

Essays on Macroeconomics in Mixed Frequency

Estimations

by

Tae Bong Kim

Department of Economics
Duke University

Date: _____

Approved:

Juan Rubio-Ramirez, Advisor

Craig Burnside

Francesco Bianchi

Tao Zha

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2011

ABSTRACT

Essays on Macroeconomics in Mixed Frequency Estimations

by

Tae Bong Kim

Department of Economics
Duke University

Date: _____

Approved:

Juan Rubio-Ramirez, Advisor

Craig Burnside

Francesco Bianchi

Tao Zha

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2011

Copyright © 2011 by Tae Bong Kim
All rights reserved except the rights granted by the
Creative Commons Attribution-Noncommercial Licence

Abstract

This dissertation asks whether frequency misspecification of a New Keynesian model results in temporal aggregation bias of the Calvo parameter. First, when a New Keynesian model is estimated at a quarterly frequency while the true data generating process is the same but at a monthly frequency, the Calvo parameter is upward biased and hence implies longer average price duration. This suggests estimating a New Keynesian model at a monthly frequency may yield different results. However, due to mixed frequency datasets in macro time series recorded at quarterly and monthly intervals, an estimation methodology is not straightforward. To accommodate mixed frequency datasets, this paper proposes a data augmentation method borrowed from Bayesian estimation literature by extending MCMC algorithm with "Rao-Blackwellization" of the posterior density. Compared to two alternative estimation methods in context of Bayesian estimation of DSGE models, this augmentation method delivers lower root mean squared errors for parameters of interest in New Keynesian model. Lastly, a medium scale New Keynesian model is brought to the actual data, and the benchmark estimation, i.e. the data augmentation method, finds that the average price duration implied by the monthly model is 5 months while that by the quarterly model is 20.7 months.

To Hyung Yun and Joshua

Contents

Abstract	iv
List of Tables	ix
Acknowledgements	x
1 Introduction	1
2 Mixed Frequency Estimation of Dynamic Stochastic General Equilibrium models : Bayesian Approach	3
2.1 Preliminaries	3
2.1.1 Generic Model	3
2.1.2 Bayesian Estimation	4
2.2 Two Stage Hybrid MCMC Algorithm	5
2.2.1 General Algorithm	5
2.2.2 Gibbs Sampling within Metropolis-Hastings Algorithm	6
2.2.3 Pseudo-Algorithm	8
2.3 Application: RBC	9
2.3.1 The model	9
2.4 Estimation Exercise	11
2.4.1 Simulation	11
2.5 Results	12
2.6 Conclusions	13

3	Temporal Aggregation Bias and Mixed Frequency Estimation of a New Keynesian model	14
3.1	Introduction	14
3.2	Temporal Aggregation Bias	18
3.2.1	AR(1)	18
3.2.2	Simple New Keynesian model	19
3.3	Mixed Frequency Estimation Strategies	25
3.3.1	Preliminary Setup	25
3.3.2	Data Augmentation Method	28
3.3.3	Stacking Method	36
3.3.4	Durbin-Koopman Method	38
3.3.5	Efficiency	39
3.4	Medium Scale New Keynesian Model	40
3.5	Estimation Exercise	42
3.5.1	Monte-Carlo Experiment	42
3.5.2	Data	45
3.5.3	Estimation Results	46
3.6	Conclusions	49
A	Appendix for Chapter 2	50
A.1	Standard Kalman Filter	50
A.2	Derivation of Target Distribution	51
A.3	Kalman Filter with Data Augmentation	52
A.4	Alternative Algorithm	53
A.5	State Space Representation of Coarse Estimation	54
A.6	State Space Representation of BMF Estimation	55

B Appendix for Chapter 3	56
B.1 Temporal Aggregation of Autoregressive Process	56
B.2 Derivation of Target Distribution	57
B.3 More Monte Carlo Results	60
B.4 Medium Scale New Keynesian Model	60
Bibliography	81
Biography	83

List of Tables

2.1	RBC True Parameters	11
2.2	RBC Simulation	12
3.1	AR(1) Example	19
3.2	Simple NK True Parameters	21
3.3	Priors	22
3.4	Estimation on Simple NK model	23
3.5	Fixed Parameters	42
3.6	True Parameters	43
3.7	MC Priors	43
3.8	Simulation Results I	44
3.9	Simulation Results II	45
3.10	Priors	47
3.11	Priors on Calvo parameters	47
3.12	Results on Price Durations	48

Acknowledgements

I would like to thank to my advisor, Juan Rubio-Ramirez, for the invaluable support and guidance. This dissertation could have not been done without his constant heedfulness. I also would like to thank to my co-advisor, Tao Zha, who has provided me a dissertation internship opportunity at Federal Reserve Bank of Atlanta. I would like to appreciate the hospitality and support from the staff at Atlanta Fed that has provided me with its computing resources to implement my estimation methodology which involves an intensive computation. Bjorn Eraker has incubated the idea of data augmentation technique in my second year study which has propagated into this dissertation package. I cannot emphasize more how grateful I am for his teaching and for the opportunity to co-work on Bayesian mixed frequency project.

I also thank Craig Burnside for suggesting me to explore the issue of a temporal aggregation bias which has been one of central pillars in my dissertation. Francesco Bianchi, though has joined my dissertation committee in the later stage of my study, has made a substantial contribution to the data augmentation methodology for its improvement and modification with his keen understanding of Bayesian statistics literature. I appreciate for Javier Garcia-Cicco's suggestion to implement the data augmentation in Kalman Filtering steps. My classmates, Andrew Foerster and Hernan Seoane, had to cope with my countless questions and cannot be forgotten for their patience and kindness. I also thank Tim Bollerslev, Daniel Waggoner, Martin Eichenbaum, Pietro Peretto, Barbara Rossi, Cosmin Ilut and many other seminar

participants at Duke University and Federal Reserve Bank of Atlanta.

Lastly, I cannot begin a word for my gratitude to my family. My wife, Hyung Yun, has together endured my seven years of the graduate study, two years of master program and five years of doctoral program, at Duke University and deserve all the credits for this final fruit of the dissertation. My son, Joshua, always contributed a positive energy to our family without which I could have not finished my dissertation.

1

Introduction

Researchers in economics have to assume a consistent decision time interval when a model of interest is brought to the data for estimation exercises. Although this may not necessarily reflect the true decision time interval of realistic agents in the model, the availability of dataset often restricts the model into a certain time unit. For example, a time unit in macroeconomic models is generally assumed to be quarterly since dataset of many macroeconomic variables are collected at quarterly frequency. Meanwhile, other variables such as inflation rate, interest rate and labor hours are collected at monthly frequency but aggregated into quarterly so that the frequency of the whole dataset is synchronized. The second chapter presents an econometric strategy for estimation of DSGE models when data is available at different time intervals. The method is based on a data augmentation technique within Bayesian estimation of structural models and allows us to jointly use data at different frequencies. The chapter works with a Real Business Cycle model and estimate structural parameters such total factor productivity, discount factor, depreciation rate of capital and capital share in production technology. This model has been chosen because it is the building block of DSGE models in macroeconomics. The third chapter is my job

market paper and thus rather a self sustaining paper of its own. It is an application of the mixed frequency estimation on a New Keynesian model to explain the gap in the Calvo parameter of which the macroevidence estimate is generally perceived as inconsistent when compared to the microevidence. Also, the data augmentation methodology has been modified from the second chapter to include the immediate future observations in the conditional information of the target density functions of missing observations. This brings an efficiency to the estimates of structural parameters in general which is shown to be consistent of the statistical theory with a Monte Carlo exercise on the New Keynesian model.

