Essays on macroeconometrics

Author: Chuanqi Zhu

Boston College Electronic Thesis or Dissertation, 2013

Boston College

The Graduate School of Arts and Sciences

Department of Economics

ESSAYS ON MACROECONOMETRICS

a dissertation

by

CHUANQI ZHU

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

May 2013

ESSAYS ON MACROECONOMETRICS

ABSTRCT

by

CHUANQI ZHU

Dissertation Committee:

Zhijie Xiao (Chair)

Peter Ireland

Georg Strasser

This dissertation contains three chapters in theoretical Macroeconometrics and applied Macroeconometrics.

This first chapter addresses the issues related to the estimation, testing and computation of ordered structural breaks in multivariate linear regressions. Unlike common breaks, ordered structural breaks are those breaks that are related across equations but not necessarily occurring at the same dates. A likelihood ratio test assuming normal errors is proposed in this chapter in order to detect the ordered structural breaks in multivariate linear regressions. The estimation of ordered structural breaks uses quasi-maximum likelihood and adopts the efficient algorithm of Bai and Perron (2003). I also provide results about the consistency and rate of convergence when searching for

ordered structural breaks. Finally, these methods are applied to one empirical example: the mean growth rate of output in three European countries and United States.

This second chapter focuses on the parameter stability of dynamic stochastic general equilibrium (DSGE) models. To this end, I solve and estimate a representative New Keynesian model using both linear and nonlinear methods. I first examine how nonlinearities affect the parameter stability of the New Keynesian model. The results show that parameter instabilities still exist even using nonlinear solutions, and also highlight differences between two nonlinear solution methods: perturbation method and projection method. In addition, I propose a sequential procedure for searching for multiple structural breaks in nonlinear models, and apply it to the New Keynesian model. Two common structural breaks among these estimated parameters are identified for all the five solutions considered in this chapter. One structural break is in the early 1970s, while another one locates around the middle 1990s.

In the third chapter, we investigate changes in long run productivity growth in the United States. In particular, we approach productivity growth from a sectoral perspective, and decompose the whole economy into two broad sectors: investment goods-producing sector and consumption goods-producing sector. Although the evidence of changes in the aggregate productivity growth is far from obvious at conventional test size, we find evidence of structural breaks in the sectoral productivity growth using both growth accounting and DSGE model based measures. There are two structural breaks in investment goods-producing sector using growth accounting measures, which indicates that the era of investment and productivity boom in the middle 1990s may have ended before the Great Recession. In addition, our results show there is one structural break in consumption goods-producing sector around the 1970s and attribute the aggregate productivity slowdown at that time to consumption goods-producing sector. These results are broadly consistent with Ireland and Schuh (2008). Our results offer up with a modestly pessimistic outlook on future productivity growth and, therefore, potential output.

ACKNOWLEDGEMENT

I am grateful to my thesis committee members Zhijie Xiao, Peter Ireland and Georg Strasser, for their kind support and guidance through my doctoral studies. I would also thank Susanto Basu, Donald Cox, Arthur Lewbel, and all other participants for their insightful comments and suggestions in the Dissertation Workshop. A specials thanks to Scott Schuh and all my colleagues at the Consumer Payments Research Center in Federal Reserve Bank of Boston.

I thank all my friends for their priceless supports and encouragement all through these years. I am also grateful to my fellow graduate classmates for their help and supports.

I dedicate this thesis to my family, including my wife Na, my parents, my grandparents, my uncles and aunts, and all my other family members, for their love, encouragement and scarifies made through my life.

Contents

Chapter 1

Estimating and Testing Ordered Structural Breaks in Multivariate Linear Regressions

1.1 Introduction

In the last two decades, there are considerable theoretical and empirical studies on testing and estimation of changes in economic and financial time series. In the theoretical econometric literature, Bai (1997) considers the least squares estimation of a single change point in multiple regression models allowing for both stationary and trending regressors. Consistency, rate of convergence, and asymptotic distributions are also provided in his paper. Bai and Perron (1998) then consider the testing and estimation of multiple structural breaks in the linear regression models estimated by least squares as well. And Bai and Perron (2003) provide an efficient dynamic programming algorithm to obtain the break date estimates. For multivariate systems, research related to structural breaks is comparatively scarce. Bai et al. (1998) first consider issues related a single common break across these equations. They show that the accuracy of break date estimates can be improved under a system of equations with common breaks. Hansen (2003) considers multiple structural changes in a co-integrated system, though his study focuses on the case of known break dates. Recently, Qu and Perron (2007) extend the testing and estimation of changes in a system of equations allowing multiple structural breaks.

The empirical motivation for this paper is based on our observations of GDP growth slowdown in the United States and Europe around the 1970s and during the current 2007 financial crisis. The common practice in modeling the breaks in GDP growth is to assume these breaks of each country occur contemporaneous. It might be due to the fact that many common factors, such as international capital flows, are driving these series. The often "*unnoticed*" underlying assumption in this manner is that these breaks across the equation are at the same date. However, several studies in the empirical literature show the GDP growth rate for the four industrialized countries including France, Germany, Italy and United States did not appear to slow down exactly at the same date.[1] This suggests that gains in precision might be achieved by relaxing this assumption of *common* breaks. In other words, the growth rates of output are modeled as changing at different dates across equations. Now these breaks are not exactly at the same dates, to some extent they seem to occur following a sort of order. Therefore, we call this sort of structural breaks in a system of equations as "ordered structural breaks". To our knowledge, few works have been proposed to address this issue.[2]

This paper, therefore, develops techniques for testing and estimation on ordered structural breaks in a system of equations. Our framework builds on Bai et al. (1998), in which they considered a single common break in a system of equations. We relax the underlying assumption on common break and put our attention on ordered

[1] For example, see Banerjee et al. (1992), and Stock and Watson (2002).

[2] One exception is Qu and Perron (2007), in which they consider the locally ordered breaks. As we will show later, this is a special case under our framework.

structural breaks. However, we only consider these systems of equations including stationary regressors. The integrated regressors and deterministically trending regressors are excluded in the current analysis because they need special treatment and add additional layer of difficulty into our framework. The null hypothesis is that no structural break occurs in a multivariate system. Under the alternative hypothesis, there is *one* single break at each equation of this system, but these break dates might not be the same. These breaks across equations may be either close to each other or be separated by a positive fraction of sample size. The statistic considered in this paper is the quasi-likelihood ratio test assuming normal errors, though as usual the limiting distribution of this test has non-standard probability distribution. The computation of estimates under our framework is not a trivial issue. In principle, a gird search can be employed but it becomes rapidly impractical since it involves the computation of maximum likelihood estimates of order $O(T^n)$. Our solution is to extend the work of Hawkins (1976), Bai and Perron (2003) and Qu and Perron (2007) and consider a dynamic programming algorithm.

Similar as many studies related to structural breaks, such as Bai et al. (1998) and Perron (1989) among others, our empirical motivation concerns breaks in the mean growth rate of output, for which the parameter describing dependent in the stochastic part of the process (the auto-regressive parameters in our case) are treated as nuisance parameters. Therefore we first present a Monte Carlo study to show the efficiency gain we could obtain if we allow for ordered structural breaks in a system of equations. We then turn to an empirical example of dating the output growth slowdown in postwar European and United States considered in Banerjee et al. (1992) and Bai et al. (1998) and advocate the evidence of order structural breaks.

The paper is organized as follows. Section 2 presents the model and the assumptions used in this paper. And it also provides an example to illustrate our framework. Section 3 considers the issues related to estimation. In particular, we provide the results on the consistency and rate of convergence of these estimates and describe the dynamic programming estimation algorithm. Section 4 contains the quasi likelihood ratio type statistic for unknown ordered structural breaks. A Monte Carlo study of the statistic and estimation for a two-equation model are provided in Section 5. Section 6 applies the test and estimation method to an empirical example: the growth slowdown in postwar European and United States output.

1.2 Model and Assumptions

1.2.1 Model and Assumptions

We first define the notation used throughout the present paper. Our framework and assumptions are similar to those in Bai et al. (1998). We have n equations and T observations excluding the initial conditions if the lagged dependent variables are used as regressors. Each equation may have *one* structural break denoted as τ_i, $i = 1, \ldots n$. These break dates are denoted by a vector $\tau = (\tau_1, \ldots, \tau_n)$. A subscript t indexes a temporal observation ($t = 1, \ldots, T$) and a subscript i indexes the equation ($i = 1, \ldots, n$) to which a scalar dependent variable y_{it} is associated.

The system of equations considered is

$$y_{it} = \mu_i + \sum_{j=1}^{p} A_j(i) y_{t-j} + \beta_i' X_t + 1[t > \tau_i] \left(\lambda_i + \sum_{j=1}^{p} B_j(i) y_{t-j} + \gamma_i' X_t \right) + \varepsilon_{it} \quad (1.2.1)$$

where y_{it}, μ_i, λ_i and ε_{it} are scalar variables; $y_{t-j} = (y_{1t-j}, \ldots, y_{nt-j})$ is $n \times 1$ vector; X_t is $k \times 1$ vector including stationary explanatory variables; β_i and γ_i are $k \times 1$ vectors including the corresponding vectors of coefficients; $A_j(i)$ and $B_j(i)$ represents the i-th row of A_j and B_j, respectively; and $\mathbf{1}[\cdot]$ is the indicator function. It is important to note that, first we assume that the regressors are the same across the equations in the current framework. As we show later, this assumption can be easily relaxed to consider the case in which each equation has different set of regressors. Second, the roots of $\{I - A(L)L\}$ and $\{I - B(L)L\}$ are outside the unit circle in which L is lag operator. Thirdly, when one break happens, say in equation i for example, we replace the i-th row of A_js with the corresponding i-th row of B_js and denote as C_js. After all the breaks occur, C_js become B_js. We assume that the roots of the sequence of $\{I - C(L)L\}$ are outside the unit circle as well.

It is convenient to write the system of equations (1.2.1) in its matrix form

$$y_t = (V_t' \otimes I)\theta + D(\tau)(V_t' \otimes I)\delta + \varepsilon_t \quad (1.2.2)$$

where $V_t' = (1,\, y_{t-1}^{'}, \ldots\, y_{t-p}^{'},\, X_t^{'})$, $\theta = vec(\mu,\, A_1, \ldots,\, A_p, \beta)$, $\delta = vec(\lambda,\, B_1, \ldots,\, B_p, \gamma)$, and $D(\tau) = diag(1[t > \tau_1], \ldots 1[t > \tau_n])$ is $n \times n$ matrix. Note model (1.2.2) is that of a full structural change in which it allows all the coefficients to change. If it is known that only a subset of coefficients such as the intercept has a possible break, a partial break structural break model is more appropriate. This leads to the consideration of a general partial structural break model

$$y_t = (V_t' \otimes I)\theta + D(\tau)(V_t' \otimes I)S'S\delta + \varepsilon_t \quad (1.2.3)$$

where S is a selection matrix, containing 0s and 1s and having full row rank. Note that $S'S$ is idempotent with non zero elements only on the diagonal. The rank of S is equal the number of coefficients that are allowed to change. For $S = I$, model (1.2.2) is obtained. For $S = s \otimes I$ with $s = (1, 0, \ldots 0)$, we have

$$y_t = (V_t' \otimes I)\theta + D(\tau)\lambda + \varepsilon_t \quad (1.2.4)$$

which has a break in the intercept only. The system (1.2.3) can be rewritten more compactly as

$$y_t = Z_t'(\tau)\beta + \varepsilon_t \tag{1.2.5}$$

where $Z_t'(\tau) = ((V_t' \otimes I),\, D(\tau)(V_t' \otimes I)S')$ and $\beta = (\theta',\, (S\delta)')'$.

As a matter of notation, we let "$\xrightarrow{p}$" denotes converge in probabilities; "$\xrightarrow{d}$" denotes converge in distribution; "$\xrightarrow{a.s.}$" denotes almost sure converge; and "$\Rightarrow$" denotes weak convergence. Our analysis is carried under the following set of assumptions:

Assumption 1: *Let ε_t be a martingale difference sequence with respect to $\mathcal{F}_{t-1} = \sigma$-field$(Z_t,\, \varepsilon_{t-1}, Z_{t-1},\, \varepsilon_{t-2}\ldots)$ satisfying, for some $\alpha > 0$, $\max_i \sup_t E(\varepsilon_{it}^{4+\alpha}) < \infty$ and $E(\varepsilon_t\varepsilon_{t-j}'|\mathcal{F}_{t-1}) = \Sigma$ for $j = 0$ and 0 otherwise.*

Assumption 2: *Suppose that $E_t X_t = \mu_x$ for all t, $\max_i \sup_t E(X_{it}^{4+\alpha}) < \infty$, $T^{-1}\sum_{t=1}^{T}(X_t - \mu_x)(X_t - \mu_x)' \xrightarrow{p} M_{xx}(0)$, $T^{-1}\sum_{t=1}^{T} X_t y_{t-j}' \xrightarrow{p} EX_t y_{t-j}' = M_{xy}(j)$, $j = -p, \ldots, p$ and $\chi_T(\cdot) = T^{-1/2}\sum_{t=1}^{[T\tau_i]}(X_t - \mu_x)$, ,$[x]$ represents the integer part of x, and $B_x(\cdot)$ is a Brownian motion with covariance matrix $M_{xx}(0)$.*

1.2.2 An Example-Two Equations Case

To illustrate the notation and the framework in the previous subsection, it is useful to consider a much simpler two equations system as follows:

$$\begin{aligned} y_{1t} &= \mu_1 + \sum_{j=1}^{p}(a_{11}^j,\, a_{12}^j)\begin{pmatrix} y_{1t-j} \\ y_{2t-j}\end{pmatrix} + (\beta_{11},\, \ldots, \beta_{1k})\begin{pmatrix} X_{1t} \\ \vdots \\ X_{kt}\end{pmatrix} + \\ &\quad 1[t > \tau_1](\lambda_1 + \sum_{j=1}^{p}(b_{11}^j,\, b_{12}^j)\begin{pmatrix} y_{1t-j} \\ y_{2t-j}\end{pmatrix} + (\gamma_{11},\, \ldots, \gamma_{1k})\begin{pmatrix} X_{1t} \\ \vdots \\ X_{kt}\end{pmatrix}) + \varepsilon_{1t} \end{aligned}$$

$$
\begin{aligned}
y_{2t} &= \mu_2 + \sum_{j=1}^{p} (a_{21}^j,\, a_{22}^j) \begin{pmatrix} y_{1t-j} \\ y_{2t-j} \end{pmatrix} + (\beta_{21},\, \ldots, \beta_{2k}) \begin{pmatrix} X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix} + \\
& \quad 1[t > \tau_2](\lambda_2 + \sum_{j=1}^{p} (b_{21}^j,\, b_{22}^j) \begin{pmatrix} y_{1t-j} \\ y_{2t-j} \end{pmatrix} + (\gamma_{21},\, \ldots, \gamma_{2k}) \begin{pmatrix} X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix}) + \varepsilon_{2t}
\end{aligned}
$$

Packing the two equations above together, we have the following equivalent expression in terms of matrix:

$$
\begin{aligned}
\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} &= \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \sum_{j=1}^{p} \begin{pmatrix} a_{11}^j & a_{12}^j \\ a_{21}^j & a_{22}^j \end{pmatrix} \begin{pmatrix} y_{1t-j} \\ y_{2t-j} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \ldots & \beta_{1k} \\ \beta_{21} & \ldots & \beta_{2k} \end{pmatrix} \begin{pmatrix} X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix} \\
& \quad + \begin{pmatrix} 1[t > \tau_1] & 0 \\ 0 & 1[t > \tau_2] \end{pmatrix} \left[\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} + \sum_{j=1}^{p} \begin{pmatrix} b_{11}^j & b_{12}^j \\ b_{21}^j & b_{22}^j \end{pmatrix} \begin{pmatrix} y_{1t-j} \\ y_{2t-j} \end{pmatrix} \right. \qquad (1.2.6) \\
& \quad \left. + \begin{pmatrix} \gamma_{11} & \ldots & \gamma_{1k} \\ \gamma_{21} & \ldots & \gamma_{2k} \end{pmatrix} \begin{pmatrix} X_{1t} \\ \vdots \\ X_{kt} \end{pmatrix} \right] + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \qquad (1.2.7)
\end{aligned}
$$

Next we define V_t, θ, and δ in the similar way as the n equations case:

$$
\begin{aligned}
V_t^{'} &= (1,\, y_{1t-1},\, y_{2t-1}, \ldots,\, y_{1t-p},\, y_{2t-p},\, X_{1t}, \ldots,\, X_{kt})_{(1+2p+k)\times 1} \\
\theta &= (\mu_1,\, \mu_2,\, a_{11}^1,\, a_{12}^1,\, a_{21}^1\, a_{22}^1, \ldots\, a_{11}^p,\, a_{12}^p,\, a_{21}^p\, a_{22}^p, \\
& \quad \beta_{11}, \ldots,\, \beta_{1k},\, \beta_{21}, \ldots,\, \beta_{2k})_{(2+4p+2k)\times 1} \\
\delta &= (\lambda_1,\, \lambda_2,\, b_{11}^1,\, b_{12}^1,\, b_{21}^1\, b_{22}^1, \ldots\, b_{11}^p,\, b_{12}^p,\, b_{21}^p\, b_{22}^p, \\
& \quad \gamma_{11}, \ldots,\, \gamma_{1k},\, \gamma_{21}, \ldots,\, \gamma_{2k})_{(2+4p+2k)\times 1}
\end{aligned}
$$

$$D(\tau) = \begin{pmatrix} 1[t > \tau_1] & 0 \\ 0 & 1[t > \tau_2] \end{pmatrix}_{2\times 2}$$

Therefore we can write equation (1.2.6) in form of (1.2.2):

$$y_{t(2\times1)} = (V'_{t(1+2p+k)\times1} \otimes I_{2\times2})\theta_{(2+4p+2k)\times1} + D(\tau)_{2\times2}(V'_{t(1+2p+k)\times1)} \otimes I_{2\times2})\delta_{(2+4p+2k)\times1} + \varepsilon_{t(2\times1)}$$

Note that this is a pure structural change model. Suppose we want to study the partial structural change model, for instance, only changes on intercept in the first equation and changes on lagged dependent variables in second equation are allowed. Now we have $(1 + 2 * p)$ coefficients are allowed to change. We define

$$S = \begin{array}{cc} & \begin{array}{ccccccccccccc} \lambda_1 & a_{11}^1 & a_{12}^1 & a_{21}^1 & a_{22}^1 & a_{11}^2 & a_{12}^2 & a_{21}^2 & a_{22}^2 & \ldots & a_{21}^p & a_{22}^p & \beta \end{array} \\ \begin{array}{c} \lambda_1 : \\ a_{21}^1 : \\ a_{22}^1 \\ a_{21}^2 \\ a_{22}^2 \\ \vdots \\ a_{21}^p \\ a_{22}^p \end{array} & \left[\begin{array}{ccccccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ldots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots & 0 & 1 & 0 \end{array}\right]_{(1+2p)\times(2+4p+2k)} \end{array}$$

The rank of S is equal to number of coefficients that are allowed to change$(1+2*p)$, and $S'S$ is idempotent with non zero elements only on the diagonal.

1.3 Estimation

1.3.1 Estimation Method

The first raised question is how one can estimate the model with unknown break dates. This problem has been well considered by various authors, using a variety of approaches. For instance, Picard (1985) provided a Gaussian maximum likelihood estimation of the break dates in the case that a univariate process follows a finite order autoregression. Recently, Qu and Perron (2007) considered the quasi-maximum likelihood estimation that assumes serially uncorrected Gaussian errors under the multivariate regressions. Our method of estimation is similar as those in Bai et al. (1998) with three notable features. First, the covariance matrix Σ of error terms is explicitly treated as unknown and estimated. Second, we only assume the error terms from a sequence of martingale differences with some moment conditions, and use quasi-Gaussian maximum likelihood estimation. Thirdly, we allow some of regression parameters to be estimated with the full sample to gain efficiency.

The method of estimation we considered here is quasi-Gaussian maximum likelihood estimation. Suppose that $||S\delta|| \neq 0$, then there indeed exists a set of ordered structural breaks. For a given combination of the break dates $\tau = (\tau_1 \dots \tau_n)$, the Gaussian quasi likelihood function is

$$L_T(\tau, \beta, \Sigma) = \prod_{t=1}^{T} f(y_t | Z_t(\tau); \tau, \beta, \Sigma) \tag{1.3.1}$$

where

$$f = (y_t | Z_t(\tau); \tau, \beta, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\{-\frac{1}{2}[y_t - Z_t'(\tau)\beta]' \Sigma^{-1} [y_t - Z_t'(\tau)\beta]\}$$

The straightforward method on estimation is based on grid search over all the combination of breaks dates. Here we impose the following assumption on the set of feasible break date.

***Assumption 3**: The maximization of $L_T(\tau, \beta, \Sigma)$ is taken over all the combinations of break dates $\tau = (\tau_1 \ldots \tau_n)$ in the following set*

$$\Lambda_\epsilon = \{(\tau_1, \ldots, \tau_i, \ldots, \tau_n) = (T\lambda_1, \ldots, T\lambda_i, \ldots, T\lambda_n); \min\{\lambda_i\} \geq \epsilon, \max\{\lambda_i\} \leq 1 - \epsilon\}$$

where ϵ is a trimming value.

ϵ represents that an initial and ending fraction of sample are trimmed. This is often taken to either 0.15 or 0.1. Therefore, we proceed the estimation as follows. First for each combination $\tau^{(i)} = (\tau_1^{(i)}, \ldots, \tau_n^{(i)})$, the associated estimates of β and Σ are obtained by maximizing the quasi likelihood function. Let $\hat{\beta}(\tau^{(i)})$, and $\hat{\Sigma}(\tau^{(i)})$ denote the resulting estimates. Substituting them in the objective function and denoting the resulting quasi likelihood functions as $L_T(\tau^{(i)}, \hat{\beta}(\tau^{(i)}), \hat{\Sigma}(\tau^{(i)}))$, the estimated break points $\hat{\tau}$ are such that

$$\hat{\tau} = arg \max_{\{\tau^i\} \in \Lambda_\epsilon} L_T(\tau^i, \hat{\beta}(\tau^i), \hat{\Sigma}(\tau^i)), \tag{1.3.2}$$

where the maximization is taken over all the combinations of $\{\tau^{(i)}\}$.[3] It is worth mentioning that the computation of maximum likelihood estimates of order $O(T^n)$. The estimated break dates $\hat{\tau}$ are therefore global maximum of the objective function. Finally, the estimated regression parameter are associated quasi maximum likelihood estimates at the estimated combination of break points $\hat{\tau}$, i.e. $\hat{\beta} = \hat{\beta}(\hat{\tau})$, $\hat{\Sigma} = \hat{\Sigma}(\hat{\tau})$.

1.3.2 Efficient Algorithm for Estimation

As we show in the previous subsection, the computation of estimates in the general framework considered in this paper is not a trivial issue. In principle, we can use

[3]For the two equations and single break problem we have $(1-2\epsilon)^2T^2$ combinations if we premise these breaks fall into $(\epsilon T, (1-\epsilon)T)$. ϵ is the trimming value.

a grid search, but this approach becomes rapidly impractical since it involves the computation of maximum likelihood estimates of order $O(T^n)$. We now consider an algorithm based on the principle of dynamic programming. Our approach is an natural extension of works by Hawkins (1976), Bai and Perron (2003) and Qu and Perron (2007). The basic idea is as follows. With any possible combination of ordered breaks, it is the case the that the overall value of the log likelihood function is the sum of the values associated with a particular combination of at most $n+1$ segments. Hence, if we have the information about the log likelihood values for all possible segments, of which there are at most $T \times (T+1)/2$, then all that is needed is a method to assess which particular combination of $n+1$ segments leads to the highest value of the likelihood function. This is achieved using a dynamic programming algorithm. More thorough details can be found in Bai and Perron (2003).

1.3.3 Statistical Properties

We now consider the statistical properties of the estimates. In order to derive the asymptotic properties of these estimates, we follow Picard (1985), Bai and Perron (1998, 2003) and Bai et al. (1998) among others, and make the following assumption on the magnitude of the shifts:

Assumption 4: *Let* $\beta^0 = (\theta^{0\prime}, (S\delta_T)^{\prime})^{\prime}$, *in which* δ_T *is a sequence such that* $\delta_T = \delta^0 v_T$. $v_t > 0$ *is a scalar satisfying* $v_T \to 0$ *and* $\sqrt{T} v_t/(\log T) \to \infty$.

This assumption implies a shrinking shifts asymptotic framework, in which the magnitudes of these shifts converges to zero as the sample size increases. One obvious reason for considering small is that, if we show that a break with a small magnitude of shift can be consistently estimated, it must be the case that we can consistently estimate a break with larger magnitude of shift, for the larger the magnitude of shifts, the easier to identify a break.

The joint behavior of $(\hat{\tau}, \hat{\beta}(\hat{\tau}), \hat{\Sigma}(\hat{\tau}))$, particularly, consistency and their rates of convergence are examined in this subsection. First, the likelihood function, such as $L(\tau,\, \beta^0 + T^{-1/2}\beta,\, \Sigma^0 + T^{-1/2}\Sigma)$, are reparameterized. These ordered break dates $\tau = (\tau_1, \ldots, \tau_n)$ are reparameterized such that $\tau_i = \tau_i(v) = \tau_i^0 + [v v_T^{-2}]$, for $v \in R$. When v varies, τ_i can take on all possible integer values. We define the likelihood function to be zero for τ_i greater than T. It is worth mentioning that maximizing the original likelihood function is equivalent to maximizing the reparameterized likelihood. We let $(\tau^0, \beta^0, \Sigma^0)$ denote the true values of these parameter in this data generating process, and construct the quasi-likelihood ratio as follows

$$\begin{aligned} LR_T &= \frac{L(\tau,\, \beta^0 + T^{-1/2}\beta,\, \Sigma^0 + T^{-1/2}\Sigma)}{L(\tau^0, \beta^0, \Sigma^0)} \\ &= \frac{\prod_{t=1}^{T} f(y_t|Z_t(\tau); \tau,\, \beta^0 + T^{-1/2}\beta,\, \Sigma^0 + T^{-1/2}\Sigma)}{\prod_{t=1}^{T} f(y_t|Z_t(\tau); \tau^0, \beta^0, \Sigma^0)} \\ &= \frac{|\Sigma^0 + T^{-1/2}\Sigma_0|^{-T/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T} \varepsilon_t(\tau_i) \left(\Sigma_0 + T^{-1/2}\Sigma\right)^{-1} \varepsilon_t(\tau_i)\right\}}{|\Sigma_0|^{-T/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T} \varepsilon_t' \Sigma_0^{-1} \varepsilon_t\right\}} \end{aligned} \tag{1.3.3}$$

where $\varepsilon_t(\tau_i) = y_t - Z_t(\tau)'(\beta_0 + T^{-1/2}\beta)$.

Theorem 1: *Under the assumptions A1-A4, for the break dates estimates* $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_n)$

$$v_T^2(\hat{\tau}_i - \tau_i^0) = O_p(1)$$

and

$$\begin{aligned} \sqrt{T}(\hat{\beta} - \beta^0) &= O_p(1) \\ \sqrt{T}(\hat{\Sigma} - \Sigma^0) &= O_p(1) \end{aligned}$$

This theorem gives the rate of convergence of the estimates. The results are the same as in most of other cases consider in the literature, see appendix A for the proof. The basic idea is as follows: First it is clear that maximizing the original likelihood function is equivalent to maximizing the likelihood ratio. Suppose that v^*, β^*, and Σ^* maximize the likelihood ratio, then $v^* = \upsilon_T^2(\hat{\tau}_i - \tau_i^0)$, $\beta^* = \sqrt{T}(\hat{\beta} - \beta^0)$, and $\Sigma^* = \sqrt{T}(\hat{\Sigma} - \Sigma^0)$. Thus to show $\upsilon_T^2(\hat{\tau}_i - \tau_i^0)$, $\sqrt{T}(\hat{\beta} - \beta^0)$, and $\sqrt{T}(\hat{\Sigma} - \Sigma^0)$ are all stochastically bounded, it is sufficient to show that v^*,β^*, and Σ^* are stochastically bounded. This, in turn is equivalent to showing that the likelihood ratio cannot achieve its maximum when any of parameters, v, β, and Σ, is too large.

