Perturbation Methods for Markov-Switching Models*

INCOMPLETE - DO NOT CIRCULATE

Andrew Foerster† Juan Rubio-Ramirez‡ Dan Waggoner§ Tao Zha¶

September 28, 2012

Abstract

Markov Switching models are a way to consider discrete changes in the economic environment, such as policy changes, and allow agents in the economy to form expectations over these changes. This paper develops a methodology for constructing approximations to the solution of Markov Switching dynamic stochastic general equilibrium (MS-DSGE) models. The method allows for changes in parameters that both do and do not affect the economy's steady state, and enables linear or higher-order approximations. In addition, the paper proves that first-order approximations to a wide class of MS-DSGE models are not certainty equivalent. The numerical procedure handles potentially large systems and considers existence and uniqueness using the concept of mean square stability. Two examples, one Real Business Cycle and one New Keynesian, illustrate the procedure and issues of certainty equivalence and mean square stability.

*The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Banks of Atlanta or Kansas City or the Federal Reserve System. We thank Seonghoon Cho, Rhys Bidder, seminar participants at Duke, the Federal Reserve Bank of St. Louis, the 2010 Society of Economic Dynamics meetings, the 2011 Federal Reserve System Committee on Business and Financial Analysis Conference, the 2012 Annual Meeting of the American Economic Association, and the 8th Dynare Conference.

†Federal Reserve Bank of Kansas City

‡Duke University, Federal Reserve Bank of Atlanta, CEPR, FEDEA, and BBVA Research

§Federal Reserve Bank of Atlanta

¶Emory University and Federal Reserve Bank of Atlanta

1 Introduction

Following the introduction of vector autoregressions (VARs) to macroeconomics by Sims (1980) it was quickly realized that it is difficult to find macroeconomic applications for which model parameters remain stable over long periods of time. This problem was not unique to reduced form representations of the data, but was also an issue when more structural approaches were considered. One way to solve the problem, pursued by Clarida *et al.* (2000) and followed up by Lubik & Schorfheide (2004), breaks the sample into sub-periods and estimates the structural models in which one or more of the model's parameters differ across sub-samples. While this approach addresses the parameter instability problem, it fails to consider that forward looking agents living in a world in which parameters are known to change occasionally would be expected to take possible parameter change into account when forming their expectations and, therefore, will affect their optimal decisions.

An alternative approach to parameter instability, suggested by the work of Hamilton (1989) and pursued in Sims & Zha (2006), is to estimate a backward-looking vector autoregression (VAR) with regime dependent parameters. This approach has its limitations since it does not allow for the presence of forward-looking components that are present in a dynamic stochastic general equilibrium (DSGE) model.

A number of authors have recently studied forward looking Markov-switching linear rational expectations (MSLRE) models. Work in this area includes papers by Leeper & Zha (2003), Svensson & Williams (2007), Blake & Zampolli (2006), Davig & Leeper (2007), and Farmer *et al.* (2009). MSLRE models are more complicated than linear rational expectations models since the agents of the model must be allowed to take account of the possibility of future regime changes when forming expectations. The MSLRE literature has made some headway in addressing questions like setting necessary and sufficient conditions to determine if the parameters of a Markov-switching rational expectations model lead to a determinate equilibrium (See Farmer *et al.* (2009)).

There are two main shortcomings with the MSLRE approach. First, most of the analyzed models do not begin from first principles. In other words, researchers consider linear rational

expectations (LRE) models where Markov-switching (MS) has been added after the model has been linearized. Second, higher order solutions are not considered. Given that MS parameters add a lot of uncertainty to the model, considering higher order approximations may be potentially important. This paper solves these two shortcomings. In particular, it shows how to use perturbation methods to solve Markov-switching rational expectations (MSRE) models - note the absence of the "linear" - starting from first principles, i.e. from the set of (non-linearized) first order conditions that define equilibrium.

Following Costa *et al.* (2005), Farmer *et al.* (2009), and Farmer *et al.* (2008), this paper uses the concept of mean square stability (MSS) to characterize stable solutions. The perturbation approach uses the theory of Gröbner Bases to find solutions, and determines existence and uniqueness of MSS solutions. It also allows for a flexible regime-switching specification, including in parameters that affect the steady state of the economy. In particular, the first order approximation of models where switching affects the steady state is not certainty equivalent.

After developing the methodology, the paper presents two example economies that illustrate the methodology and highlight the issues of mean square stability and certainty equivalence. In the first, a simple real business cycle model with stochastic drift shows how to use the methodology and the importance of certainty equivalence. The second, a New Keynesian model, adds sticky prices and a monetary authority with changes in the policy rule, and shows how mean square stability determines existence and uniqueness.

The remainder of the paper is as follows: Section 2 describes a general class of MS-DSGE models and the nature of Markov switching. Sections 3 and 4 discuss the first-order approximation, the former showing how to solve the model, and the latter highlighting the key quadratic equations and how to use Gröbner Bases to solve them. Section 6.1 has an example RBC economy, Section 6.2 has an example NK economy, and Section 7 concludes.

2 The Model

Consider a dynamic general equilibrium model in which some of the parameters follow a discrete state Markov chain indexed by s_t with transition matrix $P = (p_{s,s'})$. The element $p_{s,s'}$ represents the probability that $s_{t+1} = s'$ given $s_t = s$ for $s, s' \in \{1, \ldots, n_s\}$ where n_s is the number of regimes and when $s_t = s$ the model is said to be in regime s at time t. The vector of changing parameters θ_t has size $n_\theta \times 1$.[1] Given any $x_{t-1}, \varepsilon_t,$ and θ_t, the set of equilibrium conditions of a wide class of models can be written as

$$\mathbb{E}_t f\left(y_{t+1}, y_t, x_t, x_{t-1}, \chi\varepsilon_{t+1}, \varepsilon_t, \theta_{t+1}, \theta_t\right) = 0_{n_x+n_y} \tag{1}$$

where $\mathbb{E}_t$ denotes the mathematical expectations operator conditional on information available at time t, and 0_{m_1} denotes a $m_1 \times 1$ vector of zeros. The vector x_{t-1} of predetermined variables (endogenous and exogenous) is of size $n_x \times 1$, the vector y_t of non-predetermined variables is of size $n_y \times 1$, the vector ε_t of independent innovations to the exogenous predetermined variables with mean equal to zero is of size $n_\varepsilon \times 1$, and $\chi \in \mathbb{R}$ is the perturbation parameter. The function f maps $\mathbb{R}^{2(n_y+n_x+n_\varepsilon+n_\theta)}$ into $\mathbb{R}^{n_y+n_x}$ is the number of equations in (1). Since the parameters, θ_t, in (1) depend on the state of the Markov chain, there are n_s sets of equilibrium conditions, one for each value of the Markov chain, instead of the single set of equilibrium conditions in the constant parameter case.

The solution to the model has the form

$$y_t \equiv g\left(x_{t-1}, \varepsilon_t, \chi, s_t\right), \tag{2}$$

$$y_{t+1} \equiv g\left(x_t, \chi\varepsilon_{t+1}, \chi, s_{t+1}\right), \tag{3}$$

and

$$x_t \equiv h\left(x_{t-1}, \varepsilon_t, \chi, s_t\right) \tag{4}$$

where g maps $\mathbb{R}^{n_x+n_\varepsilon+1} \times \{1, \ldots, n_s\}$ into $\mathbb{R}^{n_y}$ and h maps $\mathbb{R}^{n_x+n_\varepsilon+1} \times \{1, \ldots, n_s\}$ into $\mathbb{R}^{n_x}$. The goal is to find the Taylor expansion of the functions g and h around the steady state.

[1] There may also be a set of non-changing parameters not included in θ_t.

The parameters θ_t depend on the regime in the following way

$$\theta_t \equiv \theta(\chi, s_t) \text{ and } \theta_{t+1} \equiv \theta(\chi, s_{t+1}) \tag{5}$$

where θ maps $\mathbb{R} \times \{1, \ldots, n_s\}$ into $\mathbb{R}^{n_\theta}$. The vector of parameters θ_t has two subvectors θ_{1t} and θ_{2t}

$$\theta_t = \begin{pmatrix} \theta_{1t}' & \theta_{2t}' \end{pmatrix}' \equiv \begin{pmatrix} \theta_1(\chi, s_t)' & \theta_2(\chi, s_t)' \end{pmatrix}', \tag{6}$$

where θ_{1t} and θ_{2t} have sizes $n_{\theta_1} \times 1$ and $n_{\theta_2} \times 1$ respectively, and

$$\theta_1(\chi, s_t) = \overline{\theta}_1 + \chi \widehat{\theta}_1(s_t) \tag{7}$$

and

$$\theta_2(\chi, s_t) = \widehat{\theta}_2(s_t). \tag{8}$$

The parameters θ_{t+1} have the same functional forms.[2] Note two things about this specification: first, $\widehat{\theta}_1(s_t)$ is the deviation of θ_{1t} from $\overline{\theta}_1$ in regime s_t and, second, θ_{2t} is not a function of the perturbation parameter χ. Hence, the perturbation parameter, χ, only affects a subset of the parameters, θ_{1t}, while θ_{2t} is not affected by the perturbation parameter. The choice of which parameters to perturb, θ_{1t}, and which ones do not perturb, θ_{2t}, is not unique, but there is one restriction. Define the steady state of the model as vectors x_{ss} and y_{ss} of size $n_x \times 1$ and $n_y \times 1$ respectively such that

$$f\left(y_{ss}, y_{ss}, x_{ss}, x_{ss}, 0_{n_\varepsilon}, 0_{n_\varepsilon}, \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2(s_{t+1})' \end{pmatrix}', \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2(s_t)' \end{pmatrix}'\right) = 0_{n_x+n_y}$$

for all s_{t+1} and s_t. Thus, the partition should be such that neither $\theta_2(0, s_{t+1}) = \widehat{\theta}_2(s_{t+1})$ nor $\theta_2(0, s_t) = \widehat{\theta}_2(s_t)$ enter in the calculation of the steady state since the last expression has to hold for all s_{t+1} and s_t.

In other words, the partition should be such that the function f, once θ_{1t} and θ_{2t} have been replaced by (7) and (8) and evaluated at $\varepsilon_t = 0_{n_\varepsilon}$ and $\chi = 0$, can be written as another function f_{ss} that only depends on y_{t+1}, y_t, x_t, x_{t-1}, and $\overline{\theta}_1$ but neither on $\widehat{\theta}_2(s_{t+1})$ nor $\widehat{\theta}_2(s_t)$, i.e.

$$f\left(y_{t+1}, y_t, x_t, x_{t-1}, 0_{n_\varepsilon}, 0_{n_\varepsilon}, \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2(s_{t+1})' \end{pmatrix}', \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2(s_t)' \end{pmatrix}'\right) = f_{ss}\left(y_{t+1}, y_t, x_t, x_{t-1}, \overline{\theta}_1, \overline{\theta}_1\right).$$

[2] These functional forms are not necessary but just convenient for the derivations; any other functional form such that $\theta_1(0, s_t) = \overline{\theta}_1$ for all s_t holds may also work.

In general, more than one partition of parameters accomplishes this objective. In any case, included in θ_{1t} is the smallest set of parameters such that the steady state is defined as described above. Since the steady state depends upon $\bar{\theta}_1$, a natural choice for this point is the mean of the ergodic distribution across θ_{1t}, but again, this selection is not unique. Sections 6 provides examples of partitions of θ_t and choices of $\bar{\theta}_1$.

Given the definition of the steady state, it is the case that

$$y_{ss} = g\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) \text{ and } x_{ss} = h\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right)$$

for all s_t and

$$y_{ss} = g\left(x_{ss}, 0_{n_\varepsilon}, 0, s_{t+1}\right) \text{ and } x_{ss} = h\left(x_{ss}, 0_{n_\varepsilon}, 0, s_{t+1}\right)$$

for all s_{t+1}.

Using equations (2), (3), (4), and (5) the function f can be written as

$$\mathcal{F}\left(x_{t-1}, \varepsilon_{t+1}, \varepsilon_t, s_{t+1}, \chi, s_t\right) = f\left(\begin{array}{c} g\left(h\left(x_{t-1}, \varepsilon_t, \chi, s_t\right), \chi\varepsilon_{t+1}, \chi, s_{t+1}\right), g\left(x_{t-1}, \varepsilon_t, \chi, s_t\right), \\ h\left(x_{t-1}, \varepsilon_t, \chi, s_t\right), x_{t-1}, \chi\varepsilon_{t+1}, \varepsilon_t, \theta\left(\chi, s_{t+1}\right), \theta\left(\chi, s_t\right) \end{array}\right)$$

for all $x_{t-1}, \varepsilon_{t+1}, \varepsilon_t, s_{t+1}$, and s_t. The function $\mathcal{F}$ maps $\mathbb{R}^{n_x+2n_\varepsilon+1} \times \{1, \ldots, n_s\} \times \{1, \ldots, n_s\}$ into $\mathbb{R}^{n_y+n_x}$.

Assuming that innovations to the exogenous predetermined variables, ε_t, are independent of the Markov chain, s_t, write (1) as

$$\mathbb{E}_t f\left(y_{t+1}, y_t, x_t, x_{t-1}, \chi\varepsilon_{t+1}, \varepsilon_t, \theta_{t+1}, \theta_t\right) = \tag{9}$$
$$\mathbb{G}\left(x_{t-1}, \varepsilon_t, \chi, s_t\right) = \sum_{s'=1}^{n_s} p_{s_t,s'} \int \mathcal{F}\left(x_{t-1}, \varepsilon', \varepsilon_t, s', \chi, s_t\right) \mu\left(\varepsilon'\right) d\varepsilon' = 0_{n_y+n_x}$$

for all x_{t-1}, ε_t, and s_t where μ is the density of the innovations. The function $\mathbb{G}$ maps $\mathbb{R}^{n_x+n_\varepsilon+1} \times \{1, \ldots, n_s\}$ into $\mathbb{R}^{n_y+n_x}$.

The remainder of the paper will use the following notation

$$\mathcal{D}\mathbb{G}\left(x_{t-1}, \varepsilon_t, \chi, s_t\right) = \left[\mathcal{D}_j \mathbb{G}^i\left(x_{t-1}, \varepsilon_t, \chi, s_t\right)\right]_{1 \le i \le n_y+n_x, 1 \le j \le n_x+n_\varepsilon+1}$$

to refer the $(n_y + n_x) \times (n_x + n_\varepsilon + 1)$ matrix of first-order partial derivatives of $\mathbb{G}$ with respect to $(x_{t-1}, \varepsilon_t, \chi)$ evaluated at $(x_{t-1}, \varepsilon_t, \chi, s_t)$ for all x_{t-1}, ε_t, and s_t. Note the absence of derivatives

with respect to s_t, since it is a discrete variable. Equivalently,

$$\mathcal{D}\mathbb{G}\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) = \left[\mathcal{D}_j \mathbb{G}^i\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right)\right]_{1 \leq i \leq n_y + n_x, 1 \leq j \leq n_x + n_\varepsilon + 1}$$

refers to the $(n_y + n_x) \times (n_x + n_\varepsilon + 1)$ matrix of first-order partial derivatives of $\mathbb{G}$ with respect to $(x_{t-1}, \varepsilon_t, \chi)$ evaluated at $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ for all s_t. To simplify notation define

$$\mathcal{D}\mathbb{G}_{ss}\left(s_t\right) \equiv \mathcal{D}\mathbb{G}\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) \text{ and } \mathcal{D}_j \mathbb{G}^i_{ss}\left(s_t\right) \equiv \mathcal{D}_j \mathbb{G}^i\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right)$$

for all i, j, and s_t. Thus,

$$\mathcal{D}\mathbb{G}_{ss}\left(s_t\right) = \left[\mathcal{D}_j \mathbb{G}^i_{ss}\left(s_t\right)\right]_{1 \leq i \leq n_y + n_x, 1 \leq j \leq n_x + n_\varepsilon + 1}$$

for all s_t. In the same way,

$$\mathcal{D}f_{ss}\left(s_{t+1}, s_t\right) =$$
$$\left[\mathcal{D}_j f^i\left(y_{ss}, y_{ss}, x_{ss}, x_{ss}, 0_{n_\varepsilon}, 0_{n_\varepsilon}, \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2\left(s_{t+1}\right)' \end{pmatrix}', \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2\left(s_t\right)' \end{pmatrix}'\right)\right]_{1 \leq i \leq n_y + n_x, 1 \leq j \leq 2(n_y + n_x + n_\varepsilon + n_\theta)}$$

is the $(n_y + n_x) \times (2\left(n_y + n_x + n_\varepsilon + n_\theta\right))$ matrix of first-order partial derivatives of f with respect to all its components evaluated at $\left(y_{ss}, y_{ss}, x_{ss}, x_{ss}, 0_{n_\varepsilon}, 0_{n_\varepsilon}, \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2\left(s_{t+1}\right)' \end{pmatrix}', \begin{pmatrix} \overline{\theta}_1' & \widehat{\theta}_2\left(s_t\right)' \end{pmatrix}'\right)$ for all s_{t+1} and s_t,

$$\mathcal{D}g\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) = \left[\mathcal{D}_j g^i\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right)\right]_{1 \leq i \leq n_y, 1 \leq j \leq n_x + n_\varepsilon + 1}$$

is the $n_y \times (n_x + n_\varepsilon + 1)$ matrix of first-order partial derivatives of g with respect to $(x_{t-1}, \varepsilon_t, \chi)$ evaluated at $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ for all s_t, and

$$\mathcal{D}h\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) = \left[\mathcal{D}_j h^i\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right)\right]_{1 \leq i \leq n_x, 1 \leq j \leq n_x + n_\varepsilon + 1}$$

is the $n_x \times (n_x + n_\varepsilon + 1)$ matrix of first-order partial derivatives of h with respect to $(x_{t-1}, \varepsilon_t, \chi)$ evaluated at $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ for all s_t. To simplify notation, define

$$\mathcal{D}g_{ss}\left(s_t\right) \equiv \mathcal{D}g\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) \text{ and } \mathcal{D}_j g^i_{ss}\left(s_t\right) \equiv \mathcal{D}_j g^i\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right),$$

for all i, j, and s_t and

$$\mathcal{D}h_{ss}\left(s_t\right) \equiv \mathcal{D}h\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right) \text{ and } \mathcal{D}_j h^i_{ss}\left(s_t\right) \equiv \mathcal{D}_j h^i\left(x_{ss}, 0_{n_\varepsilon}, 0, s_t\right)$$

for all i, j, and s_t. Thus,

$$\mathcal{D}g_{ss}(s_t) = \left[\mathcal{D}_j g_{ss}^i(s_t)\right]_{1\leq i\leq n_y+n_x, 1\leq j\leq n_x+n_\varepsilon+1}$$

and

$$\mathcal{D}h_{ss}(s_t) = \left[\mathcal{D}_j h_{ss}^i(s_t)\right]_{1\leq i\leq n_y+n_x, 1\leq j\leq n_x+n_\varepsilon+1}$$

for all s_t.

3 First-Order Approximation

This section shows how to find the first-order Taylor expansions to g and h around the point $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ of the form

$$\begin{aligned} g(x_{t-1}, \varepsilon_t, \chi, s_t) - y_{ss} &\simeq [\mathcal{D}_1 g_{ss}(s_t), \ldots, \mathcal{D}_{n_x} g_{ss}(s_t)](x_{t-1} - x_{ss}) \\ &+ [\mathcal{D}_{n_x+1} g_{ss}(s_t), \ldots, \mathcal{D}_{n_x+n_\varepsilon} g_{ss}(s_t)]\varepsilon_t + \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t)\chi \end{aligned}$$

and

$$\begin{aligned} h(x_{t-1}, \varepsilon_t, \chi, s_t) - x_{ss} &\simeq [\mathcal{D}_1 h_{ss}(s_t), \ldots, \mathcal{D}_{n_x} h_{ss}(s_t)](x_{t-1} - x_{ss}) \\ &+ [\mathcal{D}_{n_x+1} h_{ss}(s_t), \ldots, \mathcal{D}_{n_x+n_\varepsilon} h_{ss}(s_t)]\varepsilon_t + \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t)\chi \end{aligned}$$

for all s_t where $\mathcal{D}_j g_{ss}(s_t)$ is the jth column vector of $\mathcal{D}g_{ss}(s_t)$ and $\mathcal{D}_j h_{ss}(s_t)$ is the jth column vector of $\mathcal{D}h_{ss}(s_t)$. To simply notation, define

$$\mathcal{D}_{n,m} g_{ss}(s_t) \equiv [\mathcal{D}_n g_{ss}(s_t), \ldots, \mathcal{D}_m g_{ss}(s_t)] \text{ and } \mathcal{D}_{n,m} h_{ss}(s_t) \equiv [\mathcal{D}_n h_{ss}(s_t), \ldots, \mathcal{D}_m h_{ss}(s_t)]$$

for all n and m and all s_t.

