\partial \theta_i} \hat{\xi}_{t|t-1} + \Phi \frac{\partial \hat{\xi}_{t|t-1}}{\partial \theta_i} + \frac{\partial \mathbf{M}_t}{\partial \theta_i} \Omega_t^{-1} \mathbf{e}_t - \mathbf{M}_t \Omega_t^{-1} \frac{\partial \Omega_t}{\partial \theta_i} \Omega_t^{-1} \mathbf{e}_t + \mathbf{M}_t \Omega_t^{-1} \frac{\partial \mathbf{e}_t}{\partial \theta_i}, \\ \frac{\partial \mathbf{P}_{t+1|t}}{\partial \theta_i} &= \frac{\partial \Phi}{\partial \theta_i} \mathbf{P}_{t|t-1} \Phi' - \frac{\partial \mathbf{M}_t}{\partial \theta_i} \Omega_t^{-1} \mathbf{M}_t' + \mathbf{M}_t \Omega_t^{-1} \frac{\partial \Omega_t}{\partial \theta_i} \Omega_t^{-1} \mathbf{M}_t' - \mathbf{M}_t \Omega_t^{-1} \frac{\partial \mathbf{M}_t'}{\partial \theta_i} + \frac{\partial \mathbf{Q}}{\partial \theta_i}. \end{aligned}$$

By going through the Kalman recursion, we can calculate $\{\frac{\partial \hat{\xi}_{t+1|t}}{\partial \theta}\}_{t=1}^T$ and $\{\frac{\partial \hat{\xi}_{t+1|t}}{\partial \theta}\}_{t=1}^T$. Once we calculate $\frac{\partial \mathbf{e}_t}{\partial \theta_i}$ and $\frac{\partial \Omega_t}{\partial \theta_i}$, we can easily compute the Information matrix. The (i, j) element of the information matrix is given by

$$\mathbf{I}_{ij}(\boldsymbol{\theta}) = \frac{1}{2} \sum_{t=1}^T \left\{ \text{tr} \left(\Omega_t^{-1} \frac{\partial \Omega_t}{\partial \theta_i} \Omega_t^{-1} \frac{\partial \Omega_t}{\partial \theta_j} \right) \right\} + \sum_{t=1}^T \left(\frac{\partial \mathbf{e}_t}{\partial \theta_i} \right)' \Omega_t^{-1} \frac{\partial \mathbf{e}_t}{\partial \theta_j}.$$

Appendix B

Appendix to Chapter 2

B.1 Calculating the Ratio of $T/\tilde{Y}$

We can use Jorgenson's data set to obtain the ratio $T/\tilde{Y}$. In Jorgenson's data set, we have annual data on price and quantity indices of output for all industries from 1958 to 1996.

Let $P_{i,t}$ and $Y_{i,t}$ denote price and quantity indices of gross output of industry i in year t , respectively. Then the total value of gross output in the economy, which correspond to $P\tilde{Y}$ in the model economy, is given by

$$P_t\tilde{Y}_t = \sum_{i=1}^N P_{i,t}Y_{i,t}, \quad (\text{B.1})$$

where N is the number of industries in the economy.

Deflating it by P_t , we will obtain a quantity index of the aggregate gross output. Since the data set itself contains the quantity index of the transportation industry, it is easy to obtain an estimate of $T_t/\tilde{Y}_t$.

Define a Divisia price index for the aggregate gross output as:

$$\frac{dP_t}{P_t} = \sum_{i=1}^N s_{i,t} \frac{dP_{i,t}}{P_{i,t}}, \quad (\text{B.2})$$

where $s_{i,t}$ is the expenditure share of the i^{th} industry.

A discrete approximation of the above index (= Törnquist index) is then given by:

$$\Delta \log P_t = \sum_{i=1}^N \bar{s}_{i,t} \Delta \log P_{i,t}, \quad (\text{B.3})$$

where

$$\bar{s}_{i,t} = \frac{1}{2} \left\{ \frac{P_{i,t}Y_{i,t}}{P_t Y_t} + \frac{P_{i,t-1}Y_{i,t-1}}{P_{t-1}Y_{t-1}} \right\}. \quad (\text{B.4})$$

With the normalization that $P_{1992} = 1$, which is a convention in Jorgenson's data set, we can recover a sequence of $\{P_t\}$.