1.4 Test Statistics

We now consider testing for ordered structural breaks. It is important to note we focus on the changes in the coefficients of the conditional mean. Also, as we mentioned above, we can allow only a subset of coefficients to change across regimes, hence partial structural breaks are permitted. The test we proposed here is a likelihood ratio test for the null hypothesis of no change in any of the coefficients versus an alternative hypothesis with ordered structural breaks.

In order to derive the limiting distribution of the test under the null hypothesis of no structural change, we impose the following additional assumptions on the data generating.

Assumption 5: *$\tau_i^0 = [\lambda_i^0 T]$ for some $0 < \lambda_i^0 < 1$, and $[\cdot]$ is the greatest integer function. This assumes that the shift point is bounded away from the end points, which is used for asymptotic purpose.*

1.4.1 The specification of the alternative hypothesis

Under the alternative hypothesis, we consider the case that the whole coefficients to change across break dates. For instance, the break dates τ_i, can be constructed as a group of n different break dates. For a given i^{th} equation, we have

$$y_{it} = \begin{cases} \mu_i + \sum_{j=1}^{p} A_j(i) y_{t-j} + \beta_i' X_t + \varepsilon_{it} & t < \tau_i \\ \mu_i + \sum_{j=1}^{p} A_j(i) y_{t-j} + \beta_i' X_t + \left(\lambda_i + \sum_{j=1}^{p} B_j(i) y_{t-j} + \gamma_i' X_t\right) + \varepsilon_{it} & t \geq \tau_i \end{cases}$$

Hence, following the notation in the previous section, $n_y + n_x + 1$ coefficients are allowed to change. Note that this framework is such that any coefficients which is allowed to change does so simultaneously. Stacking the system equation by equation, we have

$$y_t = (V_t' \otimes I)\theta + D(\tau)(V_t' \otimes I)S'S\delta + \varepsilon_t$$

where the notations are exactly the same as in Section 2.

Under the null hypothesis of no structural change, the estimates are the values $\hat{\beta}$ and $\hat{\Sigma}$ that jointly solve the following system of equations

$$\begin{aligned} \hat{\Sigma} &= \frac{1}{T} \sum_{t=1}^{T} \left(y_t - Z_t' \hat{\beta}\right) \left(y_t - Z_t' \hat{\beta}\right)' \\ \hat{\beta} &= \left(\sum_{t=1}^{T} Z_t \hat{\Sigma}^{-1} Z_t'\right)^{-1} \left(\sum_{t=1}^{T} Z_t \hat{\Sigma}^{-1} Y_t'\right) \end{aligned}$$

with the resulting value of the log-likelihood function being

$$\log L_T = -\frac{T}{2}(\log 2\pi + 1) - \frac{T}{2} \log |\hat{\Sigma}|$$

For a given ordered structural break $\tau = (\tau_1, \ldots, \tau_n)$, the class of models described above can be estimated by quasi maximum likelihood. Denote the log-likelihood value by $\log \hat{L}(\tau_1, \ldots, \tau_n)$. The proposed test is the maximal value of the likelihood ratio test over all admissible partitions in the set Λ_ε defined by Assumption $A.4$, i.e.,

$$
\begin{aligned}
\sup LR_T &= \sup_{(\tau_1,\ldots,\tau_n)\in\Lambda_\varepsilon} 2\left[\log \hat{L}_T(\tau,\beta,\Sigma) - \log \hat{L}_T\right] \\
&= 2[\log \hat{L}_T(\hat{\tau}_1,\ldots,\hat{\tau}_n) - \log \hat{L}_T] \qquad (1.4.1)
\end{aligned}
$$

where the estimates $\hat{\tau} = (\hat{\tau}_1, \ldots, \hat{\tau}_n)$ are the QMLE obtained considering only those partitions in Λ_ε.[4] The parameter ε acts as a truncation which imposes a minimal length for each segment and will affect the limiting distribution of the test. It is also useful to describe the exact form of the log likelihood value and the estimates of the coefficients for some leading cases.

1.4.2 The limiting distribution of the test.

We now consider the limiting distribution of the $\sup LR_T$ test under the null hypothesis in the context of the class of models described in the previous subsection.

Theorem 2 *Under the assumptions A1-A*5*, with the* $\sup LR_T$ *test constructed for an alternative hypothesis in the class of models described in previous subsection, we have, as* $T \to \infty$*,*

$$
\sup LR_T \Rightarrow \sup_{(\lambda_1,\ldots,\lambda_i,\ldots,\lambda_n)\in\Lambda_\varepsilon} \sum_{i=1}^{n} \frac{\left\|\lambda_i W_n(\lambda_{i+1}) - \lambda_{i+1} W_n(\lambda_i)\right\|^2}{(\lambda_{i+1}-\lambda_i)\,\lambda_i\lambda_{i+1}} \qquad (1.4.2)
$$

where $W_n(\cdot)$ *are* n *dimensional vectors of independent Wiener processes, and* $\|\cdot\|$ *represents the Euclidean norm.*

Note that the limiting distribution of the $\sup LR_T$ not just depends on the number of coefficients are allowed to change, but also depends on the trimming values ϵ. This form of this limiting distribution is similar as the expression in Theorem 1 of Bai

[4] See Andrews (1993); Andrews and Fair (1988) for more details on how to construct hypothesis testing for parameter instability.

et al. (1998). More details can found in the appendix B, in which we utilize the proof in Bai et al. (1998) and Qu and Perron (2007).

1.5 Monte Carlo Study

In this section, we provide Monte Carlo study related to estimation method. The framework used in Monte Carlo is similar as the system (4.1) in Bai et al. (1998). For a change in the intercept in an autoregressive system, the data generating process is the system with a single break in each equation as following:

$$\begin{aligned} y_t &= (\lambda \iota_n)\, d_t(T\delta_0) + (\beta I_n) y_{t-1} + \varepsilon_t \\ \varepsilon_t &\sim i.i.d. N(0, \Sigma_\varepsilon) \end{aligned} \tag{1.5.1}$$

where ι_n is an $n-$vector of 1's and ε_t is $n \times 1$. In particular, we consider the following bivariate system with a single break in each equation:

$$\begin{aligned} y_{1t} &= \mu_1 + \lambda_1 1[t > \tau_1] + \beta_{11} y_{1t-1} + \beta_{12} y_{2t-1} + \varepsilon_{1t} \\ y_{2t} &= \mu_2 + \lambda_2 1[t > \tau_2] + \beta_{21} y_{1t-1} + \beta_{22} y_{2t-1} + \varepsilon_{2t} \end{aligned}$$

where $(\varepsilon_{1t}, \varepsilon_{2t})' \sim i.i.d.\, N(0, I_2)$. As we can see, only the intercept is allowed to change at some date τ_i for the ith equation. First, the value of μ_is is set to one. Second, we consider three values of the magnitudes of mean shift λ_i. λ_is can be one of values (0.25, 0.50, 1.00). Thirdly, the autoregressive parameters are chosen from the following three sets

$$\begin{aligned} \beta^{(1)} &= (0.10,\, 0.10,\, 0.10,\, 0.10) \\ \beta^{(2)} &= (0.50,\, 0.50,\, 0.50,\, 0.50) \\ \beta^{(3)} &= (0.90,\, 0.90,\, 0.90,\, 0.90) \end{aligned}$$

We set the number of observations T equals to 100, which is a reasonable choice given the computational cost of simulation. The last thing about the data generating process is to choose break dates. Without of loss generality, the break date in the first equation is kept fixed at $\tau_1 = 30$. The break date τ_2 in the second equation then takes values either 30, or 50, or 70. In the Monte Carlo study, we run 500 replications. Figure 1 to 3 show the results on estimate breaks.

1.6 Application

It is widely well known that there is a slowdown in the growth rate of output in European economies and United States during the postwar, particularly in the 1970s. Several statistical techniques have been proposed to identify the date of slowdown, such as structural breaks model, regime switch model among other time varying coefficient models. Obviously, our main focus is put on the task of dating this slowdown using structural break techniques. The starting point for this type of investigation is the observation by Banerjee et al. (1992). They found that output growth rate in France, Germany, and Italy each appeared to be difference stationary, but that there appeared to be a break in the mean growth rate for each country during the sample. Their analysis was based on hypothesis testing under the framework of strict univariate. As Bai et al. (1998) show that there can be substantial gains from using multivariate inference about the break dates. However, as we show in the previous section that we can also have substantial efficiency gains if we relax the assumption on common structural breaks.

For comparability to their studies, we first employ data set used in Banerjee et al. (1992) and Bai et al. (1998) data. In particular, the three European series are the logarithms of quarterly GDP for France and Italy and GNP for Germany. The logarithm of quarterly GDP for the U.S. are also included in the study. Since the data are available over different periods, the system results consider the joint behavior of output over only a short common period, 1962 : $Q1$ to 1982 : $Q4$. For mode details on the the data and data sources, refer to the descriptions in Banerjee et al. (1992).

First, we follow the common practice in the empirical literature, such as Banerjee et al. (1992) and Perron (1989) among others. We tested the null hypothesis that each of these series had a unit root, against the alternative that the series was stationary around a linear time trend, possibly with a break in the time trend at an unknown date. For all these series, it is no surprise that we find the univariate analysis of these European countries and United States output data provided no evidence against the unit root null hypothesis. Thus, under our assumption each series is $I(1)$ process, possibly with a change in drift. Differencing each of these series leads the the univariate stationary autoregressive representation. For univariate analysis, now y_t, is the growth rate of output in each country. Since there is no exogenous variables, X_t, is dropped out. In this case, the break term corresponds to a shift in the mean growth rate of output. While the series are modeled as jointly having the stationary autoregressive representation, where y_t is interpreted as the vector of growth rates of output of the various countries and X_t is omitted as well. It is still important to note that, in contrast to Bai et al. (1998) we do not impose the common break restriction on this model.

Table 2 and Table 3 present the structural breaks statistics results for three European countries and United States. As shown in Table 2 Section A, for France and Germany, treated as univariate series, both of the test statistics rejects at the 1% level; for Italy, both reject at the 5% level. The point estimates of the break date are in 1974 for France and Italy, although for Italy the estimate is imprecise. In contrast, for U.S. output the hypothesis of a constant mean growth rate cannot be rejected at the 10% level using any of the tests.

Section B in Table 2 shows the results common breaks statistics used in Bai et al. (1998). Several things need to be highlighted here. First, we take a look at the

France-Italy system, for which the univariate evidence is most consistent with a single common break date. It is not surprising that the test statistics reject the hypothesis of no break in the mean growth rate against the alternative of a break in the mean at a common break date. Second, the other bivariate systems also reject the null of no break against the common-date alternative. Bai et al. (1998) interpret these results as support for proceeding to construct interval estimates for a common break date including Germany and United States in the system. Thirdly, this multivariate analysis points to a slowdown in European and United States output. The slowdown occurred approximately simultaneously in France and Italy and, arguably, in Germany and United States as well. Of course, this dating coincides with conventional wisdom; the contribution of Bai et al. (1998) is that this date can now be associated with the formal measure of uncertainty provided by a tight confidence interval spanning slightly more than three years.

However, as the univariate evidence shows that it is clear that Germany and United States have breaks at the different time as the other countries, which motive us to apply ordered structural breaks to this example. Table 3 shows the results of ordered structural breaks. First, all the multivariate systems reject the null of no break against a set of ordered structural breaks. Second, it is worth mentioning that now we can identify one break in the systems including United States, for which the univariate analysis show no evidence of structural breaks. Finally, as we also mentioned earlier, our dating of changes in output growth is against the conventional wisdom. We argue that imposing common breaks is strong restriction, and often lead us imprecise estimates.

1.7 Conclusion

This paper provides techniques for testing, estimation, and computation of ordered structural breaks across equations in multivariate linear regressions. Our framework relaxes the often unnoticed underlying assumption of common breaks. A likelihood ratio test assuming normal errors is proposed in this paper in order to detect the ordered structural breaks in multivariate linear regressions. We take the advantage of our framework based on dynamic programming and adopt the efficient algorithm of Bai and Perron (2003). We also provide results about the consistency and rate of convergence when searching for ordered structural breaks. We finally presents the Monte Carlo study and an empirical example. It is worth mentioning two limitations in this paper. First, we only consider stationary variables as regressors. Therefore, we can not deal with cases including integrated or trending regressors. Second, we have shown how to construct confidence interval for these estimates.

1.8 Chapter 1: Appendix

1.8.1 Proof of Theorem 1

The proof proceeds similar as those in Bai et al. (1998) and Qu and Perron (2007). We first present a set of properties of the quasi-likelihood ratios. We then show that Theorem 1 can be derived as a consequence of these properties.

To begin with, we consider the model without any breaks as follows;

$$y_t = (V_t' \otimes I)\theta + \varepsilon_t \tag{1.8.1}$$

where $V_t' = (1,\, y_{t-1}^{'}, \dots\, y_{t-p}^{'},\, X_t')$, $\theta_0 = vec(\mu,\, A_1, \dots,\, A_p, \beta)$, and ε_t are martingale differences with variances Σ_0. We let (θ_0, Σ_0) denote the true parameters. Consider the quasi-likelihood ratio based on the first τ_i observations

$$\begin{aligned}\mathcal{L}(1,\tau_i;\theta,\Sigma) &= \frac{\prod_{t=1}^{\tau_i} f\left(y_t|y_{t-1},\dots,\theta_0+T^{-1/2},\Sigma_0+T^{-1/2}\Sigma\right)}{\prod_{t=1}^{\tau_i} f\left(y_t|y_{t-1},\dots,\theta_0,\Sigma_0\right)} \\ &= \frac{|\Sigma^0+T^{-1/2}\Sigma_0|^{-\tau_i/2}\exp\left\{-\frac{1}{2}\sum_{t=1}^{\tau_i}\varepsilon_t(\tau_i)\left(\Sigma_0+T^{-1/2}\Sigma\right)^{-1}\varepsilon_t(\tau_i)\right\}}{|\Sigma^0|^{-\tau_i/2}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T}\varepsilon_t(\tau_i)\Sigma_0^{-1}\varepsilon_t(\tau_i)\right\}}\end{aligned} \tag{1.8.2}$$

where

$$\varepsilon_t(\tau_i) = y_t - (V_t' \otimes I)(\theta_0 + T^{-1/2}\theta) = \varepsilon_t - T^{-1/2}(V_t' \otimes I)\theta$$

Denote by $\hat{\theta}(\tau_i)$ and $\hat{\Sigma}(\tau_i)$ as the values of θ and Σ such that $\mathcal{L}(1, \tau_i; \theta, \Sigma)$ achieves its maximum. Then we have the following properties:

Property 1: For each $\delta \in [0, 1]$

$$\sup_{T\delta \le \tau_i \le T} \mathcal{L}\left(1, \tau_i; \hat{\theta}(\tau_i), \hat{\Sigma}(\tau_i)\right) = O_p(1) \tag{1.8.3}$$

$$\sup_{T\delta \le \tau_i \le T} \left(||\hat{\theta}(\tau_i)|| + ||\hat{\Sigma}(\tau_i)||\right) = O_p(T^{-1/2}) \tag{1.8.4}$$

This property corresponds to property 1 of Bai et al. (1998) and Qu and Perron (2007). It says that the likelihood ratios and the maximum likelihood estimates are bounded in probability. The uniformity of the bound is important since we need to search over all admissible combinations to find these break dates. Since we take similar assumptions on ε_t as Bai et al. (1998), this result is a naturally consequence of the functional central limit theorem for martingale differences. The proof is omitted here. For details, see Bai et al. (1998) and Qu and Perron (2007).

Property 2: For each $\varepsilon > 0$, there exists a $B > 0$ such that for large T

$$Pr\left(\sup_{1 \le \tau_i \le T} T^{-B}\mathcal{L}\left(1, \tau_i; \hat{\theta}(\tau_i), \hat{\Sigma}(\tau_i)\right) > 1\right) < \varepsilon \tag{1.8.5}$$

This property says that the log-valued quasi-likelihood ratio has its maximum value bounded by $O_p(\log T)$, which provides a bound for the sequential quasi-likelihood function in small samples.

Proof of Property 2:

The likelihood ratio evaluated at $\hat{\theta}(\tau_i)$ and $\hat{\Sigma}(\tau_i)$ can be rewritten as

$$\log \mathcal{L}\left(1, \tau_i; \hat{\theta}(\tau_i), \hat{\Sigma}(\tau_i)\right) = -\frac{\tau_i}{2}\left(\log|\Sigma^*(\tau_i)| - \log|\Sigma_0|\right) + \frac{1}{2}\left(\sum_{t=1}^{\tau_i} \varepsilon_t' \Sigma_0^{-1} \varepsilon_t - \tau_i n\right) \tag{1.8.6}$$

where

$$\Sigma^*(\tau_i) = \frac{1}{\tau_i}\sum_{t=1}^{\tau_i} \varepsilon_t \varepsilon_t' - \left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} \varepsilon_t V_t'\right)\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} V_t V_t'\right)^{-1}\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} V_t \varepsilon_t\right)$$

thus by adding and subtracting an identity matrix, we obtain

$$\begin{aligned} -\frac{\tau_i}{2}\left(\log|\Sigma^*(\tau_i)| - \log|\Sigma_0|\right) &= -\frac{\tau_i}{2}\log\left| I + \frac{1}{\tau_i}\sum_{t=1}^{\tau_i}(\eta_t\eta_t' - I)\right. \\ &\quad \left. - \left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}\eta_t V_t'\right)\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} V_t V_t'\right)^{-1}\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} V_t \eta_t\right)\right| \end{aligned}$$

where $\eta_t = \sum_0^{-1/2} \varepsilon_t$ with $E\eta_t = 0$ and $Var(\eta_t) = I$. Applying a Taylor series expansion, we have

$$-\frac{1}{2}tr\left(\sum_{t=1}^{\tau_i}(\eta_t\eta_t' - I)\right) + \frac{1}{2}tr(\Phi_{\tau_i}) \tag{1.8.7}$$

$$+\frac{\tau_i}{4}tr\left\{\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}(\eta_t\eta_t' - I) - \frac{1}{\tau_i}\right)^2\right\} + O_p(1) \tag{1.8.8}$$

where $O_p(1)$ is uniformly in τ_i. Therefore we have the following expression of $(A.6)$[5]

$$\begin{aligned} \log \mathcal{L}\left(1, \tau_i; \hat{\theta}(\tau_i), \hat{\Sigma}(\tau_i)\right) &= \frac{1}{2}tr(\Phi_{\tau_i}) + \frac{\tau_i}{4}tr\left\{\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}(\eta_t\eta_t' - I) - \frac{1}{\tau_i}\right)^2\right\} \\ &\quad -\frac{1}{4}tr\left(\frac{1}{\tau_i}\Phi_{\tau_i}^2\right) + O_p(1) \end{aligned}$$

[5]The first term of $(A.7)$ is canceled out with the last term of $(A.8)$.

Now we need to show that

$$\frac{1}{2}tr(\Phi_{\tau_i}) + \frac{\tau_i}{4}tr\left\{\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}(\eta_t\eta_t' - I) - \frac{1}{\tau_i}\right)^2\right\} = O_p(\log T)$$

and

$$\frac{1}{4}tr\left(\frac{1}{\tau_i}\Phi_{\tau_i}^2\right) = O_p(\log T)$$

Then it suffices to show that the above is $O_p(\log T)$ uniformly in τ_i.

By the strong law of large numbers, $\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} V_t V_t'$ converges to a positive definite matrix as $\tau_i \to \infty$, this implies

$$\sup_{t \geq \tau_i}\left\|\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i} V_t V_t\right)^{-1}\right\| = O_p(1)$$

for some fixed $\tau_i > 0$.

Next,

$$\max_{1 \leq t \leq \tau_i} \log \mathcal{L}\left(1, \tau_i; \hat{\theta}(\tau_i), \hat{\Sigma}(\tau_i)\right) = O_p(1)$$

without loss of generality, we may assume $t \geq \tau_i$. By the law of iterated logarithms for martingale differences,

$$\left\|\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}\eta_t V_t\right\| = O_p((\log T)^{-1/2})$$

$$\left\|\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}(\eta_t\eta_t' - I_t)\right\| = O_p((\log T)^{-1/2})$$

uniformly in $\tau_i \in (1, T)$. Thus

$$\|\Phi_k\| = O_p(\log T)$$

uniformly in $\tau_i \in (1, T)$.

In addition, from

$$\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}\eta_t V_t = O_p(1)$$

uniformly in τ_i, we have

$$\tau_i^{-1}\Phi_{\tau_i}^2 = O_p(\log(T))$$

uniformly in $\tau_i \in (1, T)$. This proves Property 2. ■

Property 3: Let $S_T = \{(\theta, \Sigma)\}$; $||\theta|| \geq \log T$ or $||\theta|| \geq \log T$. For any $\delta > (0,1)$, $D > 0$, $\varepsilon > 0$, the following holds when T is large

$$Pr\left(\sup_{T\delta\leq\tau_i}\sup_{(\theta,\Sigma)\in S_T} T^D\mathcal{L}\left(1,\tau_i;\theta,\Sigma\right) > 1\right) < \varepsilon \tag{1.8.9}$$

This property indicates that the value of the quasi-likelihood ratio, when the parameter are away from 0, is arbitrarily small for large T.

Proof of Property 3:

Following Bai et al. (1998) and Qu and Perron (2007), the sequential log-likelihood ratio can be decomposed as

$$\log\mathcal{L}\left(1,\tau_i;\hat{\theta}(\tau_i),\hat{\Sigma}(\tau_i)\right) = \mathcal{L}_{1t} + \mathcal{L}_{2t}$$

where

$$\mathcal{L}_{1t} = -\frac{\tau_i}{2}\log|I+\Psi_T| - \frac{\tau_i}{2}\left[\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}\eta_t'(I_t+\Psi_T)^{-1}\eta_t - \frac{1}{\tau_i}\sum_{t=1}^{\tau_i}\eta_t'\eta_t\right]$$

and

$$\begin{aligned}\mathcal{L}_{2t} \quad &= \quad T^{-1/2}\theta'(I\otimes(I+\Psi_T)^{-1})\Sigma_{t=1}^{\tau_i}(V_t\otimes\eta_t)\\ &\quad -\frac{1}{2}\frac{\tau_i}{T}\theta'\left(\frac{1}{\tau_i}\sum_{t=1}^{\tau_i}V_tV_t'\otimes(I+\Psi_T)^{-1}\right)\theta\end{aligned}$$

with $\eta_t = \Sigma_0^{-1/2}\varepsilon_t$ and $\Psi_T = T^{-1/2}(\Sigma_0^{-1/2}\Sigma\Sigma_0^{-1/2})$.

Let $S_T = S_{1T} \cup S_{2T}$ with

$$S_{1T} = \{(\theta, \Sigma); ||\Sigma|| \geq \log T,\ \ \theta \text{ arbitary}\}$$

and

$$S_{2T} = \{(\theta, \Sigma); ||\Sigma|| \geq \log T, ||\theta|| \leq \log T\}$$

we then need to show that

$$Pr\left(\sup_{T\delta \leq \tau_i} \sup_{(\theta,\Sigma)\in S_{1T}} T^D \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) > 1\right) < \varepsilon$$

and

$$Pr\left(\sup_{T\delta \leq \tau_i} \sup_{(\theta,\Sigma)\in S_{2T}} T^D \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) > 1\right) < \varepsilon$$

The proof of the above two expression proceeds similarly Bai et al. (1998) and Qu and Perron (2007). The details are omitted here. ■

Property 4: For any $\varepsilon > 0$, there exists a $M > 0$ such

$$Pr\left(\sup_{T\delta \leq \tau_i} \sup_{s_M} \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) > \varepsilon\right) < \varepsilon \tag{1.8.10}$$

where s_M is defined as

$$s_M = \{(\theta, \Sigma); ||\theta|| \geq M \,\text{or}\, ||\Sigma|| \geq M\}$$

This property shows that the value of quasi-likelihood ratios is small when it is evaluated outside a bounded set. The next property is similar to Property 4, therefore the proof is shown in next property.

Property 5: Let h_T and d_T be positive sequences such that h_T is non-decreasing, $d_T \to +\infty$ and $h_t d_t^2/T \to h > 0$, where $h < \infty$. Let $S_t = \{\theta, \Sigma\}$; $||\theta|| \geq d_t$ or $||\Sigma|| \geq d_t$. Then for any $\varepsilon > 0$, there exists an $A > 0$,such that when T is large

$$Pr\left(\sup_{Ah_T \leq \tau_i} \sup_{(\theta,\Sigma)\in S_T} \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) > \varepsilon\right) < \varepsilon \tag{1.8.11}$$

This property studies the value of quasi-likelihood ratio when no positive fraction of the observations is involved. It is slightly different from that of Bai et al. (1998), in the sense that the maximum is taken over all the combinations.

Proof of Property 5:

We first define $b_t = T^{-1/2} d_t$. Then by assumption, $b_T = O_p(1)$ if h_T stays bounded and $b_T \to 0$ if $h_T \to \infty$. Furthermore, $h_T b_T^2 \to h$. As in the proof of Property 3, we decompose S_T into two subsets S_{1T} and S_{2T}, where S_{1T} and S_{2T} are defined as in the earlier proof of Property 3 with $\log T$ replaced by d_T. The reminder proof is similar as in Bai et al. (1998), which means we need to show

$$Pr\left(\sup_{Ah_T \leq \tau_i} \sup_{(\theta,\Sigma)\in S_{1T}} \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) > \varepsilon\right) < \varepsilon$$

and

$$Pr\left(\sup_{Ah_T \leq \tau_i} \sup_{(\theta,\Sigma)\in S_{2T}} \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) > \varepsilon\right) < \varepsilon$$

On S_1, all arguments in Property 3 go through if inequalities $(A.9)$ and $(A.11)$ in Bai et al. (1998) still hold true when $\tau_i \geq T\delta$ is replaced with $\tau_i \geq Ah_T$ and for the newly defined b_T.

Since the following inequality,

$$Pr\left(\sup_{Ah_T \leq \tau_i} \frac{1}{\tau_i} \left\| \sum_{t=1}^{\tau_i} (\eta_t \eta_t' - I) \right\| > \alpha b_T\right) < \frac{C}{Ah_T \alpha^2 b_T^2} < \frac{2C}{A\alpha^2 h}$$

for $C > 0$, we can apply the theorems in Hájek and Rényi (1955). The expression above is small if A is large.

Similarly, applying the inequality in Hájek and Rényi (1955). to $1/\tau_i \sum_{t=1}^{\tau_i}(V \otimes \eta_t)$ together with $H_{\tau_i}^{-1} = O_p(1)$ uniformly in large τ_i, we have, for any $\varepsilon > 0$ and $\gamma > 0$, there exists an $A > 0$ such that

$$Pr\left(\sup_{Ah_T \leq \tau_i} \left\| T^{-1/2}\left(H_k^{-1/2} \otimes I\right)\hat{\theta}(\tau_i)\right\| > \gamma b_T\right) < \varepsilon$$

Using the last two inequalities and the same arguments as in Property 3, we obtain, with probability at least $1 - 2\varepsilon$,

$$\mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) \leq -\tau_i b_T^2 C^2/8$$

for all $k \geq Ah_T$ and all $(\theta, \Sigma) \in S_{1T}$, which is further bounded by

$$-Ah_T b_T^2 C^2/8 < -AC^2 h/16 < \log \varepsilon$$

if A is large.