Hence, the above approximations are equivalent to

$$g(x_{t-1}, \varepsilon_t, \chi, s_t) - y_{ss} \simeq \mathcal{D}_{1,n_x} g_{ss}(s_t)(x_{t-1} - x_{ss}) + \mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t)\varepsilon_t + \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t)\chi$$

and

$$h(x_{t-1}, \varepsilon_t, \chi, s_t) - x_{ss} \simeq \mathcal{D}_{1,n_x} h_{ss}(s_t)(x_{t-1} - x_{ss}) + \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t)\varepsilon_t + \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t)\chi$$

The objective is now to find the coefficients of first-order partial derivatives

$$\left\{\mathcal{D}_{1,n_x} g_{ss}\left(s_t\right), \mathcal{D}_{1,n_x} h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s}, \ \left\{\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}\left(s_t\right), \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s},$$
$$\text{and } \left\{\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s_t\right), \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s}$$

of the above describe expansions, where $\mathcal{D}_{1,n_x} g_{ss}\left(s_t\right) \in \mathbb{C}^{n_y \times n_x}$, $\mathcal{D}_{1,n_x} h_{ss}\left(s_t\right) \in \mathbb{C}^{n_x \times n_x}$, $\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}\left(s_t\right) \in \mathbb{C}^{n_y \times n_\varepsilon}$, $\mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}\left(s_t\right) \in \mathbb{C}^{n_x \times n_\varepsilon}$, $\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s_t\right) \in \mathbb{C}^{n_y \times 1}$, and $\mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \in \mathbb{C}^{n_x \times 1}$ for all s_t and where we are using $\mathbb{C}^{m_1 \times m_2}$ to denote $m_1 \times m_2$ matrices over the complex numbers. The current setup requires finding a set of n_s policy functions, one for each possible value of the Markov chain, instead of the single set of policy functions in the constant parameter case.

The coefficients of these policy functions are going to be obtained by using the fact that

$$\mathbb{G}\left(x_{t-1}, \varepsilon_t, \chi, s_t\right) = 0_{n_y+n_x}$$

for all $x_{t-1}, \varepsilon_t, \chi$, and s_t and, therefore, it must be the case that

$$\mathcal{D}\mathbb{G}\left(x_{t-1}, \varepsilon_t, \chi, s_t\right) = 0_{(n_y+n_x)\times(n_x+n_\varepsilon+1)}$$

for all $x_{t-1}, \varepsilon_t, \chi$, and s_t and, in particular,

$$\mathcal{D}\mathbb{G}_{ss}\left(s_t\right) = 0_{(n_y+n_x)\times(n_x+n_\varepsilon+1)}$$

for all s_t where $0_{m_1 \times m_2}$ is a $m_1 \times m_2$ matrix of zeros. Thus,

$$\begin{gathered}\left[\mathcal{D}_1 \mathbb{G}_{ss}\left(s_t\right), \ldots, \mathcal{D}_{n_x} \mathbb{G}_{ss}\left(s_t\right)\right] = 0_{(n_y+n_x)\times n_x}, \\ \left[\mathcal{D}_{n_x+1} \mathbb{G}_{ss}\left(s_t\right), \ldots, \mathcal{D}_{n_x+n_\varepsilon} \mathbb{G}_{ss}\left(s_t\right)\right] = 0_{(n_y+n_x)\times n_\varepsilon}, \\ \text{and } \mathcal{D}_{n_x+n_\varepsilon+1} \mathbb{G}_{ss}\left(s_t\right) = 0_{n_y+n_x}\end{gathered} \tag{10}$$

for all s_t where $\mathcal{D}_j \mathbb{G}_{ss}\left(s_t\right)$ is the jth column vector of $\mathcal{D}\mathbb{G}_{ss}\left(s_t\right)$. Again, note that there are a set of n_s derivatives of $\mathbb{G}$, one for each possible value of s_t, instead of the single derivative in the constant parameter case. To simply notation, again, define

$$\mathcal{D}_{n,m} \mathbb{G}_{ss}\left(s_t\right) \equiv \left[\mathcal{D}_n \mathbb{G}_{ss}\left(s_t\right), \ldots, \mathcal{D}_m \mathbb{G}_{ss}\left(s_t\right)\right]$$

for all s_t. Therefore, expression (10) can be written as

$$\mathcal{D}_{1,n_x}\mathbb{G}_{ss}\left(s_t\right)=0_{(n_y+n_x)\times n_x},\ \mathcal{D}_{n_x+1,n_x+n_\varepsilon}\mathbb{G}_{ss}\left(s_t\right)=0_{(n_y+n_x)\times n_\varepsilon},\ \text{and}\ \mathcal{D}_{n_x+n_\varepsilon+1}\mathbb{G}_{ss}\left(s_t\right)=0_{n_y+n_x} \tag{11}$$

for all s_t.

The expression $\mathcal{D}_{1,n_x}\mathbb{G}_{ss}\left(s_t\right) = 0_{(n_y+n_x)\times n_x}$ will be used to solve for $\left\{\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right),\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s}$, while expressions $\mathcal{D}_{n_x+1,n_x+n_\varepsilon}\mathbb{G}_{ss}\left(s_t\right) = 0_{(n_y+n_x)\times n_\varepsilon}$ and $\mathcal{D}_{n_x+n_\varepsilon+1}\mathbb{G}_{ss}\left(s_t\right)=0_{n_y+n_x}$ will be used to solve for $\left\{\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right),\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s}$ and $\left\{\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right),\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s}$. In what follows, we show that $\mathcal{D}_{1,n_x}\mathbb{G}_{ss}\left(s_t\right) = 0_{(n_y+n_x)\times n_x}$ implies a quadratic system, while $\mathcal{D}_{n_x+1,n_x+n_\varepsilon}\mathbb{G}_{ss}\left(s_t\right) = 0_{(n_y+n_x)\times n_\varepsilon}$ and $\mathcal{D}_{n_x+n_\varepsilon+1}\mathbb{G}_{ss}\left(s_t\right)=0_{n_y+n_x}$ (given $\left\{\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right),\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\right\}_{s_t=1}^{n_s}$) are linear systems.

3.1 Solving for the Derivatives of x_{t-1}

Taking derivatives with respect to x_{t-1} in (9) produces the following expression

$$\mathcal{D}_{1,n_x}\mathbb{G}_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\int\begin{pmatrix}\mathcal{D}_{1,n_y}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+\mathcal{D}_{n_y+1,2n_y}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+\mathcal{D}_{2n_y+n_x+1,2(n_y+n_x)}f_{ss}\left(s',s_t\right)\end{pmatrix}\mu\left(\varepsilon'\right)d\varepsilon'$$

for all s_t. Next, taking into account that $\int\mu\left(\varepsilon'\right)d\varepsilon'=1$, this expression simplifies to

$$\mathcal{D}_{1,n_x}\mathbb{G}_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\begin{pmatrix}\mathcal{D}_{1,n_y}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+\mathcal{D}_{n_y+1,2n_y}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+\mathcal{D}_{2n_y+n_x+1,2(n_y+n_x)}f_{ss}\left(s',s_t\right)\end{pmatrix}$$

for all s_t. Now, rearranging, for each s_t

$$\mathcal{D}_{1,n_x}\mathbb{G}_{ss}\left(s_t\right)= \tag{12}$$
$$\sum_{s'=1}^{n_s}p_{s_t,s'}\begin{pmatrix}\left(\mathcal{D}_{1,n_y}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)+\mathcal{D}_{2n_y+1,2n_y+n_x}f_{ss}\left(s',s_t\right)\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+\mathcal{D}_{n_y+1,2n_y}f_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)+\mathcal{D}_{2n_y+n_x+1,2(n_y+n_x)}f_{ss}\left(s',s_t\right)\end{pmatrix}.$$

Putting together the n_s versions of (12), one for each value of s_t, and equating them to zero, as implied by (11), yields a system of $(n_y + n_x)\, n_x n_s$ quadratic equations in the same number of unknowns $\{\mathcal{D}_{1,n_x} g_{ss}(s_t), \mathcal{D}_{1,n_x} h_{ss}(s_t)\}_{s_t=1}^{n_s}$. Section 4 describes how to solve this quadratic system.

3.2 Solving for the Derivatives of ε_t and χ

This subsection shows that, after finding $\{\mathcal{D}_{1,n_x} g_{ss}(s_t), \mathcal{D}_{1,n_x} h_{ss}(s_t)\}_{s_t=1}^{n_s}$, obtaining $\{\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t), \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t)\}_{s_t=1}^{n_s}$ and $\{\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t), \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t)\}_{s_t=1}^{n_s}$ is simply solving a system of linear equations. The first step is to solve for $\{\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t), \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t)\}_{s_t=1}^{n_s}$. Then, we solve for $\{\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t), \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t)\}_{s_t=1}^{n_s}$.

3.2.1 Solving for the Derivatives of ε_t

In order to solve for $\{\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t), \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t)\}_{s_t=1}^{n_s}$, we obtain the expression for $\mathcal{D}_{n_x+1,n_x+n_\varepsilon} \mathbb{G}_{ss}(s_t)$ by taking derivatives with respect to ε_t in (9)

$$\mathcal{D}_{n_x+1,n_x+n_\varepsilon} \mathbb{G}_{ss}(s_t) =$$
$$\sum_{s'=1}^{n_s} p_{s_t,s'} \int \begin{pmatrix} \mathcal{D}_{1,n_y} f_{ss}(s', s_t)\, \mathcal{D}_{1,n_x} g_{ss}(s')\, \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t) + \\ \mathcal{D}_{n_y+1,2n_y} f_{ss}(s', s_t)\, \mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t) + \\ \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s', s_t)\, \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t) + \\ \mathcal{D}_{2(n_y+n_x)+n_\varepsilon+1,2(n_y+n_x+n_\varepsilon)} f_{ss}(s', s_t) \end{pmatrix} \mu(\varepsilon')\, d\varepsilon'$$

for all s_t. Taking into account that $\int \mu(\varepsilon')\, d\varepsilon' = 1$, this expression simplifies to

$$\mathcal{D}_{n_x+1,n_x+n_\varepsilon} \mathbb{G}_{ss}(s_t) =$$
$$\sum_{s'=1}^{n_s} p_{s_t,s'} \begin{pmatrix} \left(\mathcal{D}_{1,n_y} f_{ss}(s', s_t)\, \mathcal{D}_{1,n_x} g_{ss}(s') + \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s', s_t)\right) \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t) \\ \mathcal{D}_{n_y+1,2n_y} f_{ss}(s', s_t)\, \mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t) \\ \mathcal{D}_{2(n_y+n_x)+n_\varepsilon+1,2(n_y+n_x+n_\varepsilon)} f_{ss}(s', s_t) \end{pmatrix} \tag{13}$$

for all s_t.

Putting together the n_s versions of (13), one for each value of s_t, and equating them to zero, as implied by (11), yields a system of $(n_y + n_x)\, n_\varepsilon n_s$ equations in the same number of unknowns $\{\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s_t), \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(s_t)\}_{s_t=1}^{n_s}$. The system is linear.

The linear system can be written in matrix notation expression as

$$\begin{bmatrix} \Theta_\varepsilon & \Phi_\varepsilon \end{bmatrix} \begin{bmatrix} \mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(1) \\ \vdots \\ \mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(n_s) \\ \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(1) \\ \vdots \\ \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}(n_s) \end{bmatrix} = \Psi_\varepsilon \tag{14}$$

where

$$\Theta_\varepsilon = \sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{n_y+1,2n_y} f_{ss}(s',1) & \cdots & 0_{(n_x+n_y)\times n_y} \\ \vdots & \ddots & \vdots \\ 0_{(n_x+n_y)\times n_y} & \cdots & p_{n_s,s'} \mathcal{D}_{n_y+1,2n_y} f_{ss}(s',n_s) \end{bmatrix},$$

$$\Phi_\varepsilon = \sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{1,n_y} f_{ss}(s',1)\, \mathcal{D}_{1,n_x} g_{ss}(s') & \cdots & 0_{(n_x+n_y)\times n_x} \\ \vdots & \ddots & \vdots \\ 0_{(n_x+n_y)\times n_x} & \cdots & p_{n_s,s'} \mathcal{D}_{1,n_y} f_{ss}(s',n_s)\, \mathcal{D}_{1,n_x} g_{ss}(s') \end{bmatrix}$$
$$+ \sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s',1) & \cdots & 0_{(n_x+n_y)\times n_x} \\ \vdots & \ddots & \vdots \\ 0_{(n_x+n_y)\times n_x} & \cdots & p_{n_s,s'} \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s',n_s) \end{bmatrix},$$

and

$$\Psi_\varepsilon = -\sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{2(n_y+n_x)+n_\varepsilon+1,2(n_y+n_x+n_\varepsilon)} f_{ss}(s',1) \\ \vdots \\ p_{n_s,s'} \mathcal{D}_{2(n_y+n_x)+n_\varepsilon+1,2(n_y+n_x+n_\varepsilon)} f_{ss}(s',n_s) \end{bmatrix}.$$

Thus, given the solution for $\{\mathcal{D}_{1,n_x} g_{ss}(s_t), \mathcal{D}_{1,n_x} h_{ss}(s_t)\}_{s_t=1}^{n_s}$, expression (14) is a system of $(n_y + n_x)\, n_\varepsilon n_s$ linear equations in the same number of unknowns given by $\{\mathcal{D}_{n_x+1,n_x+n_e} g_{ss}(s_t), \mathcal{D}_{n_x+1,n_x+n_e} h_{ss}(s_t)\}_{s_t=1}^{n_s}$ that can be solved by inverting $\begin{bmatrix} \Theta_\varepsilon & \Phi_\varepsilon \end{bmatrix}$.

3.2.2 Solving for the Derivatives of χ

In order to solve for $\{\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}(s_t), \mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}(s_t)\}_{s_t=1}^{n_s}$, we obtain the expression for $\mathcal{D}_{n_x+n_\varepsilon+1}\mathbb{G}_{ss}(s_t)$ by taking derivatives with respect to χ in (9)

$$\mathcal{D}_{n_x+n_\varepsilon+1}\mathbb{G}_{ss}(s_t) =$$
$$\sum_{s'=1}^{n_s} p_{s_t,s'} \int \left(\begin{array}{c} \mathcal{D}_{1,n_y} f_{ss}(s', s_t) \left[\begin{array}{c} \mathcal{D}_{1,n_x} g_{ss}(s') \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) \\ +\mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}(s') \varepsilon' + \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s') \end{array} \right] + \\ \mathcal{D}_{n_y+1,2n_y} f_{ss}(s', s_t) \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) + \\ \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s', s_t) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) + \\ \mathcal{D}_{2n_y+n_x+1,2n_y+n_x+n_\varepsilon} f_{ss}(s', s_t) \varepsilon' + \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}(s', s_t) \mathcal{D}\theta_{ss}(s') + \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}(s', s_t) \mathcal{D}\theta_{ss}(s_t) \end{array} \right) \mu(\varepsilon') d\varepsilon'$$

for all s_t, where $\mathcal{D}\theta_{ss}(s_t)$ is the derivative of $\theta(\chi, s_t)$ with respect to χ evaluated at $\chi = 0$

$$\mathcal{D}\theta_{ss}(s_t) = \mathcal{D}\theta(0, s_t) = \left[\mathcal{D}_j^i \theta(0, s_t)\right]_{1 \le i \le n_\theta, j=1}$$

for all s_t.

Taking into account that $\int \mu(\varepsilon') d\varepsilon' = 1$ and $\int \varepsilon' \mu(\varepsilon') d\varepsilon' = 0$, the above simplifies to

$$\mathcal{D}_{n_x+n_\varepsilon+1}\mathbb{G}_{ss}(s_t) = \tag{15}$$
$$\sum_{s'=1}^{n_s} p_{s_t,s'} \left(\begin{array}{c} \mathcal{D}_{1,n_y} f_{ss}(s', s_t) \{\mathcal{D}_{1,n_x} g_{ss}(s') \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) + \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s')\} + \\ \mathcal{D}_{n_y+1,2n_y} f_{ss}(s', s_t) \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) + \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s', s_t) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) \\ +\mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}(s', s_t) \mathcal{D}\theta_{ss}(s') + \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}(s', s_t) \mathcal{D}\theta_{ss}(s_t) \end{array} \right)$$

for all s_t.

Putting together the n_s versions of (15), one for each value of s_t, and equating them to zero, as implied by (11), yields a system of $(n_y + n_x) n_s$ equations in the same number of unknowns $\{\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}(s_t), \mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}(s_t)\}_{s_t=1}^{n_s}$. This system is also linear.

The linear system can be written in matrix notation expression as

$$\begin{bmatrix} \Theta_\chi & \Phi_\chi \end{bmatrix} \begin{bmatrix} \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(1) \\ \vdots \\ \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(n_s) \\ \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(1) \\ \vdots \\ \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(n_s) \end{bmatrix} = \Psi_\chi, \tag{16}$$

where

$$\begin{aligned} \Theta_\chi &= \sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{n_y+1,2n_y} f_{ss}(s',1) & \cdots & 0_{(n_x+n_y)\times n_y} \\ \vdots & \ddots & \vdots \\ 0_{(n_x+n_y)\times n_y} & \cdots & p_{n_s,s'} \mathcal{D}_{n_y+1,2n_y} f_{ss}(s',n_s) \end{bmatrix} \\ &\quad + \begin{bmatrix} p_{1,1} \mathcal{D}_{1,n_y} f_{ss}(1,1) & \cdots & p_{1,n_s} \mathcal{D}_{1,n_y} f_{ss}(n_s,1) \\ \vdots & \ddots & \vdots \\ p_{n_s,1} \mathcal{D}_{1,n_y} f_{ss}(1,n_s) & \cdots & p_{n_s,n_s} \mathcal{D}_{1,n_y} f_{ss}(n_s,n_s) \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} \Phi_\chi &= \sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{1,n_y} f_{ss}(s',1) \mathcal{D}_{1,n_x} g_{ss}(s') & \cdots & 0_{(n_x+n_y)\times n_x} \\ \vdots & \ddots & \vdots \\ 0_{(n_x+n_y)\times n_x} & \cdots & p_{n_s,s'} \mathcal{D}_{1,n_y} f_{ss}(s',n_s) \mathcal{D}_{1,n_x} g_{ss}(s') \end{bmatrix} \\ &\quad + \sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s',1) & \cdots & 0_{(n_x+n_y)\times n_x} \\ \vdots & \ddots & \vdots \\ 0_{(n_x+n_y)\times n_x} & \cdots & p_{n_s,s'} \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s',n_s) \end{bmatrix}, \end{aligned}$$

and

$$\Psi_\chi = -\sum_{s'=1}^{n_s} \begin{bmatrix} p_{1,s'} \begin{pmatrix} \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}(s',1) \mathcal{D}\theta_{ss}(s') + \dots \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}(s',1) \mathcal{D}\theta_{ss}(1) \end{pmatrix} \\ \vdots \\ p_{n_s,s'} \begin{pmatrix} \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}(s',n_s) \mathcal{D}\theta_{ss}(s') + \dots \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}(s',n_s) \mathcal{D}\theta_{ss}(n_s) \end{pmatrix} \end{bmatrix}.$$

Thus, given the solution for $\{\mathcal{D}_{1,n_x} g_{ss}(s_t), \mathcal{D}_{1,n_x} h_{ss}(s_t)\}_{s_t=1}^{n_s}$, expression (16) is a system of $(n_y + n_x) n_s$ linear equations in the same number of unknowns given by $\{\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t), \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t)\}_{s_t=1}^{n_s}$ that can be solved by inverting $\begin{bmatrix} \Theta_\chi & \Phi_\chi \end{bmatrix}$.