Mixed Frequency Estimation of Dynamic Stochastic General Equilibrium models : Bayesian Approach

2.1 Preliminaries

2.1.1 *Generic Model*

The solution method¹ to a standard Dynamic Stochastic General Equilibrium Models is assumed to be log-linearization around a deterministic steady state. Hence, one requirement is that the steady state is well defined. The model can be characterized by a set of first order conditions as,

$$\Omega \equiv E_t [\Omega(\xi_{t+1}, \xi_t, \eta_{t+1}, \eta_t)] = 0$$

and is smooth enough to have well defined first order derivatives. And thus the solution to the optimality conditions can be expressed in the following state space representation

$$\xi_{t+1} = F\xi_t + Gv_{t+1}, v_t \sim N(0, Q)$$

$$\eta_t = H'\xi_t + u_t, u_t \sim N(0, R)$$

¹ Follows Schmitt-Grohé and Uribe (2004).

where, the first equation is the state equation and the second one is the measurement equation describing the evolution of the observables as functions of the endogenous and exogenous states. ξ_{t+1} stands for a $n_x \times 1$ vector of unobserved states and η_t for a $n_y \times 1$ vector of observed variables.

2.1.2 Bayesian Estimation

Bayesian estimation of DSGE models is about characterizing and exploring the shape of a posterior distribution of structural parameters of a underlying model. Denoting the posterior distribution by

$$p(\theta|Y) = L(\theta|Y) \pi(\theta)$$

where θ is a vector of structural parameters in F, H, G, Q and R , and Y is the data. Due to a high nonlinearity of this function, a numerical approach is used and the standard procedure is a Metropolis-Hastings algorithm with Kalman Filter. A detailed explanation of this procedure can be found in An and Schorfheide (2007). A summary of a pseudo algorithm for the Random Walk Metropolis Hastings is presented in the following.

1. Synchronize the frequency of dataset, i.e. to lower frequency
2. Initialize with $\theta^{(0)}$
3. Draw θ^* from the proposal distribution $N(\theta^{(m)}, V)$
4. Evaluate $p(\theta^*|Y)$ using standard Kalman Filter²
5. $posterior\ odds = p(\theta^*|Y) - p(\theta^{(m)}|Y) \sim Unif(0, 1)$
6. If accepted, record $\theta^{(m+1)} = \theta^*$, else $\theta^{(m+1)} = \theta^{(m)}$
7. Repeat 3 to 6 for $m = 1, \dots, m_{sim}$

² Shown in appendix

2.2 Two Stage Hybrid MCMC Algorithm

2.2.1 General Algorithm

Now consider a problem : assume Y can be partitioned into two sets of observables where z is the data, observed less frequently, and w is the data, observed more frequently. Without loss of generality, assume z is observed quarterly while w is observed monthly. Denote $w \equiv \{w\}_{t=1}^T$, $z_q \equiv \{z_1, z_4, z_7, \dots\}$ and $z_{-q} \equiv \{z_2, z_3, z_5, z_6, \dots\}$. Then, the posterior density becomes

$$p(\theta|w, z_q)$$

But due to mixed frequency, it is not clear how to evaluate this posterior of structural parameters conditional on w and z_q . Hence, instead of characterizing $p(\theta|w, z_q)$ directly, auxiliary variable, z_{-q} , can be introduced to synchronize the vector of observable variables into higher frequency and study a joint density that can be evaluated which in expectation will be the desired posterior density,

$$p(\theta|w, z_q) = E_{z_{-q}} [p(\theta, z_{-q}|w, z_q)]$$

But $p(\theta|Y)$ was already highly nonlinear and introduction of z_{-q} aggravates the bulkiness of the algorithm. To make it more tractable, this joint density can be considered as a mixture model which is constituted by two conditional densities, i.e. $f(z_{-q}|w, z_q, \theta)$ and $p(\theta|w, z_q, z_{-q})$. The common practice in statistics is to apply two stage MCMC algorithm that explores these densities which in turn allows us to characterize the joint density.

The proposed method exploits the fact that DSGE model entails the ergodic distribution of the missing data, z_{-q} , conditioning on w, z_q and some fixed θ , i.e. $f(z_{-q}|w, z_q, \theta)$. This process is known as *data augmentation*. And this augmented data will enable us to implement standard Metropolis-Hastings algorithm on the target density function of structural parameters conditional on augmented dataset.

Hence, with sufficient number of generating missing data from $f(z_{-q}|w, z_q, \theta)$, data augmentation technique will allow to identify θ even if Y is not fully observed. Succintly, the algorithm is composed of two-stage samplings in which the first step is to generate the missing data under previous draw of parameters and the second step is to draw parameters under simulated data iteratively. Hence,³

$$\hat{z}_{-q}^{(m)} \sim f\left(z_{-q}|w, z_q, \hat{\theta}^{(m)}\right)$$

$$\hat{\theta}^{(m+1)} \sim p\left(\theta|w, z_q, \hat{z}_{-q}^{(m)}\right)$$

Diebolt and Robert (1994) shows that the sequences of z and θ are ergodic Markov chains and have convergence properties for two stage Gibbs sampling. In the first stage, since the distribution of structural shocks and measurement errors are Gaussian, $z^{(m)}$ can be generated with Gibbs sampling directly from $f\left(z_{-q}|w, z_q, \hat{\theta}^{(m)}\right)$. However, in the context of DSGE estimation, Gibbs sampling is not applicable in the second stage since $p\left(\theta|w, z_q, \hat{z}_{-q}^{(m)}\right)$ does not have an analytical form. But in general MCMC algorithm does not prevent one from replacing one of Gibbs sampling step with Metropolis-Hastings sampling since the ergodic Markov chain and convergence property are still valid. This method is better known as *two stage hybrid MCMC algorithm*.