The proof on S_{2T} is almost the same as in Property 3 with only minor changes, therefore is omitted here. ■

Property 6: Under the same hypothesis as Property 5, we have for any $A > 0$

$$\sup_{Ah_T \leq \tau_i} \sup_{(\theta,\Sigma) \in S_T^C} \mathcal{L}\left(1, \tau_i; \theta, \Sigma\right) = O_p(1) \tag{1.8.12}$$

where S_T^C is the complement of S_T with S_T given in Property 5.

The last property states that the quasi-likelihood ratio is simply bounded when evaluated not far away from zero and the number of observations increasing not too fast.

Proof of Property 6:

It suffices to prove the log-valued likelihood ratio is bounded in probability. The log-likelihood ratio consists of two expressions $\mathcal{L}_{1t}$ and $\mathcal{L}_{2t}$ given in Property 3. First consider $\mathcal{L}_{2t}$. It is enough to prove the first term of $\mathcal{L}_{2t}$ is bounded because the second term of $\mathcal{L}_{2t}$ is negative. The norm of the first term is bounded by

$$T^{-1/2}\left(d_T\sqrt{Ah_T}\right)\sup_{S_T^c}\left\|(I+\Psi_T)^{-1}\right\|\sup_{\tau_i\leq Ah_T}\sum_{t=1}^{\tau_i}(V_t\otimes\eta_t)$$

Note that $\|(I+\Psi_T)^{-1}\|$ is uniformly bounded on S_T^c because $\|\Psi_T{}^{-1}\| = O(T^{-1/2}d_T) < 1$.

The second supreme is bounded by the functional central limit theorem for martingale differences. Combined with the boundedness of $T^{-1/2}(d_T\sqrt{Ah_T})$ (because its squared value is bounded by assumption), we see that the above expression is $O_p(1)$.

Next consider $\mathcal{L}_{1t}$. Because

$$(I+\Psi_T)^{-1} = I - \Psi_T + \Psi_T^2(I+\Psi_T)^{-1}$$

$\mathcal{L}_{1t}$ can be written as

$$\mathcal{L}_{1t} = -\frac{\tau_i}{2}\log|I+\Psi_T| - tr(\Psi_T) + \frac{1}{2}tr\left[\Psi_T\sum_{t=1}^{\tau_i}(\eta_t\eta_t'-I)\right] - \sum_{t=1}^{\tau_i}\eta_t'\Psi_T^2(I_t+\Psi_T)^{-1}\eta_t$$

The last term is non-positive, so it is enough to consider the first two terms on the right. The first term is equal to

$$-\frac{\tau_i}{2}\sum_{i=1}^{n}(\log(1+\lambda_i)-\lambda_i)$$

where again the $\lambda_i's$ are the eigenvalues of Ψ_T. Applying Taylor expansion, it yields

$$\begin{aligned}\frac{\tau_i}{2}\sum_{i=1}^{n}(\frac{1}{2}\lambda_i^2+o(\lambda_i^2)) &\leq \tau_i n \max\lambda_i^2 \\ &\leq \tau_i nC \max\|\Psi_T\|^2 \\ &\leq CAh_T\frac{d_T^2}{T} \text{ for all } k \leq Ah_T\end{aligned}$$

which is bounded by our assumption. We have utilized the relationship between a symmetric matrix and its eigenvalues. Next, consider the second term

$$\begin{aligned}\left\|\Psi_T\sum_{t=1}^{\tau_i}(\eta_t\eta_t'-I)\right\| &\leq \|\Psi_T\|\sqrt{Ah_T}\sup\left\|\frac{1}{\sqrt{Ah_T}}\sum_{t=1}^{\tau_i}(\eta_t\eta_t'-I)\right\| \\ &= C\left(T^{-1/2}d_T\sqrt{Ah_T}\right)O_p(1)\end{aligned}$$

which is bounded in probability. ■

Now we can use these properties to prove Theorem 1. Here, we only consider the case in which $v \leq 0$, i.e. $\tau_i \leq \tau_i^0$. The case for $v > 0$ is similar. The likelihood ratio LR_T is based on the whole sample $[1, T]$. The likelihood ratio can be rewritten as the product of likelihood ratios for three subsamples, $[1, \tau_i]$, $[\tau_i+1, \tau_i^0]$, and $[\tau_i^0+1, T]$. In this way, the likelihood ratio will have $V_t \otimes I$ rather than $Z_t(\tau)$ as regressors. Recall that $\beta = (\theta', (S\delta)')'$. Let $\Psi = \theta + S'S\delta$, which is the combination coefficients of $V_t \otimes I$ for the second regime.

The likelihood ratio LR_T can be rewritten as

$$\begin{aligned}LR_T &= \frac{L(\tau,\, \beta^0+T^{-1/2}\beta,\, \Sigma^0+T^{-1/2}\Sigma)}{L(\tau^0, \beta^0, \Sigma^0)} \\ &= \mathcal{L}(1,\tau;\theta,\Sigma)\times\mathcal{L}(\tau+1,\tau^0;\sqrt{T}S'S\delta_T+\psi,\Sigma)\times\mathcal{L}(\tau^0+1,T;\psi,\Sigma) \quad (1.8.13)\end{aligned}$$

Only the middle term of (last expression) needs some explanation. For $t \in [\tau_i +$

$1, \tau_i^0]$,

$$\begin{aligned}
\varepsilon_t(\tau_i) &= y_t - Z_t(\tau_i)'(\beta_0 + T^{-1/2}\beta) \\
&= \varepsilon_t - (V_t' \otimes I)S'S\delta - T^{-1/2}(V_t' \otimes I)(\theta + S'S\delta) \\
&= \varepsilon_t - T^{-1/2}(V_t' \otimes I)(\sqrt{T}S'S\delta_T + \psi). \qquad (1.8.14)
\end{aligned}$$

By the definition of $\mathcal{L}$, the segment $[\tau_i+1, \tau_i^0]$ involves the parameter value $\sqrt{T}S'S\delta_T+\psi$.

Let $\tau_i(v) = \tau_i^0 + [v_1 v_T^{-2}]$. For some $e_0 > 0$ and $\varepsilon_0 < \tau_0$, define

$$\begin{aligned}
B_{1,T} &= \left\{((\tau_i, \beta, \Sigma)); ||\psi|| \leq \sqrt{T}||S'S\delta_T||,\, T\varepsilon_0 \leq \tau_i \leq \tau_t(v_1)\right\} \qquad (1.8.15) \\
B_{2,T} &= \left\{((\tau_i, \beta, \Sigma)); ||\psi|| \leq \sqrt{T}||S'S\delta_T||,\, 0 \leq \tau_i \leq T\varepsilon_0\right\} \qquad (1.8.16) \\
B_{3,T} &= \left\{((\tau_i, \beta, \Sigma)); ||\psi|| \leq \sqrt{T}||S'S\delta_T||,\, T\varepsilon_0 \leq \tau_i \leq \tau_t(v_1)\right\} \qquad (1.8.17)
\end{aligned}$$

On $B_{1,T}$, both $\mathcal{L}(1, \tau; \theta, \Sigma)$ and $\mathcal{L}(\tau^0 + 1, T; \psi, \Sigma)$ are $O_p(1)$ from Property 1, since both use positive fraction of observations.

Next consider $\mathcal{L}(\tau + 1, \tau^0; \sqrt{T}S'S\delta_T + \psi, \Sigma)$ which involves $\tau_0 - \tau = -vv_T^{-2}$ observations. Since

$$\begin{aligned}
||\sqrt{T}S'S\delta_T + \psi|| &\geq ||\sqrt{T}S'S\delta_T|| - ||\psi|| \\
&\geq \frac{1}{2}||S'S\delta_T||
\end{aligned}$$

we apply property 5 with

$$\begin{aligned}
\theta &= \sqrt{T}S'S\delta_T + \psi \\
d_t &= \frac{1}{2}\sqrt{T}||S'S\delta_T|| \\
h_t &= v_t^{-2}
\end{aligned}$$

$$A = -v_1$$

to conclude that $\mathcal{L}(\tau+1, \tau^0; \sqrt{T}S'S\delta_T + \psi, \Sigma)$ can be arbitrarily small in probability if $-v_1$ is large.

We now assume that

$$||\sqrt{T}S'S\delta_T|| > \log T$$

Then on $B_{2,T}$, $\mathcal{L}(1, \tau; \theta, \Sigma)$ is less than T^B for some $B > 0$ with probability at least $1-\varepsilon$ from property 2 and $\mathcal{L}(\tau+1, \tau^0; \sqrt{T}S'S\delta_T + \psi, \Sigma)$ is $O_p(1)$ from property 1. However, by property 3 ,with $\theta = \sqrt{T}S'S\delta_T + \psi$, $\mathcal{L}(\tau+1, \tau^0; \sqrt{T}S'S\delta_T + \psi, \Sigma)$ is less than T^{-D} for any $D > 0$ with probability at least $1-\varepsilon$ when T is large. Thus the product of these three terms can be no larger than ε with probability at least $1-2\varepsilon$ when T is large.

Next on $B_{3,T}$, Property 2 is applicable to both $\mathcal{L}(1, \tau; \theta, \Sigma)$ and $\mathcal{L}(\tau+1, \tau^0; \sqrt{T}S'S\delta_T + \psi, \Sigma)$ and property 3 is applicable to $\mathcal{L}(\tau+1, \tau^0; \sqrt{T}S'S\delta_T + \psi, \Sigma)$. Thus their product can be arbitrarily small.

Then it is easy to show that

$$Pr\left(\sup_{|v|\geq v_1} \sup_{(\beta,\Sigma)\in S_T} LR_T > \varepsilon\right) < \varepsilon \tag{1.8.18}$$

and

$$Pr\left(\sup_{|v|\leq v_1} \sup_{||\beta||>M \; or \; ||\Sigma||>M} LR_T > \varepsilon\right) < \varepsilon \tag{1.8.19}$$

These two results directly give the consistency, rate of convergence of these estimates in our model.

1.8.2 Proof of Theorem 2

For a given ordered structural breaks of the sample, we have

$$LR_T(\tau_1,\ldots,\tau_n) \quad = \quad T\log|\tilde{\Sigma}| - T\log|\hat{\Sigma}|$$

where $\hat{\Sigma}$ and $\tilde{\Sigma}$ denote the covariance matrix of the errors estimated under the null and alternative hypotheses, respectively. Taking a second Taylor expansion,

$$\begin{aligned} LR_T(\tau_1,\ldots,\tau_n) &= tr\left(T\Sigma_0^{-1}(\tilde{\Sigma}-\hat{\Sigma})\right) + \frac{T}{2}tr\left(\left[(\Sigma^0)^{-1}(\hat{\Sigma}-\Sigma^0)\right]^2\right) \\ &\quad -\frac{T}{2}tr\left(\left[(\Sigma^0)^{-1}(\tilde{\Sigma}-\Sigma^0)\right]^2\right) + o_p(T^{-1}) \end{aligned}$$

First we consider the third term log

$$\begin{aligned} \left[(\Sigma^0)^{-1}(\tilde{\Sigma}-\Sigma^0)\right]^2 &= \left[(\Sigma^0)^{-1}(T^{-1}\sum_{t=1}^{T}\left(y_t - Z_t'\hat{\beta}\right)\left(y_t - Z_t'\hat{\beta}\right)' - \Sigma^0)\right]^2 \\ &= \left[(\Sigma^0)^{-1}(T^{-1}\sum_{t=1}^{T}\left(y_t - Z_t'\hat{\beta}\right)\left(y_t - Z_t'\hat{\beta}\right)' - \Sigma^0)\right]^2 \\ &= \left[(\Sigma^0)^{-1}(T^{-1}\sum_{t=1}^{T}\left(y_t - Z_t'\hat{\beta}\right)\left(y_t - Z_t'\hat{\beta}\right)' - \Sigma^0)\right]^2 + O_p(T^{-3/2}) \end{aligned}$$

where the last equality follows since $\beta^0 - \tilde{\beta} = O_p(T^{-1/2})$.

Similarly, we can show that

$$\left[(\Sigma^0)^{-1}(\tilde{\Sigma}-\Sigma^0)\right]^2 = \left[(\Sigma^0)^{-1}(T^{-1}\sum_{t=1}^{T}\left(y_t - Z_t'\hat{\beta}\right)\left(y_t - Z_t'\hat{\beta}\right)' - \Sigma^0)\right]^2$$

Hence the likelihood ratio can be simplified as

$$\begin{aligned} LR_T(\tau_1,\ldots,\tau_n) &= tr\left(T\Sigma_0^{-1}(\tilde{\Sigma}-\hat{\Sigma})\right) + \frac{T}{2}tr\left(\left[(\Sigma^0)^{-1}(\hat{\Sigma}-\Sigma^0)\right]^2\right) \\ &\quad -\frac{T}{2}tr\left(\left[(\Sigma^0)^{-1}(\tilde{\Sigma}-\Sigma^0)\right]^2\right) + o_p(T^{-1}) \end{aligned}$$

The reminder proof proceeds similar as Qu and Perron (2007).

1.9 Chapter 1: Tables and Figures

Table 1.1: Empirical Results: Output Growth in European and United States

Country	Sample	$\sup W$	$ExpW$	Break Dates
A: Breaks in Univariate				
France	64:Q2-89:Q2	23.68	9.15	74:Q2
		(0.00)	(0.00)	(72:Q4, 75:Q4)
Germany	51:Q4-89:Q2	21.68	8.28	61:Q2
		(0.00)	(0.00)	(59:Q1, 63:Q3)
Italy	53:Q2-82:Q4	10.30	2.83	74:Q3
		(0.03)	(0.03)	(70:Q2, 78:Q4)
U.S.	64:Q2-89:Q2	1.42	0.25	68:Q4
		(0.91)	(0.71)	(73:Q3, 76:Q3)
B: Common Breaks in VAR Systems				
F, G	64:Q4-89:Q2	26.00	10.14	75:Q1
		(0.00)	(0.00)	(73:Q3, 76:Q3)
F, I	64:Q4-82:Q4	17.97	6.24	73:Q4
		(0.00)	(0.00)	(72:Q1, 75:Q3)
F, U	64:Q4-89:Q2	7.43	1.28	70:Q3
		(0.07)	(0.13)	(63:Q1, 80:Q4)
G, I	53:Q4-82:Q4	14.98	5.32	74:Q1
		(0.02)	(0.01)	(71:Q1, 77:Q2)
G, U	51:Q4-89:Q2	3.21	1.09	65:Q1
		(0.35)	(0.43)	(55:Q1, 78:Q1)
I, U	53:Q4-82:Q4	4.60	0.54	69:Q3
		(0.18)	(0.51)	(63:Q1, 75:Q2)
F, G, I	64:Q4-82:Q4	19.43	6.98	73:Q4
		(0.01)	(0.00)	(72:Q2, 75:Q2)
F, G, I, U	64:Q4-82:Q4	11.47	1.49	72:Q1
		(0.06)	(0.08)	(69:Q3, 76:Q1)

Note: We employ the statistics-$\sup W$ and $ExpW$- in Bai and Perron (1998, 2003) and Bai et al. (1998). There limiting distributions and critical values are shown in these papers. The $p-$values are given in parentheses. The sample period denotes the period over which the testing were run; as convention in the literature, the trimming value ϵ is set to 0.15.

Table 1.2: Empirical Results: Output Growth in European and United States

Country	Sample	$\sup LR$	Break Dates
C: Ordered Breaks in VAR Systems			
F, G	64:Q4-89:Q2	16.70	Eq 1: 75:Q2
		(0.01)	Eq 2: 64:Q3
F, I	64:Q4-82:Q4	28.13	Eq 1: 73:Q2
		(0.00)	Eq 2: 74:Q4
F, U	64:Q4-89:Q2	13.52	Eq 1: 75:Q1
		(0.05)	Eq 2: 70:Q3
G, I	53:Q4-82:Q4	18.28	Eq 1: 63:Q3
		(0.00)	Eq 2: 71:Q2
G, U	51:Q4-89:Q2	9.36	Eq 1: 65:Q4
		(0.09)	Eq 2: 66:Q4
I, U	53:Q4-82:Q4	13.15	Eq 1: 74:Q1
		(0.05)	Eq 2: 69:Q3
F, G, I	64:Q4-82:Q4	16.32	Eq 1: 75:Q4, Eq 2: 64:Q3
		(0.01)	Eq 3: 72:Q3
F, G, I, U	64:Q4-82:Q4	12.99	Eq 1: 75:Q4, Eq 2: 64:Q3
		(0.06)	Eq 3: 72:Q3, Eq 4: 70:Q1

Note: The $p-$values are given in parentheses, and are computed using the asymptotic distributions of the test statistic $\sup LR$ in Section 4. The sample period denotes the period over which the testing were run; as convention in the literature, the trimming value ϵ is set to 0.15.

Figure 1.1: Histograms of the Estimated Break Dates for Model (5.1)

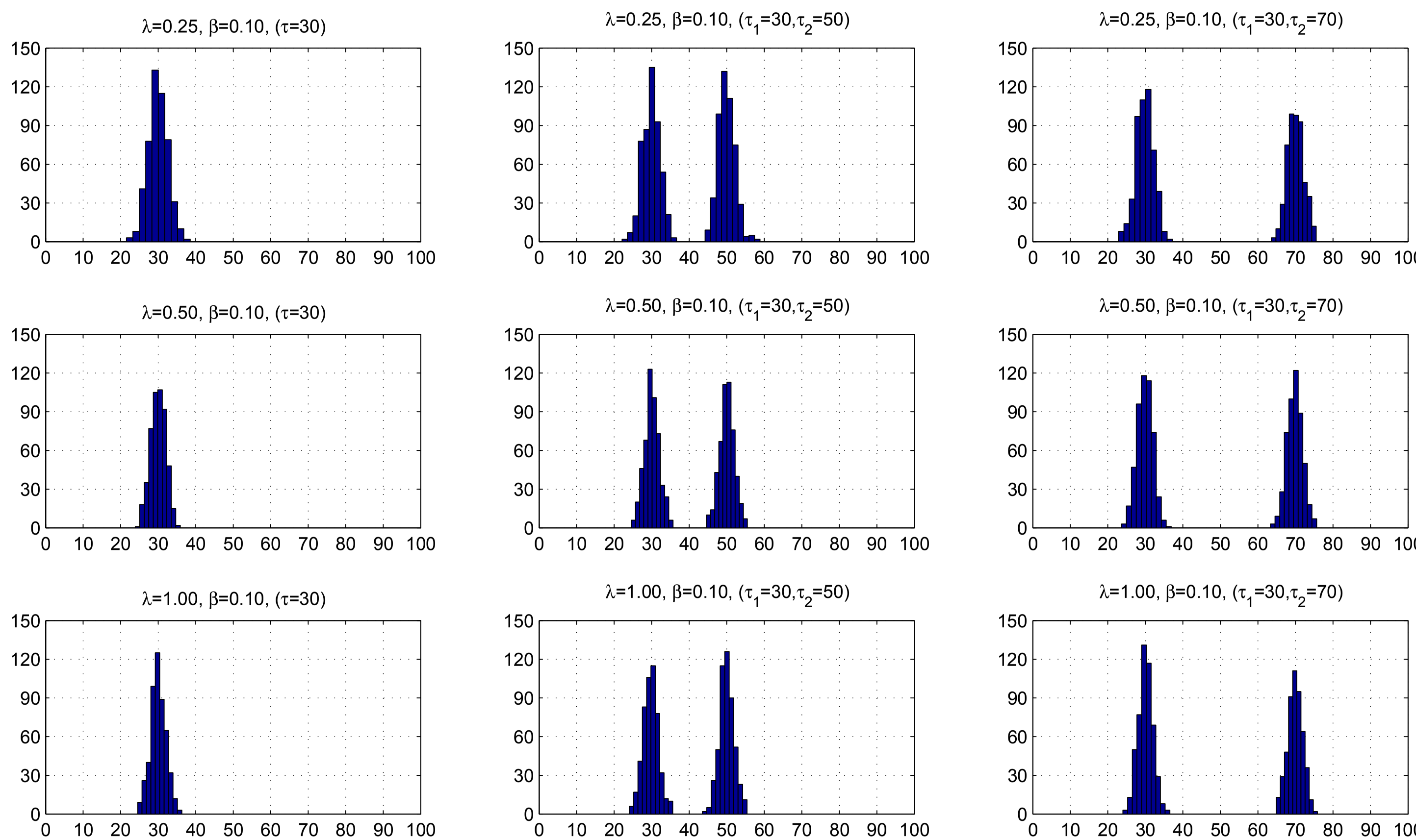


Figure 1.2: Histograms of the Estimated Break Dates for Model (5.1)

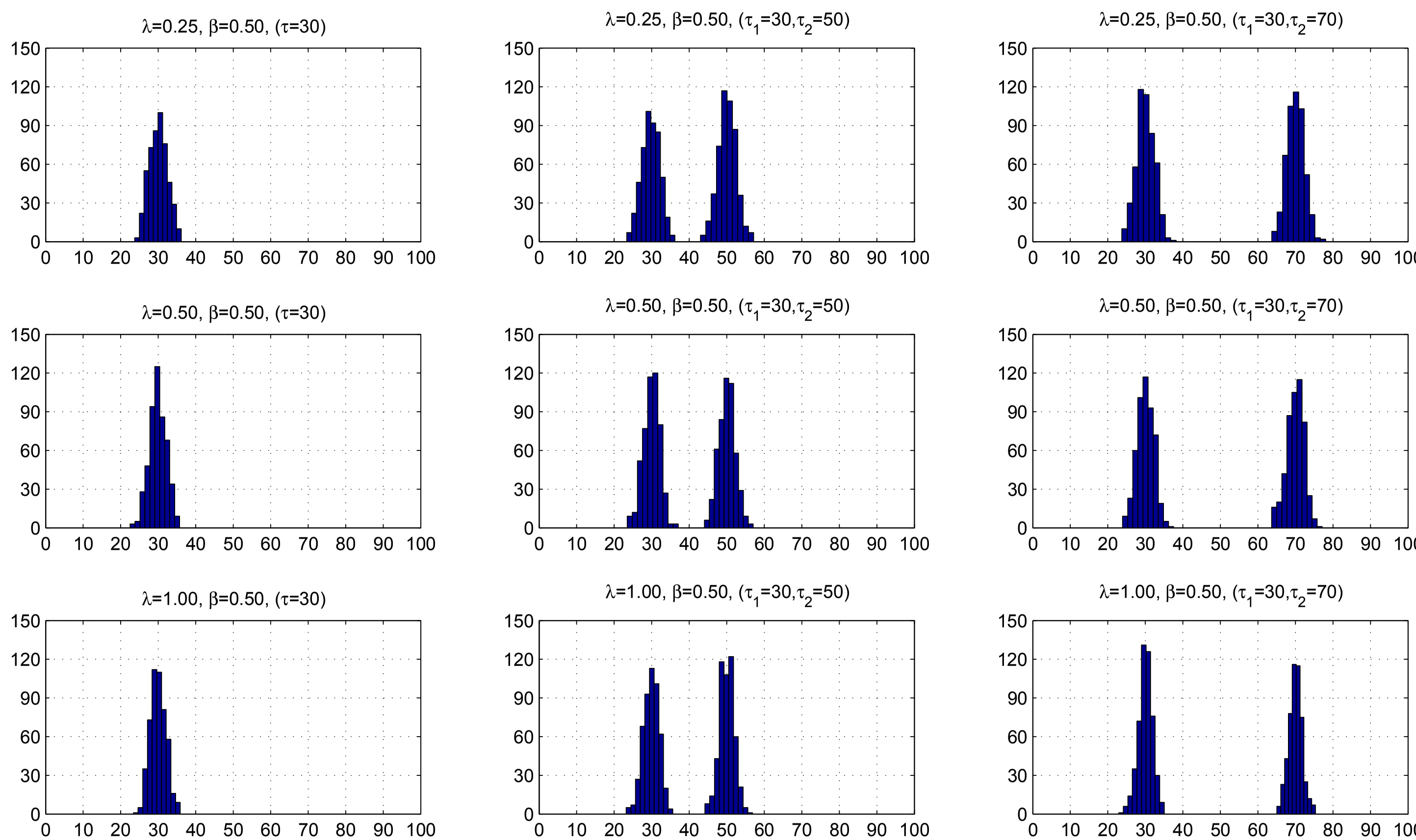
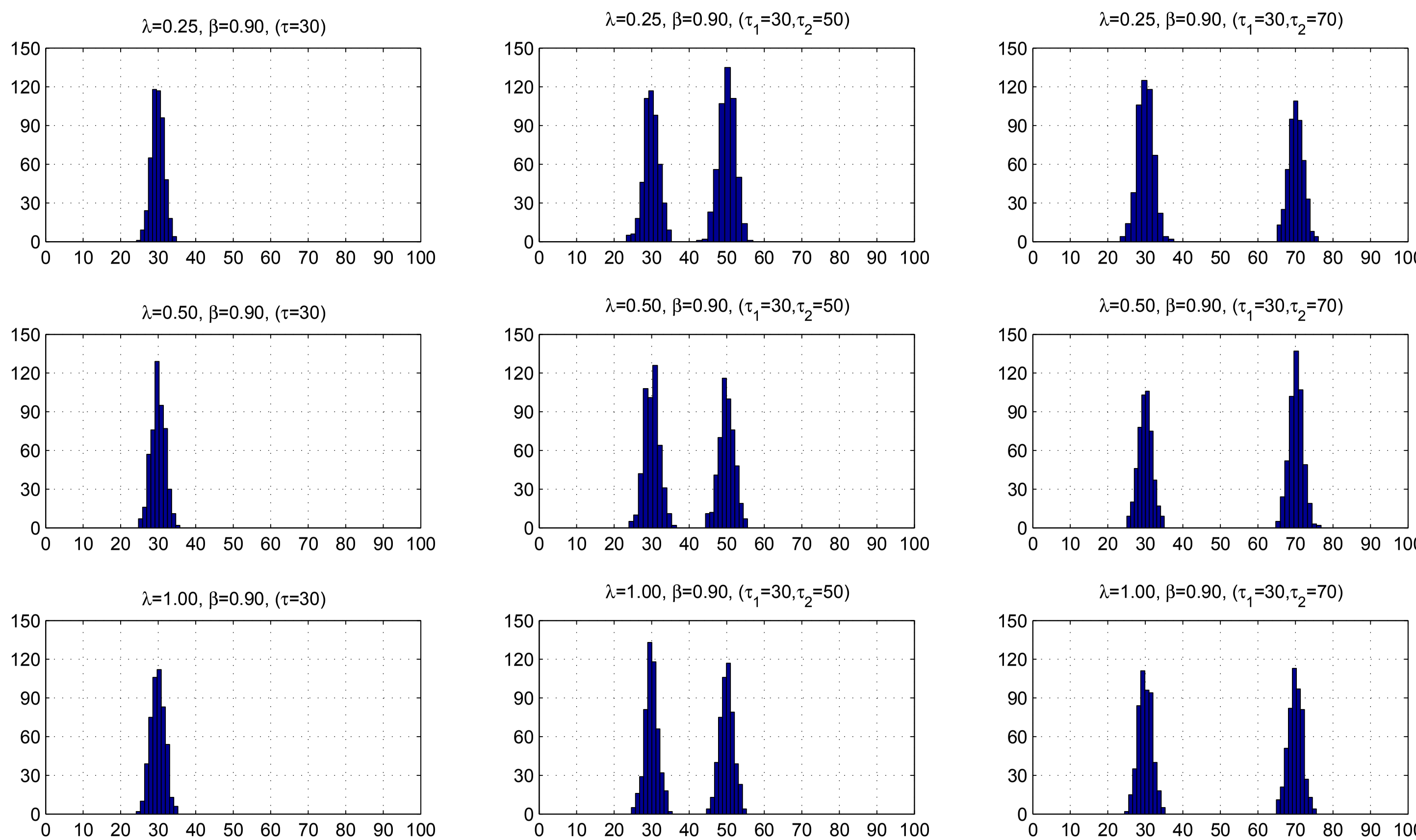


Figure 1.3: Histograms of the Estimated Break Dates for Model (5.1)



Chapter 2

Nonlinear Solutions to Dynamic Stochastic General Equilibrium Models and Parameter Stability

2.1 Introduction

This paper studies the parameter stability of dynamic stochastic general equilibrium (DSGE) models. This problem is important because DSGE models are now at the center of modern macroeconomics. Among the most powerful tools, such as VAR and structural VAR, DSGE models have been developed to match economic theory with real economic data, to help design and evaluate economic policy, and more recently to perform forecasting. They promise to be a laboratory not just for academia, but also for an increasing number of policy making institutions.[1] Furthermore, the parameters in DSGE models are defined to describe agents' preferences and technologies of the economy. As a response to Lucas (1976) critique, these parameters

[1]As we have witnessed, an increasing number of policy making institutions, such as the Federal Reserve Board, the European Central Bank and the Bank of England, the Bank of Canada, the Bank of Sweden, the Bank of Spain, and the Bank of Japan among others already employ DSGE models for policy analysis and forecasting.

have a solid micro-foundation from the perspective of economic theory, and ought to remain invariant to policy interventions. Given these reasons, one can naturally raise the question of whether these DSGE models live up to their promise of being truly "structural". In other words, how stable are these so-called deep "structural" parameters of DSGE models over time?