3.3 Non-Certainty Equivalence of First-Order Approximation

As pointed out by Schmitt-Grohe & Uribe (2004), one important feature of constant parameter models is certainty equivalence of the first-order approximation. This feature of constant parameter models implies that first-order approximations are inadequate for analyzing interesting behavior such as responses to risk because the approximated decision rules are invariant to changes in volatility. For example, van Binsbergen *et al.* (2008) and Rudebusch & Swanson (2008) note that at least second-order approximations are needed to analyze certain asset pricing implications, such as the yield curve, since second-order approximations are not certainty equivalent, and hence react to changes in volatility. Second-order approximations also imply a degree of difficulty in performing likelihood based estimation, such as Fernández-Villaverde & Rubio-Ramirez (2007) who use the particle filter for estimation. These factors mean that addressing interesting questions with second-order approximations may be necessary but difficult in constant parameter models. As shown below, first order approximations to Markov Switching models are not necessarily certainty equivalent. This nice feature opens the door to analyze risk related behaviors using linearly approximated models.

To see the certainty equivalence of the first-order approximation of constant parameter models, consider equation (16) with only one regime, so $n_s = 1$. In this case

$$\begin{bmatrix} \Theta_\chi & \Phi_\chi \end{bmatrix} \begin{bmatrix} \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(1) \\ \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(1) \end{bmatrix} = \Psi_\chi \tag{17}$$

where

$$\begin{bmatrix} \Theta_\chi & \Phi_\chi \end{bmatrix} =$$

$$\begin{bmatrix} \mathcal{D}_{1,n_y} f_{ss}(1,1)\, \mathcal{D}_{1,n_x} g_{ss}(1) + \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(1,1) & \mathcal{D}_{1,n_y} f_{ss}(1,1) + \mathcal{D}_{n_y+1,2n_y} f_{ss}(1,1) \end{bmatrix}$$

and

$$\Psi_\chi = -\left[\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}\left(1,1\right)\mathcal{D}_{ss}\theta\left(1\right)\ \right].$$

Clearly, in constant parameter models it is the case that $\theta\left(\chi,1\right)=\bar{\theta}$. Therefore, $\mathcal{D}_{ss}\theta\left(1\right)=0_{n_\theta}$, which implies $\Psi_\chi = 0_{n_x+n_y}$. Consequently the linear system (17) is homogenous. If a unique solution exists, then it is given by

$$\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(1\right)=0_{n_y} \text{ and } \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(1\right)=0_{n_x}. \tag{18}$$

Remenber that in constant parameter models the linear approximation to the policy rules imply

$$y_t - y_{ss} = \mathcal{D}_{1,n_x} g_{ss}\left(1\right)\left(x_{t-1}-x_{ss}\right) + \mathcal{D}_{n_x+1,n_x+n_\varepsilon} g_{ss}\left(1\right)\varepsilon_t + \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(1\right)$$

and

$$x_t - x_{ss} = \mathcal{D}_{1,n_x} h_{ss}\left(1\right)\left(x_{t-1}-x_{ss}\right) + \mathcal{D}_{n_x+1,n_x+n_\varepsilon} h_{ss}\left(1\right)\varepsilon_t + \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(1\right).$$

Using (18), evaluated at $x_{t-1}=x_{ss}$ and $\varepsilon_t = 0_{n_\varepsilon}$ the above approximations imply that $y_t - y_{ss} = 0_{n_y}$ and $x_t - x_{ss} = 0_{n_x}$, i.e. the linear approximation of constant parameter models is certainty equivalent.

Let us now turn to the Markov switching case. From equation (16) is clear that a necessary condition for the linear approximation not to be certainty equivalent is that $\Psi_\chi \neq 0_{n_s(n_x+n_y)}$. Let us analyze when it is the case that $\Psi_\chi \neq 0_{n_s(n_x+n_y)}$. Consider the expression for Ψ_χ

$$\Psi_\chi = -\sum_{s'=1}^{n_s}\left[\begin{array}{c} p_{1,s'}\left(\begin{array}{c}\mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}\left(s',1\right)\mathcal{D}\theta_{ss}\left(s'\right)+\ldots \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}\left(s',1\right)\mathcal{D}\theta_{ss}\left(1\right)\end{array}\right) \\ \vdots \\ p_{n_s,s'}\left(\begin{array}{c}\mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}\left(s',n_s\right)\mathcal{D}\theta_{ss}\left(s'\right)+\ldots \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}\left(s',n_s\right)\mathcal{D}\theta_{ss}\left(n_s\right)\end{array}\right)\end{array}\right].$$

Then, if $\mathcal{D}\theta_{ss}\left(s_t\right)=0_{n_\theta}$ for all s_t, it is the case that $\Psi_\chi = 0_{n_s(n_x+n_y)}$. So a necessary condition for $\Psi_\chi \neq 0_{n_s(n_x+n_y)}$ is that $\mathcal{D}\theta_{ss}\left(s_t\right)\neq 0_{n_\theta}$ for some s_t. Recalling the form of θ_t

$$\theta_1\left(\chi,s_t\right)=\bar{\theta}_1+\chi\widehat{\theta}_1\left(s_t\right)$$

and

$$\theta_2(\chi, s_t) = \widehat{\theta}_2(s_t)$$

we conclude that

$$\mathcal{D}\theta_{ss}(s_t) = \begin{bmatrix} \widehat{\theta}_1(s_t)' & 0'_{n_{\theta_2}} \end{bmatrix}'.$$

Then $\mathcal{D}\theta_{ss}(s_t) \neq 0_{n_\theta}$ for some s_t if and only if $\widehat{\theta}_1(s_t) \neq 0_{n_{\theta_1}}$ for some s_t. Hence, a necessary condition for $\Psi_\chi \neq 0_{n_s(n_x+n_y)}$ is that $\widehat{\theta}_1(s_t) \neq 0_{n_{\theta_1}}$ for some s_t.

However, the condition that $\widehat{\theta}_1(s_t) \neq 0_{n_{\theta_1}}$ for some s_t is not sufficient for $\Psi_\chi \neq 0_{n_s(n_x+n_y)}$. In addition, it must be the case that

$$\sum_{s'=1}^{n_s} p_{s_t,s'} \begin{pmatrix} \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}(s', s_t)\, \mathcal{D}\theta_{ss}(s') + \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}(s', s_t)\, \mathcal{D}\theta_{ss}(s_t) \end{pmatrix} \neq 0_{n_x+n_y}$$

for some s_t, which will be true when θ_{1t} do not enter the equilibrium conditions multiplicatively with a variable which expected value equals zero when evaluating $f(\cdot)$ in steady state. The following Proposition summarizes these results.

Proposition 1 *Let $\widehat{\theta}_1(s_t) = 0_{n_{\theta_1}}$ for all s_t. Then $\Psi_\chi = 0_{n_s(n_x+n_y)}$ and*

$$\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) = 0_{n_y} \text{ and } \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) = 0_{n_x}$$

for all s_t and the first order approximation is certainty equivalent. On the other hand, let $\widehat{\theta}_1(s_t) \neq 0_{n_{\theta_1}}$ for some s_t but let

$$\sum_{s'=1}^{n_s} p_{s_t,s'} \begin{pmatrix} \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+1,2(n_y+n_x+n_\varepsilon)+n_\theta} f_{ss}(s', s_t)\, \mathcal{D}\theta_{ss}(s') + \\ \mathcal{D}_{2(n_y+n_x+n_\varepsilon)+n_\theta+1,2(n_y+n_x+n_\varepsilon+n_\theta)} f_{ss}(s', s_t)\, \mathcal{D}\theta_{ss}(s_t) \end{pmatrix} = 0_{n_x+n_y}$$

for all s_t. Then $\Psi_\chi = 0_{n_s(n_x+n_y)}$ and

$$\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) = 0_{n_y} \text{ and } \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) = 0_{n_x}$$

for all s_t and the first order approximation is certainty equivalent.

Note that if the first-order approximation is not certainty equivalent, it means that either

$$\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) \neq 0_{n_y}$$

or

$$\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right) \neq 0_{n_x}$$

for some s_t and the linear approximation to the policy rules evaluated at $x_{t-1} = x_{ss}$ and $\varepsilon_t = 0_{n_\varepsilon}$ imply that either

$$y_t - y_{ss} = \mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right) \neq 0_{n_y}$$

or

$$x_t - x_{ss} = \mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right) \neq 0_{n_x}$$

for some s_t.

4 The Solution to the Quadratic System

As mentioned above, the n_s versions of (12) form a system of $\left(n_y + n_x\right) n_s n_x$ quadratic equations in the elements of $\left\{\mathcal{D}_{1,n_x}g_{ss}\left(s\right), \mathcal{D}_{1,n_x}h_{ss}\left(s\right)\right\}_{s=1}^{n_s}$. This section describes how to find the solution to this quadratic system. Putting (12) into matrix form produces n_s quadratic systems of the form

$$A\left(s_t\right)\begin{bmatrix} I \\ \mathcal{D}_{1,n_x}g_{ss}\left(1\right) \\ \vdots \\ \mathcal{D}_{1,n_x}g_{ss}\left(n_s\right) \end{bmatrix}\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right) = B\left(s_t\right)\begin{bmatrix} I \\ \mathcal{D}_{1,n_x}g_{ss}\left(s_t\right) \end{bmatrix} \tag{19}$$

for all s_t, where

$$A\left(s_t\right) = \begin{bmatrix} \sum_{s'=1}^{n_s} p_{s_t,s'}\mathcal{D}_{2n_y+1,2n_y+n_x}f_{ss}\left(s',s_t\right) & p_{s_t,1}\mathcal{D}_{1,n_y}f_{ss}\left(1,s_t\right) & \cdots & p_{s_t,n_s}\mathcal{D}_{1,n_y}f_{ss}\left(n_s,s_t\right) \end{bmatrix}$$

and

$$B\left(s_t\right) = -\sum_{s'=1}^{n_s} p_{s_t,s'}\begin{bmatrix} \mathcal{D}_{2n_y+n_x+1,2(n_y+n_x)}f_{ss}\left(s',s_t\right) & \mathcal{D}_{n_y+1,2n_y}f_{ss}\left(s',s_t\right) \end{bmatrix}.$$

This quadratic system is nothing else than an algebraic system of equations. In a constant regime framework, $n_s = 1$, mapping this quadratic system into a generalized eigenvalue problem allows solving it by a singular value decomposition (SVD) type of algorithm. In the case of

Markov switching, the fact that $\{\mathcal{D}_{1,n_x} g_{ss}(s)\}_{s=1}^{n_s}$ appear in every of the n_s equations described above makes it impossible to map the algebraic systems of equations into a generalized eigenvalue problem. Instead solutions are going to be found using Gröbner Bases to be introduced bellow.

Before getting into the Gröbner Bases, we need to bound the number of solutions of the quadratic system (19). At this point is important to note that any $\{\mathcal{D}_{1,n_x} g_{ss}(s), \mathcal{D}_{1,n_x} h_{ss}(s)\}_{s=1}^{n_s}$ that solves the quadratic system (19) is a minimum state variable solution. A number of other procedures, such as Newton's methods, can be used to solve this equation, with the important limitation that they cannot find all possible solutions. However, the Appendix describes a numerical procedure to find one solution.

4.1 Bounding the Number of Solutions

In order to apply Gröbner Bases to solve the quadratic system (19) we need to make sure that the system has a finite number of solutions (see below for details). In general, determining the number of solutions in a algebraic systems of equations may be difficult. However, given the structure of the quadratic system (19), it is possible to bound the number of solutions. This is going to be enough for our purposes.

Consider the fixed regime case of the quadratic system (19), when $n_s = 1$. The usual practice of solving the model, as in Schmitt-Grohe & Uribe (2004), involves constructing a single stable solution that depends upon a particular ordering of the generalized eigenvalues of the matrices $A(1)$ and $B(1)$. But the quadratic system does not have an unique solution. The full set of solutions (stable and unstable) can be found by choosing different orderings of the eigenvalues. The bound to the number of possible orderings of eigenvalues depends upon the rank of the matrix $A(1)$, the number of (endogenous and exogenous) predetermined variables, and the number of exogenous predetermined variables by the following.

Proposition 2 *Assume we are in the fixed regime case, i.e.* $n_s = 1$.. *Let* n_{exo} *denote the number of exogenous predetermined variables, so* $0 \leq n_{exo} \leq n_x$. *Then the total number of*

solutions to the quadratic system (19) when $n_s = 1$ is bounded by

$$\begin{pmatrix} rank\left[A\left(1\right)\right] - n_{exo} \\ n_x - n_{exo} \end{pmatrix} = \frac{\left(rank\left[A\left(1\right)\right] - n_{exo}\right)!}{\left(n_x - n_{exo}\right)!\left(rank\left[A\left(1\right)\right] - n_x\right)!}$$

where $rank\left[A\left(1\right)\right]$ stands for the rank of the matrix $A\left(1\right)$.

Three important remarks. First, the number in the proposition is a bound because it may be the case that we have repeated eigenvalues. If all the eigenvalues are different, then the proposition provides the exact number of solutions. Second, matrix $A\left(1\right)$ will not be of full rank when there exist redundant variables or identities that can be eliminated from the quadratic system, and consequently are linearly dependent upon another set of variables.[3] The set of exogenous predetermined variables has their set eigenvalues associated with their coefficients of autocorrelation and, conditional on the number of endogenous predetermined variables and non-predetermined variables, the number of solutions does not depend upon the number of exogenous predetermined variables.

Now consider the case of Markov Switching models, when $n_s > 1$. We now have n_s quadratic systems of the form (19). Let us first analyze the quadratic system (19) for $s_t = 1$. Given $\left\{\mathcal{D}_{1,n_x} g_{ss}\left(s_t\right), \mathcal{D}_{1,n_x} h_{ss}\left(s_t\right)\right\}_{s_t=2}^{n_s}$, the bound on the number of solutions to the quadratic system (19) for $s_t = 1$ depends upon the rank of the matrix $A\left(1\right)$, the number of (endogenous and exogenous) predetermined variables, and the number of exogenous predetermined variables as reported in Proposition 2. Since there are n_s quadratic systems of the form (19), the bound to the total number of solutions is determined by the following proposition.

Proposition 3 *Assume we are in the Markov Switching case, i.e. $n_s > 1$. Then the bound to the number of solutions to the n_s quadratic system (19) is given by*

$$\prod_{s_t=1}^{n_s} \begin{pmatrix} rank\left[A\left(s_t\right)\right] - n_{exo} \\ n_x - n_{exo} \end{pmatrix} = \left(\frac{\left(rank\left[A\left(s_t\right)\right] - n_{exo}\right)!}{\left(n_x - n_{exo}\right)!\left(rank\left[A\left(s_t\right)\right] - n_x\right)!}\right)^{n_s}$$

[3] A simple example of a system where matrix $A\left(1\right)$ will not be of full rank is a leisureless RBC model with three variables (capital, consumption, and output) and three equations (the eurler condition, the resource constraint, and the output definition). If we eliminate the output definition equation and the output variable the new matrix would be full rank.

where rank $[A(s_t)]$ *stands for the rank of the matrix* $A(s_t)$ *for all* s_t.

4.2 Gröbner Basis

A Gröbner basis for a set of polynomials is an alternative set of polynomials that possesses desirable algorithmic properties. The most important of these features for our problem at hand is that the Gröbner basis can be used to easily find all the roots of the original system of polynomials equations. We will use this technique to solve the quadratic system defined by equation (19). Every set of polynomials can be transformed into a Gröbner basis, although this transformation may not be unique. The reduced Gröbner basis is indeed unique. The transformation process generalizes the familiar techniques of Gaussian elimination for solving linear systems of equations. A careful definition of reduced Gröbner basis, and lexicographic is out of the scope of the paper. Here we only present the mathematical result (the Shape Lemma) that is useful for our purposes, i.e. solving the quadratic system defined by equation (19). In what follows we first introduce the theory around Gröbner basis. Then, we show an example. Finally, we illustrate how the Shape Lemma can be used for our purposes.

4.2.1 The Theory

Suposse a given set of polynomials $f_1, \ldots, f_m$ in $\mathbb{Q}[x_1, \ldots, x_n]$, where $\mathbb{Q}[x_1, \ldots, x_n]$ is the polynomial ring defined by the set of polynomials in n variables with coefficients in field $\mathbb{Q}$ of rational numbers. Here, we define a polynomial f in $x_1, \ldots, x_n$ with coefficients in $\mathbb{Q}$ as a finite combination of monomials

$$f = \sum_{\alpha} c_\alpha x^\alpha = \sum_{\alpha} c_\alpha x_1^{\alpha_1} \ldots x_n^{\alpha_n}$$

with $c_\alpha \in \mathbb{Q}$ and the sum is over the finite number of n-tuples $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_+^n$.

A subset I of the polynomial ring $\mathbb{Q}[x_1, \ldots, x_n]$ is called an ideal if closed under sums, $f + g \in I$ for all $f, g \in I$ and $f * h \in I$ for all $f \in I$ and $h \in \mathbb{Q}[x_1, \ldots, x_n]$. There is one particular ideal we are interested on. Let $I = \langle f_1, \ldots, f_m \rangle \subseteq \mathbb{Q}[x_1, \ldots, x_n]$ be the ideal

generated by this set of polynomials where

$$\langle f_1, \ldots, f_m \rangle = \left\{ f | f = \sum_{i=1}^{m} g_i f_i,\ g_i \in \mathbb{Q}[x_1, \ldots, x_n] \right\}.$$

Let also define the set $V(f_1, \ldots, f_m)$ to be the complex variety defined by the set of polynomials $f_1, \ldots, f_m$ where

$$V(f_1, \ldots, f_m) = \{(a_1, \ldots, a_n) \in \mathbb{C}^n | f_i(a_1, \ldots, a_n) = 0 \text{ for all } i \in \{1, \ldots, m\}\}.$$

The complex variety defines the set of common complex zeros of the polynomials $f_1, \ldots, f_m$. If we replace $f_1, \ldots, f_m$ by another set of polynomials $g_1, \ldots, g_l$ that generate the same ideal, i.e. $\langle f_1, \ldots, f_m \rangle = \langle g_1, \ldots, g_l \rangle$, the complex variety does not change, i.e. $V(f_1, \ldots, f_m) = V(g_1, \ldots, g_l)$. Any set of polynomials that generates the ideal $\langle f_1, \ldots, f_m \rangle$ is called a basis of the ideal. Thus, both $f_1, \ldots, f_m$ and $g_1, \ldots, g_l$ are basis for the ideal $\langle f_1, \ldots, f_m \rangle$. If $V(f_1, \ldots, f_m)$ is a finite set, then $\langle f_1, \ldots, f_m \rangle$ is said to be a zero-dimensional ideal.

The fact that complex varieties do not change when we change the basis is going to be crucial for our strategy. Our approach is going to ask whether we can find a basis that makes describing the solution set $V(f_1, \ldots, f_m)$ straightforward. In other words: Can we transform the original system of polynomials equations $f_1(x_1, \ldots, x_n) = \ldots = f_m(x_1, \ldots, x_n) = 0$ into a new system $g_1(x_1, \ldots, x_n) = \ldots = g_l(x_1, \ldots, x_n) = 0$ that can be easily solved? In this section, we show how reduced Gröbner basis can be used for this purpose.