2.2.2 Gibbs Sampling within Metropolis-Hastings Algorithm

The data augmentation step, $\hat{z}_{-q}^{(m)} \sim f\left(z_{-q}|w, z_q, \hat{\theta}^{(m)}\right)$, in principle can be implemented in different ways. But the dependence of the missing data on unobserved state variables implied by DSGE models hinders to simulate the whole set of missing data in one step. Instead, this paper proposes Gibbs sampling within Metropolis-Hastings algorithm : as state variables are backed out by Kalman Filter period by

³ Whenever a hat, $\hat{\cdot}$, is labeled, it means sampled variable.

period, missing data can be generated period by period. In other words, in the Kalman Filter whenever the algorithm is encountered with missing data at certain periods Gibbs sampling step is implemented from a target distribution that complies with the efficiency. To clarify this point, looking at the probabilistic interpretation of Kalman Filter can be useful. When errors are assumed to be normally distributed, Kalman Filter updates the state variables, ξ_t , at period t with

$$N\left(\begin{bmatrix} \xi_t \\ y_t \end{bmatrix} \middle| y^{t-1}\right) \rightarrow N\left(\xi_t \middle| \begin{bmatrix} y^{t-1} \\ y_t \end{bmatrix}\right) \quad (2.1)$$

Meanwhile, this also happens to be minimizing the variance-covariance matrix of forecast errors from predicting state variables. Now since y_t is not fully observed, the data augmentation step is necessary before updating with $N\left(\xi_t \middle| \begin{bmatrix} y^{t-1} \\ y_t \end{bmatrix}\right)$. When z_t is not observed, incorporating the data augmentation step changes into

$$N\left(\begin{bmatrix} \xi_t \\ z_t \\ w_t \end{bmatrix} \middle| \begin{bmatrix} \hat{z}^{t-1} \\ w^{t-1} \end{bmatrix}\right) \rightarrow N\left(\begin{bmatrix} \xi_t \\ z_t \end{bmatrix} \middle| \begin{bmatrix} \hat{z}^{t-1} \\ w^{t-1} \\ w_t \end{bmatrix}\right) \rightarrow N\left(\xi_t \middle| \begin{bmatrix} \hat{z}^{t-1} \\ w^{t-1} \\ \hat{z}_t \\ w_t \end{bmatrix}\right) \quad (2.2)$$

So while standard Kalman Filter backs out state variables as in (2.1) when z_t is observed, we add data augmentation as in (2.2) whenever z_t is not observed. This data augmentation will be implemented by Gibbs sampling from a normal distribution that exploits the contemporaneous observation, w_t , and the history of augmented data.

At period t , the history of observations, $\begin{pmatrix} \hat{z}^{t-1} \\ w^{t-1} \end{pmatrix}$, can be summarized by $\xi_{t|\widehat{t-1}}, P_{t|\widehat{t-1}}$ ⁴ where the $\widehat{\cdot}$'s represent the history of augmented data. Let $f\left(\begin{bmatrix} z_t \\ \xi_t \end{bmatrix} \middle| \xi_{t|\widehat{t-1}}^{(m)}, P_{t|\widehat{t-1}}^{(m)}, w_t, \theta^{(m)}\right)$ be the target distribution from which missing data is

⁴ $P_{t|\widehat{t-1}}$ denotes the predicting error variance-covariance matrix of the state variable.

generated in the second step of 3.2 and let $KF \left(\xi_{t+1}, P_{t+1} | \widehat{\xi}_{t|t-1}^{(m)}, \widehat{P}_{t|t-1}^{(m)}, w_t, \hat{z}_t^{(m)}, \theta^{(m)} \right)$ incorporate the prediction of state variables next period of the standard Kalman Filter as in the third step of 3.2. The pseudo-algorithm is illustrated in the next subsection and the derivation of $f \left(\begin{bmatrix} z_t \\ \xi_t \end{bmatrix} | \xi_{t|t-1}^{(m)}, P_{t|t-1}^{(m)}, w_t, \theta^{(m)} \right)$ and Kalman Filter with data augmentation steps are shown in the appendix.

2.2.3 Pseudo-Algorithm

Without loss of generality, this is the case of missing observation at period $T - 2$ and $T - 1$ but observed at T .

1. Initialize with $\theta^{(0)}$
2. Evaluate $p_1 \left(\theta^{(m)} | w, z_q, \hat{z}_{-q}^{(m)} \right)$ using KF with data augmentation⁵ and store $\hat{z}_{-q}^{(m)}$
3. Draw θ^* from the proposal distribution $N(\theta^{(m)}, V)$
4. Evaluate $p_2 \left(\theta^*; w, z_q, \hat{z}_{-q}^{(m)} \right)$ using standard KF and $\hat{z}_{-q}^{(m)}$
5. $posterior\ odds = p_2 \left(\theta^* | \hat{z}_{-q}^{(m)}, z_q, w \right) - p_1 \left(\theta^{(m)} | \hat{z}_{-q}^{(m)}, z_q, w \right) \sim Unif(0, 1)$
6. If accepted, record $\theta^{(m+1)} = \theta^*$, else $\theta^{(m+1)} = \theta^{(m)}$
7. Repeat 2 to 6 for $m = 1, \dots, m_{sim}$

⁵ Shown in appendix

2.3 Application: RBC

This section presents the standard Real Business Cycle model with labor-leisure choice.

2.3.1 The model

The model is a decentralized economy with household's utility maximization and firm's profit maximization problems. Household maximize the following lifetime utility by allocating consumption, labor supply and capital investment,

$$U_0 = \max_{\{c_t, l_t, k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t, l_t)$$

subject to $\forall t = 1, \dots, T$

$$c_t + i_t = r_t k_t + w_t l_t$$

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

And firm maximize the following profit each period facing the rental rate and wage for capital and labor demand,

$$\Pi_t = y_t - r_t k_t - w_t l_t$$

subject to Cobb-Douglas technology

$$y_t = A_t k_t^\alpha l_t^{1-\alpha}$$

with exogenous total factor productivity process

$$\ln(A_t) = \rho_z \ln(A_{t-1}) + \sigma_z \varepsilon_t$$

$$\varepsilon_t \sim iid \quad N(0, 1)$$

and the markets clear. Thus the competitive equilibrium is satisfying following conditions.

$$u'_c(c_t, l_t) = \beta E_t [u'_c(c_{t+1}, l_{t+1}) (r_{t+1} + 1 - \delta)]$$

$$-u'_l(c_t, l_t) = u'_c(c_t, l_t) w_t$$

$$r_t = \alpha A_t k_t^{\alpha-1} l_t^{1-\alpha}$$

$$w_t = (1 - \alpha) A_t k_t^\alpha l_t^{-\alpha}$$

$$c_t + k_{t+1} - (1 - \delta) k_t = y_t$$

$$y_t = A_t k_t^\alpha l_t^{1-\alpha}$$

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

$$\ln(A_t) = \rho_z \ln(A_{t-1}) + \sigma_z \varepsilon_t$$

After log-linearization of this system, the solution can be expressed in the standard State Space Representation⁶.

$$\xi_{t+1} = F\xi_t + Gv_{t+1}$$

$$\eta_t = H'\xi_t + u_t$$

$$\xi_t = \{k_t, a_t\}$$

$$\eta_t = \{w_t, z_t\}$$

$$w_t = \{r_t, wage_t, l_t\} \quad \forall t = 1, 2, 3, 4, \dots, T$$

$$z_t = \{y_t, c_t, i_t\} \quad \forall t = 1, 4, 7, 10, \dots, T$$

Consistent with the realistic setting of data availability, assume wage rate, and labor hours are observed monthly while output, consumption and investment are observed quarterly with aggregation.