As a result, we obtain the pre-1984 average of $T_t/\tilde{Y}_t$ is 0.0386 and the post-1984 average is 0.0397.

B.2 Characterizing Equilibrium

Let Ξ^Y and Ξ^T denote the Lagrange multipliers associated with (2.4) and (2.14), respectively. The symmetric equilibrium is characterized by the following system of equations.

$$\begin{aligned} \frac{1}{C_t} &= \Lambda_t \\ \theta_N \frac{N_t^{1+\varphi}}{w_t} &= \Lambda_t \left[\phi_W \left(\frac{\pi_t w_t}{w_{t-1}} - \pi \right) \frac{\pi_t w_t}{w_{t-1}} + \frac{\phi_W}{2} \left(\frac{\pi_t w_t}{w_{t-1}} - \pi \right)^2 - (1 - \theta_N) N_t \right] \\ &\quad - \beta \mathbb{E}_t \left[\Lambda_{t+1} \phi_W \left(\frac{\pi_{t+1} w_{t+1}}{w_t} - \pi \right) \left(\frac{w_{t+1}}{w_t} \right)^2 \pi_{t+1} \right] \\ m_t^{-\sigma} &= \Lambda_t - \beta \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{\pi_{t+1}} \right] \\ \Lambda_t &= \beta r_t \mathbb{E}_t \left[\frac{\Lambda_{t+1}}{\pi_{t+1}} \right] \\ Y_t - p_t^E (E_t^I + E_t^T) &= C_t + I_t + I_t^A \\ &\quad + \frac{\phi_W}{2} \left(\frac{\pi_t w_t}{w_{t-1}} - \pi \right)^2 w_t + \frac{\phi_K}{2} \left(\frac{K_{t+1}}{g K_t} - 1 \right)^2 K_t \\ &\quad + \frac{\phi_Y}{2} \left(\frac{\pi_t}{\pi_s} - 1 \right)^2 Y_t + \frac{\phi_A}{2} \left(\frac{A_{t+1}}{g A_t} - 1 \right)^2 A_t \\ 0 &= \Lambda_t [(1 - \theta) Y_t + \theta \{ \chi_Q (1 - \chi_Y) + \chi_Y p_t^T \} Y_t \\ &\quad - \phi_Y \left(\frac{\pi_t}{\pi} - 1 \right) \frac{\pi_t}{\pi} (1 - \chi_Q (1 - \chi_Y)) Y_t] \\ &\quad + \theta \Xi_t^Y (1 - \chi_Q) (1 - \chi_Y) Y_t \\ &\quad + \beta \mathbb{E}_t \left[\Lambda_{t+1} \phi_Y \left(\frac{\pi_{t+1}}{\pi} - 1 \right) \frac{\pi_{t+1}}{\pi} (1 - \chi_Q (1 - \chi_Y)) Y_{t+1} \right] \\ \Lambda_t w_t &= \Xi_t^Y \frac{(1 - \alpha)(1 - \chi_Q)(1 - \chi_Y)}{1 - \chi_Q(1 - \chi_Y)} \frac{Y_t}{N_t^I} \end{aligned}$$