Before showing how to do this we need to introduce two last concepts things. First, let us now define a monomial ordering in general and the lexicographic monomial order. A monomial ordering on $\mathbb{Q}[x_1, \ldots, x_n]$ is a relation $\succ$ on $\mathbb{Z}_+^n$ (i.e. the monomial exponents) such that (1) the relation is a total ordering, (2) if $\alpha \succ \beta$ and $\gamma \in \mathbb{Z}_+^n$, then $\alpha + \gamma \succ \beta + \gamma$, and (3) every set has a smallest element. The lexicographic term monomial order, $\succ^{lex}$, is defined as $\alpha \succ^{lex} \beta$ if the left-most nonzero entry of $\alpha - \beta$ is positive.

Second, let us define a particular class of ideals, the radical ideals. The radical ideal of an ideal I is defined as

$$\sqrt{I} = \left\{ f \subseteq \mathbb{Q}[x_1, \ldots, x_n] \,|\exists k \geq 1 \text{ such that } f^k \in I \right\}.$$

Clearly, $I \subseteq \sqrt{I}$. If $I = \sqrt{I}$, the ideal I is called a radical ideal. Why do we need our ideal to be radical? We will explain the reason below. But first, let us introduce the Shape Lemma that is going to be the crucial theoretical tool in our approach to solve the quadratic system (19).

Theorem 4 *Let $\langle f_1, \ldots, f_m \rangle$ be a zero-dimensional ideal in $\mathbb{Q}[x_1, \ldots, x_n]$ such that all d complex roots of $f_1, \ldots, f_m$ have distinct last coordinate. The reduced Gröbner basis of $\langle f_1, \ldots, f_m \rangle$ in the lexicographic term monomial order is*

$$\mathcal{G} = \{x_1 - q_1(x_1), \ldots, x_{n-1} - q_{n-1}(x_{n-1}), r(x_n)\}$$

where r is a polynomial of degree d and the q_i are polynomials of degree $\leq d-1$ for all $1 \leq i \leq n-1$.

Proof. See Becker *et al.* (1996). ■

The reduced Gröbner basis consists of one univariate equation $r(x_n)$ in the last variable and $n-1$ equations that depend only on a (different) single variable x_i and the last variable. These $n-1$ equations are linear in x_i The Shape Lemma tell us how how to find all solutions to a polynomial system of equations. We compute the reduced Gröbner basis $\mathcal{G}$. Then we need to find all solutions to $\mathcal{G}$. Remember that $\mathcal{G}$ and $f_1, \ldots, f_m$ share the same set of complex roots. So finding the set complex roots of $\mathcal{G}$ is equivalent to find the set of complex roots of $f_1, \ldots, f_m$. But finding the set of complex roots of $\mathcal{G}$ is easy. We need to solve the univariate equation $r(x_n) = 0$ in the last variable. The values for all other variables are then trivially given by the remaining linear equations. Finding all solutions to a complicated system of polynomial equations in many variables thus requires determining the reduced Gröbner basis and finding all solutions to a univariate polynomial equation. The second step is trivial.

There is a large literature on the computation of the reduced Gröbner basis. Buchberger's algorithm is the most popular one. We refer the interested reader to Cox *et al.* (1997). In this paper we use Mathematica to compute the reduced Gröbner basis.

4.2.2 Example

As an example, consider the following system of polynomial equations of four quadratic equations in four unknowns:

$$\begin{aligned} x_1x_2 + x_3x_4 + 2 &= 0, \\ x_1x_2 + x_2x_3 + 3 &= 0, \\ x_1x_3 + x_4x_1 + x_4x_2 + 6 &= 0, \text{ and} \\ x_1x_3 + 2x_1x_2 + 3 &= 0. \end{aligned}$$

A Gröbner basis, with respect to the lexicographic term monomial order is

$$\begin{aligned} &-49 - 19x_4^2 + 9x_4^4 + 3x_4^6, \\ &2x_4 + 9x_4^3 + 3x_4^5 + 14x_3, \\ &-99x_4 + 6x_4^3 + 9x_4^5 + 28x_2, \text{ and} \\ &15x_4 - 6x_4^3 - 9x_4^5 + 28x_1. \end{aligned}$$

Note that the first element of the basis is a univariate equation in x_4 only. Given a root x_4 of the first polynomial, the second polynomial is linear in x_3, the third is linear in x_2, and the last is linear in x_1. Solving the first element of the basis produces the following six solutions

$$\{x_4 = -1.55461, x_4 = -1.39592i, x_4 = 1.39592i, x_4 = -1.86232i, x_4 = 1.86232i, x_4 = 1.55461\}$$

Solving the other three basis, conditional on these solutions, gives the following roots

$$\{x_3 = 4.58328, x_3 = -0.41342i, x_3 = 0.41342i, x_3 = 0.914097i, x_3 = 0.914097i, x_3 = -4.58328\},$$

$$\{x_2 = -1.7728, x_2 = -3.81477i, x_2 = 3.81477i, x_2 = -0.768342i, x_2 = 0.768342i, x_2 = 1.7728\}$$

and

$$\{x_1 = -2.89104, x_1 = -0.37300i, x_1 = 0.37300i, x_1 = -4.81861i, x_1 = 4.81861i, x_1 = 2.89104\}.$$

It is easy to show that these roots solve the original system of four quadratic equations in four unknowns.

4.3 Mean Square Stability

The issue now is whether any of the solutions are stable, and if so, how many. In a typical fixed regime model, determinacy is easily verified by checking whether the number of eigenvalues of the quadratic system (19) inside the unit circle equals to the number of predetermined variables. In a model with Markov switching the problem is more subtle. As shown in Farmer *et al.* (2009), it is possible that the number of stable eigenvalues associated with each of the regimes is equal to the number of predetermined variables but the system, as a whole, does not have a stable solution under several concepts of stability. The good news is that the Markov switching model can be checked for mean-square stability (MSS), as defined in Costa *et al.* (2005). In particular, MSS requires checking if the following matrix has its eigenvalues inside the unit circle

$$T = \left(P' \otimes I_{n_x^2}\right) diag\left[\mathcal{D}_{1,n_x} h_{ss}(s) \otimes \mathcal{D}_{1,n_x} h_{ss}(s)\right] \tag{20}$$

where

$$diag\left[\mathcal{D}_{1,n_x} h_{ss}(s) \otimes \mathcal{D}_{1,n_x} h_{ss}(s)\right] =$$

$$\begin{bmatrix} \mathcal{D}_{1,n_x} h_{ss}(1) \otimes \mathcal{D}_{1,n_x} h_{ss}(1) & \cdots & 0_{n_x^2 \times n_x^2} \\ 0_{n_x^2 \times n_x^2} & \cdots & 0_{n_x^2 \times n_x^2} \\ \vdots & \ddots & \vdots \\ 0_{n_x^2 \times n_x^2} & \cdots & \mathcal{D}_{1,n_x} h_{ss}(n_s) \otimes \mathcal{D}_{1,n_x} h_{ss}(n_s) \end{bmatrix}.$$

Thus, with Markov switching, the policy functions $\{\mathcal{D}_{1,n_x} h_{ss}(s_t)\}_{s_t=1}^{n_s}$ associated to each of the solutions of the quadratic system (19) must be checked for MSS under (20). If only one policy function is stable then the model only has one stable solution. If more than one are stable, the model has multiple stable solutions. If none are stable, the model has no stable solutions.

5 Second Order Approximation

Having constructed the first-order approximations, this Section shows how to find the second order Taylor expansions to g and h around the point $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$. Thus, we have

$$g(x_{t-1}, \varepsilon_t, \chi, s_t) - y_{ss} \simeq \mathcal{D}g_{ss}(s_t)\,\mathcal{S}_t + \frac{1}{2}\mathcal{H}g_{ss}(s_t)\left(\mathcal{S}_t \otimes \mathcal{S}_t\right)$$

where $\mathcal{H}g_{ss}(s_t)$ is a $n_y \times (n_x + n_\varepsilon + 1)^2$ matrix related to the second-order partial derivatives of g and $\mathcal{S}_t = \begin{pmatrix} (x_{t-1} - x_{ss})' & \varepsilon_t' & \chi \end{pmatrix}'$ is a vector of size $(n_x + n_\varepsilon + 1) \times 1$ that reflects the states of the model and

$$h(x_{t-1}, \varepsilon_t, \chi, s_t) - x_{ss} \simeq \mathcal{D}h_{ss}(s_t)\mathcal{S}_t + \frac{1}{2}\mathcal{H}h_{ss}(s_t)(\mathcal{S}_t \otimes \mathcal{S}_t)$$

where $\mathcal{H}h_{ss}(s_t)$ is a $n_y \times (n_x + n_\varepsilon + 1)^2$ matrix related to the second-order partial derivatives of h.

Define

$$\mathcal{H}g_{ss}^i(s_t) = \mathcal{H}g_{ss}^i(x_{ss}, 0_{n_\varepsilon}, 0, s_t) = \left[\mathcal{D}_k\mathcal{D}_j g^i(x_{ss}, 0_{n_\varepsilon}, 0, s_t)\right]_{1 \le j,k \le n_x + n_\varepsilon + 1}$$

to be the $(n_x + n_\varepsilon + 1) \times (n_x + n_\varepsilon + 1)$ matrix of second-order partial derivatives of g^i with respect to $(x_{t-1}, \varepsilon_t, \chi)$ evaluated at $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ for all s_t and for $1 \le i \le n_x$, and

$$\mathcal{H}h_{ss}^i(s_t) = \mathcal{H}h_{ss}^i(x_{ss}, 0_{n_\varepsilon}, 0, s_t) = \left[\mathcal{D}_k\mathcal{D}_j h^i(x_{ss}, 0_{n_\varepsilon}, 0, s_t)\right]_{1 \le j,k \le n_x + n_\varepsilon + 1}$$

to be the $(n_x + n_\varepsilon + 1) \times (n_x + n_\varepsilon + 1)$ matrix of second-order partial derivatives of h^i with respect to $(x_{t-1}, \varepsilon_t, \chi)$ evaluated at $(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ for all s_t and for $1 \le i \le n_y$. Note that both $\mathcal{H}g_{ss}^i(s_t)$ and $\mathcal{H}h_{ss}^i(s_t)$ are Hessian matrices, and are hence symmetric, with the row k, column j elements of each denoted $\mathcal{D}_k\mathcal{D}_j g^i(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$ and $\mathcal{D}_k\mathcal{D}_j h^i(x_{ss}, 0_{n_\varepsilon}, 0, s_t)$.

Then the matrices of second-order partial derivatives of g and h are related to the the matrices of second-order partial derivatives of g^i and h^i by

$$\mathcal{H}g_{ss}(s_t) = \begin{bmatrix} vec\left(\mathcal{H}g_{ss}^1(s_t)\right)' \\ \vdots \\ vec\left(\mathcal{H}g_{ss}^{n_y}(s_t)\right)' \end{bmatrix}$$

and

$$\mathcal{H}h_{ss}(s_t) = \begin{bmatrix} vec\left(\mathcal{H}h_{ss}^1(s_t)\right)' \\ \vdots \\ vec\left(\mathcal{H}h_{ss}^{n_x}(s_t)\right)' \end{bmatrix}.$$

Similar to the first-order approximations, the objective is now to find the coefficients in the matrices of second-order partial derivatives

$$\left\{\left\{\mathcal{H}g_{ss}^{i}\left(s_t\right)\right\}_{i=1}^{n_y},\left\{\mathcal{H}h_{ss}^{i}\left(s_t\right)\right\}_{i=1}^{n_x}\right\}_{s_t=1}^{n_s} \tag{21}$$

in the above described expansions, where $\mathcal{H}g_{ss}^{i}\left(s_t\right)\in\mathbb{C}^{(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)}$ for all s_t and for $1\leq i\leq n_y$ and $\mathcal{H}h_{ss}^{i}\left(s_t\right)\in\mathbb{C}^{(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)}$ for all s_t and and for $1\leq i\leq n_x$, by using the second-order partial derivatives of the function $\mathbb{G}$ evaluated at steady state. The remainder of the paper will use the following notation

$$\mathcal{H}\mathbb{G}^{i}\left(x_{t-1},\varepsilon_t,\chi,s_t\right)=\left[\mathcal{D}_k\mathcal{D}_j\mathbb{G}^{i}\left(x_{t-1},\varepsilon_t,\chi,s_t\right)\right]_{1\leq j,k\leq n_x+n_\varepsilon+1}$$

to refer the $(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)$ matrix of second-order partial derivatives of $\mathbb{G}^i$ with respect to $(x_{t-1},\varepsilon_t,\chi)$ evaluated at $(x_{t-1},\varepsilon_t,\chi,s_t)$ for all x_{t-1},ε_t, and s_t and $1\leq i\leq n_y+n_x$. Equivalently,

$$\mathcal{H}\mathbb{G}^{i}\left(x_{ss},0_{n_\varepsilon},0,s_t\right)=\left[\mathcal{D}_k\mathcal{D}_j\mathbb{G}^{i}\left(x_{ss},0_{n_\varepsilon},0,s_t\right)\right]_{1\leq j,k\leq n_x+n_\varepsilon+1}$$

refers to the $(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)$ matrix of second-order partial derivatives of $\mathbb{G}^i$ with respect to $(x_{t-1},\varepsilon_t,\chi)$ evaluated at $(x_{ss},0_{n_\varepsilon},0,s_t)$ for all s_t and $1\leq i\leq n_y+n_x$. To simplify notation define

$$\mathcal{H}\mathbb{G}_{ss}^{i}\left(s_t\right)\equiv\mathcal{H}\mathbb{G}^{i}\left(x_{ss},0_{n_\varepsilon},0,s_t\right)\text{ and }\mathcal{D}_k\mathcal{D}_j\mathbb{G}_{ss}^{i}\left(s_t\right)\equiv\mathcal{D}_k\mathcal{D}_j\mathbb{G}^{i}\left(x_{ss},0_{n_\varepsilon},0,s_t\right)$$

for $1\leq i\leq n_y+n_x$, $1\leq j,k\leq n_x+n_\varepsilon+1$, and for all s_t. Thus,

$$\mathcal{H}\mathbb{G}_{ss}^{i}\left(s_t\right)=\left[\mathcal{D}_k\mathcal{D}_j\mathbb{G}_{ss}^{i}\left(s_t\right)\right]_{1\leq j,k\leq n_x+n_\varepsilon+1}$$

for all s_t and $1\leq i\leq n_y+n_x$.

The coefficients of these policy functions are going to be obtained by using the fact that

$$\mathbb{G}\left(x_{t-1},\varepsilon_t,\chi,s_t\right)=0_{n_y+n_x}$$

for all $x_{t-1},\varepsilon_t,\chi$, and s_t and, therefore, it must be the case that

$$\mathcal{H}\mathbb{G}^{i}\left(x_{t-1},\varepsilon_t,\chi,s_t\right)=0_{(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)}$$

for all $x_{t-1}, \varepsilon_t, \chi$, s_t and $1 \leq i \leq n_y + n_x$, in particular,

$$\mathcal{H}\mathbb{G}_{ss}^i = 0_{(n_x+n_\varepsilon+1)\times(n_x+n_\varepsilon+1)}$$

for all s_t and $1 \leq i \leq n_y + n_x$. Thus,

$$\begin{aligned}
\mathcal{H}_{1,n_x;1,n_x}\mathbb{G}_{ss}^i(s_t) &= 0_{n_x\times n_x},\ \mathcal{H}_{1,n_x;n_x+1,n_\varepsilon}\mathbb{G}_{ss}^i(s_t) = 0_{n_x\times n_\varepsilon}, \\
\mathcal{H}_{n_x+1,n_\varepsilon;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t) &= 0_{n_x},\ \mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon}\mathbb{G}_{ss}^i(s_t) = 0_{n_\varepsilon\times n_\varepsilon}, \\
\mathcal{H}_{n_x+1,n_\varepsilon,n_\varepsilon;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t) &= 0_{n_\varepsilon},\ \text{and}\ \mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t) = 0.
\end{aligned} \tag{22}$$

for all s_t and $1 \leq i \leq n_y + n_x$, where, to simply notation, we have used the following definition

$$\mathcal{H}_{n_1,n_2;m_1,m_2}\mathbb{G}_{ss}^i(s_t) = \begin{bmatrix} \mathcal{H}_{n_1;m_1}\mathbb{G}_{ss}^i(s_t) & \dots & \mathcal{H}_{n_1;m_2}\mathbb{G}_{ss}^i(s_t) \\ \vdots & \ddots & \vdots \\ \mathcal{H}_{n_2;m_1}\mathbb{G}_{ss}^i(s_t) & \dots & \mathcal{H}_{n_2;m_2}\mathbb{G}_{ss}^i(s_t) \end{bmatrix}$$

for all s_t and $1 \leq n_1, m_1 \leq n_2, m_2 \leq n_x + n_\varepsilon + 1$.

The appendix shows how to use equations (22) to obtain (21): $\mathcal{H}_{1,n_x;1,n_x}\mathbb{G}_{ss}^i(s_t) = 0_{n_x\times n_x}$ determines $\left\{\{\mathcal{H}_{1,n_x;1,n_x}g_{ss}^i(s_t), \mathcal{H}_{1,n_x;1,n_x}h_{ss}^i(s_t)\}_{i=1}^{n_x+n_y}\right\}_{s_t=1}^{n_s}$, $\mathcal{H}_{1,n_x;n_x+1,n_\varepsilon}\mathbb{G}_{ss}^i(s_t) = 0_{n_x\times n_\varepsilon}$ determines $\left\{\{\mathcal{H}_{1,n_x;n_x+1,n_\varepsilon}g_{ss}^i(s_t), \mathcal{H}_{1,n_x;n_x+1,n_\varepsilon}h_{ss}^i(s_t)\}_{i=1}^{n_x+n_y}\right\}_{s_t=1}^{n_s}$, $\mathcal{H}_{n_x+1,n_\varepsilon;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t) = 0_{n_x}$ determines $\left\{\{\mathcal{H}_{n_x+1,n_\varepsilon;n_x+n_\varepsilon+1}g_{ss}^i(s_t), \mathcal{H}_{n_x+1,n_\varepsilon;n_x+n_\varepsilon+1}h_{ss}^i(s_t)\}_{i=1}^{n_x+n_y}\right\}_{s_t=1}^{n_s}$, $\mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon}\mathbb{G}_{ss}^i(s_t) = 0_{n_\varepsilon\times n_\varepsilon}$ determines $\left\{\{\mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon}g_{ss}^i(s_t), \mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon}h_{ss}^i(s_t)\}_{i=1}^{n_x+n_y}\right\}_{s_t=1}^{n_s}$, $\mathcal{H}_{n_x+1,n_\varepsilon,n_\varepsilon;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t) = 0_{n_\varepsilon}$ determines $\left\{\{\mathcal{H}_{n_x+1,n_\varepsilon,n_\varepsilon;n_x+n_\varepsilon+1}g_{ss}^i(s_t), \mathcal{H}_{n_x+1,n_\varepsilon,n_\varepsilon;n_x+n_\varepsilon+1}h_{ss}^i(s_t)\}_{i=1}^{n_x+n_y}\right\}_{s_t=1}^{n_s}$, and $\mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t) = 0$ determines $\left\{\{\mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}g_{ss}^i(s_t), \mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}h_{ss}^i(s_t)\}_{i=1}^{n_x+n_y}\right\}_{s_t=1}^{n_s}$.

The equations in the appendix also imply two important results. First, conditional on obtaining the first order solution for the policy functions, each of the equations in (22) is linear in the set of unknowns. As a result, given the solution first-order approximation, the second-order approximation follows as the solution of a linear system. See appendix for details.

Second, lack of certainty equivalence matters in the second order as well. For the derivatives $\mathcal{H}_{1,n_x;n_x+n_\varepsilon+1}\mathbb{G}_{ss}^i(s_t)$ and $\mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon}\mathbb{G}_{ss}^i(s_t)$, they are homogenous systems in the unknowns

$\mathcal{H}_{n_x+1,n_\varepsilon;n_x+n_\varepsilon+1} g^i_{ss}(s_t)$ and $\mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon} h^i_{ss}(s_t)$ if $\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) = 0_{n_y}$ and $\mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) = 0_{n_x}$, which means the first order is certainty equivalent. The following Proposition summarizes this result.