⁶ From now on, the variables indicate the log deviation from the deterministic steady state.

2.4 Estimation Exercise

This section shows the benefit of data augmentation method by having an estimation exercise based on simulated data. An alternative method is to use quarterly data only to estimate the same model⁷. Since monthly data will be aggregated into quarterly data, there is information loss in principle. In contrast, data augmentation will not only maintain the monthly observed data but also will enhance the quarterly data with augmented data and thus will not suffer from the information loss. This will propagate into the efficiency gain of parameters estimates and thus the simulation exercise below is going to compare root mean squared errors of the parameters estimates between the data augmentation method and the coarse estimation.

2.4.1 Simulation

From RBC model, a vector of five observables composed by wages, labor hours, output, consumption and investment is simulated at a monthly frequency over 100 periods, which approximately reflects 8 years of a sample period, with addition of noise to observables by 20% of their own variances. Also, 100 cross section data are simulated. Calibration is as follows,

Table 2.1: RBC True Parameters

Parameter	Calibration
β	0.9992
γ	1.5
δ	0.025/3
α	0.3
ρ_z	0.9
σ_z	1

⁷ Shown in appendix

2.5 Results

Table 2.2 shows the summary posterior estimates.

Table 2.2: RBC Simulation

	$\gamma = 1.5$		$\delta = 0.0083$		$\alpha = 0.3$	
<i>statistics</i>	γ	<i>RMSE</i>	δ	<i>RMSE</i>	α	<i>RMSE</i>
<i>Augment</i>	1.5041 0.0676	0.0993	0.0090 0.0018	0.0018	0.3004 0.0069	0.0079
<i>Coarse</i>	1.5619 0.1532	0.2012	0.0099 0.0028	0.0030	0.2984 0.0149	0.0160

	$\rho_z = 0.9$		$\sigma_z = 0.01$	
<i>statistics</i>	ρ_z	<i>RMSE</i>	σ_z	<i>RMSE</i>
<i>Augment</i>	0.9000 0.0063	0.0087	0.0101 0.0008	0.0011
<i>Coarse</i>	0.8988 0.0117	0.0138	0.0103 0.0011	0.0012

For each simulated set of data, parameters have been drawn for quarterly million times. This one simulation gives one set of posterior estimates. And this whole procedure is done over 100 times since there are 100 cross sectional data. Top row has the parameters of interest with true values. For each parameter, the first column shows the cross sectional mean of posterior means from 100 simulations and the value below is the cross sectional standard deviation of posterior means. The second column shows the root mean square of these two statistics. So the smaller RMSE is, the more efficiency gain. Therefore, it is evident that all five structural parameters show clearly an efficiency gain from data augmentation compared to the coarse estimation.

2.6 Conclusions

This chapter has developed a method for estimation of DSGE models with mixed frequency data exploiting Bayesian econometric theory. As an example, the method has been applied to the standard RBC model with labor-leisure choice and the simulation exercise showed an efficiency of structural parameters' estimates.

Temporal Aggregation Bias and Mixed Frequency Estimation of a New Keynesian model

3.1 Introduction

The key feature of New Keynesian models is Calvo type of price friction which is central to understand the propagation mechanism of real activities in response to monetary policy shock. The degree of this price friction is parameterized by a probability in which a firm cannot reoptimize within one period. Due to its quantitative importance, Calvo parameter has been extensively studied from both micro-dataset and macro models but its estimates are widely dispersed. From micro dataset, Bils and Klenow (2004) finds the median price duration of firms to be 4.3 months while Nakamura and Steinsson (2008) argues 8.7 months once irrelevant pricing behaviors such as temporary salescuts are controlled. On the other hand, the average price duration implied by Calvo parameter in macro models ranges from 8 months to 24 months depending on the models and estimation strategies¹.

This discrepancy between microevidence and macro model's estimates is generally

¹ Christiano and Eichenbaum (2005), Smets and Wouters (2005), and Del Negro and Schorfheide (2008) are only a few of many other examples.

perceived as a misspecification of New Keynesian models. But this perception is neglecting the fact that microevidence is based on monthly observations while macro models are estimated at a quarterly frequency. If a temporal aggregation bias in Calvo parameter is present, modeling the pricing behavior at a coarse frequency while the true decision time interval is shorter can in fact be misleading. Thus, this paper asks whether a frequency misspecification of a New Keynesian model results in a temporal aggregation bias of the Calvo parameter. When a New Keynesian model is estimated at a quarterly frequency while the true data generating process is the same model but at a monthly frequency, the Calvo parameter is upward biased and hence implies longer average price duration. This suggests estimating a New Keynesian model at a monthly frequency may yield different results.

In order to resolve the temporal aggregation bias caused by the frequency misspecification in DSGE models, estimation of a model at the true frequency is necessary. However, when a monthly specified DSGE model brought to an estimation, a technical challenge emerges since data is available at different frequencies. For example, interest rate, inflation rate, wage rate and consumption are available at monthly frequency while GDP and investment are only at quarterly. One way is to identify the analytical mapping from a monthly specification of a model to a quarterly specification. In this way, the converted model at the quarterly frequency can be estimated with quarterly data. This conversion is well known with simple statistical models, for example, monthly AR(1) is equivalent to quarterly ARMA(1,1)². However, it is not clear how to convert a monthly specified DSGE model to a quarterly specification in general due to the forward looking nature of decision variables. Therefore, a standard estimation procedure needs to be modified to estimate monthly specified DSGE models.

² To my knowledge, Working (1960) is an early paper that illustrates this example and a more comprehensive study of temporal aggregation with various statistical models is shown in Marcellino (1999).

To accommodate mixed frequency dataset, the quarterly data can be treated as an observable variable that has missing observations. Then, "imputing some values" to these missing observations makes a complete dataset that facilitates the standard procedure of estimation. This is called a data augmentation and this paper proposes an estimation strategy using this data augmentation technique that is borrowed from Bayesian estimation literature in which "imputing some values" to missing observations is based on simulations. The simulation of missing observations is by a direct sampling from a distribution of missing observations. This distribution in general can be expressed by a marginal distribution of a joint distribution that is defined not only in terms of model's parameters but also jointly in terms of an auxiliary variable for missing observations conditional on the available data. Accordingly, MCMC algorithm can be extended to sample both parameters and missing observations similar to sampling a mixture model. MCMC theories have proven this modified algorithm converges at the geometric rate and thus central limit theorem ensures the consistency of marginal sampled estimates for parameters³. Gelfand and Smith (1990) refers this joint distribution as "Rao-Blackwellization" of an original target distribution of parameters since this is a form of a generalization of the distribution and thus incorporates richer information by allowing "imputing values" into missing observations which is in some sense an example of Rao-Blackwell Theorem. And they show theoretically that the advantage of the data augmentation is efficiency gains of parameters' estimates and this was further generalized by Liu et al. (1994).