In the existing literature, a body of evidence has been brought to document the parameter instability of estimated DSGE models, which all suggest the evolving economic environment of the U.S. has changed in fundamental ways over the last few decades.[2] For instance, Ireland (2001) estimates a DSGE model with sticky prices by maximum likelihood estimation and uses standard stability tests for a single known break date. These formal hypothesis test results show instability in the estimated parameters, particularly in estimates of the representative household's discount factor. Boivin and Giannoni (2006) also investigate the structural parameters of a New Keynesian model using minimum-distance estimation, and interpret changes in these parameter estimates from two sub-periods as evidence of the effectiveness of monetary policy in the post-1980 period. Fernandez-Villaverde and Rubio-Ramírez (2008) contribute the literature by estimating a medium-scale DSGE model directly allowing parameter drifting. They document that there is strong evidence that parameters change within their sample as well. More recently, Inoue and Rossi (2011) use a novel approach to directly investigate the source of instability and seek to find which structural parameters are truly "structural".

However, there are three limitations in the existing empirical findings. First of all, these studies aforementioned, except Fernandez-Villaverde and Rubio-Ramírez (2008), rely on the linear solution methods for the DSGE models since the linearization or log-linearization approach is appealing from both the econometric perspec-

[2]In the discussion of Great Moderation, a number of empirical studies use VAR models with time-varying parameters. Example includes Uhlig (1997), Stock and Watson (2002), Primiceri (2005), Cogley and Sargent (2005), and Sims and Zha (2006).

tive and the computational perspective. Over the last two decades, a number of nonlinear solution methods for DSGE models have been proposed as alternatives to more traditional linear solution methods. These new methods have promised superior performance on the long experience of mathematics and in a growing economic literature that emphasizes the role of nonlinearities in dynamic equilibrium economies. Fernandez-Villaverde and Rubio-Ramírez (2005) and Fernandez-Villaverde et al. (2006) point out that estimating DSGE models based on their linear solutions will generally lead to biases, which will not consequently generate correct inference—stability tests in this case. Second, the common practice in the existing literature is to divide the sample into two subsamples and to construct classical structural break tests typically attributed to Chow (1960) and the recent treatment of Andrews and Fair (1988). The limitation of the Chow test, however, is the break date must be known as a priori. In the literature, one has to either pick an arbitrary candidate break date—usually around 1979, or pick a break date based on some known feature of data—for instance, the sharp decline in the volatility of output documented by Stock and Watson (2002) among others. The Chow test may be uninformative in the first case, since the true break date might be missed. In the second case, the Chow test might be misleading, as the candidate break date is endogenous—it is correlated with the data. Therefore, different studies can easily reach quite distinct conclusions, since the results can be highly sensitive to these arbitrary choices. Third, most of the literature has focused on a single structural break in these parameters, whereas allowing multiple structural breaks might be more suitable given the substantial changes of economic structure, technological innovations, and historical events over the last few decades.

This paper addresses these limitations and completes the literature in the following ways. First, this paper solves and estimates a small-scale DSGE model—a representative New Keynesian model in this case—using two main numerical methods: perturbation and projection methods. Within perturbation, I consider first, second, and third order approximations. Note the first order perturbation is equivalent to the traditional log-linearized solution since variables in the model are transformed in logarithm. Within projection, I use second and third order Chebyshev polynomials approximations. These parameter estimates and stability tests from nonlinear solutions offer a comparison to results from log-linearized solution, and can be used to investigate the effect of nonlinearities on parameter stability. Second, this paper explicitly treats the break date(s) as unknown. Although the statistics of testing unknown structural break(s) have been set up by Quandt (1960) and their asymptotic properties have been derived in Andrews (1993) and Andrews and Ploberger (1994) among others, they have been surprisingly rarely incorporated into empirical studies under a more structural framework, like DSGE models.[3] Therefore, this paper tries to fill part of this gap by treating the break date(s) as unknown. Third, this paper considers the possibility of multiple structural breaks in DSGE models. Here I propose a sequential procedure for multiple structural breaks in nonlinear models. The procedure was originally developed by Bai (1997) for testing multiple breaks in linear regressions. The advantage of this sequential procedure is to avoid the computation complexity when estimating multiple structural breaks simultaneously, and also to circumvent the challenge of unavailable asymptotic properties on statistics for multiple structural breaks in nonlinear models.

The main finding of this paper is that there is strong evidence of parameter instabilities of DSGE models – a representative New Keynesian model using U.S. data in this case. In particular, the results first show that the presence of parameter instability does not result from linearization or log-linearization to the DSGE model. Neither linear nor nonlinear solutions could support null hypothesis of parameter stability, which indicates some more fundamental changes of economy structure. Also, this paper documents the two *common* structural breaks among these parameters of

[3]One exception is Estrella and Fuhrer (2003), in which they examine the stability of a forward-looking monetary policy model.

this DSGE model. These breaks occur in the early 1970s, and the middle 1990s, corresponding to when fundamental changes in U.S. are widely believed to have occurred. Finally, the empirical results in this paper suggest that parameter instabilities are not only due to changes in monetary policy reactions, but also to changes in agents' preferences and technology as well as changes in shocks volatility. It is broadly consistent with findings of Ireland (2001), Fernandez-Villaverde and Rubio-Ramírez (2008), and Inoue and Rossi (2011).

These results shown in this paper are important for the following reasons. First, the evidence of parameter instabilities convey the message, in the sense of Lucas (1976) critique, that structural parameters in DSGE models ought to be policy invariant, but not necessarily time invariant. And they also highlight the importance of applying stability tests to so-called "structural" macroeconomic models, like DSGE models. Second, the identification of timing of structural breaks is very informative regarding understanding the instability of macroeconomic fluctuation. More importantly, one should incorporate the information of break dates when considering to use DSGE models to perform policy analysis and forecasting. Third, the comparison between linear solution to non-linear solution emphasizes that exploiting nonlinearities allows for more accurate estimation and inference. Nonlinear models can offer better explanation of economic dynamics, for instance, zero lower bound constraint confronted by the central banks of U.S. and Europe. Finally, this paper makes a methodological contribution by introducing sequential procedure for multiple breaks into DSGE models. Based on the Monte Carlo simulation, such a procedure can be directly applied to other non-linear models, and it will deliver consistent results.

The rest of the paper is organized as follows. Section 2 describes the representative DSGE model and its equilibrium conditions. Section 3 outlines the solutions to the DSGE model. Section 4 describes the estimations of these linear and nonlinear solutions, and proposes two structural break tests and a sequential procedure for multiple breaks. Section 5 contains data, computational implication, and results. Finally, section 6 concludes.

2.2 The DSGE Model

The DSGE model in this paper consists of a representative household, a representative final goods-producing firm, a continuum of intermediate goods-producing firms and a central bank. This model is a small-scale version of the New Keynesian model, which has been developed by Ireland (2001, 2004, 2007) and Woodford (2003) among others for the analysis of monetary and fiscal policy. More elaborate versions can be found in Christiano et al. (2005) and Smets and Wouters (2003, 2007).

2.2.1 The Household

At the beginning of each period $t = 0, 1, 2, \ldots$, the representative household enters with M_{t-1} units of money, B_{t-1} units of bonds and K_t units of capital. Meanwhile, the household receives a lump-sum nominal transfer T_t from the central bank. In addition, the household's bonds mature, providing B_{t-1} additional unites of money. During period t, the household supplies $H_t(i)$ units of labor and $K_t(i)$ to the various intermediate goods-producing firms, taking the nominal factor prices W_t and Q_t as given, where $i \in [0,\ 1]$ indicates each intermediate goods-producing firm. Here the model denotes $H_t = \int_0^1 H_t(i)di$ as the total amount of labor supplied and $K_t = \int_0^1 K_t(i)di$ as the total amount of capital supplied. Thus, the household receives total nominal factor payments $W_tH_t + Q_tK_t$. At last, the household receives nominal profits $D_t = \int_0^1 D_t(i)di$ from the intermediate goods-producing firms.

The household expenditures are characterized as the following. First, the household purchases the final goods from the representative final goods-producing firm at the nominal price P_t. This purchase is divided into C_t units of consumption and I_t

units of investment. In order to transform final goods to productive investment, the household must pay an adjustment cost. It is measured in terms of the final goods and given by

$$\frac{\phi_k}{2}\left(\frac{K_{t+1}}{K_t}-1\right)^2 K_t$$

where $\phi_k > 0$ measures the magnitude of the capital adjustment cost, and the capital accumulation follows

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

where the depreciate rate is $\delta \in (0,\, 1)$. Also, The household uses some of its funds to purchase B_t new bonds at price of $1/R_t$, where R_t denotes the gross nominal interest rate between t and $t+1$. The household then carries B_t units of bonds, and K_{t+1} units of capital, and M_t units of remaining money into period $t+1$. Therefore the budget constraint for the household is

$$C_t+I_t+\frac{\phi_k}{2}\left(\frac{K_{t+1}}{K_t}-1\right)^2 K_t+\frac{B_t/R_t+M_t}{P_t} \leq \frac{M_{t-1}+B_{t-1}+T_t+W_tH_t+Q_tK_t+D_t}{P_t}$$

The household aims to maximize its expected utility, given by

$$\max E\sum_{t=0}^{\infty}\beta^t u\left(C_t,\, \frac{M_t}{P_t},\, H_t\right)$$

where the discount factor is $\beta \in (0,\, 1)$. Assume that the instantaneous utility function takes the form as

$$u\left(C_t,\, \frac{M_t}{P_t},\, H_t\right) = a_t\frac{C_t^{1-\gamma}-1}{1-\gamma}+\chi_m \log\left(\frac{M_t}{P_t}\right)-\chi_h\frac{H_t^{1+\nu}-1}{1+\nu}$$

where γ is the inverse of the elasticity of substitution between current and future consumption, and ν is the inverse of the Frisch labor supply elasticity. χ_m and χ_h are weights associated with utility from real money balances and disutility from worked hours. The preference shock a_t follows the stationary autoregressive process

$$\ln(a_t) = \rho_a \ln(a_{t-1}) + \varepsilon_{at}$$

with $\rho_a \in (0,\, 1)$, where the serially uncorrected innovation ε_{at} has the normal distribution with zero mean and standard deviation σ_a.

2.2.2 The Final Goods-Producing Firm

During each period t, the representative final goods-producing firm purchase $Y_t(i)$ units of each intermediate good $i \in [0,1]$ at the nominal price $P_t(i)$ to produce Y_t units of the final goods according to constant-return-to-scale technology described by

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

where $\theta > 1$ measures the constant elasticity of demand for each intermediate good. Thus, the final goods-producing firm seeks to choose Y_t and $Y_t(i)$ for all $i \in [0,1]$ to maximize its profits given the nominal price P_t of final goods and the nominal price $P_t(i)$ of intermediate goods. Perfect competition in the final goods market drives the firm's profits in equilibrium to zero. This zero profit condition leads to determine the equilibrium price P_t as

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

Also, the demand for each intermediate good i turns out to be

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

2.2.3 The Intermediate Goods-Producing Firms

The representative intermediate goods-producing firm hires $H_t(i)$ units of labor and rents $K_t(i)$ units of capital from the household to produce $Y_t(i)$ units of intermediate good i during period t. The constant returns to scale production technology is described by

$$Y_t(i) = K_t(i)^{\alpha}[z_t H_t(i)]^{1-\alpha}$$

where $\alpha \in (0, 1)$ is capital's share in production function. Here the aggregate technology shock z_t follows a first order autoregressive process

$$\ln(z_t) = (1 - \rho_z)\ln(z) + \rho_z \ln(z_{t-1}) + \varepsilon_{zt}$$

with $z > 0$ and $\rho_z \in (0, 1)$, where the serially uncorrelated innovation ε_{zt} has the normal distribution with mean zero and standard deviation σ_z.

Since the representative intermediate goods-producing firm can sell its output in a monopolistically competitive market; during period t, the firm sets the nominal price $P_t(i)$ for its output, subject to the requirement that it satisfy the final goods-producing firm's demand at that price. In addition, following Rotemberg (1982), the intermediate goods-producing firm faces a quadratic cost of adjusting its nominal price between periods, measured in terms of the final goods and given by

$$\frac{\phi_p}{2}\left[\frac{P_t(i)}{\pi P_{t-1}(i)} - 1\right]^2 Y_t$$

where $\phi_p \geq 0$ gives the magnitude of the price adjustment cost, and $\pi \geq 1$ measures the gross steady-state rate of inflation.

The intermediate goods-producing firm must choose $H_t(i)$, $K_t(i)$, $Y_t(i)$, and $P_t(i)$

to maximize its total expected real market value, given by

$$\max E \sum_{t=0}^{\infty} \beta^t \Lambda_t \left[\frac{D_t(i)}{P_t} \right]$$

where $\beta^t \Lambda_t$ measures the marginal utility value to the representative household of an additional unit of real profits received in the form of dividends during the period t and where

$$\frac{D_t(i)}{P_t} = \frac{P_t(i)Y_t(i)}{P_t} - \frac{W_t H_t(i) + Q_t K_t(i)}{P_t} - \frac{\phi_p}{2} \left[\frac{P_t(i)}{\pi P_{t-1}(i)} - 1 \right]^2 Y_t$$

measures the firm's real profits during the same period t.

2.2.4 The Central Bank

The central bank conducts monetary policy by adjusting short term nominal R_t according to the following conventional rule:

$$\ln \left(\frac{R_t}{R} \right) = \rho_R \ln \left(\frac{R_{t-1}}{R} \right) + \rho_\pi \ln \left(\frac{\pi_t}{\pi} \right) + \rho_y \ln \left(\frac{Y_t}{Y} \right) + \varepsilon_{Rt}$$

where R, π, and Y denote the target values of the respective variables. R_t follows a version of Taylor (1993) rule that depends on the lagged interest rate, the deviation of inflation with respect to its target, and the output gap. The central bank can choose the level of one of these target variables, as well as the parameters ρ_R, ρ_y, and ρ_π. The term ε_{Rt} denotes monetary policy shock with mean zero and standard deviation σ_R.

2.2.5 The Equilibrium Conditions

In a symmetric equilibrium, all intermediate goods-producing firms make identical decisions, so that $Y_t(i) = Y_t$, $H_t(i) = H_t$, $D_t(i) = D_t$, $K_t(i) = K_t$, and $P_t(i) = P_t$ for all $i \in [0, 1]$. In addition, the market clearing conditions $M_t = M_{t-1} + T_t$ and $B_t = B_{t-1} = 0$ must hold. Letting $c_t = C_t$, $k_t = K_t$, $h_t = H_t$, $m_t = M_t/P_t$, $w_t = W_t/P_t$, and $q_t = Q_t/P_t$ and $\pi_t = P_t/P_{t-1}$, the equilibrium conditions can be summarized as the following.[4]

$$\chi_m m_t^{-1} - a_t c_t^{-\gamma}(1 - R_t^{-1}) = 0 \quad (2.2.1)$$

$$\chi_h h_t^{\nu} - a_t c_t^{-\gamma} w_t = 0 \quad (2.2.2)$$

$$\frac{a_t c_t^{-\gamma}}{R_t} - \beta E_t \frac{a_{t+1} c_{t+1}^{-\gamma}}{\pi_{t+1}} = 0 \quad (2.2.3)$$

$$\begin{aligned} -a_t c_t^{-\gamma}\left[1 + \phi_k\left(\frac{k_{t+1}}{k_t} - 1\right)\right] + \beta E_t \Big\{ a_{t+1} c_{t+1}^{-\gamma} \Big[(q_{t+1} + 1 - \delta) \\ -\frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t} - 1\right)^2 + \phi_k\left(\frac{k_{t+2}}{k_{t+1}} - 1\right)\left(\frac{k_{t+2}}{k_{t+1}}\right)\Big]\Big\} &= 0 \quad (2.2.4) \end{aligned}$$

$$y_t - k_t^{\alpha}[z_t h_t]^{1-\alpha} = 0 \quad (2.2.5)$$

$$y_t - w_t h_t - q_t k_t - \frac{\phi_p}{2}\left(\frac{\pi_t}{\pi} - 1\right)^2 y_t - d_t = 0 \quad (2.2.6)$$

$$a_t w_t h_t - (1 - \alpha) q_t k_t = 0 \quad (2.2.7)$$

$$\begin{aligned} a_t c_t^{-\gamma}\left[1 - \theta + \theta \frac{w_t h_t}{(1-\alpha) y_t} - \phi_p\left(\frac{\pi_t}{\pi} - 1\right)\frac{\pi_t}{\pi}\right] \\ + \beta \phi_p E_t \left\{ a_{t+1} c_{t+1}^{-\gamma}\left(\frac{\pi_{t+1}}{\pi} - 1\right)\frac{\pi_{t+1}}{\pi}\frac{y_{t+1}}{y_t}\right\} &= 0 \quad (2.2.8) \end{aligned}$$

$$c_t + i_t + \frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t} - 1\right)^2 k_t + \frac{\phi_p}{2}\left(\frac{\pi_t}{\pi} - 1\right)^2 y_t - y_t = 0 \quad (2.2.9)$$

$$k_{t+1} - (1 - \delta) k_t - i_t = 0 \quad (2.2.10)$$

$$\ln\left(\frac{R_t}{R}\right) - \left[\rho_R \ln\left(\frac{R_{t-1}}{R}\right) + \rho_\pi \ln\left(\frac{\pi_t}{\pi}\right) + \rho_y \ln\left(\frac{y_t}{y}\right) + \varepsilon_{Rt}\right] = 0 \quad (2.2.11)$$

$$\ln(a_t) - \rho_a \ln(a_{t-1}) - \varepsilon_{at} = 0 \quad (2.2.12)$$

[4]In appendix A, I provide a complete derivation of these equilibrium conditions and steady states.

$$\ln(z_t) - [(1-\rho_z)\ln(z) + \rho_z \ln(z_{t-1}) + \varepsilon_{zt}] \;=\; 0 \quad (2.2.13)$$

It is worth noting that the first four equations above describe the representative household's optimization decisions: equation (1) gives the household's demand for real balance; equation (2) equates the marginal rate of substitution between labor and consumption to the real wage; equation (3) describes the household's indifference between consumption and bond holdings; equation (4) states that, in equilibrium, the marginal utility cost of one unit of additional investment at time t equals the discounted expected marginal utility value of its return in period $t+1$. Moreover, equations (5) through (10) come from the production side of the DSGE model: (5) gives the aggregate production function; (6) characterizes the intermediate firm's budget constraint; (7) computes the marginal products of labor and capital to their respective factor prices; and (8) describes the price-setting behavior of firms; equation (9) denotes the aggregate resource constraint in the economy; equation (10) defines investment. Finally, equation (11) describes the monetary policy rule, and equation (12) and (13) characterize the evolution of the exogenous state variables.

2.3 Solutions to the DSGE Model

The set of equilibrium conditions (1) – (13) derived in the previous section forms a nonlinear rational expectation system, which can be written as

$$\mathbf{E_t f}\left(\mathbf{y_{t+1}},\, \mathbf{y_t},\, \mathbf{x_{t+1}},\, \mathbf{x_t},\, \varepsilon_{\mathbf{t+1}},\, \Theta\right) = \mathbf{0} \tag{2.3.1}$$

where E_t denotes the conditional expectation given all the information available at time t;

$$\mathbf{x_t} = (k_t,\, R_{t-1},\, a_t,\, z_t,\, \varepsilon_{R,t})$$

includes all the state (predetermined) variables;

$$\mathbf{y_t} = (c_t,\, y_t,\, k_{t+1},\, w_t,\, h_t,\, q_t,\, d_t,\, \pi_t,\, m_t)$$

consists of all the control (non-predetermined) variables;[5]

$$\varepsilon_{\mathbf{t+1}} = (\varepsilon_{a,t+1},\, \varepsilon_{z,t+1},\, \varepsilon_{R,t+1})$$

collects the exogenous innovations; and finally, $\mathbf{\Theta}$ includes all the structural parameters in the model.

$$\mathbf{\Theta} = (\beta,\, \gamma,\, \nu,\, \chi_m,\, \chi_h, \delta,\, \theta, \alpha,\, \phi_k,\, \phi_p,\, \pi,\, z,\, \rho_a,\, \sigma_a,\, \rho_z,\, \sigma_z,\, \rho_R,\, \sigma_R,\, \rho_y,\, \rho_\pi)$$

This nonlinear rational expectation system has to be solved before the DSGE model can be estimated. Like most DSGE models, this model does not have a "paper and pencil" solution. Hence, numerical methods are employed to solve the equilibrium dynamics of the model. Two numerical solution methods are considered here: perturbation methods, which find solutions locally using Taylor expansions of equilibrium conditions; and projection methods, which approximate solutions on a per-specified domain using function basis. As shown in Aruoba et al. (2006) among others, these two methods have their relative advantages and drawbacks. More specifically, perturbation methods are easily put in practice for large-scale models, but the range of their accuracy is uncertain; on the other hand, projection methods are accurate and fast when applied to models with few state variables, however, their computation costs often increase rapidly when the number of state variables increase. The rest of this section briefly describes how I apply each of these solution methods to the model.

[5] Here I combine equation (9) and (10), then replace i_t by k_{t+1}.

2.3.1 The Perturbation Method

In principle, perturbation methods approximate the equilibrium conditions around the non-stochastic steady states using Taylor expansion. Specifically, adding a perturbation parameter σ into Equation (14) yields

$$\mathbf{E_t f}\left(\mathbf{y_{t+1}},\, \mathbf{y_t},\, \mathbf{x_{t+1}},\, \mathbf{x_t},\, \sigma\varepsilon_{\mathbf{t+1}},\, \Theta\right)= \mathbf{0} \tag{2.3.2}$$

Here the known parameter $\sigma \geq 0$ determines the distance from the non-stochastic steady states. When σ is equal to zero, the model corresponds to the non-stochastic steady states. As observed by Schmitt-Grohé and Uribe (2004), the exact solution to this system is given by

$$\mathbf{y_t}= \mathbf{g}(\mathbf{x_t},\, \sigma) \tag{2.3.3}$$

$$\mathbf{x_{t+1}}=\mathbf{h}(\mathbf{x_t},\, \sigma) + \sigma\eta\varepsilon_{\mathbf{t+1}} \tag{2.3.4}$$

where nonlinear policy function $\mathbf{g}(\cdot)$ maps $\mathbb{R}^5 \times \mathbb{R}^+$ into $\mathbb{R}^9$, and nonlinear function $\mathbf{h}(\cdot)$ describes transitions of the 5 state variables in the model. As convention, I am interested in the percentage deviation of a variable z_t from its steady state z; thus all the variables in this model are taken natural logarithm. Let $\hat{z}_t = \ln(z_t/z)$, the first order perturbation solution is equivalent to the log-linearized solution. A number of solution methods have been proposed in the literature to obtain the first order approximation to $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$, including Blanchard and Kahn (1980), Uhlig (1999), Klein (2000) and Sims (2002). Here, I use the algorithm provided by Klein (2000). The solution is taken in the following form

$$\mathbf{\hat{y}_t}= \mathbf{g_x}\mathbf{\hat{x}_t}$$

$$\mathbf{\hat{x}}_{t+1}= \mathbf{h_x}\mathbf{\hat{x}_t} + \sigma\eta\varepsilon_{\mathbf{t+1}}$$

where $\mathbf{g_x}$ and $\mathbf{h_x}$ are the first derivatives of $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$ with respect to $\mathbf{x_t}$ at the steady state. All the elements in matrices $\mathbf{g_x}$ and $\mathbf{h_x}$ are functions of structural parameters $\mathbf{\Theta}$.

Since one of the main goals in this paper is to explore nonlinearities embedding in DSGE models, higher order approximations are also considered here. Indeed, it is very straightforward to extend Taylor expansion to their higher order approximations. Several algorithms for computing such solutions have been developed by Judd (1998), Schmitt-Grohé and Uribe (2004), Kim et al. (2005), and Andreasen (2011). For the present analysis, I use the second-order approximation method derived by Schmitt-Grohé and Uribe (2004), and rely on Andreasen (2011) to find the third-order approximation solution. Following the representation as Gomme and Klein (2010), the resulting second-order approximate solution takes the form as follows:

$$\begin{aligned}\mathbf{\hat{y}_t} &\approx \frac{\mathbf{1}}{\mathbf{2}}\mathbf{g}_{\sigma\sigma}\sigma^{\mathbf{2}} + \mathbf{g_x}\mathbf{\hat{x}_t} + \frac{\mathbf{1}}{\mathbf{2}}\left(\mathbf{I_9}\otimes\mathbf{\hat{x}_t}\right)'\mathbf{g_{xx}}\mathbf{\hat{x}_t} \\ \mathbf{\hat{x}_{t+1}} &\approx \frac{\mathbf{1}}{\mathbf{2}}\mathbf{h}_{\sigma\sigma}\sigma^{\mathbf{2}} + \mathbf{h_x}\mathbf{\hat{x}_t} + \frac{\mathbf{1}}{\mathbf{2}}\left(\mathbf{I_5}\otimes\mathbf{\hat{x}_t}\right)'\mathbf{h_{xx}}\mathbf{\hat{x}_t} + \sigma\eta\varepsilon_{\mathbf{t+1}}\end{aligned}$$

where $\mathbf{g}_{\sigma\sigma}$ and $\mathbf{h}_{\sigma\sigma}$ are the second derivatives of $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$, with respect to σ, respectively; and $\mathbf{g_{xx}}$ and $\mathbf{h_{xx}}$ are the second derivatives of $\mathbf{g}$ and $\mathbf{h}$ with respect to $\mathbf{x}_t$, respectively; and all the elements in matrices $\mathbf{g_{xx}}$ and $\mathbf{h_{xx}}$ are functions of these structural parameters $\mathbf{\Theta}$. The third order approximation is following the similar form with additional terms of the third derivatives of $\mathbf{g}(\cdot)$ and $\mathbf{h}(\cdot)$.

Finally, it is worth mentioning that there are three possible results depending on parametrization of this DSGE model: no stable rational expectations solution exists; the stable solution is unique (determinacy); or there are multiple stable solutions (indeterminacy). In this paper, I focus on the case of the determinacy and restrict the parameter space accordingly.

2.3.2 The Projection Method

As described in Judd (1992, 1998), projection methods solve the DSGE model by proceeding the following steps. In the first step, the policy functions derived from DSGE model's optimization decisions are approximated by polynomial functions with unknown coefficients. Then, when equilibrium conditions involve the conditional expectations, often seen in the model's inter-temporal equilibrium conditions, numerical integration method is employed. Third, a set of grid points in the state space is chosen and approximation error (residual) from the model's equilibrium conditions is calculated at each grid point. Finally, these unknown coefficients associated with the approximating polynomial functions are determined by minimizing the residuals subject to some loss criterion.