Proposition 5 *Let $\widehat{\theta}_1(s_t) = 0_{n_{\theta_1}}$ for all s_t. Then, by Proposition 1,*

$$\mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}(s_t) = 0_{n_y} \text{ and } \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}(s_t) = 0_{n_x}$$

for all s_t and the first order approximation is certainty equivalent. In addition

$$\mathcal{H}^i_{1,n_x;n_x+n_\varepsilon+1} \mathbb{G}_{ss}(s_t) = 0_{n_x} \text{ and } \mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon} \mathbb{G}^i_{ss}(s_t) = 0_{n_\varepsilon}$$

are homogenous systems in the unknowns, and so

$$\mathcal{H}_{n_x+1,n_\varepsilon;n_x+n_\varepsilon+1} g^i_{ss}(s_t) = 0_{n_x} \text{ and } \mathcal{H}_{n_x+1,n_\varepsilon;n_x+1,n_\varepsilon} h^i_{ss}(s_t) = 0_{n_\varepsilon}.$$

6 Examples

6.1 Example 1: The RBC Model

This section presents a simple exercise to illustrate the theoretical framework at hand. The perfect vehicle for such pedagogical effort is the real business cycle model. There are two reasons. First, the stochastic neoclassical growth model is the foundation of modern macroeconomics. Even the more complicated New Keynesian models, such as those in Woodford (2003) or Christiano *et al.* (2005), are built around the core of the neoclassical growth model augmented with nominal and real rigidities. Thus, after understanding how to deal with Markov switching in this prototype economy, it will be rather straightforward to extend it to richer environments such as the ones commonly used for policy analysis. Second, the model is so well known, its working so well understood, and its computation so thoroughly explored that the role of time-varying volatility in it will be staggeringly transparent.

Consider a real business cycle model where the TFP process follows a random walk in logs with drift. The drift will take one of two levels, high and low, depending on the value of

the Markov process. The random walk specification helps simplify the number of variables considered in a stationary equilibrium, and is hence the most parsimonious illustrative example. The specification of two regimes will allow a succinct discussion of the methodology, but, as mentioned above, more regimes can be handled easily within the framework.

To get into the substantive questions as soon as possible, the description of the standard features of the prototype economy will be limited to fix notation. There is a representative household in the economy, whose preferences over stochastic sequences of consumption, c_t, are represented by a utility function:

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \log c_t$$

where $\beta \in (0,1)$. The resource constraint is

$$c_t + k_t = z_t k_{t-1}^{\alpha} + (1-\delta) k_{t-1}$$

where k_t is capital and the technological change, z_t, proceeds according to a random walk in logs with drift where the Markov switching is in the drift, i.e.

$$\log z_t = \mu_t + \log z_{t-1} + \sigma \varepsilon_t$$

where the drift, μ_t, takes two values depending on the value of the Markov process, s_t.

For this model it is natural to work with the solution to the social planner's problem. The optimality conditions are standard:

$$\frac{1}{c_t} = \beta \mathbb{E}_t \frac{1}{c_{t+1}} \left(\alpha z_{t+1} k_t^{\alpha-1} + 1 - \delta \right)$$

and

$$c_t + k_t = z_t k_{t-1}^{\alpha} + (1-\delta) k_{t-1}.$$

Due to the unit root the economy is non-stationary. Thus, define $\omega_t = z_{t-1}^{\frac{1}{1-\alpha}}$, and let $\tilde{c}_t = \frac{c_t}{\omega_t}$, $\tilde{k}_{t-1} = \frac{k_{t-1}}{\omega_t}$, $\tilde{z}_t = \frac{z_t}{z_{t-1}}$. Then the re-scaled equilibrium conditions are

$$\frac{1}{\tilde{c}_t} = \beta \mathbb{E}_t \frac{\tilde{z}_t^{\frac{1}{\alpha-1}}}{\tilde{c}_{t+1}} \left(\alpha \tilde{z}_{t+1} \tilde{k}_t^{\alpha-1} + 1 - \delta \right),$$

$$\tilde{c}_t + \tilde{k}_t \tilde{z}_t^{\frac{1}{1-a}} = \tilde{z}_t \tilde{k}_{t-1}^{\alpha} + (1-\delta)\,\tilde{k}_{t-1},$$

and,

$$\log \tilde{z}_t = \mu_t + \sigma\varepsilon_t.$$

Substituting the expression for $\tilde{z}_t$, the conditions are then

$$\frac{1}{\tilde{c}_t} = \beta \mathbb{E}_t \frac{1}{\tilde{c}_{t+1}} \exp\left(\frac{\mu_t + \sigma\varepsilon_t}{\alpha - 1}\right)\left(\alpha \exp\left(\mu_{t+1} + \sigma\varepsilon_{t+1}\right)\tilde{k}_t^{\alpha-1} + 1 - \delta\right)$$

and

$$\tilde{c}_t + \tilde{k}_t \exp\left(\frac{\mu_t + \sigma\varepsilon_t}{1-\alpha}\right) = \exp\left(\mu_t + \sigma\varepsilon_t\right)\tilde{k}_{t-1}^{\alpha} + (1-\delta)\,\tilde{k}_{t-1}.$$

Using the notation in Section 2, $x_{t-1} = \tilde{k}_{t-1}$, $y_t = \tilde{c}_t$, and $\theta_t = \theta_{1t} = \mu_t$, so

$$f\left(y_{t+1}, y_t, x_t, x_{t-1}, \chi\varepsilon_{t+1}, \varepsilon_t, \theta_{t+1}, \theta_t\right) =$$
$$\begin{bmatrix} \frac{1}{\tilde{c}_t} - \beta \frac{1}{\tilde{c}_{t+1}} \exp\left(\frac{\mu_t + \sigma\varepsilon_t}{\alpha-1}\right)\left(\alpha \exp\left(\mu_{t+1} + \chi\sigma\varepsilon_{t+1}\right)\tilde{k}_t^{\alpha-1} + 1 - \delta\right) \\ \tilde{c}_t + \tilde{k}_t \exp\left(\frac{\mu_t + \sigma\varepsilon_t}{1-\alpha}\right) - \exp\left(\mu_t + \sigma_t\varepsilon_t\right)\tilde{k}_{t-1}^{\alpha} - (1-\delta)\,\tilde{k}_{t-1} \end{bmatrix}.$$

Clearly,

$$\tilde{c}_t = g\left(\tilde{k}_{t-1}, \varepsilon_t, \chi, s_t\right),$$
$$\tilde{c}_{t+1} = g\left(\tilde{k}_t, \chi\varepsilon_{t+1}, \chi, s_{t+1}\right),$$
$$\tilde{k}_t = h\left(\tilde{k}_{t-1}, \varepsilon_t, \chi, s_t\right),$$

and the drift, μ_t, has the following law of motion that depends on the Markov process

$$\mu_{t+1} = \mu\left(\chi, s_{t+1}\right) = \overline{\mu} + \chi\widehat{\mu}\left(s_{t+1}\right)$$

where $s_t \in \{1, 2\}$.

6.1.1 Solving the RBC Model

The following subsections show how to solve the model using a second-order approximation. The first step is to find the steady state. The second is to define the matrices of first-order partial derivatives of f with respect to all its components evaluated at steady state that we

use to solve for first-order approximation to the policy functions. Third, we find the second-order approximation to the policy functions (we omit the matrices that we use to solve for the second-order terms to simplify notation). We will consider two parametrization. The first parametrization has symmetric regimes. The second parametrization has asymmetric regimes.

6.1.2 Steady State

In order to calculate steady state, we set $\chi = 0$ and, hence, $\mu_t = \mu_{t+1} = \overline{\mu}$. Therefore, the steady state of the model are values $\tilde{c}_{ss}$ and $\tilde{k}_{ss}$ such that

$$\begin{bmatrix} \frac{1}{c_{ss}} - \beta \frac{1}{c_{ss}} \exp\left(\frac{\bar{\mu}}{\alpha-1}\right) \left(\alpha \exp\left(\bar{\mu}\right) \tilde{k}_{ss}^{\alpha-1} + 1 - \delta\right) \\ \tilde{c}_{ss} + \tilde{k}_{ss} \exp\left(\frac{\bar{\mu}}{1-\alpha}\right) - \exp\left(\bar{\mu}\right) \tilde{k}_{ss}^{\alpha} - \left(1-\delta\right) \tilde{k}_{ss} \end{bmatrix} = 0.$$

These two equations produce the steady state values

$$\tilde{k}_{ss} = \left(\frac{1}{\alpha \exp\left(\bar{\mu}\right)} \left(\frac{1}{\beta \exp\left(\frac{\bar{\mu}}{\alpha-1}\right)} - 1 + \delta\right)\right)^{\frac{1}{\alpha-1}}$$

and

$$\tilde{c}_{ss} = \exp\left(\bar{\mu}\right) \tilde{k}_{ss}^{\alpha} + \left(1-\delta\right) \tilde{k}_{ss} - \tilde{k}_{ss} \exp\left(\frac{\bar{\mu}}{1-\alpha}\right).$$

6.1.3 The Matrices

The next step is to define the matrices of partial derivatives of f with respect to all its components evaluated at steady state as introduced in Section in () and that we use to solve for first-order approximation to the policy functions. Recall in this example that $n_y = 1$, $n_x = 1$, $n_\varepsilon = 1$, and $n_\theta = 1$. Thus, the matrices are

$$\mathcal{D}_{1,1} f_{ss}\left(s', s\right) = \begin{bmatrix} \frac{1}{c_{ss}^2} \\ 0 \end{bmatrix}, \mathcal{D}_{2,2} f_{ss}\left(s', s\right) = \begin{bmatrix} -\frac{1}{c_{ss}^2} \\ 1 \end{bmatrix},$$

$$\mathcal{D}_{3,3} f_{ss}\left(s', s\right) = \begin{bmatrix} \left(1-\alpha\right) \alpha\beta \exp\left(\frac{\alpha\bar{\mu}}{\alpha-1}\right) \frac{k_{ss}^{\alpha-2}}{c_{ss}} \\ \exp\left(\frac{\bar{\mu}}{1-\alpha}\right) \end{bmatrix}, \mathcal{D}_{4,4} f_{ss}\left(s', s\right) = \begin{bmatrix} 0 \\ -\frac{\exp\left(\frac{\bar{\mu}}{1-\alpha}\right)}{\beta} \end{bmatrix},$$

$$\mathcal{D}_{5,5}f_{ss}\left(s',s\right)=\left[\begin{array}{c}-\alpha\beta\exp\left(\frac{\alpha\bar{\mu}}{\alpha-1}\right)\frac{k_{ss}^{\alpha-1}}{c_{ss}}\sigma\\0\end{array}\right],\mathcal{D}_{6,6}f_{ss}\left(s',s\right)=\left[\begin{array}{c}\frac{\sigma}{(1-\alpha)c_{ss}}\\\left(\frac{\exp\left(\frac{\bar{\mu}}{1-\alpha}\right)k_{ss}}{1-\alpha}-\exp\left(\bar{\mu}\right)k_{ss}^{\alpha}\right)\sigma\end{array}\right]$$

$$\mathcal{D}_{7,7}f_{ss}\left(s',s\right)=\left[\begin{array}{c}-\alpha\beta\exp\left(\frac{\bar{\mu}\alpha}{\alpha-1}\right)\frac{k_{ss}^{\alpha-1}}{c_{ss}}\\0\end{array}\right],\mathcal{D}_{8,8}f_{ss}\left(s',s\right)=\left[\begin{array}{c}\frac{1}{c_{ss}}\\\exp\left(\frac{\bar{\mu}}{1-\alpha}\right)\frac{k_{ss}}{1-\alpha}-\exp\left(\bar{\mu}\right)k_{ss}^{\alpha}\end{array}\right].$$

6.1.4 Symmetric Regimes

Consider first the following parametrization with symmetric regimes:

α	β	δ	σ	$\bar{\mu}$	$\mu(1)$	$\mu(2)$	$p_{1,1}$	$p_{2,2}$
0.3300	0.9976	0.0250	0.0002	0.00333	0.00167	−0.00163	0.90	0.90

The growth rates $\bar{\mu}+\mu(1)$ and $\bar{\mu}+\mu(2)$ correspond to regimes where the annual output growth rate is three or one percent, β produces a risk free rate of three percent annually, and σ is set to match the total volatility of TFP growth as estimated in Fernández-Villaverde & Rubio-Ramirez (2007). The standard calibration of α implies a capital share of one-third, and δ implies an annual depreciation rate of approximately ten percent. The steady state values of capital and consumption are $k_{ss}=32.0986$ and $c_{ss}=2.18946$. Consequently the numerical values of the matrices of first-order partial derivatives of f with respect to all its components evaluated at steady state are

$$\mathcal{D}_{1,1}f_{ss}\left(s',s\right)=\left[\begin{array}{c}0.20861\\0\end{array}\right],\ \mathcal{D}_{2,2}f_{ss}\left(s',s\right)=\left[\begin{array}{c}-0.2086\\1\end{array}\right],\ \mathcal{D}_{3,3}f_{ss}\left(s',s\right)=\left[\begin{array}{c}0.00031\\1.00499\end{array}\right]$$

$$\mathcal{D}_{4,4}f_{ss}\left(s',s\right)=\left[\begin{array}{c}0\\-1.0074\end{array}\right],\ \mathcal{D}_{5,5}f_{ss}\left(s',s\right)=\left[\begin{array}{c}0\\0\end{array}\right],\ \mathcal{D}_{6,6}f_{ss}\left(s',s\right)=\left[\begin{array}{c}0.00014\\0.00900\end{array}\right]$$

$$\mathcal{D}_{7,7}f_{ss}\left(s',s\right)=\left[\begin{array}{c}-0.0147\\0\end{array}\right],\ \text{and}\ \mathcal{D}_{8,8}f_{ss}\left(s',s\right)=\left[\begin{array}{c}0.68169\\44.9953\end{array}\right].$$

Using these matrices, we can solve the quadratic system (19) to get the following set of solutions for $\{\mathcal{D}_{1,1}g_{ss}\left(s_t\right),\mathcal{D}_{1,1}h_{ss}\left(s_t\right)\}_{s_t=1}^{n_s}$. In this example, since $n_x=1$ and there are no

exogenous variables, $n_y = 1$, and $n_s = 2$, Proposition (3) indicates that, at most, there are four total solutions. As shown below, there are, indeed, four solutions

	$D_{1,1}h_{ss}(1)$	$D_{1,1}g_{ss}(1)$	$D_{1,1}h_{ss}(2)$	$D_{1,1}g_{ss}(2)$
(1)	0.96364	0.03896	0.96364	0.03896
(2)	1.04023	-0.0380	1.04023	-0.0380
(3)	$1.11326 + 0.116871i$	$-0.1114 - 0.11745i$	$1.11326 + 0.11687i$	$-0.1114 - 0.11745i$
(4)	$1.11326 - 0.116871i$	$-0.1114 + 0.11745i$	$1.11326 - 0.11687i$	$-0.1114 + 0.11745i$

The only MSS solution is (1), hence the model is determined. Given solution (1) can find $\{\mathcal{D}_{2,2}g_{ss}(s_t), \mathcal{D}_{2,2}h_{ss}(s_t)\}_{s_t=1}^{n_s}$ and $\{\mathcal{D}_3 g_{ss}(s_t), \mathcal{D}_3 h_{ss}(s_t)\}_{s_t=1}^{n_s}$ for the first order approximation and $\left\{\{\mathcal{H}_{1,3;1,3}g_{ss}^i(s_t), \mathcal{H}_{1,3;1,3}h_{ss}^i(s_t)\}_{i=1}^2\right\}_{s_t=1}^{n_s}$ for the second-order approximation. Let $\mathcal{S}_t$ denote the state for time t. Then the second-order solution is, for $s_t = 1$:

$$\begin{bmatrix} \hat{c}_t \\ \hat{k}_t \end{bmatrix} = \begin{bmatrix} 0.03896 & 0.00028 & 0.00972 \\ 0.96364 & -0.0092 & -0.0843 \end{bmatrix} \mathcal{S}_t$$

$$+\frac{1}{2}\begin{bmatrix} -0.0004 & 4\times10^{-6} & 0.00016 & 4\times10^{-6} & 4\times10^{-8} & 1\times10^{-6} & 0.00016 & 1\times10^{-6} & -0.0003 \\ -0.0002 & -0.0003 & -0.0025 & -0.0003 & 2.7\times10^{-6} & 0.00002 & -0.0025 & 0.00002 & 0.00057 \end{bmatrix} (\mathcal{S}_t \otimes \mathcal{S}_t)$$

and for $s_t = 2$:

$$\begin{bmatrix} \hat{c}_t \\ \hat{k}_t \end{bmatrix} = \begin{bmatrix} 0.03896 & 0.00028 & -0.00972 \\ 0.96364 & -0.0092 & 0.0843 \end{bmatrix} \mathcal{S}_t$$

$$+\frac{1}{2}\begin{bmatrix} -0.0004 & 4\times10^{-6} & -0.0002 & 4\times10^{-6} & 4\times10^{-8} & -1\times10^{-6} & -0.0002 & -1\times10^{-6} & -0.0003 \\ -0.0002 & -0.0002 & 0.00251 & -0.0003 & 3\times10^{-6} & -2\times10^{-5} & 0.00251 & -2\times10^{-5} & 0.00057 \end{bmatrix} (\mathcal{S}_t \otimes \mathcal{S}_t)$$

This solution highlights the results of Propositions 1 and 5. Let us first analyze Proposition 1 that indicates that the first-order approximation is not certainty equivalent. Since $\widehat{\mu}(s_t) \neq 0$ for all s_t and

$$\sum_{s'=1}^{n_s} p_{s_t,s'} \left(\mathcal{D}_{7,7} f_{ss}(s', s_t)\, \mathcal{D}\theta_{ss}(s') + \mathcal{D}_{8,8} f_{ss}(s', s_t)\, \mathcal{D}\theta_{ss}(s_t)\right) \neq 0_2$$

for all s_t. As a result, the first-order partial derivatives of g and h with respect to χ are nonzero: $\mathcal{D}_3 g_{ss}(s_t) \neq 0$ and $\mathcal{D}_3 h_{ss}(s_t) \neq 0$.

Let us now analyze Proposition 5. This Proposition states that the cross-derivatives of χ with the other state variables in the second-order approximation are nonzero: $\mathcal{D}_i \mathcal{D}_3 g_{ss}(s_t) = \mathcal{D}_3 \mathcal{D}_i g_{ss}(s_t)\, g_{ss}(s_t) \neq 0$ and $\mathcal{D}_i \mathcal{D}_3 h_{ss}(s_t) = \mathcal{D}_3 \mathcal{D}_i h_{ss}(s_t) \neq 0$ for all $s_t = 1, 2$ and $1 \leq i \leq 2$.