A sampling scheme for the data augmentation procedure is not unique at least in the context of estimating DSGE models since missing observations are multiple periods. However, due to the general structure of DSGE models, it is not feasible to sample the whole missing observations in one step which would have been the most efficient method. Instead, this paper chooses to sample missing observa-

³ Diebolt & Robert (1994)

tions sequentially period by period using information from adjacent periods since an analytical distribution of missing observations for each period can be derived from marginal distributions used in Kalman Filter updating step. This sampling scheme is similar to Elerian et al. (2001) and Eraker (2001) in which missing observations are sampled from a marginal distribution conditioning on observations of two closest periods back and forth. Alternative to the data augmentation, Kalman Filter can be modified in two ways to evaluate the likelihood of proposed values for parameters under mixed frequency dataset without the data augmentation procedure. Thus, in order to demonstrate the advantage of the data augmentation over these alternatives, a Monte Carlo experiment on a medium scale New Keynesian model is presented. The second main finding of this paper is that data augmentation estimation delivers lower root mean squared errors for parameters of interest in a medium scale New Keynesian model.

Lastly, the medium scale New Keynesian model is brought to the actual data, and with the benchmark estimation method the average price duration implied by the monthly model is 5 months while that by the quarterly model is 20.7 months.

3.2 Temporal Aggregation Bias

This section shows the temporal aggregation bias issue with AR(1) model and with 3 equation New Keynesian model due to a frequency misspecification.

3.2.1 AR(1)

A monthly AR(1) process is converted into an ARMA(1,1) process when aggregated into quarterly frequency⁴. Given this conversion, this section demonstrates Monte-Carlo simulation results and also the time aggregation bias when this conversion is ignored and estimated simply with quarterly AR(1) specification. Assume that the true model is

$$a_t = \rho a_{t-1} + \sigma \varepsilon_t, \forall t = 1, 2, \dots, T$$

$$\varepsilon_t \sim iid \quad N(0, 1)$$

Monthly observations are simulated from this model for $T = 300$ and with 10,000 MC simulations. And suppose an econometrician observes the aggregated data at quarterly frequency with following aggregation scheme.

$$\tilde{a}_t = a_t + a_{t-1} + a_{t-2}$$

For each dataset of 10,000 simulations, the econometrician can estimate quarterly AR(1), i.e. a misspecified model, or quarterly ARMA(1,1) if the true model is known to the econometrician. So

$$\tilde{a}_t = \tilde{\rho}_q \tilde{a}_{t-3} + \tilde{\sigma}_1 \tilde{\nu}_t$$

or

$$\tilde{a}_t = \rho_q \tilde{a}_{t-3} + \sigma_1 \nu_t + \sigma_2 \nu_{t-3}$$

Thus, given the exact conversion, the persistence parameter should be $\rho_q \equiv \rho^3$. The next table shows the estimation results of this persistence parameter. Assume $\rho = 0.9$

⁴ Derivation of a quarterly specification from aggregation of monthly AR(1) model is shown in the appendix.

is the true value with the monthly model. $\tilde{\rho}_q$ is an estimate with quarterly AR(1) while $\hat{\rho}_q$ is an estimate with quarterly ARMA(1,1). These estimates are an average of point estimates over MC simulations. And the value below is the standard deviation of those point estimates.

Table 3.1: AR(1) Example

ρ	ρ^3	$\tilde{\rho}_q$	$\hat{\rho}_q$
0.9	0.7290	0.8011 0.0285	0.7236 0.0469

It clearly shows upward bias on $\hat{\rho}_q$ with AR(1) specification and this is due to the misspecification. When estimated with a correct model, ARMA(1,1), it gives a value close to the truth. With this simple statistical model that has backward looking variable, the true model can be retrieved even when data is aggregated over time because the exact conversion from a monthly specification to a quarterly specification is known. However, this exact conversion will not be apparent with DSGE models where forward looking variables are present and thus aggregation of variables with expectations into coarser time interval is not clear.

3.2.2 Simple New Keynesian model

A parsimonious New Keynesian model with Calvo pricing feature is as follows in log-linearized form.

$$\pi_t = \beta E_t \pi_{t+1} + \kappa y_t + \varepsilon_{\mu,t}$$

$$E_t y_{t+1} = y_t + R_t - E_t \pi_{t+1}$$

$$R_t = \gamma_{\Pi} \pi_t + \gamma_y (y_t - y_{t-1}) + \varepsilon_{m,t}$$

where

$$\varepsilon_{m,t} = \sigma_m \eta_{m,t}$$

$$\varepsilon_{\mu,t} = \rho_{\mu}\varepsilon_{\mu,t-1} + \sigma_{\mu}\eta_{\mu,t}$$

$$\kappa \equiv \frac{(1 - \beta\theta)(1 - \theta)}{\theta}$$

For simplicity, log-utility and inelastic labor supply are assumed and the monetary authority targets the interest rate following Taylor rule that responds to inflation and to growth rate of output. And the source of uncertainty is price markup shock $\varepsilon_{\mu,t}$ with AR(1) and monetary shock $\varepsilon_{m,t}$ with iid process. The monetary policy in this model is no longer neutral due to the price friction and thus causes a reaction of real activities. And it is well known that the degree of Calvo parameter, θ , determines the length of propagation of the real activities in response to the monetary policy shock. The observables are interest rate, inflation rate and quarterly growth rate of output. The reason why quarterly measure for output is used is to mimic the estimation with real data in which monthly growth rate of output is not observed but only quarterly. Suppose for now the quarterly growth rate of output is $y_t^Q - y_{t-3}^Q$. Two measurement errors are added to observables so that stochastic singularity is avoided⁵. So the observation equation of this model can be expressed as

$$\begin{bmatrix} R_t \\ \pi_t \\ y_t^Q - y_{t-3}^Q \end{bmatrix} = H'\xi_t + \begin{bmatrix} 0 \\ \sigma_{\nu_1}\nu_{1t} \\ \sigma_{\nu_2}\nu_{2t} \end{bmatrix}$$

where H is derived from a solution of the model and state variable, $\xi_t \equiv [x_t, \varepsilon_{\mu,t}, \varepsilon_{m,t}]'$ with x_t being a vector of predetermined variables⁶. Calibrated parameters are shown in the next table⁷, and simulated this model with $T = 100$ and with 100 MC simulations.

⁵ To avoid stochastic singularity, only one measurement error is needed but I added one more because on rare cases a numerical singularity could arise.

⁶ Given the aggregation scheme explained in the following paragraph $x_t \equiv [y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, \pi_{t-4}]'$

⁷ I also checked with different degree of calibrated Calvo parameter, $\theta \in \{0.85, 0.75\}$ and find temporal aggregation biases. Also different values of Taylor rule parameters have not affected these findings with Calvo parameters.