$$\begin{aligned}
\Lambda_t \left\{ 1 + \phi_K \left(\frac{K_{t+1}}{K_t} - 1 \right) \right\} &= \beta \mathbb{E}_t \left[\Lambda_{t+1} \left\{ (1 - \delta_{t+1}) - p_{t+1}^E \frac{(u_{t+1})^\nu}{\nu} \right. \right. \\
&\quad \left. \left. + \phi_K \left(\frac{K_{t+2}}{K_{t+1}} - 1 \right) \left(\frac{K_{t+2}}{K_{t+1}} \right) - \frac{\phi_K}{2} \left(\frac{K_{t+2}}{K_{t+1}} - 1 \right)^2 \right\} \right. \\
&\quad \left. + \Xi_{t+1}^Y \frac{\alpha(1 - \chi_Q)(1 - \chi_Y)}{1 - \chi_Q(1 - \chi_Y)} \frac{Y_{t+1}}{K_{t+1}} \right] \\
\Xi_t^Y \frac{\alpha(1 - \chi_Q)(1 - \chi_Y)}{1 - \chi_Q(1 - \chi_Y)} \frac{Y_t}{u_t} &= \Lambda_t \left\{ (u_t)^{\kappa-1} K_t + p_t^E (u_t)^{\nu-1} K_t \right\} \\
\frac{(1 - \chi_Q)(1 - \chi_Y)}{1 - \chi_Q(1 - \chi_Y)} Y_t &= [u_t K_t]^\alpha [N_t^I]^{1-\alpha} \\
K_{t+1} &= (1 - \delta_t) K_t + I_t \\
\delta_t &= \frac{(u_t)^\kappa}{\kappa} \\
\frac{E_t^I}{K_t} &= \frac{(u_t)^\nu}{\nu} \\
T_t &= \frac{\chi_Y}{1 - \chi_Q(1 - \chi_Y)} Y_t \\
p_t^T &= \frac{\theta_T}{(\theta_T - 1)} \left\{ \frac{\Xi_t^T}{\Lambda_t} + \frac{p_t^E}{\omega} \right\} \\
w_t &= \frac{\Xi_t^T}{\Lambda_t} (T_t)^{1/\gamma} (1 - a) (N_t^T)^{-1/\gamma} \\
\Lambda_t \left\{ 1 + \phi_A \left(\frac{A_{t+1}}{A_t} - 1 \right) \right\} &= \beta \mathbb{E}_t \left[\Lambda_{t+1} \left\{ (1 - \delta_A) + \phi_A \left(\frac{A_{t+2}}{A_{t+1}} - 1 \right) \left(\frac{A_{t+2}}{A_{t+1}} \right) \right. \right. \\
&\quad \left. \left. - \frac{\phi_A}{2} \left(\frac{A_{t+2}}{A_{t+1}} - 1 \right)^2 \right\} + \Xi_{t+1}^T (T_{t+1})^{1/\gamma} a (A_{t+1})^{-1/\gamma} \right] \\
T_t &= \left[a A_t^{(\gamma-1)/\gamma} + (1 - a) N_t^{T(\gamma-1)/\gamma} \right]^{\gamma/(\gamma-1)} \\
T_t &= \omega E_t^T \\
A_{t+1} &= (1 - \delta_A) A_t + I_t^A \\
\log(r_t/r_s) &= \rho_y \log(Y_t/Y_s) + \rho_\pi \log(\pi_t/\pi_s) + \rho_\mu \log(\mu_t/\mu_s) + \varepsilon_{r,t} \\
\log(p_t^E) &= (1 - \rho_{p^E}) \log(p^E) + \rho_{p^E} \log(p_{t-1}^E) + \varepsilon_{p^E,t} \\
N_t &= N_t^I + N_t^T \\
\mu_t &= \left(\frac{m_t}{m_{t-1}} \right) \pi_t
\end{aligned}$$

Appendix C

Appendix to Chapter 3

C.1 Detailed Expression for (3.14)

Let ς denote a scale parameter of the Logistic distribution, such that its variance equals $\frac{\pi^2}{3}\varsigma^2$, and also let $z = x/\varsigma$. Then, the GEE model (3.14) can be expressed as:

$$\begin{aligned}\tilde{F}(x) = & \frac{\exp(z)}{1 + \exp(z)} + \frac{0.1 \exp(z) \{-1 + \exp(z)\}}{\varsigma^2 \{1 + \exp(z)\}^3} + \frac{\gamma_1 \exp(z) \{-1 + 4 \exp(z) - \exp(2z)\}}{6 \varsigma^3 \{1 + \exp(z)\}^4} \\ & - \left(\frac{\gamma_2}{24} - 0.067 \right) \frac{\exp(z) \{-1 + 11 (\exp(z) - \exp(2z)) + \exp(3z)\}}{\varsigma^4 \{1 + \exp(z)\}^5}\end{aligned}\tag{C.1}$$

C.2 Parameter Estimates based on the Recent Data

Parameter estimates based on the data from 1961:04 to 2008:03 are presented in Table C.1. The first two columns are those from the GEE models. For reference, the last two columns also report estimates based on the Laplace and GEV3, respectively.

Table C.1: Parameter Estimates based on the Recent Data

	GEE		Laplace	GEV3
Constant	−0.1821	−1.2764	−0.1443	−2.2979
TS10Y3M	−0.5257	−0.4704	−1.2593	−0.7643
SP500	−0.0426	−0.0445	−0.0997	−0.0774
EMP	−1.3923		−3.2104	
A		0.1061		0.2467
γ_1	1.5470	1.5670	n.a.	n.a.
γ_2	3.1069	3.0722	n.a.	n.a.