6.1.5 Asymmetric Regimes

As an alternative parameterization, consider the same parameters above, but with $p_{1,1} = 0.5$, $\bar{\mu} = 0.00222$, $\mu(1) = 0.00278$, and $\mu(2) = 0.00052$. In this example, regime 1 has a shorter expected duration, and regime 2 occurs more frequently in the ergodic distribution. The steady state is $k_{ss} = 34.6774$, and $c_{ss} = 2.24769$. Also, with this alternative parametrization, the numerical values of the matrices of first-order partial derivatives of f will change with respect to the ones from the symmetric regimes presented above (in the interest of space we do not present them here). Using these new matrices, we can solve the quadratic system (19) for the alternative parameterization. Again, since $n_x = 1$ and there are no exogenous variables, $n_y = 1$, and $n_s = 2$, Proposition 3 indicates that, at most, there are four total solutions. As shown below, there are, indeed, four solutions

	$D_{1,1}h_{ss}(1)$	$D_{1,1}g_{ss}(1)$	$D_{1,1}h_{ss}(2)$	$D_{1,1}g_{ss}(2)$
(1)	0.96545	0.0370821	0.96545	0.0370821
(2)	1.03828	-0.035996	1.03828	-0.035996
(3)	$2.00373 - 0.7042i$	$-1.00465 + 0.70654i$	$1.11318 + 0.39122i$	$-0.111145 - 0.39252i$
(4)	$2.00373 + 0.7042i$	$-1.00465 - 0.70654i$	$1.11318 - 0.39122i$	$-0.111145 + 0.39252i$

The only MSS solution is (1), hence the model is determined. Given solution (1) can find the rest of the matrices needed to build the second-order approximation. We have, for $s_t = 1$

$$\begin{bmatrix} \hat{c}_t \\ \hat{k}_t \end{bmatrix} = \begin{bmatrix} 0.03708 & 0.00029 & 0.00637 \\ 0.96545 & -0.0100 & -0.1412 \end{bmatrix} \mathcal{S}_t$$
$$+\frac{1}{2}\begin{bmatrix} -0.0004 & 4\times10^{-6} & 0.00009 & 4\times10^{-6} & 5\times10^{-8} & 1\times10^{-6} & 0.00009 & 1\times10^{-6} & -5\times10^{-6} \\ -0.0002 & -0.0003 & -0.0040 & -0.0003 & 3\times10^{-6} & 0.00004 & -0.0040 & 0.00004 & 0.00065 \end{bmatrix} (\mathcal{S}_t \otimes \mathcal{S}_t)$$

and for $s_t = 2$

$$\begin{bmatrix} \hat{c}_t \\ \hat{k}_t \end{bmatrix} = \begin{bmatrix} 0.03708 & 0.00029 & -0.0013 \\ 0.96545 & -0.0100 & 0.02823 \end{bmatrix} \mathcal{S}_t$$
$$+\frac{1}{2}\begin{bmatrix} -0.0004 & 4\times10^{-6} & -1\times10^{-5} & 4\times10^{-6} & 5\times10^{-8} & -2\times10^{-7} & -1\times10^{-5} & -2\times10^{-7} & -6\times10^{-5} \\ -0.0002 & -0.0003 & 0.00080 & -0.0003 & 3\times10^{-6} & -8\times10^{-6} & 0.00080 & -8\times10^{-6} & 0.00009 \end{bmatrix} (\mathcal{S}_t \otimes \mathcal{S}_t)$$

As in the symmetric regime case, the first-order derivatives in the first-order approximation and the cross-derivatives in the second-order approximation with respect to χ are non-zero. However, given the asymmetry of regimes the terms reflecting the lack of certainty equivalence (first-order derivatives with respect to χ) are not simply negated across regimes.

6.1.6 Accuracy: Euler Equation Errors

In this section, we will compare the accuracy of the first-order and second-order approximations. For that purpose we will use the Euler equation errors as reported in Judd (1998) and Aruoba *et al.* (2006). The Euler equation error at point $\left(\tilde{k}_{t-1}, \varepsilon_t; s_t\right)$ for an approximation order $order \in \{\text{first,second}\}$:

$$EE^{order}\left(\tilde{k}_{t-1}, \varepsilon_t, \chi = 1; s_t\right) =$$

$$1 - \beta \int \sum_{s'=1}^{2} p_{s_t,s'} \left(\begin{array}{c} \frac{\tilde{c}^{order}\left(\tilde{k}_{t-1},\varepsilon_t,\chi=1;s_t\right)}{\tilde{c}^{order}\left(\tilde{k}^{order}\left(\tilde{k}_{t-1},\varepsilon_t,\chi=1;s_t\right),\varepsilon';1\right)} \exp\left(\frac{\mu(s_t)+\sigma\varepsilon_t}{\alpha-1}\right) \\ \times \left(\alpha \exp\left(\mu\left(s'\right) + \sigma\varepsilon'\right) \tilde{k}^{order}\left(\tilde{k}_{t-1}, \varepsilon_t, \chi = 1; s_t\right)^{\alpha-1} + 1 - \delta\right) \end{array} \right) \mu\left(\varepsilon'\right) d\varepsilon'$$

where $\tilde{c}^{order}\left(\tilde{k}_{t-1}, \varepsilon_t, \chi = 1; s_t\right)$ and $\tilde{k}^{order}\left(\tilde{k}_{t-1}, \varepsilon_t, \chi = 1; s_t\right)$ are the approximations to the policy functions. The unconditional absolute Euler equation error is

$$EE^{order} = \int \left| EE^{order}\left(\tilde{k}_{t-1}, \varepsilon_t, \chi = 1; s_t\right) \right| \mu^{order}\left(\tilde{k}_{t-1}, \varepsilon_t, \chi = 1; s_t\right) d\tilde{k}_{t-1} d\varepsilon_t ds_t$$

where μ^{order} is the unconditional distribution of the variables implied by the approximation. The table shows the base-10 logarithm of the absolute Euler equation errors for the first and second order approximations in both the symmetric and asymmetric examples. Both the first and second order approximations produce a high degree of accuracy: a value of -5 implies an error of \$1 for each \$100,000. In addition, the second order achieves an even higher degree of accuracy.

Unconditional Absolute Euler Equation Errors ($\log_{10}$)

	$p_{1,1} = 0.9$	$p_{1,1} = 0.5$
EE^{first}	-4.6099	-5.3929
EE^{second}	-5.4158	-6.1983

6.2 Example 2: NK Model

This section presents a second example: a simple New Keynesian model to highlight the issue of determinacy and mean square stability.

The model is a model with quadratic price adjustment costs where the monetary authority follows a Taylor Rule that changes according to a Markov Process. The reaction coefficient of monetary policy switches with the regime, which Davig & Leeper (2007), Farmer *et al.* (2008), and Bianchi (2009), among others, have argued captures the changing stance of policy in the United States.

A representative consumer maximizes expected lifetime utility over consumption C_t and hours worked H_t

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log C_t - H_t\right)$$

subject to the budget constraint

$$C_t + \frac{B_t}{P_t} = W_t H_t + R_{t-1} \frac{B_{t-1}}{P_t} + T_t + D_t$$

where B_t is next period's nominal bonds, W_t is the real wage, R_{t-1} is the nominal return on bonds, T_t is lump-sum transfers, and D_t is profits from firms.

A competitive final goods producer combines a continuum of intermediate goods $Y_{j,t}$ into a final good Y_t by a CES aggregator

$$Y_t = \left(\int_0^1 Y_{j,t}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}}$$

Intermediate goods firms take the wage and their demand function

$$Y_{j,t} = \left(\frac{P_{j,t}}{P_t} \right)^{-\eta} Y_t$$

as given and set their price $P_{j,t}$ demand hours $H_{j,t}$ to produce according to

$$Y_{j,t} = A_t H_{j,t}$$

where total factor productivity follows

$$\log A_t = \mu_t + \log A_{t-1}$$

where, similar to the RBC model in Section 6.1, the drift can take two values

$$\mu_t = \mu\left(s_t\right),\ s_t \in \{1,2\}.$$

These firms face quadratic price adjustment costs according to

$$AC_{j,t} = \frac{\kappa}{2}\left(\frac{P_{j,t}}{P_{j,t-1}} - 1\right)^2.$$

The monetary authority sets prices by a Taylor rule where the coefficient varies over time

$$\frac{R_t}{R_{ss}} = \left(\frac{R_{t-1}}{R_{ss}}\right)^{\rho} \Pi_t^{(1-\rho)\psi_t} \exp\left(\sigma\varepsilon_t\right)$$

In a symmetric equilibrium $P_{j,t} = P_t$, $Y_{j,t} = Y_t$, and $H_{j,t} = H_t$ for all j, and market clearing implies

$$Y_t = C_t + \frac{\kappa}{2}\left(\Pi_t - 1\right)^2 Y_t.$$

Using the notation in Section 2, $y_t = [\Pi_t, Y_t]'$, $x_{t-1} = R_{t-1}$, $\theta_{1t} = \mu_t$, and $\theta_{2t} = \psi_t$. Then the stationary equilibrium is expressed as

$$f\left(y_{t+1}, y_t, x_t, x_{t-1}, \chi\varepsilon_{t+1}, \varepsilon_t, \theta_{t+1}, \theta_t\right) =$$
$$\begin{bmatrix} 1 - \beta\frac{\left(1-\frac{\kappa}{2}(\Pi_t-1)^2\right)\tilde{Y}_t}{\left(1-\frac{\kappa}{2}(\Pi_{t+1}-1)^2\right)\tilde{Y}_{t+1}}\frac{1}{\exp\left(\mu_{t+1}\right)}\frac{R_t}{\Pi_{t+1}} \\ (1-\eta) + \eta\left(1 - \frac{\kappa}{2}\left(\Pi_t - 1\right)^2\right)\tilde{Y}_t + \beta\kappa\frac{\left(1-\frac{\kappa}{2}(\Pi_t-1)^2\right)}{\left(1-\frac{\kappa}{2}(\Pi_{t+1}-1)^2\right)}\left(\Pi_{t+1} - 1\right)\Pi_{t+1} - \kappa\left(\Pi_t - 1\right)\Pi_t \\ \left(\frac{R_{t-1}}{R_{ss}}\right)^{\rho}\Pi_t^{(1-\rho)\psi_t}\exp\left(\sigma\varepsilon_t\right) - \frac{R_t}{R_{ss}} \end{bmatrix}$$

6.2.1 Solving the NK Model

Similar to the RBC example, the following subsections show how to solve the model: find the steady state, define the matrices of first-order partial derivatives of f with respect to all its components evaluated at steady state that we use to solve for first-order approximation to the policy functions find the second-order approximation to the policy functions. Two parameterizations show how using Gröbner basis and MSS allows checking for determinacy.

6.2.2 Steady State

In order to calculate steady state, set $\chi = 0$. Therefore, $\Pi_t = \Pi_{t+1} = \Pi_{ss}$, $\tilde{Y}_t = \tilde{Y}_{t+1} = \tilde{Y}_{ss}$, $R_t = R_{t-1} = R_{ss}$, and $\mu_{t+1} = \mu_t = \bar{\mu}$. So the equilibrium conditions in steady state are

$$\begin{bmatrix} 1 - \beta \frac{\left(1-\frac{\kappa}{2}(\Pi_{ss}-1)^2\right)\tilde{Y}_{ss}}{\left(1-\frac{\kappa}{2}(\Pi_{ss}-1)^2\right)\tilde{Y}_{ss}} \frac{1}{\exp(\bar{\mu})} \frac{R_{ss}}{\Pi_{ss}} \\ (1-\eta) + \eta\left(1 - \frac{\kappa}{2}\left(\Pi_{ss}-1\right)^2\right)\tilde{Y}_{ss} + \beta\kappa \frac{\left(1-\frac{\kappa}{2}(\Pi_{ss}-1)^2\right)}{\left(1-\frac{\kappa}{2}(\Pi_{ss}-1)^2\right)}\left(\Pi_{ss}-1\right)\Pi_{ss} - \kappa\left(\Pi_{ss}-1\right)\Pi_{ss} \\ \left(\frac{R_{ss}}{R_{ss}}\right)^{\rho} \Pi_{ss}^{(1-\rho)\psi_t} - \frac{R_{ss}}{R_{ss}} \end{bmatrix} = 0_{3\times 1}$$

Using the assumption $\Pi_{ss} = 1$, and solving these produces the steady state values

$$R_{ss} = \frac{\exp\left(\bar{\mu}\right)}{\beta},$$

and

$$\tilde{Y}_{ss} = \frac{\eta - 1}{\eta}.$$

Note that $\bar{\mu}$ affects the steady state, but $\psi\left(s\right)$ does not, demonstrating the partition of the switching vector $\theta_t = [\theta_{1t}, \theta_{2t}]$.

6.2.3 The Matrices

The next step is to define the matrices in expression (19), which depend on the derivatives of the function f evaluated at the steady state. Recall in this example that $n_y = 2$, $n_x = 1$, $n_\varepsilon = 1$, and $n_\theta = 2$. The necessary matrices are

$$\mathcal{D}_{1,2} f_{ss}\left(s', s\right) = \begin{bmatrix} \frac{\eta}{\eta-1} & 1 \\ 0 & \beta\kappa \\ 0 & 0 \end{bmatrix}, \mathcal{D}_{3,4} f_{ss}\left(s', s\right) = \begin{bmatrix} \frac{\eta}{1-\eta} & 0 \\ \eta & -\kappa \\ 0 & \left(1-\rho\right)\psi\left(s\right) \end{bmatrix}$$

$$\mathcal{D}_{5,5} f_{ss}\left(s', s\right) = \begin{bmatrix} -\beta e^{-\bar{\mu}} \\ 0 \\ -\beta e^{-\bar{\mu}} \end{bmatrix}, \mathcal{D}_{6,6} f_{ss}\left(s', s\right) = \begin{bmatrix} 0 \\ 0 \\ \rho\beta e^{-\bar{\mu}} \end{bmatrix}, \mathcal{D}_{7,7} f_{ss}\left(s', s\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathcal{D}_{8,8} f_{ss}\left(s', s\right) = \begin{bmatrix} 0 \\ 0 \\ \sigma \end{bmatrix}, \mathcal{D}_{9,10} f_{ss}\left(s', s\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathcal{D}_{11,12} f_{ss}\left(s', s\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

6.2.4 Determinate Solution

Consider the following parameterization

β	κ	η	ρ	σ	$p_{1,1}$	$p_{2,2}$	$\bar{\mu}$	$\hat{\mu}(1)$	$\hat{\mu}(2)$	$\psi(1)$	$\psi(2)$
0.9976	162	10	0.8	0.0025	0.90	0.90	0.005	0.0025	−0.0025	3.1	0.9

The growth rates $\bar{\mu} + \hat{\mu}(1)$ and $\bar{\mu} + \hat{\mu}(2)$ again correspond to regimes where annual growth rate is three or one percent, β produces a risk-free rate of three percent annually, η produces a steady state markup of 11 percent, and ρ and σ match estimates in Fernandez-Villaverde *et al.* (2009). The two monetary policy parameters $\psi(1)$ and $\psi(2)$ are such that, in two fixed regime models, regime 1 would imply a determinate equilibrium, and regime 2 would imply indeterminacy. The steady state values of the nominal rate and output are $R_{ss} = 1.0074$ and $\tilde{Y}_{ss} = 0.90$. Consequently the numerical values of the matrices of first-order partial derivatives of f with respect to all its components evaluated at steady state are

$$\mathcal{D}_{1,2} f_{ss}(s', s) = \begin{bmatrix} 1.11111 & 1 \\ 0 & 160.614 \\ 0 & 0 \end{bmatrix}, \mathcal{D}_{3,4} f_{ss}(s', s) = \begin{bmatrix} -1.11111 & 0 \\ 10 & -161. \\ 0 & 0.2\psi(s) \end{bmatrix},$$

$$\mathcal{D}_{5,5} f_{ss}(s', s) = \begin{bmatrix} -0.9926 \\ 0 \\ -0.9926 \end{bmatrix}, \mathcal{D}_{6,6} f_{ss}(s', s) = \begin{bmatrix} 0 \\ 0 \\ 0.77941 \end{bmatrix}, \mathcal{D}_{7,7} f_{ss}(s', s) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$$\mathcal{D}_{8,8} f_{ss}(s', s) = \begin{bmatrix} 0 \\ 0 \\ 0.0025 \end{bmatrix}, \mathcal{D}_{9,10} f_{ss}(s', s) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \mathcal{D}_{11,12} f_{ss}(s', s) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Using these matrices, we can solve the quadratic system to get the following set of solutions for $\{\mathcal{D}_{1,1} g_{ss}(s_t), \mathcal{D}_{1,1} h_{ss}(s_t)\}_{s_t=1}^{n_s}$. In this example, since $n_x = 1$ and there are no exogenous variables, $n_y = 2$, and $n_s = 2$, Proposition 3 indicates that, at most, there are 9 total solutions.

As shown below, there are exactly 9 solutions

	$D_{1,1}h_{ss}(1)$	$D_{1,1}g_{ss}(1)'$		$D_{1,1}h_{ss}(2)$	$D_{1,1}g_{ss}(2)'$	
(1)	0.59517	-1.92815	-0.327932	0.699414	-2.9541	-0.554689
(2)	0.77508	-3.64018	-0.0398952	1.3018	-7.43725	2.76721
(3)	0.79559	-1.82393	-0.00706061	1.05423	1.21892	1.40196
(4)	$1.0939-0.4363i$	$-0.8264+4.2641i$	$0.4706-0.6986i$	$1.3311+0.0574i$	$-10.008-1.9739i$	$2.9287+0.3165i$
(5)	$1.0939+0.4363i$	$-0.8264-4.2641i$	$0.4706+0.6986i$	$1.3311-0.0574i$	$-10.008+1.9739i$	$2.9287-0.3165i$
(6)	$1.0952-0.2105i$	$-0.9833+1.9595i$	$0.4727-0.3370i$	$1.0240-0.0200i$	$0.8689+0.7833i$	$1.2351-0.1103i$
(7)	$1.0952+0.2105i$	$-0.9833-1.9595i$	$0.4727+0.3370i$	$1.0240+0.0200i$	$0.8689-0.7833i$	$1.2351+0.1103i$
(8)	$1.2360-0.2511i$	$0.7554+3.0821i$	$0.6980-0.4020i$	$0.7507+0.0047i$	$-2.2696+0.6345i$	$-0.2718+0.0260i$
(9)	$1.2360+0.2511i$	$0.7554-3.0821i$	$0.6980+0.4020i$	$0.7507-0.0047i$	$-2.2696-0.6345i$	$-0.2718-0.0260i$

The only MSS solution is (1), hence the model is determined. Given solution (1), we can find $\{\mathcal{D}_{2,2}g_{ss}(s_t), \mathcal{D}_{2,2}h_{ss}(s_t)\}_{s_t=1}^{n_s}$ and $\{\mathcal{D}_{3,3}g_{ss}(s_t), \mathcal{D}_{3,3}h_{ss}(s_t)\}_{s_t=1}^{n_s}$ for the first-order approximation and $\left\{\{\mathcal{H}_{1,3;1,3}g_{ss}^i(s_t), \mathcal{H}_{1,3;1,3}h_{ss}^i(s_t)\}_{i=1}^{3}\right\}_{s_t=1}^{n_s}$ for the second-order approximation. Letting $\mathcal{S}_t$ denote the state for time t, the second-order solution is, for $s_t = 1$:

$$\begin{bmatrix} \hat{\tilde{Y}}_t \\ \hat{\Pi}_t \\ \hat{R}_t \end{bmatrix} = \begin{bmatrix} -1.9282 & -0.0062 & 0.00481 \\ -0.3279 & -0.0011 & 0.00014 \\ 0.59517 & 0.00191 & 0.00008 \end{bmatrix} \mathcal{S}_t$$

$$+ \begin{bmatrix} 21.3771 & 0.06247 & -0.0188 & 0.06247 & 0.00020 & -0.0001 & -0.0188 & 0.00020 & -0.0004 \\ 0.49793 & 0.00056 & -0.0008 & 0.00056 & 2\times10^{-6} & -2\times10^{-6} & -0.0008 & 2\times10^{-6} & -0.0003 \\ -0.1986 & 0.00124 & -0.0004 & 0.00124 & 4\times10^{-6} & -1\times10^{-6} & -0.0004 & 4\times10^{-6} & -0.0002 \end{bmatrix} (\mathcal{S}_t \otimes \mathcal{S}_t)$$

and for $s_t = 2$:

$$\begin{bmatrix} \hat{\tilde{Y}}_t \\ \hat{\Pi}_t \\ \hat{R}_t \end{bmatrix} = \begin{bmatrix} -2.9541 & -0.0090 & -0.0094 \\ -0.5547 & -0.0017 & -0.0026 \\ 0.69941 & 0.00214 & -0.0006 \end{bmatrix} \mathcal{S}_t$$

$$+ \begin{bmatrix} 56.9733 & 0.16487 & 0.23174 & 0.16487 & 0.00050 & 0.00071 & 0.23174 & 0.00050 & -0.0016 \\ 0.99333 & 0.00136 & 0.00033 & 0.00136 & 4\times10^{-6} & 1\times10^{-6} & 0.00033 & 4\times10^{-6} & -0.0010 \\ -0.1842 & 0.00160 & -0.0005 & 0.00160 & 5\times10^{-6} & -2\times10^{-6} & -0.0005 & 5\times10^{-6} & -0.0002 \end{bmatrix} (\mathcal{S}_t \otimes \mathcal{S}_t)$$

As in the RBC example, the first-order derivatives in the first-order approximation and the cross-derivatives in the second-order approximation with respect to χ are non-zero. However, in this example, the derivatives with respect to R_{t-1} and ε_t are different across regimes, which is a result of the fact that $\psi(s_t)$ affects these derivatives.