Since different policy rules usually show various degrees of nonlinearities, it is sensitive to choose the set of policy functions when applying projection method. After some experimentation, I choose to approximate the policy functions for c_t, k_{t+1}, and π_t , and denote them as functions of the 5 state variables $\mathbf{x_t}$ in the present analysis:

$$c_t = f^c(\mathbf{x_t}) \tag{2.3.5}$$

$$k_{t+1} = f^k(\mathbf{x_t}) \tag{2.3.6}$$

$$\pi_t = f^\pi(\mathbf{x_t}) \tag{2.3.7}$$

where $f^i(\cdot) : \mathbb{R}^5 \to \mathbb{R}$ for $i \in \{c, k, \pi\}$. As shown in the beginning of this section, the 5 state variables are the capital stock k_t, the lagged short term interest rate R_{t-1}, the preference shocks a_t, the productivity shocks z_t and the monetary shocks $\varepsilon_{R,t}$. Note that only the first two variables are endogenous state variables, and the other 3 variables are exogenous state variables.

Since the function forms of f^i are unknown, they need to be approximated before fitting into estimation procedure. Here, I choose Chebyshev polynomial functions as basis functions and consider the second and third degree approximation to $f^i : \mathbb{R}^5 \to \mathbb{R}$ for $i \in \{c, k, \pi\}$. It yields

$$\begin{aligned} c_t &\approx \hat{f}^c(\mathbf{x_t}, \mathbf{B^c}) \\ k_{t+1} &\approx \hat{f}^k(\mathbf{x_t}, \mathbf{B^k}) \\ \pi_t &\approx \hat{f}^\pi(\mathbf{x_t}, \mathbf{B}^\pi) \end{aligned}$$

where $\hat{f}^i(\cdot) : \mathbb{R}^5 \to \mathbb{R}$ for $i \in \{c, k, \pi\}$ are Chebyshev polynomials, and $\mathbf{B^i}$ for $i \in \{c, k, \pi\}$ are unknown coefficients associated with these polynomials. Univariate Chebyshev polynomials are a family of orthogonal polynomials on the interval $[-1, 1]$, and multivariate Chebyshev polynomials can easily be constructed as the products of these univariated polynomials.[6] However, it is extremely computationally expensive to apply conventional projection method using the tensor product, even for medium-dimensional models like this current one. Therefore, I follow a more convenient approach – using complete polynomial basis – suggested by Judd (1992) and Gaspar and Judd (1997).[7]

In order to avoid the curse of dimensionality, I apply a monomial rule Galerkin method proposed by Pichler (2011) to this model. The key feature of this Galerkin method is to use non-product monomial cubature rules for computing conditional expectations and weighted residuals. The basic structure is as follows. In order to compute these unknown coefficients in $\hat{f}^i(\cdot)$, $i \in \{c, k, \pi\}$, I first define a hypercube

[6] Chebyshev polynomials are defined over $[-1, 1]$ by the formula $T_n(x) \equiv \cos(n \arccos(x))$. They are generated by the recursion scheme

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

, which is initialized by $T_0(x) = 1$ and $T_1(x) = x$.

[7] The complete set of polynomials of total degree k is defined by

$$\Psi = \left\{ \prod_{j=1}^{n_x} T_{i_j}(x_j) | \sum_{j=1}^{n_x} \leq k,\, 0 \leq i_1, \ldots, i_{n_x} \right\}$$

.

for the state variables and rely on non-product monomial rules to obtain grid points within the hypercube.[8] Then I use the basis functions of $\hat{f}^i(\cdot)$, $i \in \{c,\, k,\, \pi\}$ – Chebyshev polynomials – as weighting functions $\omega_i(\mathbf{x_t})$ to compute the weighted residuals. Finally, I search for the values for these unknown coefficients by equating all weighted residuals to zero. For more details see the technical appendix.

Once these policy functions $\hat{f}^i(\cdot)$, $i \in \{c,\, k,\, \pi\}$ are obtained, the expressions for the remaining endogenous variables can be easily found using the equilibrium conditions. In particular, take output y_t as an example,

$$
\begin{aligned}
y_t &= \left[1 - \frac{\phi_p}{2}\left(\frac{\pi_t}{\pi} - 1\right)\right]^{-1}\left[k_{t+1} + \frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t} - 1\right)^2 k_t - (1-\delta)k_t\right] \\
&\approx \left[1 - \frac{\phi_p}{2}\left(\frac{\hat{f}^\pi(\mathbf{x_t},\, \mathbf{B}^\pi)}{\pi} - 1\right)\right]^{-1}\left[\hat{f}^k(\mathbf{x_t},\, \mathbf{B^k}) + \frac{\phi_k}{2}\left(\frac{\hat{f}^k(\mathbf{x_t},\, \mathbf{B^k})}{k_t} - 1\right)^2 k_t - (1-\delta)k_t\right] \\
&\equiv \hat{f}^y(\mathbf{x_t},\, \mathbf{B^y})
\end{aligned}
$$

2.4 Estimation and Stability Tests

This section first describes the estimation of the structural parameters by maximum likelihood method, then illustrates how to construct structural break tests to investigate parameter stability. Equipped with the solutions in previous section, the state space representation is completed by specifying the measurement equations. It yields

$$\mathbf{x_{t+1}} = \mathbf{H}(\mathbf{x_t},\, \varepsilon_{\mathbf{t+1}},\, \mathbf{\Theta}) \tag{2.4.1}$$

$$\mathcal{Y}_t = \mathbf{G}(\mathbf{x_t},\, \mathbf{u_t},\, \mathbf{\Theta}) \tag{2.4.2}$$

[8]Another possible solution is the Smolyak's algorithm presented by Malin et al. (2011). However, the Smolyak algorithm suffers from its lack of universal applicability.

where equation (21) called transition equation provides the law of motion for 5 state variables, and equation (22) is the measurement equation, in which $\mathcal{Y}_t$ is the subset of imperfectly observable variables of $\mathbf{y_t}$. Particularly, I assume only output, inflation, and nominal interest rates are observables; measurement errors $\mathbf{u_t}$ are also added into equation (22) and are assumed normally distributed and uncorrelated, i.e. $\mathbf{u_t} \sim \mathbf{N}(\mathbf{0}, \mathbf{\Sigma_u})$.

2.4.1 Evaluation of the Likelihood Function

In order to evaluate the likelihood function, I take advantage of the hidden Markov structure of the state space representation. In principle, the likelihood function can be written as

$$L(\mathbf{\Theta}|\mathcal{Y}^T) = p(\mathcal{Y}_1|\mathbf{\Theta}) \prod_{t=2}^{T} p(\mathcal{Y}_t|\mathcal{Y}^{T-1}, \mathbf{\Theta}) \tag{2.4.3}$$

where $\mathcal{Y}^T = \{\mathcal{Y}_1, \ldots, \mathcal{Y}_T\}$. For the first-order perturbation (log-linearized) approximation of the model, the transition and measurement equation are linear and the shocks are normally distributed. So the Kalman filter (see Hamilton (1994)) is applied to construct the likelihood function. Unfortunately, linearity and Gaussian assumptions are no longer satisfied for these nonlinear solutions to DSGE models. The Kalman filter can not be used to compute the likelihood function.

Clearly, conditional distributions of state variables do not belong to, in general, any known distribution family. The evaluation of the likelihood function is forced to resort to some type of simulation: a particular example of sequential Monte Carlo methods, also known as the particle filter. The main idea is extremely simple. Now the conditional distribution $\{p(\mathbf{x_t}|\mathcal{Y}^{t-1}, \mathbf{\Theta})\}_{t=1}^T$ is approximated by an empirical distribution of N draws $\{\{\mathbf{x^i_{t|t-1}}\}_{i=1}^N\}_{t=1}^T$(a swarm of particles) from the sequence $\{p(\mathbf{x_t}|\mathcal{Y}^{T-1}, \mathbf{\Theta})\}_{t=1}^T$ generated by simulation. Then from the law of the large numbers, the likelihood function is obtained

$$L(\mathbf{\Theta}|\mathcal{Y}^T) \approx \frac{1}{N}\sum_{i=1}^{N} p(\mathcal{Y}_1|\mathbf{x^i_{0|0}}, \mathbf{\Theta}) \prod_{t=2}^{T} \frac{1}{N}\sum_{i=1}^{N} p(\mathcal{Y}_t|\mathbf{x^i_{t|t-1}}, \mathbf{\Theta})$$

A brief description of the procedure can be found in the appendix. For more details on these methods see Doucet et al. (2001) and Arulampalam et al. (2002). In the literature on estimation of DSGE models, An and Schorfheide (2007) and Fernandez-Villaverde and Rubio-Ramírez (2007) have shown that the particle filter deliver better estimation of DSGE models.

2.4.2 Maximum Likelihood Estimation

In contrast to Bayesian estimation, I perform likelihood-based estimation from a classical perspective. In other words, parameters are interpreted as fixed but unknown, and the data are interpreted as the realization of a random drawing of data generating process, the present DSGE model. Therefore, once the likelihood function is constructed from either Kalman filter or Particle filter, parameter estimates $\hat{\Theta}$ are chosen from parameter space to maximize the likelihood function:

$$\hat{\Theta} = \arg\max_{\Theta \in \mathbf{\Theta}} L(\mathbf{\Theta}|\mathcal{Y}^T) \tag{2.4.4}$$

However, maximum likelihood estimation (MLE) of DSGE models, even small-scale ones as this model, is very challenging. The main source of challenges arises from the fact the likelihood $L(\mathbf{\Theta}|\mathcal{Y}^T)$ is parametrized in terms of state-space model (21) and (22), and these state-space models are complex and highly nonlinear function of Θ. Consequently, an ill-behaved likelihood surface may exist. Three classes of problems are usually seen in practice. First, the likelihood surface contains discontinuities, so that a small change in parameter value leads a jump in the value of the likelihood function. At such a point the likelihood function is not differentiable, so it makes derivative-based optimization methods not feasible. Second, the likelihood function usually has many local maxima so that estimates $\hat{\Theta}$ may not be a global maximum. This is not unusual for likelihood-based estimations, and it obviously makes the maximization task difficult. Finally, as documented by Canova and Sala (2009), lack of identification exists in estimation of DSGE models. Again, this model is no exception. The problem arises when the likelihood function is almost flat along some parameters space. One may raise the question of whether nonlinear solutions, like higher order perturbation methods, help with identification. That question is beyond the current range, and should be explored in the future work.

To address these problems aforementioned, I use the following strategies in this paper. First, to help with identification, I calibrate several parameters in line with the literature rather than estimate them via MLE. In particular, these are the parameters associated with investment, leisure, and real balances, which are hard to pin down without data on the respective variables. Second, I follow Fernandez-Villaverde and Rubio-Ramírez (2007) and use a simulated annealing approach instead of gradient-based methods for maximizing the likelihood function. It allows us to deal with a discontinuous likelihood function. Finally, I choose various sets of initial parameter values to deal with the presence of local maxima.

2.4.3 Parameter Stability Tests

One great strength of DSGE models is that they are supposed to be structural: these models have solid microfundations and include rational expectations, which imply that parameters of these models describing private agent's tastes and technologies ought to remain constant, even across periods when monetary and fiscal policy regimes change. In order to study the parameter stability, I take two different treatments on break dates. The first one is to estimate this model over two disjoint sub-samples. The break date is chosen at 1979 in line with the literature, which corresponds to a date around when there are major changes in US monetary and fiscal policies. Another one is to treat the break dates as unknown, and to search for the break dates.

2.4.3.1 Structural Break Test with Known Break Date

Let the vector Θ^1 and Θ^2 denote the estimated parameters from two disjoint subsamples: pre-1979 and post-1979. The null hypothesis can be specified as H_0 : $\Theta^1 = \Theta^2$. This classical structural break test goes back to Chow (1960), and is extended by Andrews and Fair (1988). Therefore, I follow the treatment in Andrews and Fair (1988) and construct the likelihood ratio statistic as

$$LR = 2[\ln L(\mathbf{\Theta}^1|\mathcal{Y}^{1T}) + \ln L(\mathbf{\Theta}^2|\mathcal{Y}^{2T}) - \ln L(\mathbf{\Theta}|\mathcal{Y}^{T})] \tag{2.4.5}$$

where $L(\mathbf{\Theta}^1|\mathcal{Y}^{1T})$ and $L(\mathbf{\Theta}^2|\mathcal{Y}^{2T})$ are the maximized log-likelihood functions for the first subsample and the second subsample, respectively. Andrews and Fair (1988) has shown that this LR statistic is asymptotically distributed as χ^2 random variable with q degrees of freedom under the null hypothesis, where q is the number of estimated parameters allowed to change.

2.4.3.2 Structural Break Test with a Single Unknown Break Date

One of the main assumptions of the Chow type LR test above is that one can "arbitrarily" choose the break date. However, this information on the break date is often unknown, or unobservable. As Hansen (2001) points out that the Chow type test will be either uninformative or misleading, this paper, therefore, first consider a single structural break with unknown date. The idea was originally proposed by Quandt (1960), and recently Andrews (1993) and Andrews and Ploberger (1994) generalized the structural break tests with unknown change point in nonlinear parametric models. Their proposed statistics are designed for one-time change in the value of a parameter vector. Tests are considered for both the case of pure structural change and the case of partial structural change.

Now the null hypothesis of interest here is one of parameter stability: $H_0 : \Theta_t = \Theta$ for all t. The alternative hypothesis of interest is a one-time structural change with break date $\lambda_\epsilon \in (\epsilon, 1-\epsilon)$, where ϵ is a trimming parameter. The one time change alternative hypothesis with break date $[\lambda_\epsilon T]$ is given by

$$H_1(\lambda_\epsilon) : \Theta_t = \begin{cases} \Theta_1 & \text{for } t = 1, \ldots, \lambda_\epsilon T \\ \Theta_2 & \text{for } t = \lambda_\epsilon T + 1, \ldots, T \end{cases}$$

The basic idea is to use the maximum of the likelihood ratio test over all possible break dates. In this case of a single unknown break, this translates into the following statistic

$$\sup_{\lambda_\epsilon \in (\epsilon, 1-\epsilon)} LR_t(\lambda_\epsilon) \tag{2.4.6}$$

where the likelihood ratio test LR_t is constructed the same as the previous subsection. One can rejects H_0 for large values of $\sup_{\lambda \in \Lambda} LR_t(\lambda_\epsilon)$. The limiting distribution is given by Andrews (1993)

$$\sup_{\lambda_\epsilon \in (\epsilon, 1-\epsilon)} LR_t(\lambda_\epsilon) \Rightarrow \sup_{\lambda_\epsilon \in (\epsilon, 1-\epsilon)} G_q(\lambda_\epsilon) \tag{2.4.7}$$
$$\text{where } G_q(\lambda_\epsilon) = \frac{\lambda_\epsilon \left[W_q(1) - W_q(\lambda_\epsilon)\right]' \left[W_q(1) - W_q(\lambda_\epsilon)\right]}{\lambda_\epsilon(1-\lambda_\epsilon)}$$

where $W_q(\lambda_\epsilon)$ is a vector of independent Wiener processes of dimension q , the number of coefficients that are allowed to change. Note the limiting distribution depends on

q but it also depends on Λ_ϵ.[9]

2.4.3.3 Structural Breaks Test with Multiple Unknown Break Dates

While the $\sup LR$ test is primarily designed to test for a single structural break, multiple structural breaks may exist. However, the literature on tests for multiple structural breaks is relatively scarce, except of Bai and Perron (1998) and Qu and Perron (2007) for linear regressions. To this date, little has been known for multiple breaks in nonlinear models. In this paper, I propose a sequential procedure originally developed by Bai (1997) for testing multiple breaks in linear regressions. The advantage of this sequential procedure is to avoid the computation complexity when estimating multiple structural breaks simultaneously, and also to circumvent the challenge of nonexistence of asymptotic properties on statistics for multiple structural breaks in nonlinear models.

The basic structure is as follows. First, I test for a structural break using the $\sup LR_t(\lambda_\epsilon)$ statistic for the full sample of data. If there is significant evidence of a structural break over the full sample according to the $\sup LR_t(\lambda_\epsilon)$, I then calculate the $\sup LR_t(\lambda_\epsilon)$ for each of the two subsamples defined by the full-sample break date. If I fail to find evidence of a structural break using the $\sup LR_t(\lambda_\epsilon)$ statistic for each of the two subsamples, I can conclude that there is a single break. If there is significant evidence of a structural break in either of the two subsamples, I compute the $\sup LR_t(\lambda_\epsilon)$ statistic for each of the new subsamples defined by the new break date. I proceed in this manner until all of the subsamples defined by any significant break date have insignificant $\sup LR_t(\lambda_\epsilon)$ or the number of breaks reaches the maximum of allowed number of breaks, in this case, I choose at most 3 structural breaks.

[9]It is worth noting that the search for a maximum value be restricted is not simply a technical requirement. It affects the properties of the test. As Andrews (1993) shows, if $\epsilon = 0$ so that no restrictions are imposed, the $\sup LR$ test diverges to infinity under the null hypothesis, which means the power of the test decreases as ϵ get smaller. Hence, the trimming value should be large enough for the test to retain descent power, yet small enough to include break dates that are potential candidates. In the single break case, a popular choice is $\epsilon = 0.15$.

One may naturally wonder at the consistency of these estimated breaks from sequential procedure. In other words, are these estimated break dates consistent with those unknown true break dates? Since this paper does not provide any asymptotic properties, which is beyond the range of the paper and will be explored in my future research. However, I can still investigate finite-sample properties of the proposed sequential procedure by conducting a small Monte carol analysis. As shown in the appendix, I use a small non-linear state space model, which still keeps most ingredients from these more complicated DSGE models. The Monte Carlo experiment shows that this sequential procedure method seems to work well, which implies that the sequential procedure can deliver the "true" breaks.

2.5 Results

2.5.1 Data

This paper uses quarterly macroeconomic data for the United States to study parameter stability of different solutions to the DSGE model. In the data, output is measured by real gross domestic product (GDP), where I remove a linear trend from the logged GDP series. Inflation is based on changes in GDP deflator, and the nominal interest rate is measured by the rate on three-month Treasury bills. All these series are extracted from FRED2 database maintained by the Federal Reserve Bank of St. Louis. Except for the interest rate, all are seasonally adjusted. Also, the series for output is expressed in per-capita terms by dividing by the civilian, non-institutional population, age 16 and above. I focus on the sample from 1959:Q1 to 2007:Q4 for two reasons. First, the results here can be compared directly with those in Ireland (2001) and Inoue and Rossi (2011); the other reason is to avoid the constraint of zero lower bound for interest rates.[10]

2.5.2 Implication Issues

The first step in constructing stability tests for the DSGE model is to estimate the structural parameters together with the measurement error variances $\Sigma_{\mathbf{u}}$. This is done by using maximum likelihood methods as outlined in Section 4. However, for reasons that were also discussed in this section, I calibrate some parameters prior to estimation. As in Ireland (2001), I find it is difficult to estimate α, δ, and ϕ_k without data on the capital stock or investment. I thus set these parameters to $\alpha = 0.36$, $\delta = 0.025$, and $\phi_k = 10$. For similar reasons, I calibrate the mark-up parameter to $\theta = 6$. Furthermore, I choose χ_h such that the household spends 30% of its time working in the steady state, and χ_m to match the steady state ratio between real balances and quarterly output. Finally, as in An and Schorfheide (2007), the measurement error variances are calibrated rather than estimated. I set these variances equal to 10% of the variance of the respective data series. The remaining 14 parameters are estimated via maximum likelihood, either using the linear model together with the Kalman filter, or by using the nonlinear model together with the particle filter. In the latter case, I use $100,000$ particles for estimation.

2.5.3 Stability Tests Results

2.5.3.1 Known Break Date

The LR test results for the five different solutions to the DSGE model are reported in Table 1. Here the data are splited into two subsamples, in which the first subsample covers the periods ending in 1979:Q2 and the second subsample starts in 1979:Q3. This break date corresponds to the beginning of Paul Volker's chairmanship at the

[10]In the appendix, I discuss how to deal with zero lower bound using projection method. That will be extension of the current paper.

Federal Reserve System, when a major and fundamental change in monetary policy of the U.S. has occurred. This break date is also broadly in line with the literature, such as Clarida et al. (2000) and Ireland (2001). The test procedure has been described in section 4 and the corresponding LR test is asymptotically distributed as a χ^2 random variables with q degrees of freedom. It is important to mention this paper use asymptotic χ^2 distribution. Although a limited study found that finite sample distributions of LR tests for these solutions generally shows a shift to the right relative to the asymptotic χ^2 distribution, the use of asymptotic distribution can still lead to same conclusions as the empirical distribution.

Column 2 in Table 1 presents the LR test results for a pure structural break, in which all the 14 estimated structural parameters are allowed to change. Thus the corresponding LR test is asymptotically distributed as a χ^2 random variable with 14 degrees of freedom. The 1% and 5% critical values for χ^2 with 14 degrees of freedom are 29.1 and 23.7, respectively. Two observations can be drawn from column 2 in Table 1. On one hand, the null hypothesis that estimated parameters are stable are rejected by LR test results for all these five solutions. This finding supports the previous studies based on log-linearized solutions to DSGE models, and would not be considered as surprising. One the other hand, the strengths of the rejection of stability for these solutions are different, which is more interesting than the previous finding. Two points are worth pointing out. First, the null hypothesis of stability is rejected at 1% significant level for all the three perturbation solutions, while it is rejected at 5% significant level for the two solutions using projection method. Second, a closer examination of the LR test results reveals that the strength of rejection of stability for first-order perturbation (log-linearized) solutions is stronger than the other two nonlinear perturbation solutions. One possible interpretation of different strengths of rejection indicates the effect of misspecification. Since these misspecication can also manifest themselves in the form of time-varying parameters, log-linearization solutions to DSGE models would be more likely rejected by classical stability tests. The slight difference of rejection of stability between perturbation and projection methods might also highlight the fact that the solutions from projection methods are more accurate than those from perturbation methods.

However, it is important to note that the parameter instability detected by allowing all these parameters to change may reflect instability in policy rather than instability in the parameter describing tastes and technologies of the economy. In order to diagnose the possible source of instability, two additional LR tests are considered in this paper by allowing only subsets of these parameters to change. The 14 estimated structural parameters are divided into two groups: the first group consists of ρ_R, σ_R, ρ_y, and ρ_π, which represent policy reaction function; the second group includes that the 10 estimated parameters describe household preferences and firm behaviors, such as β, and γ among others.

Column 3 in Table 1 reports the LR test results for the null hypothesis of stability of monetary policy reaction function. In this test I do not assume the other 10 parameters are stable across subsamples, which help to avoid the traditional problem of standard stability tests mentioned by Inoue and Rossi (2011). Similar as the pure structural break tests, the null hypotheses of stability are strongly rejected for all five solutions. This is broadly consistent with empirical studies focus on just monetary policy rules, such as Clarida et al. (2000), Estrella and Fuhrer (2003) and Boivin and Giannoni (2006). However, this result is different from the findings of Ireland (2001), in which he considers the stability tests for each of these policy parameters and fail to reject the null hypotheses. Column 4 in Table 1 shows the LR test results for the null hypothesis of the parameters in the second group has remained stable. The LR tests reject the null hypothesis for all these perturbation solutions, but fail to reject them for the two solutions using projection method.

2.5.3.2 A Single Unknown Break

The $\sup LR$ tests for a single unknown break are constructed as one described in section 4 for all these five different solutions. Here the constrained model holds all parameters fixed across the entire sample, whereas the unconstrained model allows all parameters to vary across a certain candidate break date. The trimming value ϵ is set as 0.15, which means the length of any regimes should be greater than $[0.15 * T]$ periods. T is the sample size of 196 in this case. Table 2 provides all $\sup LR$ test results and associated estimated break dates.

This paper uses the asymptotic values of $\sup LR$ test provided by Andrews (1993). The 1% critical values for this asymptotic distribution with 14 degrees of freedom and trimming value of 0.15 is 39.2. As column 2 in Table 2 shows, the null hypothesis of stability is overwhelmingly rejected for all 5 solutions. First, the largest value of the $\sup LR$ test for log-linearization solution is recorded in 1984:Q1. This break date is in line with Estrella and Fuhrer (2003) in which they employ forward looking monetary policy models. It is also broadly consistent with Stock and Watson (2002), in which they apply $\sup Wald$ test in Bai et al. (1998) to VAR models. Secondly, for second and third order perturbation solutions, the largest values of $\sup LR$ tests shift backward in the late 1970s or the very early 1980s. These break dates are roughly similar as those break dates picked in the literature and the finding of Moreno (2004) using VAR models. Third, the estimated break dates for projection solutions are quite different from those for perturbation methods. The $\sup LR$ tests for the second and third projection solutions take their maxima in the early 1970s. This break date is surprisingly consistent with Zhu (2012), in which I apply latest structural break tests of Qu and Perron (2007) to both backward-looking and forward-looking monetary models. Overall, Table 2 shows that, once again, there is strong evidence of parameter instability in this DSGE model. In particular, the results from nonlinear solutions play the role of robustness analysis, and complement the existing findings of Ireland (2001) and Boivin and Giannoni (2006) among others.

The most important finding from Table 2 is that the estimated break dates differ from perturbation and projection solutions. There are several possible explanations for the difference. The first explanation may be related with the difference between perturbation method and projection method. The first one approximates locally the policy functions while the latter builds approximated functions globally. However, little has been known about the estimation and inference from different solution methods to this date. This paper is among very few attempts to estimate DSGE models solved by projection methods.[11] Further investigation needs to do in the future work. Another possible way to explain the difference is that there might be multiple breaks in this DSGE model. Given the substantial changes of economic structure over the last few decades, it is natural to allow the possibility of multiple structural breaks. Figure 2 plots the sequence of LR tests as a function of all these candidate break dates for three solutions: the log-linearized solution, the second order perturbation solution and the second order projection solution. The candidate break dates are along the x-axis; the values of the LR tests on the y-axis. Figure 2 shows that there are considerable variations of the LR test sequences across candidate break dates. Most importantly, it presents that there are few local maxima in these LR test sequences, which implies the possibility of multiple structural breaks.

2.5.3.3 Multiple Structural Breaks

Figure 2 has already shown the possibility of multiple structural breaks in this DSGE model, thus this section presents the results of searching for multiple breaks in the model using sequential procedure. The $\sup LR$ tests are constructed as similar as those for a single unknown break and the trimming value ϵ is still set as 0.15. The only modification is to redefine subsample ranges if a second round of searching

[11] Fernandez-Villaverde and Rubio-Ramírez (2005) compare estimates of a neoclassical growth model solved by both linear and nonlinear solution methods.

needed. One can also recycle these LR test sequences in the previous subsection, which help to reduce computation burden.

Table 3 reports the results from sequential procedure for all the five solutions. First, the sequential procedure shows that there are three structural breaks for perturbation solutions to this DSGE model. The first breaks come from the first round of searching, and they are the same as estimated break dates shown in Table 2, which occur in the late 1970s and the early 1980s. Two additional breaks are found in the second round of searching. One is in the early 1970s, and another one locates roughly in the middle 1990s. Second, there are only two breaks detected for projection solutions. The sequential procedure could not identify a third break around the late 1970s and the early 1980s. Figure 3 visualizes these breaks in Table 3. Evidently, although all these break dates are not exactly the same with one another, one can still identify two break dates are shared by all these five solutions: one in the early 1970s, another in the middle 1990s. To this date, this paper appears to the first investigation on multiple breaks in DSGE models. Unlike these breaks found using VAR and Structural VAR models, these common break dates in this paper are informative for calibration and estimation of DSGE models. Furthermore, one may also relate these breaks to economic theory or history events. For instance, the first break of the early 1970s may correspond to either starting of oil price shocks or changes in aggregate productivity. To identify the source of these instabilities, one must choose accurate break dates, otherwise, the conclusions might be misguided.