Note: Estimates are based on the data from 1961:04 to 2008:03. γ_1 and γ_2 denote skewness and excess kurtosis in the GEE model, respectively.

References

- AGUIAR-CONRARIA, L., AND Y. WEN (2007): “Understanding the Large Negative Impact of Oil Shocks,” *Journal of Money, Credit and Banking*, 39(4), 925–944.
- BARSKY, R. B., AND L. KILIAN (2002): “Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative,” in *NBER Macroeconomic Annual*, ed. by B. Bernanke, and K. Rogoff, vol. 2001, pp. 115–134. MIT Press, Cambridge, MA.
- (2004): “Oil and the Macroeconomy since the 1970s,” *Journal of Economic Perspectives*, 18(4), 115–134.
- BARSKY, R. B., AND J. A. MIRON (1989): “The Seasonal Cycle and the Business Cycle,” *Journal of Political Economy*, 97(3), 503–534.
- BASU, S. (1996): “Procyclical Productivity: Increasing Returns or Cyclical Utilization,” *Quarterly Journal of Economics*, 111(3), 719–751.
- BASU, S., AND J. G. FERNALD (1995): “Are Apparent Productive Spillovers a Figment of Specification Error?,” *Journal of Monetary Economics*, 36(1), 165–188.
- (2001): “Why is Productivity Procyclical? Why Do We Care?,” in *New Developments in Productivity Analysis*, ed. by C. R. Hulten, E. R. Dean, and M. J. Harper, vol. 63 of *Studies in Income and Wealth*, chap. 7, pp. 225–301. University of Chicago Press.
- BASU, S., J. G. FERNALD, AND M. S. KIMBALL (2004): “Are Technology Improvements Contractionary?,” *NBER Working Paper*, No. 10592.
- BASU, S., J. G. FERNALD, AND M. D. SHAPIRO (2001): “Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?,” *Carnegie-Rochester Conference Series on Public Policy*, 55, 117–165.
- BASU, S., AND M. S. KIMBALL (1997): “Cyclical Productivity with Unobserved Input Variation,” *NBER Working Paper*, No. 5915.
- BEAULIEU, J. J., AND J. MATTEY (1998): “The Workweek of Capital and Capital Utilization in Manufacturing,” *Journal of Productivity Analysis*, 10(2), 199–223.
- BERNANKE, B. S., M. GERTLER, AND M. WATSON (1997): “Systematic Monetary Policy and the Effects of Oil Price Shocks,” *Brookings Papers on Economic Activity*, 1997(1), 91–142.

- BERNANKE, B. S., M. GERTLER, AND M. W. WATSON (2004): "Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy. A Reply," *Journal of Money, Credit and Banking*, 36(2), 287–291.
- BILS, M., AND P. J. KLENOW (2004): "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112(5), 947–985.
- BIRCHENHALL, C. R., H. JESSEN, D. R. OSBORN, AND P. SIMPSON (1999): "Predicting U.S. Business-Cycle Regimes," *Journal of Business Economic Statistics*, 17(3), 313–323.
- BLANCHARD, O. J., AND J. GALÍ (2007): "The Macroeconomic Effects of Oil Shocks. Why are the 2000s So Different from the 1970s?," Working Paper 13368, National Bureau of Economic Research.
- BURBIDGE, J., AND A. HARRISON (1984): "Testing for the Effects of Oil-Price Rises using Vector Autoregressions," *International Economic Review*, 25(2), 459–484.
- BURMEISTER, E., K. D. WALL, AND J. D. HAMILTON (1986): "Estimation of Unobserved Expected Monthly Inflation Using Kalman Filter," *Journal of Business & Economic Statistics*, 4(2), 147–160.
- BURNSIDE, C., AND M. EICHENBAUM (1996): "Factor-Hording and the Propagation of Business-Cycle Shocks," *American Economic Review*, 86(5), 1154–1174.
- BURNSIDE, C., M. EICHENBAUM, AND S. REBELO (1995): "Capital Utilization and Returns to Scale," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke, and J. J. Rotemberg, pp. 67–110. MIT Press.
- (1996): "Sectoral Solow Residuals," *European Economic Review*, 40(3-5), 861–869.
- CARLSTROM, C. T., AND T. S. FUERST (2006): "Oil Prices, Monetary Policy, and Counterfactual Experiments," *Journal of Money, Credit and Banking*, 38(7), 1945–1958.
- CARRUTH, A. A., M. A. HOOKER, AND A. J. OSWALD (1998): "Unemployment Equilibria and Input Prices: Theory and Evidence from the United States," *Review of Economics and Statistics*, 80(4), 621–628.
- CAVALLO, M., AND T. WU (2007): "Measuring Oil-price Shocks Using Market-based Information," *Federal Reserve Bank of San Francisco Working Paper*.
- CHAUVET, M., AND J. D. HAMILTON (2006): "Dating Business Cycle Turning Points," in *Nonlinear Time Series Analysis of Business Cycles*, ed. by C. Milas, P. Rothman, and D. van Dijk, vol. 276 of *Contributions to Economic Analysis*, chap. 1, pp. 1–54. Elsevier.
- CHAUVET, M., AND S. POTTER (2005): "Forecasting Recessions Using the Yield Curve," *Journal of Forecasting*, 24(2), 77–103.

- CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.
- CLARIDA, R., J. GALÍ, AND M. GERTLER (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115(1), 147–180.
- CLEMENTS, M. P., AND A. B. GALVÃO (2006): "Combining Predictors & Combining Information in Modelling: Forecasting US Recession Probabilities and Output Growth," in *Nonlinear Time Series Analysis of Business Cycles*, ed. by C. Milas, P. Rothman, and D. van Dijk, vol. 276 of *Contributions to Economic Analysis*, chap. 2, pp. 55 – 74. Elsevier.
- CONGRESSIONAL BUDGET OFFICE (2006): "The Economic Effects Recent Increases in Energy Prices," A CBO Paper.
- DANIEL, B. C. (1997): "International interdependence of national growth rates: A structural trends analysis," *Journal of Monetary Economics*, 40(1), 73–96.
- DAVIS, S. J., AND J. HALTIWANGER (2001): "Sectoral Job Creation and Destruction Responses to Oil Price Changes," *Journal of Monetary Economics*, 48, 465–512.
- DEPARTMENT OF ENERGY, ENERGY INFORMATION ADMINISTRATION (1995): "Measuring Energy Efficiency In The United States' Economy: A Beginning," U.S. Government Printing Office.
- (2006): "Annual Energy Review 2005," U.S. Government Printing Office.
- DOMAR, E. D. (1961): "On the Measurement of Technological Change," *Economic Journal*, 71(284), 709–729.
- DOTSEY, M. (1998): "The Predictive Content of the Interest Rate Term Spread for Future Economic Growth," *Federal Reserve Bank of Richmond Economic Quarterly*, 84(3), 31–51.
- DOTSEY, M., AND R. G. KING (2006): "Pricing, Production, and Persistence," *Journal of European Economic Association*, 4(5), 893–928.
- DUEKER, M. (2005): "Dynamic Forecasts of Qualitative Variables: A Qual VAR Model of U.S. Recessions," *Journal of Business & Economic Statistics*, 23(1), 96–104.
- ESTRELLA, A., AND F. S. MISHKIN (1998): "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *Review of Economics and Statistics*, 80(1), 45–61.
- ESTRELLA, A., A. P. RODRIGUES, AND S. SCHICH (2003): "How Stable is the Predictive Power of the Yield Curve? Evidence from Germany and the United States," *Review of Economics and Statistics*, 85(3), 629–644.
- EVANS, C. L. (1992): "Productivity Shocks and Real Business Cycles," *Journal of Monetary Economics*, 29(2), 191–208.