6.2.5 Indeterminate Solution

As an alternative parameterization, consider the same parameters above, but with $\psi(2) = 0.7$. In this example, the steady state is identical to the previous case, and the first regime would imply determinacy and the second indeterminacy if considered in isolation. However, the second regime now has a slightly lower response by the monetary authority to inflation. The numerical values of the first-order partial derivatives are largely unchanged, with the exception of

$$\mathcal{D}_{3,4} f_{ss}(s', s) = \begin{bmatrix} -1.11111 & 0 \\ 10 & -161. \\ 0 & 0.2\psi(s) \end{bmatrix}$$

which depends on $\psi(s)$. As shown below, there are nine total solutions to the quadratic system.

	$D_{1,1}h_{ss}(1)$	$D_{1,1}g_{ss}(1)'$		$D_{1,1}h_{ss}(2)$	$D_{1,1}g_{ss}(2)'$	
(1)	0.59067	-1.9452	-0.3351	0.71244	-3.2185	-0.6209
(2)	0.79733	-4.7813	-0.0043	1.32443	-11.313	3.71833
(3)	0.85231	-1.7727	0.08374	1.01525	2.03718	1.52618
(4)	$1.0444 - 0.4989i$	$-0.7000 + 4.6327i$	$0.3912 - 0.7987i$	$1.3523 + 0.0442i$	$-14.210 - 1.8733i$	$3.9161 + 0.3132i$
(5)	$1.0444 + 0.4989i$	$-0.7000 - 4.6327i$	$0.3912 + 0.7987i$	$1.3523 - 0.0442i$	$-14.210 + 1.8733i$	$3.9161 - 0.3132i$
(6)	$1.0670 - 0.1894i$	$-1.2845 + 1.5032i$	$0.4274 - 0.3033i$	$0.9995 - 0.0089i$	$1.6635 + 0.5703i$	$1.4141 - 0.0629i$
(7)	$1.0670 + 0.1894i$	$-1.2845 - 1.5032i$	$0.4274 + 0.3033i$	$0.9995 + 0.0089i$	$1.6635 - 0.5703i$	$1.4141 + 0.0629i$
(8)	$1.2374 - 0.2527i$	$0.8018 + 3.0980i$	$0.7004 - 0.4046i$	$0.7582 + 0.0045i$	$-2.3764 + 0.6699i$	$-0.2963 + 0.0317i$
(9)	$1.2374 + 0.2527i$	$0.8018 - 3.0980i$	$0.7004 + 0.4046i$	$0.7582 - 0.0045i$	$-2.3764 - 0.6699i$	$-0.2963 - 0.0317i$

Now there are two MSS solutions – (1) and (3) – and so the model is indeterminate. This example then shows how the combination of Gröbner basis to find all possible solutions, and then checking these solutions for MSS can check for determinacy.

6.3 Example 3: NK Model with Habits

Now consider a slight variant of the previously discussed New Keynesian model, but with households that have habit formation. In this case, they maximize

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\log\left(C_t - \varphi C_{t-1}\right) - H_t\right)$$

where φ denotes habit persistence. For simplicity, assume TFP is constant

$$A_t = 1$$

and that there is no interest rate smoothing, so $\rho = 0$:

$$\frac{R_t}{R_{ss}} = \Pi_t^{\psi_t} \exp\left(\sigma_r \varepsilon_{r,t}\right).$$

With habits, consumption appears dated as C_{t-1}, C_t, and C_{t+1} in the equilibrium conditions, which are

$$\lambda_t = \frac{1}{C_t - \varphi C_{t-1}} - \beta \mathbb{E}_t \frac{\varphi}{C_{t+1} - \varphi C_t}$$

$$\lambda_t = \beta \mathbb{E}_t \lambda_{t+1} \frac{R_t}{\Pi_{t+1}}$$

$$\kappa \lambda_t \left(\Pi_t - 1\right) \Pi_t = \left(1 - \eta\right) \lambda_t + \eta + \beta \kappa \mathbb{E}_t \lambda_{t+1} \left(\Pi_{t+1} - 1\right) \Pi_{t+1} \frac{Y_{t+1}}{Y_t}$$

$$\frac{R_t}{R_{ss}} = \Pi_t^{\psi_t} \exp\left(\sigma_r \varepsilon_{r,t}\right)$$

$$Y_t = C_t + \frac{\kappa}{2} \left(\Pi_t - 1\right)^2 Y_t$$

Consequently, define the auxiliary variable $\tilde{C}_t = C_t$, so $\tilde{C}_{t+1} = C_{t+1}$. Substituting out R_t and Y_t to simplify the equilibrium conditions, now gives $y_t = \left[\Pi_t, \tilde{C}_t, \lambda_t\right]'$, $y_{t+1} = \left[\Pi_{t+1}, \tilde{C}_{t+1}, \lambda_{t+1}\right]$, $x_t = [C_t]$, $x_{t-1} = [C_{t-1}]'$, and $\theta_{2t} = \psi_t$. Then the equilibrium is expressed as

$$f\left(y_{t+1}, y_t, x_t, x_{t-1}, \chi \varepsilon_{t+1}, \varepsilon_t, \theta_{t+1}, \theta_t\right) =$$

$$\begin{bmatrix} \frac{1}{C_t - \varphi C_{t-1}} - \beta \mathbb{E}_t \frac{\varphi}{\tilde{C}_{t+1} - \varphi C_t} - \lambda_t \\ \beta \mathbb{E}_t \lambda_{t+1} \frac{R_{ss} \Pi_t^{\psi_t} \exp(\sigma_r \varepsilon_{r,t})}{\Pi_{t+1}} - \lambda_t \\ \left(1 - \eta\right) \lambda_t + \eta + \beta \kappa \mathbb{E}_t \lambda_{t+1} \left(\Pi_{t+1} - 1\right) \Pi_{t+1} \frac{C_{t+1}}{C_t} \frac{1 - \frac{\kappa}{2}\left(\Pi_t - 1\right)^2}{1 - \frac{\kappa}{2}\left(\Pi_{t+1} - 1\right)^2} - \kappa \lambda_t \left(\Pi_t - 1\right) \Pi_t \\ \tilde{C}_t - C_t \end{bmatrix}$$

6.3.1 Solving the Habit Model

The following subsections show how to solve the model.

6.3.2 Steady State

Calculating the steady state involves setting $\chi = 0$, so $\Pi_t = \Pi_{t+1} = \Pi_{ss}$, $\tilde{C}_t = \tilde{C}_{t+1} = C_{ss}$, and $C_t = C_{t-1} = C_{ss}$. The equilibrium conditions in steady state are

$$\begin{bmatrix} \frac{1}{C_{ss}-\varphi C_{ss}} - \beta \frac{\varphi}{\tilde{C}_{ss}-\varphi C_{ss}} - \lambda_{ss} \\ \beta \lambda_{ss} \frac{R_{ss}\Pi_{ss}^{\psi_t} \exp(\sigma_r \varepsilon_{r,t})}{\Pi_{ss}} - \lambda_{ss} \\ (1-\eta)\lambda_{ss} + \eta + \beta\kappa\lambda_{ss}(\Pi_{ss}-1)\Pi_{ss}\frac{C_{ss}}{C_{ss}}\frac{1-\frac{\kappa}{2}(\Pi_{ss}-1)^2}{1-\frac{\kappa}{2}(\Pi_{ss}-1)^2} - \kappa\lambda_{ss}(\Pi_{ss}-1)\Pi_{ss} \\ \tilde{C}_{ss} - C_{ss} \end{bmatrix} = 0_{4\times 1}$$

Assuming $\Pi_{ss} = 1$, the steady state satisfies

$$\lambda_{ss} = \frac{\eta}{\eta - 1}$$

$$C_{ss} = \frac{1-\beta\varphi}{1-\varphi}\frac{\eta-1}{\eta}$$

$$\tilde{C}_{ss} = C_{ss}$$

6.3.3 The Matrices

Now to define the matrices in Section, which are the derivatives of the function f evaluated at steady state. In this example, $n_y = 3$, $n_x = 1$, $n_\varepsilon = 1$, and $n_\theta = 2$. The necessary matrices are

$$\mathcal{D}_{1,3} f_{ss}(s', s) = \begin{bmatrix} 0 & \frac{\varphi\beta}{c_{ss}^2(1-\varphi)} & 0 \\ -\lambda_{ss} & 0 & 1 \\ \beta\kappa\lambda_{ss} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathcal{D}_{4,6} f_{ss}(s', s) = \begin{bmatrix} 0 & 0 & -1 \\ \lambda_{ss}\psi(s) & 0 & -1 \\ -\kappa\lambda_{ss} & 0 & 1-\eta \\ 0 & 1 & 0 \end{bmatrix}$$

$$\mathcal{D}_{7,7} f_{ss}(s', s) = \begin{bmatrix} -\frac{1+\beta\varphi^2}{c_{ss}^2(\varphi-1)^2} \\ 0 \\ 0 \\ -1 \end{bmatrix}, \mathcal{D}_{8,8} f_{ss}(s', s) = \begin{bmatrix} \frac{\varphi}{c_{ss}^2(1-\varphi)^2} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathcal{D}_{9,9} f_{ss}(s', s) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathcal{D}_{10,10}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ \lambda_{ss}\sigma\\ 0\\ 0\end{bmatrix},\mathcal{D}_{11,11}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix},\mathcal{D}_{12,12}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}$$

6.3.4 Determinate Solution

Consider first the following parametrization

β	κ	η	φ	σ	$p_{1,1}$	$p_{2,2}$	$\psi(1)$	$\psi(2)$
0.993	161	10	0.95	0.0025	0.90	0.90	2.1	0.9

The parameters are similar to those in the previous NK example, with the exception that β is changed to reflect that there is no growth in the economy, and φ dictates a very high degree of habit formation. The steady state values are $R_{ss}=1.007$, $\tilde{C}_{ss}=C_{ss}=1.0197$, and $\lambda_{ss}=1.11111$. Consequently the numerical values of the matrices of first-order partial derivatives of f with respect to all its components evaluated at steady state are

$$\mathcal{D}_{1,3}f_{ss}\left(s',s\right)=\begin{bmatrix}0 & 362.901 & 0\\ -1.1111 & 0 & 1\\ 177.637 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},\ \mathcal{D}_{4,6}f_{ss}\left(s',s\right)=\begin{bmatrix}0 & 0 & -1\\ 1.11111\psi\left(s\right) & 0 & -1\\ 178.889 & 0 & -9\\ 0 & 1 & 0\end{bmatrix}$$

$$\mathcal{D}_{7,7}f_{ss}\left(s',s\right)=\begin{bmatrix}-729.45\\ 0\\ 0\\ -1\end{bmatrix},\ \mathcal{D}_{8,8}f_{ss}\left(s',s\right)=\begin{bmatrix}365.459\\ 0\\ 0\\ 0\end{bmatrix},\mathcal{D}_{9,9}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}$$

$$\mathcal{D}_{10,10}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ 0.00278\\ 0\\ 0\end{bmatrix},\mathcal{D}_{11,11}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix},\mathcal{D}_{12,12}f_{ss}\left(s',s\right)=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}$$

Using these matrices, we can solve the quadratic system to get the following set of solutions for $\{\mathcal{D}_{1,1}g_{ss}(s_t), \mathcal{D}_{1,1}h_{ss}(s_t)\}_{s_t=1}^{n_s}$. In this example, since $n_x = 1$ and there are no exogenous variables, $n_y = 3$, and $n_s = 2$, Proposition 3 indicates that, at most, there are 16 total solutions. As shown below, there are exactly 16 solutions. For brevity, we only present $\mathcal{D}_{1,1}h_{ss}(s_t)$

	$D_{1,1}h_{ss}(1)$	$D_{1,1}h_{ss}(2)$
(1)	0.95	0.95
(2)	1.06005	1.06005
(3)	0.91787	1.04360
(4)	1.04128	1.12014
(5)	1.03760	1.13389
(6)	1.35512	1.20296
(7)	$1.1167 - 0.10864i$	$1.1167 + 0.10864i$
(8)	$1.1167 + 0.10864i$	$1.1167 - 0.10864i$
(9)	$1.11048 + 0.19451i$	$0.96460 + 0.17167i$
(10)	$1.11048 - 0.19451i$	$0.96460 - 0.17167i$
(11)	$1.14351 - 0.30537i$	$1.00574 + 0.16203i$
(12)	$1.14351 + 0.30537i$	$1.00574 - 0.16203i$
(13)	$1.16037 - 0.16517i$	$1.14822 + 0.07969i$
(14)	$1.16037 + 0.16517i$	$1.14822 - 0.07969i$
(15)	$1.16056 + 0.34768i$	$1.34334 - 0.11370i$
(16)	$1.16056 - 0.34768i$	$1.34334 + 0.11370i$

The only MSS solution is (1), hence the model is determinate.

6.3.5 Indeterminate Solution

Now, consider an alternative parameterization where the degree of habit persistence is lower, $\varphi = 0.7$. The steady state under these parameters is $R_{ss} = 1.007$, $\tilde{C}_{ss} = C_{ss} = 0.9147$, and $\lambda_{ss} = 1.11111$. A similar set of numerical matrices as above can be used to find the 16 total

solutions to the quadratic equation. For brevity, we only present $\mathcal{D}_{1,1}h_{ss}(s_t)$

	$D_{1,1}h_{ss}(1)$	$D_{1,1}h_{ss}(2)$
(1)	0.7	0.7
(2)	1.43864	1.43864
(3)	0.79717	1.57910
(4)	1.57910	0.79717
(5)	0.69800	1.02883
(6)	1.62054	1.09263
(7)	$0.89005 - 0.10590i$	$1.20624 + 0.02242i$
(8)	$0.89005 + 0.10590i$	$1.20624 - 0.02242i$
(9)	$1.11412 - 0.31013i$	$0.95067 + 0.10869i$
(10)	$1.11412 + 0.31013i$	$0.95067 - 0.10869i$
(11)	$1.42678 - 0.06046i$	$1.24724 - 0.01867i$
(12)	$1.42678 + 0.06046i$	$1.24724 + 0.01867i$
(13)	$1.11018 + 0.24344i$	$0.80144 + 0.07992i$
(14)	$1.11018 - 0.24344i$	$0.80144 - 0.07992i$
(15)	$1.19334 + 0.28773i$	$1.58875 - 0.07190i$
(16)	$1.19334 - 0.28773i$	$1.58875 + 0.07190i$

Now there are two MSS solutions – (1) and (5) – again showing how the combination of Gröbner basis to find all possible solutions, and then checking these solutions for MSS can check for determinacy.

7 Conclusion

This paper developed a perturbation method for constructing approximations to the solutions to Markov switching DSGE models. The framework allows introducing switching from first principles, including when switching affects the steady state of the economy. While not pursued here, second- or higher-order approximations are straightforward, and follow the single-regime

case studied by Schmitt-Grohe & Uribe (2004). Using Gröbner bases to solve the system and mean square stability to check for existence and uniqueness of stable solutions, the method handles a wide variety of models, and shows that switching in parameters that affect the steady state implies that first order approximations are not certainty equivalent.

References

Aruoba, S.B., Fernandez-Villaverde, J.F., & Rubio-Ramirez, J.F. 2006. Comparing Solution Methods for dynamic equilibrium economies. *Journal of Economic Dynamics and Control*, **30**(12), 2477–2508.

Becker, E., Marianari, M.G., Mora, T., & Treverso, C. 1996. The Shape of the Shape Lemma. *International Symposium on Symbolic and Algebraic Computation*, 129–133.

Bianchi, F. 2009. Regime Switches, Agents' Beliefs, and Post-World War II US Macroeconomic Dynamics. *Working Paper*.

Blake, A., & Zampolli, F. 2006. Optimal Monetary Policy in Markov-Switching Models with Rational Expectations Agents. *Bank of England Working Paper*, **298**.

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

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

Costa, O., Fragoso, M., & Marques, R. 2005. *Discrete-Time Markov Jump Linear Systems*. Springer.

Cox, D.A., Little, J.B., & O'Shea, D.B. 1997. *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics, Springer-Verlag.

Davig, T., & Leeper, E. 2007. Generalizing the Taylor Principle. *American Economic Review*, **97**(3), 607–635.

Farmer, R, Waggoner, D., & Zha, T. 2008. Minimal State Variable Solutions to Markov-Switching Rational Expectations Models. *Journal of Economic Dynamics and Control*, **35**(12), 2150–2166.

Farmer, R, Waggoner, D., & Zha, T. 2009. Understanding Markov-Switching Rational Expectations Models. *Journal of Economic Theory*, **144**(5), 1849–1867.

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

Fernandez-Villaverde, J., Guerron-Quintana, P., & Rubio-Ramirez, J. 2009. Fortune or Virtue: Time-Variant Volatilites versus Parameter Drifting in U.S. Data. *Working Paper*.

Hamilton, J. 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, **57**(2), 357–384.

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

Leeper, E., & Zha, T. 2003. Modest Policy Interventions. *Journal of Monetary Economics*, **50**(8), 1673–1700.

Lubik, T.A., & Schorfheide, F. 2004. Testing for Indeterminacy: An Application to US Monetary Policy. *American Economic Review*, **94**(1), 190–217.

Rudebusch, G., & Swanson, E. 2008. Examining the Bond Premium Puzzle with a DSGE Model. *Journal of Monetary Economics*, **55**(S1), S111–S126.

Schmitt-Grohe, S., & Uribe, M. 2004. Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. *Journal of Economic Dynamics and Control*, **28**(4), 755–775.

Sims, C. 1980. Macroeconomics and Reality. *Econometrica*, **48**(1), 1–48.

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

Svensson, L., & Williams, N. 2007. Monetary Policy with Model Uncertainty: Distribution Forecast Targeting. *CEPR Discussion Paper*, **6331**.

van Binsbergen, J.H., Fernandez-Villaverde, J., Koijen, R., & Rubio-Ramirez, J. 2008. Likelihood estimation of DSGE models with Epstein-Zin preferences. *Working Paper*.

Woodford, M. 2003. *Interest and Prices*. Princeton UP.

8 Appendix A: Iterative Algorithm

This section describes an iterative procedure to potentially find one solution of the quadratic system (19). This method cannot be used to check for determinacy, but can be used to find one solution and then check that solution for MSS. Recall the quadratic system to be solved is

$$A(s_t)\begin{bmatrix} I \\ \mathcal{D}_{1,n_x} g_{ss}(1) \\ \vdots \\ \mathcal{D}_{1,n_x} g_{ss}(n_s) \end{bmatrix} \mathcal{D}_{1,n_x} h_{ss}(s_t) = B(s_t)\begin{bmatrix} I \\ \mathcal{D}_{1,n_x} g_{ss}(s_t) \end{bmatrix} \tag{23}$$

for all s_t, where

$$A(s_t) = \begin{bmatrix} \sum_{s'=1}^{n_s} p_{s_t,s'} \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s', s_t) & p_{s_t,1}\mathcal{D}_{1,n_y} f_{ss}(1, s_t) & \cdots & p_{s_t,n_s}\mathcal{D}_{1,n_y} f_{ss}(n_s, s_t) \end{bmatrix}$$

and

$$B(s_t) = -\sum_{s'=1}^{n_s} p_{s_t,s'} \begin{bmatrix} \mathcal{D}_{2n_y+n_x+1,2(n_y+n_x)} f_{ss}(s', s_t) & \mathcal{D}_{n_y+1,2n_y} f_{ss}(s', s_t) \end{bmatrix}.$$

The idea of the algorithm is to guess at the policy functions, and solve for each regime's policy functions as in the fixed regime case. When the solutions are close to the guesses, then a solution has been found.