Table 3.2: Simple NK True Parameters

β	θ	γ_y	γ_Π	ρ_μ	$\sigma_\mu, \sigma_m, \sigma_\nu$
0.9992	0.9	0.15	1.5	0.9	0.01

In order to estimate the model at quarterly frequency, aggregation schemes for each observable is necessary and they generally differ depending on whether the variable is flow or stock. Interest rate and inflation rate are growth rates of stock variables and thus it is relatively easier to aggregate compared to flow variables such as GDP. Since quarterly interest rate is a three months of compounded monthly interest rates,

$$R_t^Q = R_t + R_{t-1} + R_{t-2}$$

Similarly, quarterly inflation rate is inflation rate from three months prior to current month ⁸ so,

$$\pi_t^Q = \pi_t + \pi_{t-1} + \pi_{t-2}$$

And I follow NIPA convention of GDP aggregation which sums monthly nominal GDPs. So in log-linearized form, the real output would be

$$y_t^Q \equiv \frac{1}{3} [y_t + (y_{t-1} - \pi_t) + (y_{t-2} - \pi_t - \pi_{t-1})]$$

Since the original monthly model has already generated quarterly growth rate for output, observations from every last month of quarters can simply be collected to construct the quarterly dataset. Given these aggregation scheme, the quarterly ob-

⁸ I also experimented having simply

$$\pi_t^Q \equiv \frac{1}{3} [(\pi_t + \pi_{t-1} + \pi_{t-2}) + (\pi_{t-1} + \pi_{t-2} + \pi_{t-3}) + (\pi_{t-2} + \pi_{t-3} + \pi_{t-4})]$$

which did not affect the results.

servables are

$$\begin{bmatrix} R_t^Q \\ \pi_t^Q \\ y_t^Q - y_{t-3}^Q \end{bmatrix}$$

Measurement errors are attached to all three observables reflecting the fact that this quarterly specified model might be potentially misspecified. Following the standard Bayesian technique (An and Schorfheide (2007)), New Keynesian model which is analytically same as above is estimated based on those aggregated observables. The discount factor is calibrated by having $\tilde{\beta}_q = \beta^3$ so that the steady states of interest rates are consistent across two frequencies. The diffuse priors are set including Calvo parameters except Taylor rule parameters⁹ are set to have reasonable acceptance rate and desirable convergence of MC chains.

Table 3.3: Priors

$\tilde{\theta}_q$	$\tilde{\gamma}_y$	$\tilde{\gamma}_\pi$	$\tilde{\rho}_\mu$	$\tilde{\sigma}_m, \tilde{\sigma}_\mu, \tilde{\sigma}_\nu$
<i>Unif</i> (0, 1)	<i>N</i> (0.15, 0.1)	<i>N</i> (1.5, 0.2)	<i>Unif</i> (0, 1)	<i>IG</i> (0.02, 2)

In addition to the standard quarterly model, a quarterly model in which price markup shock has ARMA(1,1) process and monetary policy shock has MA(1) process is also estimated. Although this does not necessarily have a theoretical justification, it is worth to examine whether additional MA terms can correct the biases following Smets and Wouters (2005). For the comparison of Calvo parameter specified at different time frequencies, an envelope calculation is needed. Because this parameter is a probability that the monopolistic competitive firm cannot reoptimize their prices within its specified decision time interval, the average price duration for those

⁹ Under various calibration schemes, Taylor rule parameters had generally bad identifications in quarterly estimation but this anomaly did not affect Calvo estimates.

firms can be computed and further the implied probability under different decision time interval, say coarser interval, can be backed out from this price duration. For example, suppose $\theta = 0.9$ in the monthly model, then the average price duration of firms is 10 months¹⁰ which is equivalent to $\frac{10}{3}$ quarters. So the implied probability¹¹ at quarterly frequency would then be $\theta_q = 0.7$. In principle, if there were no temporal aggregation bias on this parameter, the quarterly estimation results would be expected to have $\theta_q = 0.7$. The results are shown in the following table.

Table 3.4: Estimation on Simple NK model

methods	θ	γ_y	γ_Π	ρ_μ	$\log\sigma_{\mu_2}$	$\log\sigma_\mu$	$\log\sigma_m$
True	0.9 (0.7)	0.15	1.5	0.9 (0.729)	NA	-4.6052	-4.6052
Q-AR	0.8977 0.0173	0.0784 0.0154	1.4997 0.0221	0.6771 0.0142	NA	-2.5082 0.1540	-3.3380 0.3725
Q-ARMA	0.8642 0.0529	0.1207 0.0182	1.5273 0.0279	0.5959 0.0274	-0.2306 0.3832	-2.8447 0.2802	-3.0052 0.2671

The values in the parenthesis next to true values in the first row are the quarterly parameter values that are implied by our conversion schemes with Calvo parameter and persistence of AR(1). The second row shows substantial upward aggregation bias with respect to both Calvo parameter and persistence parameter of markup shock. Including moving average terms to the exogenous processes mitigates the bias on persistence parameter but does not eliminate the bias completely and moreover does not improve Calvo parameter estimate.

This is quite a surprising outcome relative to AR(1) example but can lead to an interesting conclusion that specifying a rational expectation model at different frequencies can produce a different estimates of certain structural parameters which cannot be resolved by simply adding MA terms. Hence, this difference is difficult to

¹⁰ average price duration = $\sum_{j=0}^{\infty} \theta^j = \frac{1}{1-\theta}$

¹¹ $\tilde{\theta} = 1 - \frac{1}{\text{avg.price dura.}}$

be reconciled when rational expectation models are specified at different frequencies unless the model is estimated at its true frequency. So if a research believes in a model that is at higher frequency than he or she observes, this results show that modeling at the lower frequency due to data availability can be problematic.

3.3 Mixed Frequency Estimation Strategies

This section shifts the focus to an estimation strategy under a mixed frequency dataset without specifying DSGE model at a lower frequency that gives rise to a time aggregation bias. This section provides three different alternatives of mixed frequency estimation. Although these strategies are all correct in a sense of the convergence of the markov chains and the asymptotic consistency of estimates but the data augmentation shows advantage in the efficiency of parameter estimates. The detailed algorithms are explained in this section and the efficiency performance of estimation methodologies are compared by root mean squared errors of parameter estimates of a medium scale New Keynesian model.

3.3.1 Preliminary Setup

Given equilibrium conditions of a model, those conditions are log-linearized around a deterministic steady state and can be summarized by state-space representation that allows Bayesian estimation framework to be implemented.

$$\xi_{t+1} = F(\theta) \xi_t + v_{t+1}, v_t \sim N(0, Q(\theta))$$

$$\eta_t = H(\theta)' \xi_t + u_t, u_t \sim N(0, R(\theta))$$

where the first equation is the state equation and the second one is the measurement equation describing the evolution of the observables as functions of the endogenous and exogenous states. ξ_{t+1} is an $n_x \times 1$ vector of a latent state and η_t is an observed variable with $n_y \times 1$. Bayesian estimation is to maximize a posterior density function of state-space equations constructed by setting priors on parameters of a model and by the likelihood function.

$$p(\theta|\mathbf{y}^T) = \pi(\theta) L(\theta|\mathbf{y}^T)$$

Prediction error decomposition (Harvey (1991)) facilitates the evaluation of the likelihood function period by period using Kalman Filter that optimally estimates latent variables.

$$L(\theta; \mathbf{y}^T) = \prod_{t=1}^T \ell(\theta; y_t, \hat{x}_{t|t-1})$$

Since the parameters of a DSGE model are highly nonlinear, Metropolis-Hastings algorithm is applied to numerically explore the shape of the posterior with parameter values. However, when data are available at multiple frequencies, this standard procedure of Bayesian estimation is not straightforward. For the observation variable, η_t , is incomplete and has missing observations which prevents from evaluating likelihood function with standard Kalman Filter.