- FINN, M. G. (2000): "Perfect Competition and the Effects of Energy Price Increases on Economics Activity," *Journal of Money, Credit and Banking*, 32(3), 400–416.
- GISSER, M., AND T. H. GOODWIN (1986): "Crude Oil and the Macroeconomy: Tests of Some Popular Notions: Note," *Journal of Money, Credit and Banking*, 18(1), 95–103.
- GREENWOOD, J., Z. HERCOWITZ, AND G. W. HUFFMAN (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.
- HAMILTON, J. D. (1983): "Oil and the Macroeconomy since World War II," *Journal of Political Economy*, 91(2), 228–248.
- (1985): "Uncovering Financial Market Expectations of Inflation," *Journal of Political Economy*, 93(6), 1224–1241.
- (1994): *Time Series Analysis*. Princeton University Press, Princeton.
- (1996): "This is what happened to the oil price-macroeconomy relationship;," *Journal of Monetary Economics*, 38(2), 215–220.
- (2003): "What is an oil shock?," *Journal of Econometrics*, 113(2), 363–398.
- HAMILTON, J. D., AND A. M. HERRERA (2004): "Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy: A Comment," *Journal of Money, Credit and Banking*, 36(2), 265–286.
- HANSEN, C. B., J. B. McDONALD, AND P. THEODOSSIOU (2007): "Some Flexible Parametric Models for Partially Adaptive Estimators of Econometric Models," Economics Discussion Papers 2007-13, Kiel Institute for the World Economy.
- HANSEN, P. R. (2008): "In-Sample and Out-of-Sample Fit: Their Joint Distribution and its Implications for Model Selection," Working paper, Stanford University.
- HARVEY, A. C. (1989): *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge.
- HARVEY, A. C., AND R. G. PIERSE (1984): "Estimating Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, 79(385), 125–131.
- HOOKE, M. A. (1996): "What happened to the oil price-macroeconomy relationship?;," *Journal of Monetary Economics*, 38(2), 195–213.
- HUANG, K. X., AND Z. LIU (2004): "Input-output Structure and Nominal Rigidity: The Persistence Problem Revisited," *Macroeconomic Dynamics*, 8, 188–206.
- HULTEN, C. R. (2001): "Total Factor Productivity: A Short Biography," in *New Developments in Productivity Analysis*, ed. by C. R. Hulten, E. R. Dean, and M. J. Harper, vol. 63 of *Studies in Income and Wealth*, chap. 1, pp. 1–47. University of Chicago Press.

- IRELAND, P. N. (2001): "Sticky-price Models of the Business Cycle: Specification and Stability," *Journal of Monetary Economics*, 47, 3–18.
- (2004): "A Method for Taking Models to the Data," *Journal of Economic Dynamics & Control*, 28(6), 1205–1226.
- JARROW, R., AND A. RUDD (1982): "Approximate Option Valuation for Arbitrary Stochastic Processes," *Journal of Financial Economics*, 10(3), 347–369.
- JORGENSON, D. W. (1990): "Productivity and Economic Growth," in *Fifty Years of Economic Measurement*, ed. by E. R. Berndt, and J. E. Triplett, vol. 54 of *Studies in Income and Wealth*, chap. 3, pp. 19–118. University of Chicago Press.
- JORGENSON, D. W., F. M. GOLLAP, AND B. M. FRAUMENI (1987): *Productivity and U.S. Economic Growth*. Harvard University Press, Cambridge.
- JORGENSON, D. W., AND K. J. STIROH (2000): "Raising the Speed Limit: U.S. Economic Growth in the Information Age," *Brookings Papers on Economic Activity*, 2000(1), 125–211.
- KANE, T. (2008): "Employment Numbers as Recession Indicators," Discussion paper, Joint Economic Committee, U.S. Congress.
- KAUPPI, H., AND P. SAIKKONEN (2007): "Predicting U.S. Recessions with Dynamic Binary Response Models," Working Paper, Forthcoming in Review of Economics and Statistics.
- KILIAN, L. (2007): "The Economic Effects of Energy Price Shocks," Working paper, University of Michigan.
- KIM, I.-M., AND P. LOUNGANI (1991): "The Role of Energy in Real Business Cycle Models," *Journal of Monetary Economics*, 29, 173–189.
- KIM, J. (2000): "Constructing and Estimating a Realistic Optimizing Model of Monetary Policy," *Journal of Monetary Economics*, 45, 329–359.
- KING, T. B., A. T. LEVIN, AND R. PERLI (2007): "Financial Market Perceptions of Recession Risk," Finance and economics discussion series, Board of Governors of the Federal Reserve System.
- KLEIN, P. (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.
- LEDUC, S., AND K. SILL (2004): "A Quantitative Analysis of Oil-price Shocks, Systematic Monetary Policy, and Economic Downturns," *Journal of Monetary Economics*, 51, 781–808.
- LEVIN, A. T., A. ONATSKI, J. C. WILLIAMS, AND N. WILLIAMS (2006): "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models," in *NBER Macroeconomics Annual 2005*, ed. by M. Gertler, and K. Rogoff, pp. 229–287. MIT Press.