Algorithm 6 *Let* $\left\{\mathcal{D}_{1,n_x} g_{ss}^{(j)}(s_t), \mathcal{D}_{1,n_x} h_{ss}^{(j)}(s_t)\right\}$ *denote solution at iteration* j.

1. *Set* $j = 0$ *and initialize* $\mathcal{D}_{1,n_x} g_{ss}^{(0)}(s_t) = 0_{n_y \times n_x}$ *and* $\mathcal{D}_{1,n_x} h_{ss}^{(0)}(s_t) = 0_{n_x \times n_x}$.

2. *Set* $j = 1$. *For each* s_t, *construct the transformed system*

$$A\left(s_t, \left\{\mathcal{D}_{1,n_x} g_{ss}^{(j-1)}(s')\right\}_{s'=1, s' \neq s_t}^{n_s}\right) \begin{bmatrix} I \\ \mathcal{D}_{1,n_x} g_{ss}^{(j)}(s_t) \end{bmatrix} \mathcal{D}_{1,n_x} h_{ss}^{(j)}(s_t) \tag{24}$$
$$= B(s_t) \begin{bmatrix} I \\ \mathcal{D}_{1,n_x} g_{ss}^{(j)}(s_t) \end{bmatrix}$$

where

$$A\left(s_t, \left\{\mathcal{D}_{1,n_x} g_{ss}^{(j)}(s')\right\}_{s'=1, s' \neq s_t}^{n_s}\right) =$$
$$\begin{bmatrix} \begin{pmatrix} \sum_{s'=1}^{n_s} p_{s_t,s'} \mathcal{D}_{2n_y+1,2n_y+n_x} f_{ss}(s', s_t) \\ + \sum_{s'=1, s' \neq s_t}^{n_s} p_{s_t,s'} \mathcal{D}_{1,n_y} f_{ss}(s', s_t) \mathcal{D}_{1,n_x} g_{ss}^{(0)}(s') \end{pmatrix} & p_{s_t,s_t} \mathcal{D}_{1,n_y} f_{ss}(s_t, s_t) \end{bmatrix}$$

3. *The set of systems (24) are identical to a fixed regime problem, and can be solved using standard singular value decomposition (SVD) algorithm. When constructing* $\mathcal{D}_{1,n_x} h_{ss}^{(j)}(s_t)$, *use the "most stable" generalized eigenvalues of* $\left(A\left(s_t, \left\{\mathcal{D}_{1,n_x} g_{ss}^{(j-1)}(s')\right\}_{s'=1, s' \neq s_t}^{n_s}\right), B(s_t)\right)$ *– those with the smallest modulus.*

4. *Check* $\max\left\{\left\|\mathcal{D}_{1,n_x} h_{ss}^{(j)}(s_t) - \mathcal{D}_{1,n_x} h_{ss}^{(j-1)}(s_t)\right\|, \left\|\mathcal{D}_{1,n_x} g_{ss}^{(j)}(s_t) - \mathcal{D}_{1,n_x} g_{ss}^{(j-1)}(s_t)\right\|\right\} < crit$. *If yes, then stop and check for MSS. If no, then set* $j \implies j+1$ *and return to step* 2.

Note in step (2), when $s_t > 1$, then $\left\{g_{ss}^{(j)}(s')\right\}_{s'=1}^{s_t-1}$ can be used instead of $\left\{g_{ss}^{(j-1)}(s')\right\}_{s'=1}^{s_t-1}$.

9 Appendix B: Second Order Derivatives

This appendix contains the forms of the second derivatives in (22).

The first equation is the derivative with respect to x_{t-1} twice.

$$H_{1,n_x;1,n_x}G^i_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\times$$

$$\left(\begin{array}{c}
\left[\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\right]^{\intercal}\left[\begin{array}{c}\left[\begin{array}{c}\mathcal{H}_{1,n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+2\mathcal{H}_{1,n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{1,n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)\\+2\mathcal{H}_{1,n_y;2n_y+n_x+1,2n}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}\mathcal{H}_{n_y+1,2n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)\\+2\mathcal{H}_{n_y+1,2n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{n_y+1,2n_y;2n_y+n_x+1,2n}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}\left[\begin{array}{c}\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\\+\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;1,n_x}g_{ss}\left(s'\right)\end{array}\right]\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+n_x+1,2n}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\left[\begin{array}{c}\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{H}_{1,n_x;1,n_x}h_{ss}\left(s_t\right)\\
+\mathcal{D}_{n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;1,n_x}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n_y+n_x+1,2n;2n_y+n_x+1,2n}f^i_{ss}\left(s',s_t\right)
\end{array}\right)$$

The second equation is the derivative with respect to x_{t-1} and ε_t.

$$\mathcal{H}_{1,n_x;n_x+1,n_x+n_\varepsilon}\mathbb{G}^i_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\times$$

$$\left(\begin{array}{c}
\left[\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)\right]^{\intercal}\left[\begin{array}{c}\left[\begin{array}{c}\mathcal{H}_{1,n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+2\mathcal{H}_{1,n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\+\mathcal{H}_{1,n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\+\mathcal{H}_{1,n_y;2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\mathcal{D}_{1,n_x}g_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}\left[\begin{array}{c}\mathcal{H}_{n_y+1,2n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{H}_{n_y+1,2n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\+\mathcal{H}_{n_y+1,2n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\+\mathcal{H}_{n_y+1,2n_y;2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\mathcal{D}_{1,n_x}h_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}\mathcal{H}_{2n_y+1,2n_y+n_x;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\+\mathcal{H}_{2n_y+1,2n_y+n_x;2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)\\+\left[\begin{array}{c}\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\\+\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;1,n_x}g_{ss}\left(s'\right)\end{array}\right]\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\end{array}\right]\\
+\left[\begin{array}{c}\mathcal{H}_{2n_y+n_x+1,2(n_x+n_y);1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{H}_{2n_y+n_x+1,2n;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\
+\left[\begin{array}{c}\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{H}_{1,n_x;n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n_y+n_x+1,2n;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\
+\mathcal{D}_{n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n_y+n_x+1,2(n_x+n_y);2(n_x+n_y)+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)
\end{array}\right)$$

The third equation is the derivative with respect to x_{t-1} and χ.

$$\mathcal{H}_{1,n_x;n_x+n_\varepsilon+1}\mathbb{G}^i_{ss}\left(s_t\right) = \sum_{s'=1}^{n_s} p_{s_t,s'} \times$$

$$\left(\begin{array}{c}
\left[\mathcal{D}_{1,n_x} g_{ss}\left(s'\right) \mathcal{D}_{1,n_x} h_{ss}\left(s_t\right)\right]^{\intercal} \left[\begin{array}{c}
\mathcal{H}_{1,n_y;1,n_y} f^i_{ss}\left(s',s_t\right) \left[\begin{array}{c} \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s'\right) \\ +\mathcal{D}_{1,n_x} g_{ss}\left(s'\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \end{array}\right] \\
+\mathcal{H}_{1,n_y;n_y+1,2n_y} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s_t\right) \\
+\mathcal{H}_{1,n_y;2n_y+1,2n_y+n_x} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \\
+\mathcal{H}_{1,n_y;2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s'\right) \\
+\mathcal{H}_{1,n_y;2(n_x+n_y+n_\varepsilon+n_\theta)+1,2(n_x+n_y+n_\varepsilon+n_\theta)} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s_t\right)
\end{array}\right] \\
+\mathcal{D}_{1,n_x} h_{ss}\left(s_t\right)^{\intercal} \left[\begin{array}{c}
\mathcal{H}_{2n_y+1,2n_y+n_x;1,n_y} f^i_{ss}\left(s',s_t\right) \left[\begin{array}{c} \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s'\right) \\ +\mathcal{D}_{1,n_x} g_{ss}\left(s'\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \end{array}\right] \\
+\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+1,2n_y+n_x} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \\
+\mathcal{H}_{2n_y+1,2n_y+n_x;n_y+1,2n_y} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s_t\right) \\
+\mathcal{H}_{2n_y+1,2n_y+n_x;2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s'\right) \\
+\mathcal{H}_{2n_y+1,2n_y+n_x;2(n_x+n_y+n_\varepsilon)+n_\theta+1,2(n_x+n_y+n_\varepsilon+n_\theta)} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s_t\right) \\
\mathcal{D}_{1,n_y} f^i_{ss}\left(s',s_t\right) \left[\begin{array}{c} \mathcal{H}_{1,n_x;n_x+n_\varepsilon+1} g_{ss}\left(s'\right) \\ +\mathcal{H}_{1,n_x;1,n_x} g_{ss}\left(s'\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \end{array}\right]
\end{array}\right] \\
+\mathcal{D}_{1,n_x} g_{ss}\left(s_t\right)^{\intercal} \left[\begin{array}{c}
\mathcal{H}_{n_y+1,2n_y;1,n_y} f^i_{ss}\left(s',s_t\right) \left[\begin{array}{c} \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s'\right) \\ +\mathcal{D}_{1,n_x} g_{ss}\left(s'\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \end{array}\right] \\
+\mathcal{H}_{n_y+1,2n_y;n_y+1,2n_y} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s_t\right) \\
+\mathcal{H}_{n_y+1,2n_y;2n_y+1,2n_y+n_x} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \\
+\mathcal{H}_{n_y+1,2n_y;2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s'\right) \\
+\mathcal{H}_{n_y+1,2n_y;2(n_x+n_y+n_\varepsilon+n_\theta)+1,2(n_x+n_y+n_\varepsilon+n_\theta)} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s_t\right)
\end{array}\right] \\
+\mathcal{H}_{2n_y+n_x+1,2(n_x+n_y);1,n_y} f^i_{ss}\left(s',s_t\right) \left[\begin{array}{c} \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s'\right) \\ +\mathcal{D}_{1,n_x} g_{ss}\left(s'\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \end{array}\right] \\
+\mathcal{H}_{2n_y+n_x+1,2n;n_y+1,2n_y} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} g_{ss}\left(s_t\right) \\
+\mathcal{H}_{2n_y+n_x+1,2n;2n_y+1,2n_y+n_x} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{n_x+n_\varepsilon+1} h_{ss}\left(s_t\right) \\
+\mathcal{H}_{2n_y+n_x+1,2(n_x+n_y);2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta} f^i_{ss}\left(s',s_t\right) \mathcal{D}\theta_{ss}\left(s'\right) \\
+\mathcal{H}_{2n_y+n_x+1,2(n_x+n_y);2(n_x+n_y+n_\varepsilon)+n_\theta+1,2(n_x+n_y+n_\varepsilon+n_\theta)} f^i_{ss}\left(s',s_t\right) \theta_\chi\left(s\right) \\
+\left[\begin{array}{c} \mathcal{D}_{1,n_y} f^i_{ss}\left(s',s_t\right) \mathcal{D}_{1,n_x} g_{ss}\left(s'\right) \\ +\mathcal{D}_{2n_y+1,2n_y+n_x} f^i_{ss}\left(s',s_t\right) \end{array}\right] \mathcal{H}_{1,n_x;n_x+n_\varepsilon+2} h_{ss}\left(s_t\right) \\
+\mathcal{D}_{n_y+1,2n_y} f^i_{ss}\left(s',s_t\right) \mathcal{H}_{1,n_x;n_x+n_\varepsilon+1} g_{ss}\left(s_t\right)
\end{array}\right)$$

The fourth equation is the derivative with respect to ε_{t-1} twice.

$$\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+1,n_x+n_\varepsilon}\mathbb{G}^i_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\times$$

$$\left(\begin{array}{c}
\left[\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\right]^{\intercal}\left[\begin{array}{c}\left[\begin{array}{c}\mathcal{H}_{1,n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+2\mathcal{H}_{1,n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{1,n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\+2\mathcal{H}_{1,n_y;2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}\left[\begin{array}{c}\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\\+\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;1,n_x}g_{ss}\left(s'\right)\end{array}\right]\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{2n_y+1,2n_y+n_x;2(n_x+n_y)+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}\mathcal{H}_{n_y+1,2n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\+2\mathcal{H}_{n_y+1,2n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{n_y+1,2n_y;2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)\end{array}\right]\\
+\left[\begin{array}{c}\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\\
+\mathcal{D}_{n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon);2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon)}f^i_{ss}\left(s',s_t\right)
\end{array}\right)$$

The fifth equation is the derivative with respect to ε_{t-1} and χ.

$$\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+n_\varepsilon+1}\mathbb{G}^i_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\times$$

$$\left(\begin{array}{c}
\left[\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)\right]^{\intercal}\left[\begin{array}{c}
\mathcal{H}_{1,n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\left[\begin{array}{c}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{array}\right]\\
+\mathcal{H}_{1,n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{1,n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\
+\mathcal{H}_{1,n_y;2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\
+\mathcal{H}_{1,n_y;2(n_x+n_y+n_\varepsilon+n_\theta)+1,2(n_x+n_y+n_\varepsilon+n_\theta)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)
\end{array}\right]\\
+\mathcal{D}_{n_x+1,n_x+n_\varepsilon}h_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}
\mathcal{H}_{2n_y+1,2n_y+n_x;1,n_y}f^i_{ss}\left(s',s_t\right)\left[\begin{array}{c}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{array}\right]\\
+\mathcal{H}_{2n_y+1,2n_y+n_x;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n_y+1,2n_y+n_x;2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\
+\mathcal{H}_{2n_y+1,2n_y+n_x;2(n_x+n_y+n_\varepsilon)+n_\theta+1,2(n_x+n_y+n_\varepsilon+n_\theta)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\\
\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\left[\begin{array}{c}\mathcal{H}_{1,n_x;n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{H}_{1,n_x;1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{array}\right]
\end{array}\right]\\
+\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s_t\right)^{\intercal}\left[\begin{array}{c}
\mathcal{H}_{n_y+1,2n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\left[\begin{array}{c}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{array}\right]\\
+\mathcal{H}_{n_y+1,2n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{n_y+1,2n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\
+\mathcal{H}_{n_y+1,2n_y;2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\
+\mathcal{H}_{n_y+1,2n_y;2(n_x+n_y+n_\varepsilon+n_\theta)+1,2(n_x+n_y+n_\varepsilon+n_\theta)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)
\end{array}\right]\\
+\mathcal{H}_{2(n_x+n_y)+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon);1,n_y}f^i_{ss}\left(s',s_t\right)\left[\begin{array}{c}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{array}\right]\\
+\mathcal{H}_{2(n_x+n_y)+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon);n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\
+\mathcal{H}_{2(n_x+n_y)+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon);2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\
+\mathcal{H}_{2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon);2(n_x+n_y+n_\varepsilon)+1,2(n_x+n_y+n_\varepsilon)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\
+\mathcal{H}_{2n+n_\varepsilon+1,2(n_x+n_y+n_\varepsilon);2(n_x+n_y+n_\varepsilon)+n_\theta+1,2(n_x+n_y+n_\varepsilon+n_\theta)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\\
+\left[\begin{array}{c}\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{array}\right]\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\
+\mathcal{D}_{n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)
\end{array}\right)$$

The sixth equation is the derivative with respect to χ twice.

$$\mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}\mathbb{G}^i_{ss}\left(s_t\right)=\sum_{s'=1}^{n_s}p_{s_t,s'}\times$$

$$\left(\begin{array}{c}
\begin{bmatrix}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{bmatrix}^{\intercal}\mathcal{H}_{1,n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\begin{bmatrix}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{bmatrix}\\
+2\begin{bmatrix}\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\+\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\end{bmatrix}^{\intercal}\begin{bmatrix}\mathcal{H}_{1,n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\+\mathcal{H}_{1,n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\+\mathcal{H}_{1,n_y;2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\+\mathcal{H}_{1,n_y;2\left(n_x+n_y+n_\varepsilon+n_\theta\right)+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\end{bmatrix}\\
+\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)^{\intercal}\begin{bmatrix}\mathcal{H}_{n_y+1,2n_y;n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\+2\mathcal{H}_{n_y+1,2n_y;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{n_y+1,2n_y;2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\+2\mathcal{H}_{n_y+1,2n_y;2\left(n_x+n_y+n_\varepsilon+n_\theta\right)+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\end{bmatrix}\\
+\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)^{\intercal}\begin{bmatrix}\mathcal{H}_{2n_y+1,2n_y+n_x;2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\+2\mathcal{H}_{2n_y+1,2n_y+n_x;2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\+2\mathcal{H}_{2n_y+1,2n_y+n_x;2\left(n_x+n_y+n_\varepsilon\right)+n_\theta+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\\+\left[\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;1,n_x}g_{ss}\left(s'\right)\right]\mathcal{D}_{n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\+2\left[\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{1,n_x;n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\right]\end{bmatrix}\\
+\left(\varepsilon'\right)^{\intercal}\begin{bmatrix}\mathcal{H}_{2\left(n_x+n_y\right)+1,2\left(n_x+n_y\right)+n_\varepsilon;2\left(n_x+n_y\right)+1,2\left(n_x+n_y\right)+n_\varepsilon}f^i_{ss}\left(s',s_t\right)\\+\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{n_x+1,n_x+n_\varepsilon;n_x+1,n_x+n_\varepsilon}g_{ss}\left(s'\right)\\+\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s'\right)'\mathcal{H}_{1,n_y;1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s'\right)\\+2\mathcal{D}_{n_x+1,n_x+n_\varepsilon}g_{ss}\left(s'\right)'\mathcal{H}_{1,n_y;2n+1,2n+n_\varepsilon}f^i_{ss}\left(s',s_t\right)\end{bmatrix}\varepsilon'\\
+\mathcal{D}\theta_{ss}\left(s'\right)^{\intercal}\begin{bmatrix}\mathcal{H}_{2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta;2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s'\right)\\+2\mathcal{H}_{2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta;2\left(n_x+n_y+n_\varepsilon\right)+n_\theta+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\end{bmatrix}\\
+\mathcal{D}\theta_{ss}\left(s_t\right)^{\intercal}\mathcal{H}_{2\left(n_x+n_y+n_\varepsilon\right)+n_\theta+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right);2\left(n_x+n_y+n_\varepsilon\right)+n_\theta+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right)}f^i_{ss}\left(s',s_t\right)\mathcal{D}\theta_{ss}\left(s_t\right)\\
+\begin{bmatrix}\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{D}_{1,n_x}g_{ss}\left(s'\right)\\+\mathcal{D}_{2n_y+1,2n_y+n_x}f^i_{ss}\left(s',s_t\right)\end{bmatrix}\mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}h_{ss}\left(s_t\right)\\
+\mathcal{D}_{1,n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}g_{ss}\left(s'\right)\\
+\mathcal{D}_{n_y+1,2n_y}f^i_{ss}\left(s',s_t\right)\mathcal{H}_{n_x+n_\varepsilon+1;n_x+n_\varepsilon+1}g_{ss}\left(s_t\right)\\
+\mathcal{D}_{2\left(n_x+n_y+n_\varepsilon\right)+1,2\left(n_x+n_y+n_\varepsilon\right)+n_\theta}f^i_{ss}\left(s',s_t\right)\mathcal{H}\theta_{ss}\left(s'\right)\\
+\mathcal{D}_{2\left(n_x+n_y+n_\varepsilon\right)+n_\theta+1,2\left(n_x+n_y+n_\varepsilon+n_\theta\right)}f^i_{ss}\left(s',s_t\right)\mathcal{H}\theta_{ss}\left(s_t\right)
\end{array}\right)$$