An example with more notations need to be introduced to accommodate this mixed frequency dataset. As a practical purpose¹², this paper is restricting to a case where data is combined with monthly and quarterly time series and estimating monthly model. Hence, η_t can be partitioned into two variables,

$$\eta_t = \begin{bmatrix} z_t \\ w_t \end{bmatrix}$$

where w_t is a monthly observable while z_t is a quarterly observable. Then, the state space representation can be rewritten

$$\xi_{t+1} = F\xi_t + v_{t+1}$$

$$\begin{bmatrix} z_t \\ w_t \end{bmatrix} = \begin{bmatrix} H'_z \\ H'_w \end{bmatrix} \xi_t + \begin{bmatrix} u_t^z \\ u_t^w \end{bmatrix}$$

and note that $t = 1, 2, 3, \dots, T$ is a sequence of months¹³. In order to disentangle z_t into one that is observed and one missing, subsequence notations for time is necessary.

¹² At least two of the following methods can also deal a situation where dataset is constructed by more than two frequencies in principle. However, practicality of estimating under this circumstance is questionable.

¹³ For the consistency of notations, let T be the last month of the last quarter.

Let q_i be the last month of every quarter in which data for both variables, z_t and w_t , are collected and $q_i - 1$ and $q_i - 2$ be the months in which only data for w_t are available. So when $Q \equiv \frac{T}{3}$

$$\{q_i\}_{i=1}^Q \equiv \{3, 6, 9, \dots, T - 3, T\}$$

$$\{q_i - 1\}_{i=1}^Q \equiv \{2, 5, \dots, T - 1\}$$

$$\{q_i - 2\}_{i=1}^Q \equiv \{1, 4, \dots, T - 2\}$$

The history notations are as follows, the monthly variables will be

$$\mathbf{w}^t = \{w_\tau\}_{\tau=0}^t$$

while the quarterly observed variable is

$$\mathbf{z}^t = \tilde{\mathbf{z}}^{q_i} \cup \tilde{\mathbf{z}}^{q_i-1} \cup \tilde{\mathbf{z}}^{q_i-2}$$

where

$$i \in \{i : q_i \leq t\}$$

tildes denote the collection of only one months from each quarter. Given above notations, the posterior density function of interest under mixed frequency dataset becomes

$$p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q})$$

In principle, the mixed frequency estimation strategies will differ by how to evaluate likelihood function that constitutes this posterior density. In short, the stacking method will redefine the state space representation at quarterly frequency while the underlying model is monthly, Durbin-Koopman method will modify Kalman Filter in which the dimension of Kalman Filter gain changes consistent with the dimension of available data in each period, and the data augmentation method simulates the missing observations to fill the gap by demarginalizing the above posterior density

so that it transforms into a joint density which is in terms of not only parameters but also of missing observations.

3.3.2 Data Augmentation Method

Data augmentation method is based on sampling from a joint posterior density that is constructed by "Rao-Blackwellization" or "demarginalization" of an original posterior density. In other words, an auxiliary variable that stands for missing observations is introduced and filling this variable with a proxy value can complete the mixed frequency dataset and thus allows to evaluate the posterior density under a complete dataset. And this proxy will be simulated at every iteration of MCMC algorithm from a tractable distribution that is derived from the model under certain parameter values. So the joint density is

$$p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) = \int_{\mathbf{z}} \tilde{p}(\theta, \tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) d\mathbf{z}$$

$$\tilde{p}(\theta, \tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) = \ell(\theta, \tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) \pi(\theta)$$

The original posterior density is a marginal density of this joint density by integrating out the auxiliary variable, missing observations. Thus, this artificial extension of a posterior density function is only for the computational device and does not invalidate the inference on the structural parameters, θ , as will be shown below. So the objective is to sample the parameters and missing observations¹⁴ jointly but this would be feasible by separating this joint density function into two stage samplings.

$$\{\hat{\mathbf{z}}^{q_Q-1}, \hat{\mathbf{z}}^{q_Q-2}\}^{(m)} \sim f\left(\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \hat{\theta}^{(m)}\right)$$

$$\hat{\theta}^{(m+1)} \sim p\left(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \{\hat{\mathbf{z}}^{q_Q-1}, \hat{\mathbf{z}}^{q_Q-2}\}^{(m)}\right)$$

¹⁴ Whenever a hat, $\hat{\cdot}$, is labeled, it means sampled values for the variables.

This is a well known strategy in statistics when the joint density function of two variables is complicated. In this scheme, one variable can be easily sampled by having other variable as a condition and vice versa. The simplest example would be a mixture model of two normally distributed random variable conditional on the other variable. The bivariate distribution of this example is hardly tractable which makes difficult to sample two variables in one step but becomes much easier with two stage Gibbs sampler algorithm. And this alternating samplings will eventually converge to the desired joint distribution of interest. This convergence is proved for more general cases by Diebolt and Robert (1994) and here it merely repeats the theorem of convergence.

Corollary 1. *The sequences $\{(\hat{\mathbf{z}}^{qQ-1}, \hat{\mathbf{z}}^{qQ-2})^{(m)}\}$ and $\{\hat{\theta}^{(m)}\}$ are ergodic Markov chains with respective invariant distributions $\int_{\theta} \tilde{p}(\theta, \tilde{\mathbf{z}}^{qQ-1}, \tilde{\mathbf{z}}^{qQ-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) d\theta$ and $\int_{\mathbf{z}} \tilde{p}(\theta, \tilde{\mathbf{z}}^{qQ-1}, \tilde{\mathbf{z}}^{qQ-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) d\mathbf{z}$. Moreover, the convergence is uniformly geometric, i.e. there exists $0 < \rho < 1$ and $C > 0$ such that*

$$\int_{\theta} |p^m(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) - p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ})| d\theta \leq C\rho^m$$

where

$$p^{(m)}(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) \equiv \int_{\theta} K_T(\theta^{(m)} | \theta^{(m-1)}) p^{(m-1)}(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) d\theta$$

and

$$K_T(\theta^{(m)} | \theta^{(m-1)}) \equiv \int_{\mathbf{z}} p(\theta^{(m)} | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}, \{\hat{\mathbf{z}}^{qQ-1}, \hat{\mathbf{z}}^{qQ-2}\}^{(m-1)}) f(\tilde{\mathbf{z}}^{qQ-1}, \tilde{\mathbf{z}}^{qQ-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}, \hat{\theta}^{(m-1)}) d\mathbf{z}$$

and

$$p(\theta | \mathbf{w}^T, \mathbf{z}^{qQ}) = \int_{\mathbf{z}} \tilde{p}(\theta, \tilde{\mathbf{z}}^{qQ-1}, \tilde{\mathbf{z}}^{qQ-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) d\mathbf{z}$$