- MATTEY, J., AND S. STRONGIN (1997): "Factor Utilization and Margins for Adjusting Output: Evidence from Manufacturing Plants," *Federal Reserve Bank of San Francisco Economic Review*, (2), 3–17.
- MCCONNELL, M. M., AND G. PEREZ-QUIROS (2000): "Output Fluctuations in the United States: What Has Changed Since the Early 1980's?," *American Economic Review*, 90(5), 1464–1476.
- MORK, K. A. (1989): "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results," *Journal of Political Economy*, 97(3), 740–744.
- NEBESKY, W., B. S. MCMULLEN, AND M.-K. LEE (1995): "Testing for Market Power in the U.S. Motor Carrier Industry," *Review of Industrial Organization*, 10, 559–576.
- OLSON, J. E. (1972): "Price Discrimination by Regulated Motor Carriers," *The American Economic Review*, 62(3), 395–402.
- PAQUET, A., AND B. ROBIDOUX (2001): "Issues on the Measurement of the Solow Residual and the Testing of its exogeneity: Evidence for Canada," *Journal of Monetary Economics*, 47(3), 595–612.
- ROTEMBERG, J. J., AND M. WOODFORD (1995): "Dynamic General Equilibrium Models with Imperfectly Competitive Product Markets," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, pp. 243–293. Princeton University Press.
- (1996): "Imperfect Competition and the Effects of Energy Price Increases on Economic Activity," *Journal of Money, Credit and Banking*, 28(4), 549–577.
- SATO, K. (1975): *Production Functions and Aggregation*. North-Holland, Amsterdam.
- (1976): "The Meaning and Measurement of the Real Value Added Index," *Review of Economics and Statistics*, 58(4), 434–442.
- SHAPIRO, M. D. (1989): "Assessing the Federal Reserve's Measures of Capacity and Utilization," *Brookings Papers on Economic Activity*, (1), 181–225.
- (1996): "Macroeconomic Implications of Variation in the Workweek of Capital," *Brookings Papers on Economic Activity*, (2), 79–119.
- SHEA, J. (1993): "The Input-Output Approach to Instrument Selection," *Journal of Business & Economic Statistics*, 11(2), 145–155.
- SIMS, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1), 1–20.
- SMETS, F., AND R. WOUTERS (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123–1175.
- SNOW, J. W., AND S. SOBOTKA (1977): "Certificate Values," in *Regulation of Entry and Pricing in Truck Transportation*, ed. by P. W. MacAvoy, and J. W. Snow, chap. 6, pp. 153–156. American Enterprise Institute for Public Policy Research.

- STOCK, J. H., AND M. W. WATSON (1989): "New Indexes of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual*, 4, 351–394.
- (2003): "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature*, 41(3), 788–829.
- TAYLOR, J. B. (1993): "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.
- (1999): "Staggered Price and Wage Setting in Macroeconomics," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1B. Elsevier.
- THEODOSSIOU, P. (1998): "Financial Data and the Skewed Generalized t Distribution," *Management Science*, 44(12), 1650–1661.
- U.S. CONGRESS, SENATE (1980): *Federal Restraints on Competition in the Trucking Industry: Antitrust Immunity and Economic Regulation: Report of the Committee on Judiciary*, vol. 96th Congress, 2d Session. U.S. Government Printing Office, Washington.
- WINSTON, C. (1981): "The Welfare Effects of ICC Rate Regulation Revisited," *Bell Journal of Economics*, 12(1), 232–244.
- WINSTON, C., T. M. CORSI, C. M. GRIMM, AND C. A. EVANS (1990): *The Economic Effects of Surface Freight Deregulation*. Brookings Institution Press.
- WRIGHT, C. L. (1983): "Issues in Freight Transportation Regulation: Comment," *American Journal of Agricultural Economics*, 65(1), 162–166.
- WRIGHT, J. H. (2006): "The Yield Curve and Predicting Recessions," Finance and Economics Discussion Series.
- YING, J. S., AND T. E. KEELER (1991): "Pricing in a Deregulated Environment: The Motor Carrier Experience," *RAND Journal of Economics*, 22(2), 264–273.