Once again, the puzzling thing is that the different results from perturbation solutions and projection solutions. The projection solutions fail to identify the breaks around in the late 1970s and the early 1980s. It is also in contrast with these existing findings in the literature, and even more surprising considering that Paul Volker started his Chairmanship at the Federal Reserve System in 1979, subsequently the U.S. were confronted some fundamental changes in monetary and fiscal policies. One possible explanation might be that projection methods approximate these policy functions globally, and the solutions would be less mis-specificed than those locally approximation methods, such as perturbation methods.

2.5.3.4 Discussion

Having documented the strong evidence of parameter instability of the DSGE model, I proceed to discuss few implications of this exercise. First, one might argue that the strong evidence of parameter instabilities of DSGE models, regardless of what solution methods used, as the evidence that those models, like current one, just don't fit the data well and fail to response to Lucas (1976) critique. And therefore one might attempt to formulate that DSGE models are misguided. However, as Inoue and Rossi (2011) also point out, the structural parameters are just policy-invariant, not necessarily time-invariant. As the economic structures change, these structural parameters may vary as well. The Lucas (1976) critique plays like a warning sign which highlights the importance of applying stability tests to macroeconomic models, even on the so-called structural DSGE models. Second, since the timing of structural breaks from linear and nonlinear solutions is quite similar, one might conclude that the nonlinearities have no effect on parameter stability of DSGE models. I must caution that the current New Keynesian model is not a fundamental one. In a model with stochastic volatility (see Justiniano and Primiceri (2008) among others), nonlinear policy rules (see Eggertsson and Woodford (2003) among others), or Epstein-Zin preference (see van Binsbergen et al. (2012) among others), nonlinearities may lead to different conclusions. For instance, uncertainty shocks can not enter into solutions if one only considers linearization or log-linearization solutions. As the literature has documented, uncertainty shocks are important to interpreting the macroeconomic fluctuations (see Basu and Bundick (2012) among others). Third, the searching for multiple structural breaks has connections with the literature which deals with DSGE models with a Markov-switching process in different aspects of the environment, such as monetary or fiscal policy (see Sims and Zha (2006), Liu et al. (2011) among others). The proposed sequential procedure offers an alternative way to help us understand the dynamics of economies.

2.6 Conclusion

In this paper, I employ a representative New Keynesian model to study the parameter stability of DSGE models. The New Keynesian model is solved and estimated by both linear and nonlinear solution methods, particularly perturbation methods and projection methods. The hypothesis test results show that there is strong evidence of parameter instability in this New Keynesian model. The comparison among these different solutions first supports the findings using log-linearization solutions in the existing literature. It indicates that these parameter instabilities documented in the existing literature are not all due to linearization. Also, the comparison highlights the different effects of perturbation methods and projection methods on parameter stability. Furthermore, this paper documents two *common* structural breaks among these parameters for all these five solutions. The timing of structural breaks are informative, and should be considered to incorporate into any DSGE modeling.

The purpose of this paper, however, is not to argue that DSGE models are misguided. I would emphasize that, these results shown in this paper, once again, highlight the message conveyed by Lucas (1976). The so-called structural parameters in macroeconomic models are just policy invariant, but not necessarily time invariant. Therefore, one should be encouraged to apply stability tests to these macroeconomic models. In light of these results and methods, this paper raises a number of questions to be explored for future research. Perhaps the most pressing one is that this paper has not identified specific source of parameter instability in the detail. Recently Inoue and Rossi (2011) has proposed a procedure to search for set of stable parameters. This new econometric technique allows to address the stability properties of each single parameter in DSGE models. It can be extended to these nonlinear solutions in this paper. Another area of future research will be the introduction stochastic volatility in the representative New Keynesian model in this paper, as the literature has shown stochastic volatility may account for large part of macroeconomic fluctuations, and instabilities in the parameter estimates. Finally, the sequential procedure proposed in this paper needs more rigorous econometric treatments, which will complete the econometric literature on estimation and testing of structural breaks in nonlinear models.

2.7 Chapter 2: Appendix

2.7.1 Derivation of Equilibrium Conditions

This section documents how to derive the symmetric equilibrium conditions of the New Keynesian model and calculate the steady state values.

2.7.1.1 Household

Denote $c_t = C_t$, $k_t = K_t$ $m_t = \frac{M_t}{P_t}$, $b_t = \frac{B_t}{P_t}$, $w_t = \frac{W_t}{P_t}$, $q_t = \frac{Q_t}{P_t}$ and $\pi_t = \frac{P_t}{P_{t-1}}$, I first compute the first order conditions from the household's optimization problem by constructing the following Lagrangian:

$$
\begin{aligned}
L_H \quad = \quad & E\sum_{t=0}^{\infty}\beta^t\left\{a_t\frac{c_t^{1-\gamma}-1}{1-\gamma}+\chi_m\log(m_t)-\chi_h\frac{h_t^{1+\nu}-1}{1+\nu}\right. \\
& +\Lambda_t\left[\frac{m_{t-1}+b_{t-1}}{\pi_t}+w_th_t+q_tk_t+d_t+\tau_t-c_t-k_{t+1}+(1-\delta)k_t\right.
\end{aligned}
$$

$$-\frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t}-1\right)^2 k_t - \frac{b_t}{R_t} - m_t\Bigg]\Bigg\}$$

where Λ_t is the multiplier associated with the household's budget constraint. In order to calculate the first-order conditions, the Lagrangian is differentiated with respect to the choice variables c_t, m_t, h_t, b_t, k_{t+1}, and Λ_t:

$$\begin{aligned}
a_t c_t^{-\gamma} - \Lambda_t &= 0 && (2.7.1)\\
\chi_m m_t^{-1} - \Lambda_t + \beta E \frac{\Lambda_{t+1}}{\pi_{t+1}} &= 0 && (2.7.2)\\
\chi_h h_t^{\nu} - \Lambda_t w_t &= 0 && (2.7.3)\\
\frac{\Lambda_t}{R_t} - \beta E_t \frac{\Lambda_{t+1}}{\pi_{t+1}} &= 0 && (2.7.4)\\
-\Lambda_t\left[1+\phi_k\left(\frac{k_{t+1}}{k_t}-1\right)\right] + \beta E_t \Lambda_{t+1}\big[(q_{t+1}+1-\delta) & && \\
-\frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t}-1\right)^2 + \phi_k\left(\frac{k_{t+2}}{k_{t+1}}-1\right)\left(\frac{k_{t+2}}{k_{t+1}}\right) &= 0 && (2.7.5)\\
\frac{m_{t-1}+b_{t-1}}{\pi_t} + w_t h_t + q_t k_t + d_t + \tau_t - c_t - k_{t+1} + (1-\delta)k_t & && \\
-\frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t}-1\right)^2 k_t - \frac{b_t}{R_t} - m_t &= 0 && (2.7.6)
\end{aligned}$$

The multiplier Λ_t can be eliminated using (A.1), then

$$\begin{aligned}
\chi_m m_t^{-1} - a_t c_t^{-\gamma} + \beta E \frac{a_{t+1} c_{t+1}^{-\gamma}}{\pi_{t+1}} &= 0 && (2.7.7)\\
\chi_h h_t^{\nu} - a_t c_t^{-\gamma} w_t &= 0 && (2.7.8)\\
\frac{a_t c_t^{-\gamma}}{R_t} - \beta E_t \frac{a_{t+1} c_{t+1}^{-\gamma}}{\pi_{t+1}} &= 0 && (2.7.9)\\
-a_t c_t^{-\gamma}\left[1+\phi_k\left(\frac{k_{t+1}}{k_t}-1\right)\right] + \beta E_t\left\{a_{t+1} c_{t+1}^{-\gamma}\big[(q_{t+1}+1-\delta)\right. & && \\
\left.-\frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t}-1\right)^2 + \phi_k\left(\frac{k_{t+2}}{k_{t+1}}-1\right)\left(\frac{k_{t+2}}{k_{t+1}}\right)\Big]\right\} &= 0 && (2.7.10)\\
\frac{m_{t-1}+b_{t-1}}{\pi_t} + w_t h_t + q_t k_t + d_t + \tau_t - c_t - k_{t+1} + (1-\delta)k_t & &&
\end{aligned}$$

$$-\frac{\phi_k}{2}\left(\frac{k_{t+1}}{k_t}-1\right)^2 k_t-\frac{b_t}{R_t}-m_t \quad = \quad 0 \quad (2.7.11)$$

Note that the first four equations stated above correspond to equilibrium conditions $(1)-(4)$ in section 2. As symmetric equilibrium is considered in this model, market clears such that $b_t = b_{t-1} = 0$ and $\tau_t = m_t - m_{t-1}/\pi_t$ for equation (A.11), which will yield the equilibrium condition (9).

2.7.1.2 The Intermediate Goods-producing firms

Denote $y_t = Y_t$, I then consider the optimization problem confronted the intermediate goods-producing firm i, $i \in [0, 1]$ which is also described by the following Lagrangian:

$$\begin{aligned} L_M^i \quad = \quad & E\sum_{t=0}^{\infty}\beta^t\left\{\Lambda_t\left[w_tH_t(i)+q_tK_t(i)+\frac{\phi_p}{2}\left[\frac{P_t(i)}{\pi P_{t-1}(i)}-1\right]^2 y_t-\left(\frac{P_t(i)}{P_t}\right)^{1-\theta}y_t\right]\right. \\ & +\Xi_t(i)\left[K_t(i)^{\alpha}[z_tH_t(i)]^{1-\alpha}-\left(\frac{P_t(i)}{P_t}\right)^{-\theta}y_t\right] \end{aligned}$$

where $\Xi_t(i)$ denotes the multiple associated with the firm's production constraint. As well, L_M^i is differentiated with respect to $H_t(i)$, $K_t(i)$, $P_t(i)$, and multiplier, $\Xi_t(i)$. The first order conditions are represented as follows:

$$\Lambda_t w_t-\Xi_t(i)(1-\alpha)K_t(i)^{\alpha}z_t^{1-\alpha}H_t(i)^{-\alpha} \quad = \quad 0 \qquad (2.7.12)$$

$$\Lambda_t q_t-\Xi_t(i)\alpha K_t(i)^{\alpha-1}z_t^{1-\alpha}H_t(i)^{1-\alpha} \quad = \quad 0 \qquad (2.7.13)$$

$$\begin{aligned} (1-\theta)\Lambda_t\left(\frac{P_t(i)}{P_t}\right)^{-\theta}\frac{y_t}{P_t}+\theta\Xi_t(i)\left(\frac{P_t(i)}{P_t}\right)^{-\theta-1}\frac{y_t}{P_t} & \\ -\left(\frac{P_t(i)}{\pi P_{t-1}(i)}-1\right)\frac{\Lambda_t\phi_p y_t}{\pi P_{t-1}(i)} & \\ +\beta\phi_p E_t\left\{\Lambda_{t+1}\left(\frac{P_t(i)}{\pi P_{t-1}(i)}-1\right)\left(\frac{P_{t+1}(i)y_{t+1}}{\pi P_t(i)^2}\right)\right\} & \quad = \quad 0 \qquad (2.7.14) \end{aligned}$$

$$K_t(i)^\alpha [z_t H_t(i)]^{1-\alpha} - \left(\frac{P_t(i)}{P_t}\right)^{-\theta} y_t = 0 \tag{2.7.15}$$

As well as the previous subsection, a symmetric equilibrium is considered in this model, which implies that the intermediate goods-producing firms make identical choices such that $h_t = H_t = H_t(i)$, $k_t = K_t = K_t(i)$, $P_t = P_t(i)$, and $\Xi_t = \Xi_t(i)$.

$$\Lambda_t w_t h_t - \Xi_t(1-\alpha)y_t = 0 \tag{2.7.16}$$

$$\Lambda_t q_t k_t - \Xi_t \alpha y_t = 0 \tag{2.7.17}$$

$$\begin{aligned}(1-\theta)\Lambda_t + \theta\Xi_t - \Lambda_t\phi_p\left(\frac{\pi_t}{\pi} - 1\right)\frac{\pi_t}{\pi} & \\ +\beta\phi_p E_t\left\{\Lambda_{t+1}\left(\frac{\pi_{t+1}}{\pi} - 1\right)\frac{\pi_{t+1}}{\pi}\frac{y_{t+1}}{y_t}\right\} &= 0\end{aligned} \tag{2.7.18}$$

$$k_t^\alpha [z_t h_t(i)]^{1-\alpha} - y_t = 0 \tag{2.7.19}$$

Here Ξ_t is eliminated and Λ_t is replaced using (A.1). Finally, it is worth noting that the central bank's short term interest rate rule must be satisfied and all the exogenous variables must follow their respective laws of motion. In a symmetric equilibrium, adding these equations complete the equilibrium conditions for the model.

2.7.1.3 The Steady State

Without any disturbances in the model, the economy will converge to a non-stochastic steady state in which all of the variables remain constant over time. First, the steady state value of the preference shock a_t is equals 1, and the steady state value of the technology shock z_t correspond to the parameter z. Also, the steady state value of the inflation rate π_t corresponds to the parameter π. The remaining 9 steady state values of y, c, m, h, w, q, k, d, and R are then uniquely determined as

$$R = \frac{\pi}{\beta} \tag{2.7.20}$$

$$q = \frac{1}{\beta} - 1 + \delta \tag{2.7.21}$$

$$w = \frac{(1-\alpha)}{\alpha} qz \left(\frac{\alpha(\theta - 1)}{q\theta} \right)^{1/(1-\alpha)} \tag{2.7.22}$$

$$h = \left\{ \frac{1}{\chi_h} \left[\left(\frac{q\theta}{\theta - 1} - \delta \right) z \left(\frac{\alpha(\theta - 1)}{q\theta} \right)^{1/(1-\alpha)} \right]^{-\gamma} w \right\}^{\frac{1}{\nu+\gamma}} \tag{2.7.23}$$

$$k = zh \left(\frac{\alpha(\theta - 1)}{q\theta} \right)^{1/(1-\alpha)} \tag{2.7.24}$$

$$c = \left(\frac{q\theta}{\theta - 1} - \delta \right) zh \left(\frac{\alpha(\theta - 1)}{q\theta} \right)^{1/(1-\alpha)} \tag{2.7.25}$$

$$y = \frac{q\theta}{\alpha(\theta - 1)} k \tag{2.7.26}$$

$$m = c^{\gamma} \frac{\chi_m}{(1 - \beta/\pi)} \tag{2.7.27}$$

$$d = y - wh - qk \tag{2.7.28}$$

2.7.2 Solution to the Model using Projection Method

This section describes how to solve the model using projection method. In particular, I apply the monomial-rule Galerkin method proposed by Pichler (2011) to the model in order to reduce the computation burden.

As described in section 3.2, the first step before applying any projection methods is to choose a set of policy functions. Here I seek to approximate the policy functions for consumption c_t, capital stock of next period k_{t+1}, and inflation rate π_t. These policy functions are functions of the 5 state variables $\mathbf{x_t}$ which consists of the capital stock k_t, the lagged short term interest rate R_{t-1}, the preference shocks a_t, the productivity shocks z_t and the monetary shocks $\varepsilon_{R,t}$.

$$c_t = f^c(\mathbf{x_t}) \tag{2.7.29}$$

$$k_{t+1} = f^k(\mathbf{x_t}) \tag{2.7.30}$$

$$\pi_t = f^\pi(\mathbf{x_t}) \tag{2.7.31}$$

Since the function forms of f^i are unknown, they need to be approximated using function basis. Here I use Chebyshev polynomials and consider both second and third order approximations, in which the order refers to the degree of polynomial function. Taking second order approximation as an example, if tensor product were used, the number of coefficients associated with these policy functions for this 5 dimensional state space is equal to $3^5 = 243$. This exponential growth of the number of coefficients becomes a computational burden. Thus, I follow Judd (1992) and Gaspar and Judd (1997) to use complete polynomials instead. In this case, the number of coefficients only amounts to 21 for second order approximation. Thus, the next period's capital stocks, for instance, is approximated by the following form

$$k_{t+1} \approx \hat{f}^k(\mathbf{x_t}, \mathbf{B^k}) = \sum_{j=1}^{n_{B^k}} B_j \phi_j(\xi(\mathbf{x}^i))$$

where B_j is a coefficient vector of size n_{B^k}, and $\xi(\mathbf{x}^i)$ maps the state space $\mathbb{R}^5$ into the unit hypercube $[-1, 1]^5$. Here the standard linear transformation is given by

$$\xi(\mathbf{x}^i) = 2\frac{\mathbf{x_t^i} - \underline{\mathbf{x}}^\mathbf{i}}{\mathbf{x_t^i} - \overline{\mathbf{x}}^\mathbf{i}} - 1 \tag{2.7.32}$$

where $\underline{\mathbf{x}}^\mathbf{i}$ and $\overline{\mathbf{x}}^\mathbf{i}$ are the lower and upper bound for each state variable, respectively. Following suggestions in Pichler (2011), I choose a symmetric interval around $\pm 10\%$ of steady states for k_t and R_{t-1}, and for these three endogenous state variables, I choose the interval of ± 3 standard deviation around steady state values.

After all these preparations, the remaining steps are similar as the conventional Galerkin method except that I use more efficient multi-dimentional integration techniques, i.e. no-product monomial cubature rules: First, (Initial) values are assigned to structural parameters. Note, during the estimation process, these parameters should also satisfy their theoretical constraints. For example, the discount factor β should always lie between 0 and 1. Second, I can calculate the model's steady state, and build the associated state space according to the rules aforementioned. The Galerkin method determines the coefficients $\mathbf{B^i}$ of the approximating polynomial by constructing a system of weighted sums over residuals and equating these sums to zero. Thus, the third step is to select the grid-points and construct their associated weights. For second-order approximation, I use a non-product monomial rule of degree 5 – Rule $[C_n d5]$ as described in Pichler (2011). For third-order projection method, I use a non-product monomial rule of degree 7 – Rule $[C_n d7]$ as described in Pichler (2011) since Rule $[C_n d5]$ only allows for first and second order approximations. The advantage of non-product monomial rule is that it uses less grid points than Gauss-Chevbyshev approach. For example, only 51 grid points need to be used by Rule $[C_n d5]$ compared to 243 grid points by the Gauss-Chevbyshev approach. Fourth, the nodes and weights used to compute conditional expectations are selected. Again, I use a non-product monomial rule of degree 3 – Rule $[E_n^{r^2} d3]$ as described in Pichler (2011). Finally, The model is solved by equating weighted residuals to zero. Here I use the coefficients from the first-order perturbation solution as an initial guess.

Once the solution for the three policy functions are obtained, the remaining variables can be derived from equilibrium conditions. One example for output y_t is shown in section 3.

2.7.3 Dealing with Zero Lower Bound Constraint

When zero lower bound constraint is imposed on nominal interest rates, the central bank adjusts the short term nominal interest rate according to

$$R_t = \max[\tilde{R}_t,\ 1] \tag{2.7.33}$$

where the first term $\tilde{R}_t$ still follows the same conventional Taylor rule as (11), and the second term means that the gross nominal interest rate can not be lower than 1. Since this constraint brings a kink into the model, it prevents us from using perturbation methods. Fortunately, projection methods can handle this zero lower bound constraint without modifying the procedure stated above.[12] In application, once these three policy functions $\hat{f}^i$ are still solved as usual, then one can just substitute them into interest rate rules with zero lower bound. It yields

$$\begin{aligned} R_t &= \max\left[\exp^{\rho_R \ln R_{t-1} + (1-\rho_R)\ln R + \rho_\pi \ln\left(\frac{\pi_t}{\pi}\right) + \rho_y \ln\left(\frac{y_t}{y}\right) + \varepsilon_{Rt}},\ 1\right] \\ &\approx \max\left[\exp^{\rho_R \ln R_{t-1} + (1-\rho_R)\ln R + \rho_\pi \ln\left(\frac{\hat{f}^c(\mathbf{x_t},\, \mathbf{B}^\pi)}{\pi}\right) + \rho_y \ln\left(\frac{\hat{f}^y(\mathbf{x_t},\, \mathbf{B^y})}{y}\right) + \varepsilon_{Rt}},\ 1\right] \\ &\equiv \hat{f^R}(\mathbf{x_t},\ \mathbf{B^R}) \end{aligned}$$

As Fernandez-Villaverde. et. al. (2012) points out, this approach deals the kink of R_t at 1 in the Taylor rule without any approximation conditional on c_t, k_{t+1}, and π_t. Here R_t comes from application of Taylor rules directly.

2.7.4 Particle Filter Algorithm

As pointed out in Section 4, the Kalman filter can not be used to compute the likelihood function for nonlinear models, I employ the particle filter for these nonlinear

[12]Recent works such as Judd et al. (2010) and Fernandez-Villaverde et. al (2012) take similar strategy.

solutions to the DSGE model. The algorithm can be described as follows:

1. Initialization: Draw N particles $\{\mathbf{x_{0|0}}^{\mathbf{i}}\}$, $i = 1, \ldots, N$ by using the initial distribution of $p\left(\mathbf{x_0}|\Theta\right)$. In period t, I start with the particles $\{\mathbf{x_{t-1|t-1}}^{\mathbf{i}}\}$, $i = 1, \ldots, N$, which are randomly sampled from the discrete approximation of the true filtering density $p\left(\mathbf{x_{t-1}}|\mathcal{Y}^{t-1}, \Theta\right)$.

2. Prediction: Draw one-step ahead forecasted particles $\{\mathbf{x_{t|t-1}}^{\mathbf{i}}\}$, $i = 1, \ldots, N$ from $p\left(\mathbf{x_t}|\mathcal{Y}^{t-1}, \Theta\right)$ for each i. Note that

$$
\begin{aligned}
p\left(\mathbf{x_t}|\mathcal{Y}^{t-1}, \Theta\right) &= \int p\left(\mathbf{x_t}|\mathbf{x_{t-1}}, \Theta\right) p\left(\mathbf{x_{t-1}}|\mathcal{Y}^{t-1}, \Theta\right) d\mathbf{x_{t-1}} \\
&\approx \frac{1}{N}\sum_{i=1}^{N} p\left(\mathbf{x_{t|t-1}}^{\mathbf{i}}, \Theta\right)
\end{aligned}
$$

   Thus, one can draw N particles from $p\left(\mathbf{x_t}|\mathcal{Y}^{t-1}, \Theta\right)$by generating one particle from $\left(\mathbf{x_t}|\mathbf{x_{t-1}}^{\mathbf{i}}, \Theta\right)$ for each i.

3. Updating: From Bayes theorem,

$$
\begin{aligned}
p\left(\mathbf{x_t}|\mathcal{Y}^{t}, \Theta\right) &= \frac{p\left(\mathcal{Y}_\mathbf{t}|\mathbf{x_t}, \Theta\right) p\left(\mathbf{x_t}|\mathcal{Y}^{t-1}, \Theta\right)}{p\left(\mathcal{Y}_\mathbf{t}|\mathcal{Y}^{t-1}, \Theta\right)} \\
&\propto p\left(\mathcal{Y}_\mathbf{t}|\mathbf{x_t}, \Theta\right) p\left(\mathbf{x_t}|\mathcal{Y}^{t-1}, \Theta\right)
\end{aligned}
$$

   since $\{\mathbf{x_{t|t-1}}^{\mathbf{i}}\}$, $i = 1, \ldots, N$ are generated from the approximated $p\left(\mathbf{x_t}|\mathcal{Y}^{t-1}, \Theta\right)$, the approximation of the filtering density reduces to adjusting to probability weights assigned to $\{\mathbf{x_{t|t-1}}^{\mathbf{i}}\}$,$i = 1, \ldots, N$ according to $\{\hat{\mathbf{w}}_\mathbf{t}^\mathbf{i}\} = p\left(\mathcal{Y}_\mathbf{t}|\hat{\mathbf{x}}_\mathbf{t}^\mathbf{i}, \Theta\right)$. I normalized $\{\hat{\mathbf{w}}_\mathbf{t}^\mathbf{i}\}$, $i = 1, \ldots, N$ as follows:

$$
\mathbf{w}_\mathbf{t}^\mathbf{i} = \frac{\hat{\mathbf{w}}_\mathbf{t}^\mathbf{i}}{\sum_{j=1}^{N} \hat{\mathbf{w}}_\mathbf{t}^\mathbf{j}}
$$

   and note that the resulting sampler $\{\mathbf{x_{t|t-1}}^{\mathbf{i}}, \mathbf{w}_\mathbf{t}^\mathbf{i}\}$, $i = 1, \ldots, N$ approximates the true filtering density $p\left(\mathbf{x_t}|\mathcal{Y}^t,\,\Theta\right)$.

4. Resampling: The above samplers are undesirable because after a few iterations, most particles will have negligible weights and the accuracy of Monet Carlo approximation of the integral in step 2 and step 3 would deteriorate. To overcome this problem, I generate a new swarm of particles $\mathbf{x_{t|t}}^{\mathbf{i}}$ such that

$$\Pr\{\mathbf{x_{t|t}}^{\mathbf{i}} = \mathbf{x_{t|t-1}}^{\mathbf{i}}\} = \mathbf{w_t^i}$$

   the resulting sample is indeed a random sample from the discrete approximation of the filtering density $p\left(\mathbf{x_t}|\mathcal{Y}^t,\,\Theta\right)$, and hence is equally weighted.

5. Likelihood Evaluation. The log-likelihood can be approximated by using the average of unnormalized weights

$$\begin{aligned}\ln L(\boldsymbol{\Theta}|\mathcal{Y}^T) &\approx \frac{1}{N}\sum_{i=1}^{N} p(\mathcal{Y}_1|\mathbf{x_{0|0}^i},\boldsymbol{\Theta})\prod_{t=2}^{T}\frac{1}{N}\sum_{i=1}^{N} p(\mathcal{Y}_t|\mathbf{x_{t|t-1}^i},\boldsymbol{\Theta})\\ &= \sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{w}}_\mathbf{t}^\mathbf{i}\right)\end{aligned}$$

In my implementation of the above algorithm, several remarks are necessary to consider. First, since the state variables include pre-determined endogenous variables as well as structural shocks which follow linear processes, it is not obvious to get initial values for them. As in An and Schorfheide (2007), I draw the initial structural shocks from their unconditional distributions and generate the initial values of pre-determined endogenous state variables from putting the previous period's values to their steady state values. Second, I choose the number of particles based on the evaluation of the log-likelihood across 40 different random seeds. The standard deviation of the likelihood values changes only by a small amount after 100,000 particles. That motivates our choice of 100,000 particles.

2.7.5 Sequential Procedure for Multiple Breaks: A Small Monte Carlo Analysis

Since this paper does not provide the asymptotic properties on sequential procedure for multiple breaks, I conduct a small Monte Carlo simulation to investigate finite-sample properties of the proposed sequential procedure. First, the data-generating process is borrowed from Fernández-Villaverde and Rubio-Ramírez's short note on sequential Monte Carlo filter. The nonlinear state space is given by

$$x_t = \alpha + \beta \frac{x_{t-1}}{1 + x_{t-1}^2} + w_t \tag{2.7.34}$$
$$y_t = \delta x_t + v_t$$

where $w_t \sim N(0, \sigma_w)$ and $v_t \sim N(0, \sigma_v)$. α, β, δ, σ_w and σ_v are unknown parameter and need to be estimated. Assume there are two breaks in some elements of the coefficients in this model. Here I focus on estimates of α, β, and σ_v , and their true values are given by

$$\begin{cases} \alpha = 0.5,\ \beta = 0.3,\ \sigma_w = 1, & t \leq [0.3T] \\ \alpha = 1.0,\ \beta = 0.6,\ \sigma_w = 2, & [0.3T] < t \leq [0.7T] \\ \alpha = 0.5,\ \beta = 0.3,\ \sigma_w = 1, & t > [0.7T] \end{cases} \tag{2.7.35}$$

and I fixed $\delta = 1$ and $\sigma_v = 1$. The sample size T is taken with 100, and the "true" break points are at 30 and 70. In order to keep the features of DSGE models, I assume that only y_t is observable. Thus particle filter is employed to construct likelihood function. The estimation procedure and sequential procedure for multiple breaks are described in the Section 4. All simulation reports are based on 500 replications.

Figure 3 displays the estimated break points for this nonlinear state space model.

This Monte Carlo simulation shows that the sequential procedure can identify the number of breaks, and also indicates that these breaks are consistently estimated. The asymmetry shown in the distribution of the estimated break points may be due to numerical error. Overall, the sequential procedure can deliver consistent estimates of break points in nonlinear models.

2.8 Chapter 2: Tables and Figures

Table 2.1: Andrews and Fair (1988) LR test results for known break date (1979:Q2)

Solutions	Full Set	Monetary Policy	Private Sectors
Log Linearized	62.13***	18.71***	24.68***
Second Order Perturbation	50.01***	14.62***	17.83*
Third Order Perturbation	45.65***	15.00***	18.21**
Second Order Projection	24.83**	14.42***	14.84
Third Order Projection	23.74**	15.38***	15.05

Note: LR denotes the likelihood ratio statistic for testing the null hypothesis of parameter stability with known break date at 1979:Q2. This statistic is asymptotically distributed as χ_q^2 with q degrees of freedom, where q is the number of parameters which are allowed to change. $***$, $**$, and $*$ represent significant level at 1%, 5%, and 10%, respectively. For the full set case in column 2, all the 14 estimated parameters are allowed to change. In column 3 and column 4, only 4 and 10 parameters are allowed to change for the monetary policy case and the private sectors case, respectively.

Table 2.2: Andrews (1993) $\sup LR$ test results for a single unknown break

Solutions	$\sup LR$	Break Dates
Log Linearized	82.50***	$1984:Q1$
Second Order Perturbation	64.15***	$1978:Q3$
Third Order Perturbation	47.48***	$1980:Q2$
Second Order Projection	61.82***	$1974:Q2$
Third Order Projection	45.46***	$1971:Q1$

Note: $\sup LR$ denotes the statistic for testing the null hypothesis of parameter stability with a single unknown break. This statistic is asymptotically distributed as

$$\sup_{\lambda_\epsilon \in (\epsilon, 1-\epsilon)} \frac{\lambda_\epsilon \left[W_q(1) - W_q(\lambda_\epsilon)\right]' \left[W_q(1) - W_q(\lambda_\epsilon)\right]}{\lambda_\epsilon(1-\lambda_\epsilon)}$$

where q is the number of parameters which are allowed to change. The asymptotic critical values for the $\sup LR$ test are shown in Andrews (1993)(P.840). $***$, $**$, and $*$ represent significant level at 1%, 5%, and 10%, respectively. As convention in the literature, the trimming value ϵ is set to 0.15.

Table 2.3: Results on multiple structural breaks using sequential procedure

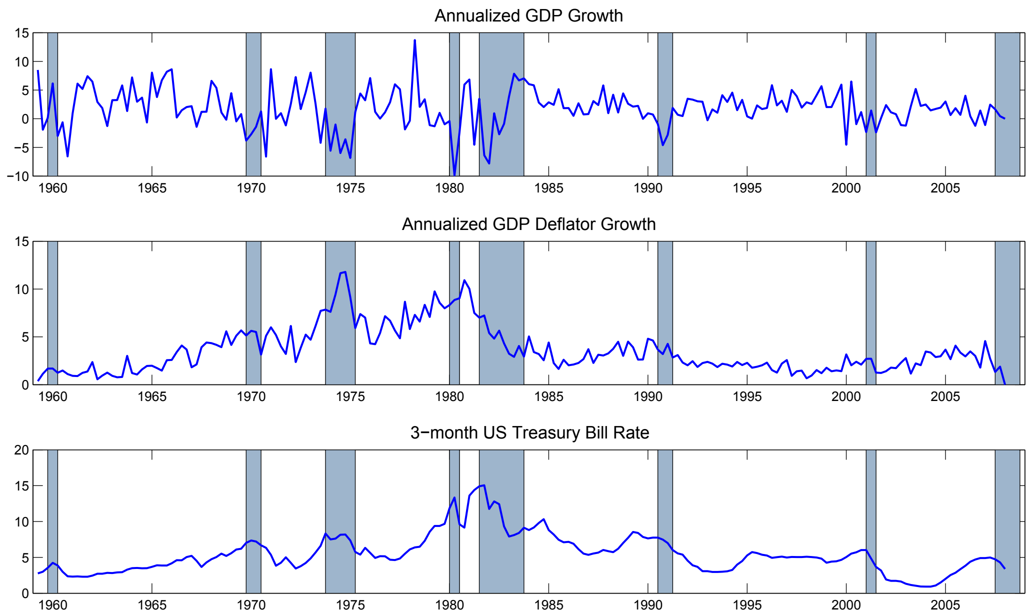
Solution	Sample Range	$\sup LR$	Break Dates
Log-Linearized	1959 : Q1 − 2007 : Q4	82.50***	1984 : $Q1$
	1959 : Q1 − 1984 : Q1	30.83*	1970 : $Q4$
	1984 : Q2 − 2007 : Q4	40.32***	1998 : $Q2$
Second Order Perturbation	1959 : Q1 − 2007 : Q4	64.15***	1978 : $Q3$
	1959 : Q1 − 1978 : Q3	38.41**	1972 : $Q4$
	1978 : Q4 − 2007 : Q4	41.07***	1995 : $Q2$
Third Order Perturbation	1959 : Q1 − 2007 : Q4	47.48***	1980 : $Q2$
	1959 : Q1 − 1980 : Q2	33.40*	1970 : $Q3$
	1980 : Q3 − 2007 : Q4	37.59**	1994 : $Q1$
Second Order Projection	1959 : Q1 − 2007 : Q4	61.82***	1974 : $Q2$
	1959 : Q1 − 1974 : Q2	−	−
	1974 : Q3 − 2007 : Q4	46.69***	1994 : $Q4$
Third Order Projection	1959 : Q1 − 2007 : Q4	45.46***	1971 : $Q1$
	1959 : Q1 − 1971 : Q4	−	−
	1971 : Q1 − 2007 : Q4	49.72***	1996 : $Q3$

Note: The sequential procedure is described in section 4. Here $\sup LR$ denotes the statistic for testing the null hypothesis of parameter stability with a single unknown break in the (sub)sample. This statistic is asymptotically distributed as

$$\sup_{\lambda_\epsilon \in (\epsilon, 1-\epsilon)} \frac{\lambda_\epsilon \left[W_q(1) - W_q(\lambda_\epsilon)\right]' \left[W_q(1) - W_q(\lambda_\epsilon)\right]}{\lambda_\epsilon(1-\lambda_\epsilon)}$$

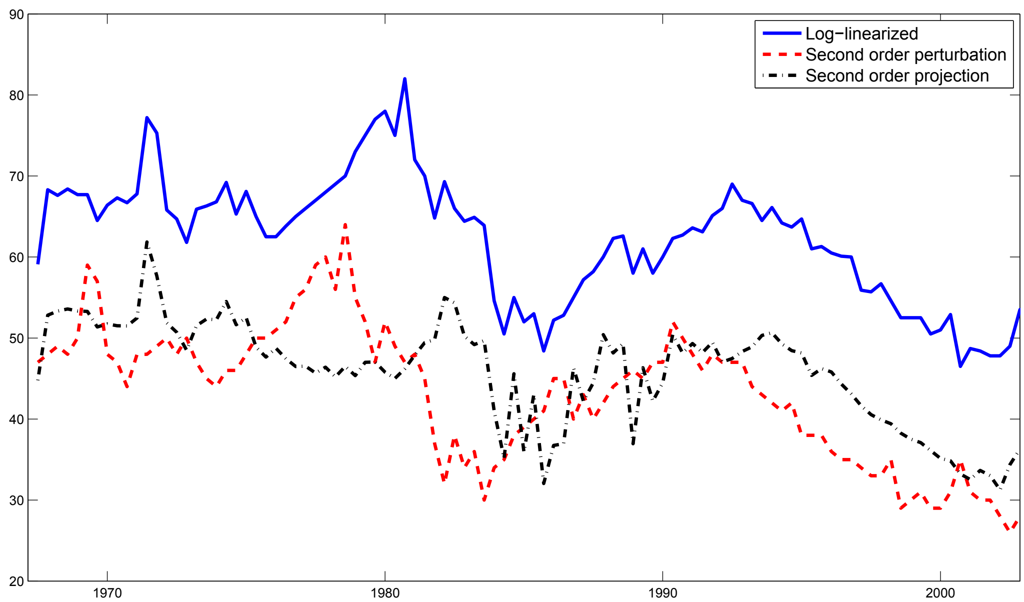
where q is the number of parameters which are allowed to change. The asymptotic critical values for the $\sup LR$ test are shown in Andrews (1993)(P.840). $***$, $**$, and $*$ represent significant level at 1%, 5%, and 10%, respectively. As convention in the literature, the trimming value ϵ is set to 0.15. $-$ denotes that no break is detected.

Figure 2.1: GDP growth, inflation, and short term interest rate (1959:Q1-2007:Q4) in percentage



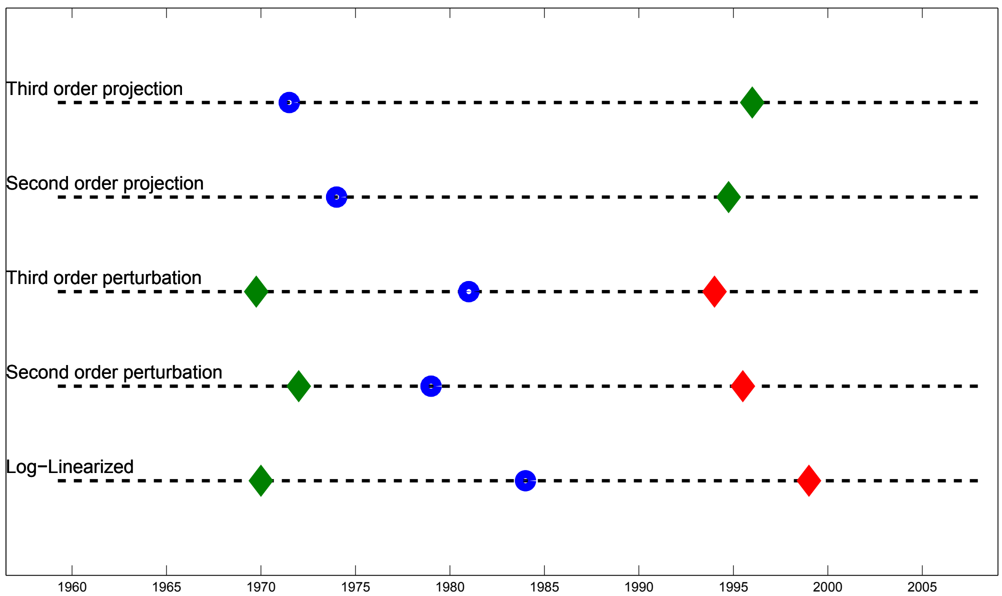
Note: Figure 1 shows the series for output growth, inflation and three month U.S. treasury bill rate for the period 1959 :Q1-2007:Q4. The data are extracted from FRED2 database maintained by the Federal Reserve Bank of St. Louis. The shaded areas represent the NBER recessions.

Figure 2.2: LR test sequences from three solution methods



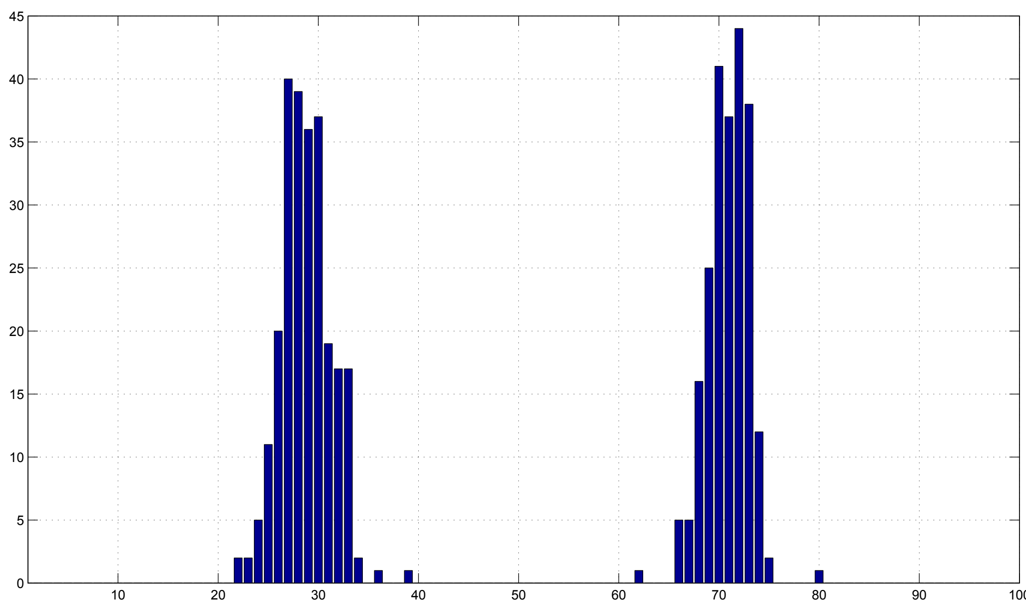
Note: Figure 2 plots the series of LR test for three different solution methods: log-linearization (blue solid line), second order perturbation (red dashed line), and second order projection (black dash-dot line). Here all 14 estimated parameters are allowed to change.

Figure 2.3: Multiple breaks using sequential procedure for five solutions



Note: Figure 3 visualizes the results shown in Table 3. Here the dashed lines represent the time horizon line. Blue spots represent the first structural breaks detected using sequential procedure, while diamonds represent the second break (green) or third break (red) found in each subsample.

Figure 2.4: Histograms of the estimated break points from Monte Carlo simulation



Note: Figure 4 displays Monte Carlo simulation using sequential procedure for the nonlinear state space model (D.1), which captures most features of nonlinear solutions to DSGE models. The sample size T is taken with 100. The true break points are at 30 and 70. The number of replications is 500.

Chapter 3

Is the Productivity Boom Over?

3.1 Introduction

Under the backdrop of the Great Recession, it appears to be extremely important to to identify changes of long-run trend in productivity growth, since it provides a useful guide on resource allocations for policy makers. The existing empirical literature has put focus on assessing dynamics of aggregate productivity growth, such as labor productivity and total factor productivity (TFP). In terms of structural breaks in productivity growth in the U.S., these results depend closely on the data source and sample scope.[1] Hence, researchers have proposed differing interpretations for the aggregate productivity slowdown in the 1970s and productivity rebound in the middle 1990s.

In this paper we approach the changes in long run productivity from a sectoral perspective. In particular, we decompose the whole economy into two broad sectors: investment goods-producing sector and consumption goods-producing sectors, and investigate structural breaks using sectoral productivity growth measures. To some

[1] For instance, while Benati (2007) failed to identify any structural break for labor productivity growth using data from Bureau of Labor Statistics (BLS) over 1947:Q1-2005:Q4, Fernald (2007) found two breaks, with a slowdown after 1973:Q1 and a speedup after 1997:Q2, for private business sector labor productivity growth over 1950:Q2-2004:Q2.

extent, our exercise plays the empirical role corresponding to those theoretical studies on investment-specific technology initiated by Greenwood et al. (1997). Figure 1 compares the evolutions of TFP in investment goods-producing sector and consumption goods-producing sector with the aggregate labor productivity.[2] Clearly, this figure highlights the fact that the TFP growth in investment goods-producing sector has outpaced TFP growth in consumption goods-producing sector, which Greenwood et al. (1997) interpret as the evidence for investment-specific or capital-emobided technological shocks. Also, it shows that labor productivity growth appears to slow down since the 1970s, and TFP growth in consumption goods-producing sector seems to be the principle source of the aggregate productivity slowdown. All these observations provide us the very first impression of sectoral analysis on productivity growth.

Furthermore, another goal of our exercise is to answer whether the era of productivity boom accompanied with "new-economy" since the middle 1990s is over. As shown in Figure 2, the labor productivity plays as a perfect illustration. Here we consider the mean of labor productivity in splited samples since 1948. As a vast of literature have documented, productivity growth slowed down since the early 1970s, and revived since the middle 1990s. The productivity resurgence, average 2.7 percent at annually, is attributed to information and communication technology. However, the productivity growth appear to decelerate to 1.7 percent. Through sectoral analysis on productivity growth, we could identify whether the productivity growth return back its conventional path.

Related to works as Hansen (2001), Fernald (2007) and Benati (2007), we contribute two new elements into the literature. First, it extends data time coverage including data in the current Great Recession. Second, in contrast to a vast of other studies focusing on the drifts and breaks in aggregate productivity growth, we extend our scope to sectoral productivity growth. Our results are able to shed some

[2]The two TFP measures are provided and updated by the Federal Reserve Bank of San Francisco, which are utilization adjusted follows ?.

light on the questions raised above. Although the evidence of structural break in the aggregate productivity growth is far from obvious at conventional test size, we find the evidence of structural breaks in the sectoral productivity growth. There are two structural breaks in investment goods-producing sector, which indicates that the investment boom accompanied with new economy in the middel 1990s has already ended. We also find there is one structural break in consumption goods-producing sector around the 1970s. Our results support the findings of Ireland and Schuh (2008) and Ireland (2011), in which they estimated a two-sector real business cycles model.

The paper is organized as follows. The next section first briefly describes econometric methodology of Bai and Perron (1998, 2003) and the data. In Section 3 we discuss the empirical results and implications. Section 4 concludes.

3.2 Econometric Methodology and Data Source

3.2.1 Econometric Methodology

We first consider the simplest dynamic model, the first order autoregressive model, for Δy_t with m breaks ($m+1$ regimes):

$$\begin{aligned} \Delta y_t &= \alpha_j + \rho_j \Delta y_{t-1} + \varepsilon_t, \; t = T_{j-1}+1, \ldots T_j \\ \varepsilon_t &\sim N(0, \sigma_j^2) \end{aligned}$$

for $j = 1, \ldots, m+1$, where Δy_t represents productivity growth measures in period t, and α_j and ρ_j are the corresponding coefficients of the first order atuoregressive model. The $m-$partition, $(T_1, \ldots, T_m)$, indicates the unknown break dates (Here we use the convention of $T_0 = 0$ and $T_{m+1} = T$). The dynamic properties of productivity growth would vary whenever any of the three parameters, α_j, ρ_j, and σ_j^2 changes. Since the key issue concerns the changes in long-run productivity growth, we focus on permanent shifts in the constant term α_j or the autoregressive parameter ρ_j.

We employ Bai and Perron (1998, 2003) methodology as our econometric approach, in which they use least squares method to estimate multiple unknown structural breaks. While we do not reproduce their global minimization algorithm for the break dates estimation, it is necessary to introduce the various test hypotheses and statistics. First, to identify the number of structural breaks (m), Bai and Perron (1998, 2003) begin with specifying a type of maximum $F-$statistic for testing the null hypothesis of no structural breaks against the alternative that there are $m = b$ breaks.

$$\sup F_T(b) = \sup F_T(\hat{\lambda}_1, \ldots, \hat{\lambda}_b)$$

where $\hat{\lambda}_i = [\frac{\hat{T}_i}{T}]$, $i = 1, \ldots, b$ are the break fractions and minimize the global sum of squared residuals. This statistic is a generalized version of the $\sup F$ test proposed by Andrews (1993). Bai and Perron (1998, 2003) then consider two double maximum statistics, for testing null hypothesis of no structural breaks against the alternative hypothesis of an unknown number of breaks given an upper bound, M. The first double maximum statistic is given by

$$UD\max(a_1, \ldots, a_M) = \max_{1 \leq m \leq M} a_M \sup F_T(m)$$

where $\{a_1, \ldots, a_M\}$ are some fixed weights and set equal to unity. The second maximum statistic, $WD\max$, applies different weights to the individual $\sup F_T(m)$ statistics so that the marginal $p-$values are equal across values of m. Finally, Bai and Perron (1998, 2003) consider a test of the null hypothesis of l breaks against the alternative hypothesis of $l+1$ breaks. The $\sup F_T(l+1|l)$ statistic is used to test whether the additional break leads a significant reduction in the sum of squared residuals.

In our application below, we follow the practical recommendations of Bai and Perron (2003), which have been shown to be adequate in an extensive simulation analysis in Bai and Perron (2004). We start with looking at $UD\max$ and/or $WD\max$ to determine if at least one structural break is present. If the double maximum statistics indicate the presence of structural breaks, i.e. being significant at the 10% level, we go on to decide the number of breaks by sequentially examining the $\sup F(l+1|l)$ test statistics, starting from the $\sup F(1|0)$. Finally, we set the maximum allowed number of structural breaks m equals to 4 and choose a trimming value, the minimal length of possible regimes, as 0.10.

3.2.2 The Data

The sample on the various productivity measures used in our application are from 1948:Q1 to 2012:Q1. As for aggregate productivity growth, we first consider labor productivity growth rate, ΔLP, measured by the quarterly growth rate of GDP per hour worked provided by Bureau of Labor Statistics. Also, we employ a real-time, quarterly series on total factor productivity (TFP) for the U.S. business sector from the Federal Reserve Bank of San Francisco.[3] The series also include TFP measures adjusted for variations in factor utilization labor effort and capital's workweek follows Basu et al. (2006) (thereafter BFK TFP measures). Therefore, we have two aggregate TFP measures: ΔTFP and ΔTFP^{util}. For our sectoral analysis, the first group of measures also obtain from the BFK TFP series. By using relative prices and input-output, BFK series are decomposed into separate measures of TFP for equipment investment good-producing sector (including consumer durables) , ΔTFP_I, and consumption goods-producing sector (defined as business output less equipment and consumer durables), ΔTFP_c. We also consider the utilization adjusted TFP measures ΔTFP_I^{util} and ΔTFP_c^{util}.[4] Since BFK series are based on the conventional growth accounting exercise, they correspond to several theoretical model based mea-

[3]Data source: http://www.frbsf.org/csip/tfp.php
[4]For more details, see Fernald (2012a).

sures developed by Greenwood et al. (1997), Marquis and Trehan (2008), and Ireland and Schuh (2008) among others. For comparison, we re-estimated the two sector real business model in Ireland and Schuh (2008) using extending sample period data, and constructed corresponding growth rates of TFP in the investment goods producing sector ΔZ_i and consumption goods producing ΔZ_c (thereafter IS TFP measures).

3.3 Results and Implications

3.3.1 Structural Break Hypothesis Tests

Table 2 presents Bai and Perron (1998, 2003) statistics results for tests of structural breaks of the nine productivity measures in our sample. For the aggregate productivity measures, we first find that both double maximum statistics are significant at conventional significant levels only for labor productivity. In contrast, only $WD\max$ statistics are significant at 10% level for BFK TFP and utilization adjusted BFK TFP, while $UD\max$ statistics are insignificant. Although the labor productivity growth has shown evidence of structural breaks, $\sup F(1|0)$ is still not significant at any convention levels. These results are likely partly due to potentially low power of the $\sup F$ statistics, and indicate once again that we need to analyze productivity growth from a sectoral perspective.

The remaining six rows in Table 2 report structural breaks results on sectoral productivity measures. We find that both double maximum statistics are significant at conventional significant levels for all these six sectoral productivity measures. In particular, we first find that utilization adjustment does not affect our test results for BFK TFP measures. The $\sup F(2|1)$ statistic is significant at the 10% level or higher, while $\sup F(3|2)$ is insignificant for TFP measures in investment goods-producing sector. It suggests two structural breaks (three regimes) for these two measures in investment goods-producing sector. The $\sup F(1|0)$ is significant at 10% level or higher for BFK TFP and utilization adjusted BFK TFP in consumption goods-producing sector, while the $\sup F(2|1)$ is insignificant. This indicates one structural break (two regimes) for these consumption goods-producing sector measures. The test results from IS TFP measures are quite different from those from BFK TFP measures. We only find that $\sup F(1|0)$ is significant at 5% for both investment goods-producing sector and consumption goods-producing sector, while $\sup F(2|1)$ is insignificant. However, we find that $\sup F(3|2)$ is significant at 10% for consumption sector, which may indicate there are three structural breaks in this measure. It is worth to note that the difference from BFK measure and IS measure is most likely due to that BFK TFP investment measures only consider the equipment investment and durable goods while IS TFP investment measure also include residential investment.

3.3.2 Structural Break Test Results

Table 3 reports the break dates and their 90% confidence intervals for each of the nine productivity measures, in addition to regression coefficients and the mean growth rate for each regime. We observe that the three aggregate productivity measures have structural breaks occurring between the late 1960s to the early 1970s. The finding is line with widely accepted productivity slowdown in the 1970s in the literature. For instance, recent works as Benati (2007) and Fernald (2007) find one structural break for labor productivity growth in 1973. However, we could not identify the second structural break around the middle 1990s in aggregate productivity using the extended sample. This might imply that the acceleration in aggregate productivity growth during the late 1990s would be transitory.

For BFK TFP in investment goods-producing sector, the first break is around the middle 1990s and the second break is in the first half year of 2005. We observe a inverted U-shaped evolution of mean productivity growth in investment sector, first accelerating from 2% to 5% then falling back to almost 3%. These break dates are in line with dates found by other researchers using different methods and productivity measures. Our finding of the first break date is similar to Jorgenson (2001) and Oliner and Sichel (2000)'s finding of a growth resurgence in the U.S. beginning in 1995, which they link to information technology in general. However, our results show no evidence of a structural break in investment-goods producing sector in the 1970s. Unlike Marquis and Trehan (2008) found the capital sector productivity growth rate actually slowed down over this period and Greenwood et al. (1997) and Greenwood and Yorukoglu (1997) concluded that capital-specific productivity accelerated in the early 1970s at about the same time that aggregate productivity slowed down. In addition, we found a second break around 2005 using the expanded sample including the recent recession. This finding is also consistent with many observations about the investment boom. For TFP in consumption-goods producing sector, there is evidence for one break in our sample period for both utilization adjusted and unadjusted series, which are located in the late 1960s. Although our results suggest the productivity slowdown might be earlier than we thought, the associated interval stretches from the early 1960s to the early 1970s, which is roughly in line with other studies.

The last two rows in Table 3 present the results for the structural breaks for IS TFP measures. We only find only one structural break in 2005 in investment goods-producing sector sector. The productivity growth in investment sector slows down since the middle 2000s. The timing is close to the second break we found using BFK TFP measures. In addition, we find two more structural breaks in consumption sector. One is in the early 1980s and another is in early this 2000s. Even the difference between these two measures, the results from two group of measures, broadly speaking, leads to two similiar findings. The first is that the main contribution to the productivity growth slowdown is consumption goods-producing sector. The other one is that there appears to exist deceleration in investment goods-producing sector during the middle 2000s.

3.3.3 Implications

Even though our approach to testing for trend breaks in present paper is strictly statistical, we can still draw several key implications. First of all, broadly consistent with the results derived by Basu et al. (2006), Ireland and Schuh (2008) , and Marquis and Trehan (2008), our results show, once again, attribute most of the productivity slowdown of the 1970s to the consumption goods-producing sector. Here, in particular, the BFK TFP growth in the consumption goods-producing sector remains essentially unchanged since the late 1960s, only at 0.28 percent annual rate. Whereas Basu et al. (2006) estimates suggest that productivity slowdown occurred contemporaneously across both two sectors in the 1970s, here we could not find a break in the 1970s, and we attribute the more recent productivity revival that accompanied the long economic expansion of 1990s in the US rapid investment-specific technological change. Overall, our empirical results echoed with the observations in Ireland (2011) obtained from an estimated two sector real business cycles model.

Second, our results provide answer to the question raised in the introduction: Is the era of rapid productivity growth over? Most likely, Yes. Our results show the more recent episode of robust investment and investment-specific technological change during the 1990s largely as they fall back to their unexceptional longer-run averages, about at 2.9 percent annually. As also documented by Fernald (2012b), the slowdown preceded the current recession and is consistent with an apparent reducing in intangible organizational investment associated with information and communications technology. He argues that ICT has had a broad-based and pervasive effect on measured TFP through its role as a general purpose technology that fostered complementary innovations, including business reorganization.

Thirdly, our results, therefore offer up a pessimistic view of the future. The not-so-good news is that the results show the more recent episode of robust growth in investment and investment-specific productivity as largely representing a catch-up in levels after the previous productivity slowdown—hence, the results predict that this recent episode of unusual strength is unlikely to persist or to be repeated anytime soon. Thus, we would expect the potential output will be lower than our conventional projections.

Finally, our results highlight the importance of sectoral analysis on productivity. As these tables have shown, the statistic results on aggregate productivity measures are far from clear to identify structural break. Benati (2007) accounted for this "puzzling" results partly is that the change in labor productivity growth may have simple been too gradual to be detected via a crude structural break tests. Our sectoral analysis supports, in some sense, this explanation since the effects from these two sectors are mixed up when we consider aggregate productivity, which reinforce our motivation on sectoral analysis.

3.4 Conclusion

In this paper, we apply Bai and Perron (1998, 2003) methodology to examine structural breaks in productivity growth in U.S. over the postwar era from the perspective of sectoral analysis. We decompose the economy into two broad sectors: investment goods-producing sector and consumption goods-producing sector. Although the evidence of structural break in the aggregate productivity growth is far from obvious at conventional test sizes, we found the evidence of structural breaks in the sectoral productivity growth. Our results are closely consistent with Ireland and Schuh (2008) and Ireland (2011), in which they study two different technology shocks: investment goods-producing technology and consumption goods-producing technology. Our structural breaks results echo with their implication of the consumption goods-producing sector as the principal source of the productivity slowdown of the 1970s. And we also show evidence of a productivity slowdown in the investment goods-producing sector in the middle 2000s, which we conclude that the era of productivity growth due to new economy is over. Viewed against this broader backdrop, we confirms Ireland and Schuh (2008)'s projection that the accelerated growth in investment and investment specific technological change appears largely as a snap-back in levels to a long-run deterministic trend rather than a persistent shift in growth rates. Therefore, our results offer up a pessimistic outlook for the future. The productivity slowdown of the 1970s has not ended. It also suggests the future productivity growth rates in investment sectors that will match their healthy but unexceptional longer-run averages before the latest resurgence. In turn. our results suggest the potential output growth will be likely to be lower than the conventional projections, and we will see a longer and slower recovery from the current recession.

3.5 Chapter 3: Tables and Figures

Table 3.1: Descriptive Statistics (1948:Q1-2012:Q1)

	Variables	Description	Mean	S.D.	Median	Min	Max	AC(1)	AC(2)
Aggregate	ΔLP	Business Sector Labor Productivity	2.41	3.53	2.06	−7.94	17.74	0.02	0.13
	ΔTFP	Business Sector TFP	1.30	3.82	1.12	−9.31	16.80	0.11	0.17
	$\Delta\text{TFP}^{\text{Util}}$	Utilization-adjusted TFP	1.32	3.56	1.32	−9.89	15.99	0.00	0.15
Investment Goods-producing Sector	ΔTFP_I	TFP in Equipment and Consumer Durables	2.91	4.38	3.09	−10.09	15.63	0.19	0.17
	$\Delta\text{TFP}_I^{\text{Util}}$	Utilization-adjusted TFP in Producing Equipment and Consumer Durables	2.91	4.19	1.94	−10.02	16.51	0.13	0.17
	$\Delta\hat{Z}_{it}$	Ireland-Schuh TFP in Investment-goods Producing	2.18	23.86	1.90	−22.42	63.14	0.08	0.13
Consumption Goods-producing Sector	ΔTFP_C	TFP in non-equipment Business Output	0.86	3.85	0.65	−9.22	17.07	0.12	0.19
	$\Delta\text{TFP}_C^{\text{Util}}$	Utilization-adjusted TFP in Producing Non-equipment Output	0.89	3.63	0.81	−9.84	15.89	0.03	0.18
	$\Delta\hat{Z}_{ct}$	Ireland-Schuh TFP in Consumption-goods Producing	1.80	4.62	1.92	−10.74	22.14	0.04	0.15

Notes: Columns (3) and (4) report the mean and standard deviation of the variable in column (1). The first and second order autocorrelations are reported in columns (8) and (9). Data Source: Bureau of Labor Statistics, Federal Reserve Bank of San Francisco, and author's calculations.

Table 3.2: Bai and Perron (1998, 2003) Statistics for Tests of Multiple Structural Breaks in AR(1) Model (1948:Q1-2012:Q1)

	Variables	$UDmax^a$	$WDmax^b$	$\sup F_T(1\|0)^c$	$\sup F_T(2\|1)^d$	$\sup F_T(3\|2)^e$	$\sup F_T(4\|3)^f$
Aggregate	ΔLP	12.03*	15.87**	9.54	4.18	4.68	4.00
	ΔTFP	10.50	12.82*	6.60	2.91	7.96	2.18
	$\Delta \text{TFP}^{\text{Util}}$	10.18	12.50*	10.18*	6.01	6.01	5.01
Investment	ΔTFP_I	23.99***	30.11***	11.59**	13.76**	2.74	3.76
Goods-	$\Delta \text{TFP}_I^{\text{Util}}$	12.78**	14.80**	12.62**	13.09*	5.23	4.48
producing Sector	$\Delta \hat{Z}_{it}$	14.14**	14.14**	14.14**	1.35	5.12	4.22
Consumption	ΔTFP_C	11.15*	14.70**	11.08**	2.71	1.80	2.65
Goods-	$\Delta \text{TFP}_C^{\text{Util}}$	17.48***	17.48**	17.48***	7.56	7.56	7.83
producing Sector	$\Delta \hat{Z}_{ct}$	18.16***	18.16***	18.16***	4.97	12.29*	1.79

Notes: *, **, *** indicate significance at the 10%, 5% and 1% levels, respectively according to the critical values in Bai and Perron (2003); a- One-sided (upper-tail) test of the null hypothesis of 0 breaks against the alternative hypothesis of an unknown number of breaks given an upper bound of 4. b- One-sided (upper-tail) test of the null hypothesis of 0 breaks against the alternative hypothesis of an unknown number of breaks given an upper bound of 4. c- One-sided (upper-tail) test of the null hypothesis of 0 breaks against the alternative hypothesis of 1 break. d- One-sided (upper-tail) test of the null hypothesis of 1 break against the alternative hypothesis of 2 breaks. e- One-sided (upper-tail) test of the null hypothesis of 2 breaks against the alternative hypothesis of 3 breaks. f- One-sided (upper-tail) test of the null hypothesis of 3 breaks against the alternative hypothesis of 4 breaks.

Table 3.3: Bai and Perron(1998, 2003) Multiple Regime AR(1) Model Estimation Results (1948:Q1-2012:Q1)

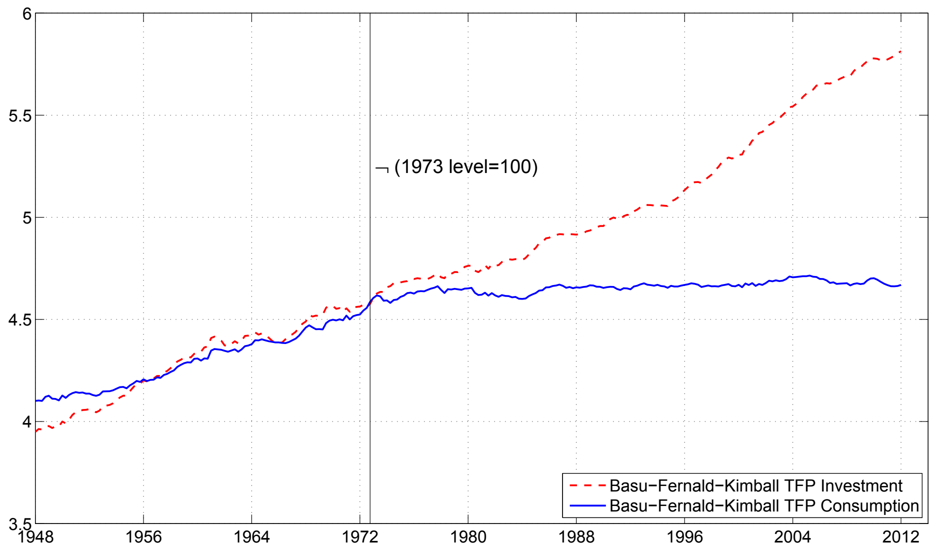
		Regime 1				Regime 2				Regime 3				Regime 4		
	Variables	$\hat{\alpha}$	$\hat{\rho}$	mean	Break Date	$\hat{\alpha}$	$\hat{\rho}$	mean	Break Date	$\hat{\alpha}$	$\hat{\rho}$	mean	Break Date	$\hat{\alpha}$	$\hat{\rho}$	mean
	ΔLP	3.49	-0.09	3.23	1973:Q2	1.73	0.07	1.89								
		(0.44)	(0.09)		[1967:Q2-1989:Q2]	(0.33)	(0.09)									
Aggregate	ΔTFP	2.03	0.05	2.14	1973:Q2	0.64	0.10	0.76								
		(0.42)	(0.09)		[1962:Q3-1994:Q1]	(0.31)	(0.09)									
	$\Delta\text{TFP}^{\text{Util}}$	2.28	0.02	2.92	1968:Q3	0.89	-0.01	0.86								
		(0.45)	(0.01)		[1961:Q2-1979:Q3]	(0.27)	(0.08)									
	ΔTFP_I	1.95	0.16	2.83	1995:Q3	7.32	-0.38	5.21	2005:Q2	1.42	0.48	2.90				
Investment		(0.34)	(0.07)		[1994:Q2-1997:Q3]	(1.31)	(0.22)		[2004:Q1-2006:Q4]	(1.10)	(0.24)					
Goods-	$\Delta\text{TFP}_I^{\text{Util}}$	2.10	0.09	2.38	1994:Q1	6.07	-0.14	5.17	2005:Q1	2.18	0.18	2.90				
producing		(0.34)	(0.07)		[1992:Q2-1996:Q3]	(1.13)	(0.18)		[2003:Q2-2007:Q3]	(1.11)	(0.27)					
Sector	$\Delta\hat{Z}_{it}$	2.54	-0.01	2.52	2005:Q3	-0.81	0.65	-0.38								
		(0.44)	(0.06)		[2002:Q1-2009:Q2]	(1.39)	(0.23)									
	ΔTFP_C	2.16	0.08	2.36	1966:Q3,	0.23	0.05	0.28								
Consumption		(0.50	(0.10)		[1961:Q2-1974:Q1]	(0.28)	(0.08)									
Goods-	$\Delta\text{TFP}_C^{\text{Util}}$	2.22	0.02	2.34	1968:Q3,	0.26	-0.14	0.28								
producing		(0.44)	(0.02)		[1965:Q1-1974:Q2]	(0.27)	(0.09)									
Sector	$\Delta\hat{Z}_{ct}$	3.14	-0.24	2.58	1971:Q4	-0.06	0.41	0.09	1981:Q2	2.29	-0.03	2.14	2001:Q3	0.32	0.44	0.88
		(0.12)	(0.08)		[1969:Q4-1972:Q4]	(0.01)	(0.14)		[1976:Q3-1985:Q1]	(0.09)	(0.13)		[1998:Q4-2008:Q2]	(0.01)	(0.26)	

Notes: $\hat{\alpha}$ in column (2) , (6), (10) and (14), and $\hat{\rho}$ in columns (3), (7), (11) and (15) are the least squares estimates of the following first-order autoregression model:

$$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$$

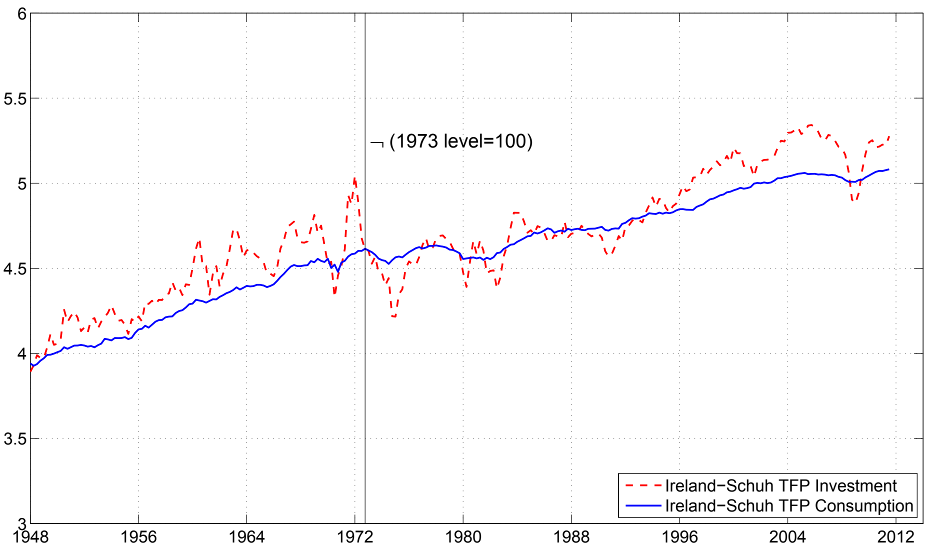
where y_t are the 9 producitivity measures listed in the column (1), whose descriptions are the same as those in Table 1; standard errors are given in the parentheses. The breakdatesare reported in columns (5), (9), and (13); and their confidence interval are included in brackets.

Figure 3.1: Logs of Sectoral Productivity Level in the U.S. (Basu-Fernald-Kimball TFP Measurements)



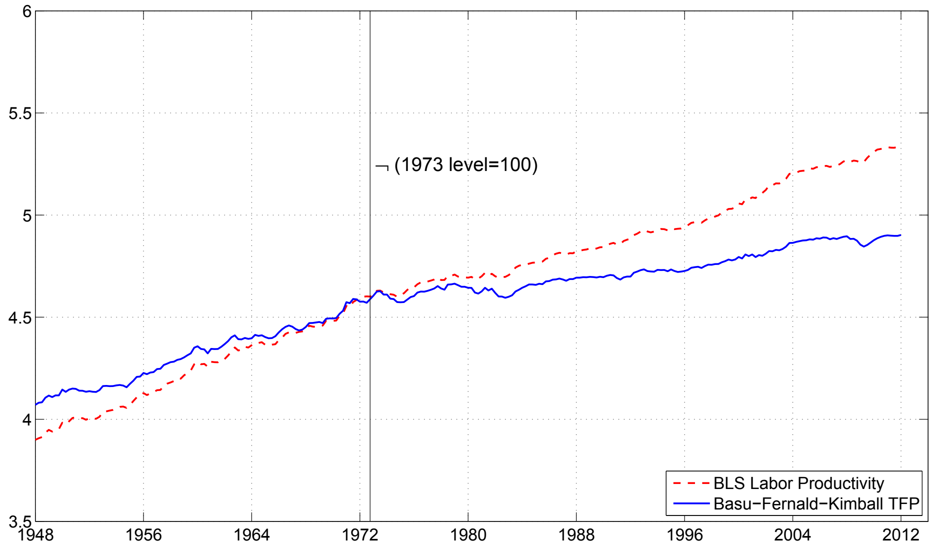
Source: Federal Reserve Bank of San Francisco, Authors' calculation.

Figure 3.2: Logs of Sectoral Productivity Level in the U.S. (Ireland-Schuh TFP Measurements)



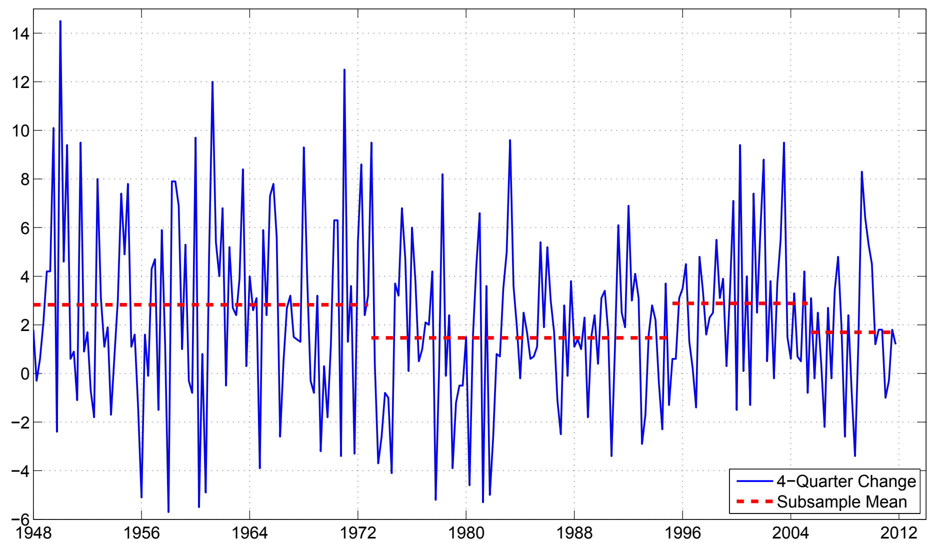
Source: Ireland and Schuh (2008), Authors' calculation.

Figure 3.3: Logs of Aggreagte Productivity Level in the U.S.



Source: Bureau of Labor Statistics, Federal Reserve Bank of San Francisco, Authors' calculation.

Figure 3.4: Labor Productivity Growth in the U.S. Business Sector



Source: Bureau of Labor Statistics, Authors' calculation.

Bibliography

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews*, 26(2-4):113–172.

Andreasen, M. (2011). How non-gaussian shocks affect risk premia in non-linear dsge models. Bank of England Working Paper, No. 47.

Andrews, D. W. K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica*, 61(4):821–856.

Andrews, D. W. K. and Fair, R. C. (1988). Inference in Nonlinear Econometric Models with Structural Change. *Review of Economic Studies*, 55(4):615–639.

Andrews, D. W. K. and Ploberger, W. (1994). Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative. *Econometrica*, 62(6):1383–1414.

Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking. *IEEE Transactions on Signal Processing*, 50(2):173–188.

Aruoba, S. B., Fernández-Villaverde, J., and Rubio-Ramírez, J. F. (2006). Comparing Solution Methods for Dynamic Equilibrium Economies. *Journal of Economic Dynamics and Control*, 30:2477–2508.

Bai, J. (1997). Estimating Multiple Breaks One at a Time. *Econometric Theory*, 13(3):315–352.

Bai, J., Lumsdaine, R. L., and Stock, J. H. (1998). Testing for and Dating Common Breaks in Multivariate Time Series. *Review of Economic Studies*, 65:395–432.

Bai, J. and Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, 66(1):47–78.

Bai, J. and Perron, P. (2003). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, 18(1):1–22.

Bai, J. and Perron, P. (2004). *Multiple Structural Change Models: A Simulation Analysis*, pages 212–238. Cambridge University Press.

Banerjee, A., Lumsdaine, R. L., and Stock, J. H. (1992). Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence. *Journal of Business and Economic Statistics*, 10(3):271–287.

Basu, S. and Bundick, B. (2012). Uncertainy Shocks in a Model of Effective Demand. Manuscript.

Basu, S., Fernald, J. G., and Kimball, M. S. (2006). Are Technology Improvements Contractionary? *American Economic Review*, 96(5):1418–1448.

Benati, L. (2007). Drift and Breaks in Labor Productivity. *Journal of Economic Dynamics and Control*, 31(8):2847–2877.

Blanchard, O. J. and Kahn, C. M. (1980). The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, 48(5):1305–1311.

Boivin, J. and Giannoni, M. P. (2006). Has Monetary Policy Become More Effective? *Review of Economics and Statistics*, 88(3):445–462.

Canova, F. and Sala, L. (2009). Back to Square One: Identification Issues in DSGE Models. *Journal of Monetary Economics*, 56(4):431–449.

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, 28(3):591–605.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, 113(1):1–45.

Clarida, R., Galí, J., and Gertler, M. (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *The Quarterly Journal of Economics*, 115(1):147–180.

Cogley, T. and Sargent, T. J. (2005). Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics*, 8:262–302.

Doucet, A., de Freitas, N., and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice*. Springer.

Eggertsson, G. B. and Woodford, M. (2003). The Zero Bound on Interest Rates and Optimal Monetary Policy. *Brookings Papers on Economic Activity*, 34(1):139–235.

Estrella, A. and Fuhrer, J. C. (2003). Monetary Policy Shifts and the Stability of Monetary Policy Models. *Review of Economics and Statistics*, 85(1):94–104.

Fernald, J. (2007). Trend Breaks, Long-run Restrictions, and Contractionary Technology Improvements. *Journal of Monetary Economics*, 54(8):2467–2485.

Fernald, J. (2012a). A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Manuscript.

Fernald, J. (2012b). Productivity and Potential Output before, during, and after the Great Recession. Manuscript.

Fernandez-Villaverde, J. and Rubio-Ramírez, J. F. (2005). Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood. *Journal of Applied Econometrics*, 20:891–910.

Fernandez-Villaverde, J. and Rubio-Ramírez, J. F. (2007). Estimating Macroeconomic Models: A Likelihood Approach. *Review of Economic Studies*, 74(4):1059–1087.

Fernandez-Villaverde, J. and Rubio-Ramírez, J. F. (2008). How Structural are Structural Parameters? *NBER Macroeconomics Annual*, pages 83–137.

Fernandez-Villaverde, J., Rubio-Ramírez, J. F., and Santos, M. S. (2006). Convergence Properties of the Likelihood of Computed Dynamic Models. *Econometrica*, 74(1):93–119.

Gaspar, J. and Judd, K. L. (1997). Solving Large-Scale Rational-Expectations Models. *Macroeconomic Dynamics*, 1(1):45–75.

Gomme, P. and Klein, P. (2010). Second-order Approximation of Dynamic Models without the Use of Tensors. Manuscript.

Greenwood, J., Hercowitz, Z., and Krusell, P. (1997). Long-run Implications of Investment Specific Technological Change. *American Economic Review*, 87(3):342–362.

Greenwood, J. and Yorukoglu, M. (1997). 1974. *Carnegie-Rochester Conference Series on Public Policy*, 46(1):49–95.

Hájek, J. and Rényi, A. (1955). Generalization of an Inequality of Kolmogorov. *Acta Mathematica Academiae Scientiarum Hungarica*, 6:281–283.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hansen, B. E. (2001). The New Econometrics of Structural Change: Dating Breaks in U.S. Labour Productivity. *Journal of Economic Perspectives*, 15(4):117–128.

Hansen, P. R. (2003). Structural Changes in the Cointegrated Vector Autoregressive Model. *Journal of Econometrics*, 114:261–295.

Hawkins, D. M. (1976). Point Estimation of the Parameters of Piecewise Regression Models. *Applied Statistics*, 25:51–57.

Inoue, A. and Rossi, B. (2011). Identifying the Sources of Instabilities in Macroeconomic Fluctuations. *Review of Economics and Statistics*, 93(4):1186–1204.

Ireland, P. N. (2001). Sticky-price Models of the Business Cycle: Specification and Stability. *Journal of Monetary Economics*, 47:3–18.

Ireland, P. N. (2004). Technology Shocks in the New Keynesian Model. *Review of Economics and Statistics*, 86(4):923–936.

Ireland, P. N. (2007). Changes in the Federal Reserve's Inflation Target: Causes and Consequences. *Journal of Money, Credit and Banking*, 39(8):1851–1882.

Ireland, P. N. (2011). Stochastic Growth in the United States and Euro Area. Manuscript.

Ireland, P. N. and Schuh, S. (2008). Productivity and U.S. Macroeconomic Performance: Interpreting the Past and Predicting the Future with a Two-Sector Real Business Cycle Model. *Review of Economic Dynamics*, 11(3):473–492.

Jorgenson, D. W. (2001). Information Technology and the U.S. Economy. *American Economic Review*, 91(1):1–32.

Judd, K. L. (1992). Projection Methods for Solving Aggregate Growth Models. *Journal of Economic Theory*, 58(2):410–452.

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT Press.

Judd, K. L., Maliar, L., and Maliar, S. (2010). A Cluster-Grid Projection Method: Solving Problems with High Dimensionality. *NBER Working Paper*, No. 15965.

Justiniano, A. and Primiceri, G. E. (2008). The Time-Varying Volatility of Macroeconomic Fluctuations. *American Economic Review*, 98(3):604–641.

Kim, J., Kim, S., Schaumburg, E., and Sims, C. (2005). Calculating and Using Second-order Accurate Solutions of Discrete Time Dynamic Equilibrium Models. Manuscript.

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.

Liu, Z., Waggoner, D. F., and Zha, T. (2011). Sources of Macroeconomic Fluctuations: A Regime-Switching DSGE Approach. *Quantitative Economics*, 2(2):251–301.

Lucas, R. E. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy*, 1:19–46.

Malin, B. A., Krueger, D., and Kubler, F. (2011). Solving the Multi-country Real Business Cycle Model Using a Smolyak-collocation Method. *Journal of Economic Dynamics and Control*, 35:229–239.

Marquis, M. H. and Trehan, B. (2008). On Using Relative Prices to Measure Capital-specific Technological Progress. *Journal of Macroeconomics*, 30(4):1390–1406.

Moreno, A. (2004). Reaching Inflation Stability. *Journal of Money, Credit and Banking*, 36(4):801–825.

Oliner, S. D. and Sichel, D. E. (2000). The Resurgence of Growth in the late 1990s: Is information technology the story? *Journal of Economic Perspectives*, 14(4):3–22.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica*, 57(6):1361–1401.

Picard, D. (1985). Testing and Estimating Change-Points in Time Series. *Advance in Applied Probability*, 176:841–867.

Pichler, P. (2011). Solving the Multi-country Real Business Cycle Model Using a Monomial Rule Galerkin Method. *Journal of Economic Dynamics and Control*, 35(2):240–251.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *Review of Economic Studies*, 72(3):821–852.

Qu, Z. and Perron, P. (2007). Estimating and Testing Structural Changes in Multivariate Regressions. *Econometrica*, 75(2):459–502.

Quandt, R. E. (1960). Tests of the Hypothesis that a Linear Regression Obeys Two Separate Regimes. *Journal of the American Statistical Association*, 55:324–330.

Rotemberg, J. J. (1982). Sticky Prices in the United States. *Journal of Political Economy*, 90(6):1187–1211.

Schmitt-Grohé, S. and Uribe, M. (2004). Solving Dynamic General Equilibrium Models Using a Second-order Approximation to the Policy Function. *Journal of Economic Dynamics and Control*, 28:755–775.

Sims, C. (2002). Solving Linear Rational Expectations Models. *Computational Economics*, 20(1-2):1–20.

Sims, C. A. and Zha, T. (2006). Were There Regime Switches in U.S. Monetary Policy? *American Economic Review*, 96(1):54–81.

Smets, F. and Wouters, R. (2003). An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *Journal of the European Economic Association*, 1(5):1123–1175.

Smets, F. and Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 97(3):586–606.

Stock, J. H. and Watson, M. W. (2002). Has the Business Cycle Changed and Why? *NBER Macroeconomics Annual*.

Taylor, J. B. (1993). Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy*, 1:195–214.

Uhlig, H. (1997). Bayesian Vector Autoregressions with Stochastic Volatility. *Econometrica*, 65:59–73.

Uhlig, H. (1999). *A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily*. Oxford University Press.

van Binsbergen, J., Fernandez-Villaverde, J., Koijen, R. S., and Rubio-Ramírez, J. F. (2012). The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences. *Journal of Monetary Economics*, page forthcoming.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Zhu, C. (2012). Monetary Policy Shifts and the Stability of Monetary Policy Models: An International Comparison. Manuscript.