Given this geometric convergence and additionally a finite variance of θ , the Central Limit Theorem can be applied to ensure the asymptotic consistency of parameter

estimates which is an average of $\hat{\theta}$. So

Corollary 2. *The Central Limit Theorem holds, i.e. the sample estimates are consistent with finite variance,*

$$\bar{\theta} = \frac{1}{\sqrt{M}} \sum_{m=1}^M (\theta^{(m)} - E_p [\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}]) \xrightarrow{\mathcal{L}} N(0, V)$$

where

$$V = \text{var}_p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{qQ}) < \infty$$

Sampling Scheme for Data Augmentation

The data augmentation step, $\{\{\hat{\mathbf{z}}^{q_i-1}, \hat{\mathbf{z}}^{q_i-2}\}_{i=1}^Q\}^{(m)} \sim f(\{\mathbf{z}^{q_i-1}, \mathbf{z}^{q_i-2}\}_{i=1}^Q | \mathbf{w}^T, \mathbf{z}^{qQ}, \hat{\theta}^{(m)})$, in general can be implemented in different ways. One way is to simulate the whole set of missing observations at once from a distribution implied by a model. This was shown in Eraker et al. (2008) with VAR(1) model. However, this was only feasible when the target distribution from which missing observations are drawn can be derived analytically in terms of observed data and parameters of a model. But in DSGE model's estimation in which prediction errors are estimated sequentially period by period due to existence of latent variables, simulating missing observations in one step is not feasible¹⁵ since the target distribution of a whole set of missing observations cannot be derived in general. But similar to a conditional distribution of a state variable in Kalman Filter, a target distribution for missing observation in one period can be derived in terms of parameters and observed data. Hence, data augmentation can be done by Gibbs sampling from a target distribution sequentially period by period in the first stage of MCMC algorithm. However, since data augmentation typically involves observations not only of past but also of future, a predictive distribution from Kalman Filter cannot simply used in this context. Instead, the

¹⁵ There are some cases when DSGE model can be transformed into VAR(2) model under certain circumstances and therefore this method could be feasible. See Ravenna(2006) for more detail with this transformation. However, this paper focuses on mixed frequency estimation of DSGE models in a general framework.

state space form is redefined into a companion form and this facilitates derivation of distribution of missing observations in terms of observations in adjacent periods including both past and future. To illustrate this point, the standard Kalman Filter is demonstrated first and then the target distribution of data augmentation is discussed.

Define the prediction error variance-covariance of ξ_t and that of η_t .

$$\Sigma_{t|t-1} \equiv E (\xi_t - \xi_{t|t-1}) (\xi_t - \xi_{t|t-1})'$$

$$\Omega_{t|t-1} \equiv E (\eta_t - \eta_{t|t-1}) (\eta_t - \eta_{t|t-1})'$$

Kalman Filter gain, K_t , minimizes $\Sigma_{t|t}$ so that $\xi_{t|t}$ is optimally estimated. So Kalman Filter is

1. Starting with given $\xi_{t|t-1}$ and $\Sigma_{t|t-1}$
2. $\Omega_{t|t-1} = H' \Sigma_{t|t-1} H + R$
3. $\eta_{t|t-1} = H' \xi_{t|t-1}$
4. $K_t = \Sigma_{t|t-1} H (H' \Sigma_{t|t-1} H + R)^{-1}$
5. $\Sigma_{t|t} = \Sigma_{t|t-1} - K_t H' \Sigma_{t|t-1}$
6. $\xi_{t|t} = \xi_{t|t-1} + K_t (\eta_t - \eta_{t|t-1})$
7. $\Sigma_{t+1|t} = F \Sigma_{t|t} F' + Q$
8. $\xi_{t+1|t} = F \xi_{t|t}$

The probabilistic interpretation of Kalman Filter, complementary to the minimizing variance of prediction errors of the state variables, says that the state variable, ξ_t , is updated by the conditional normal distribution with a new observation η_t . Hence,

$$N \left(\begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix} \middle| \eta^{t-1} \right) \rightarrow N \left(\xi_t \middle| \begin{bmatrix} \eta^{t-1} \\ \eta_t \end{bmatrix} \right)$$

The normality is due to the assumption on normally distributed structural shocks and measurement errors. Hence, given the history of observations up to period t , $\xi_{t|t}$ is optimally estimated by a conditional mean of the latter distribution. The missing observation can be treated like ξ_t and a conditional distribution can be similarly derived but instead a proxy value is sampled from this distribution instead of the conditional mean. But since observations of periods ahead are necessary in the conditional information, a trick is necessary to derive the desired distribution.

Now, redefine the state space representation in a companion form.

$$\tilde{\eta}_{q_i} = \tilde{H}\tilde{\xi}_{q_i} + \tilde{u}_{q_i}$$

$$\tilde{\xi}_{q_i} = \tilde{F}\tilde{\xi}_{q_i-1} + \tilde{v}_{q_i}$$

where

$$\tilde{\eta}_{q_i} = \begin{bmatrix} z_{q_i-2} \\ z_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix}, \tilde{\xi}_{q_i} = \begin{bmatrix} \xi_{q_i} \\ \xi_{q_i-1} \\ \xi_{q_i-2} \end{bmatrix}$$

This form naturally entails a joint distribution of three periods observations, so if η_{q_i} is fully observed the updating would have been

$$N \left(\begin{bmatrix} \tilde{\xi}_{q_i} \\ z_{q_i-2} \\ z_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} \middle| \tilde{\eta}^{q_i-1} \right) \rightarrow N \left(\begin{bmatrix} \tilde{\eta}^{q_i-1} \\ z_{q_i-2} \\ z_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} \middle| \tilde{\xi}_{q_i} \right) \quad (3.1)$$

But since $\tilde{\eta}_{q_i}$ is not fully observed, i.e. $\{z_{q_i-2}, z_{q_i-1}\}$ is not observed, the updating the distribution of state variables and missing observations conditioning on all observed

data in a quarter q_i can be

$$N \left(\begin{bmatrix} \tilde{\xi}_{q_i} \\ z_{q_i-2} \\ z_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} \middle| \tilde{\eta}^{q_i-1} \right) \rightarrow N \left(\begin{matrix} \tilde{\xi}_{q_i} \\ z_{q_i-2} \\ z_{q_i-1} \end{matrix} \middle| \begin{bmatrix} \tilde{\eta}^{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} \right) \quad (3.2)$$

Hence, missing observations can be simulated from this conditional density. This distribution will still be normal and can be derived analytically¹⁶. So the standard Kalman Filter would have estimated state variables as in (3.1) when $\tilde{\eta}_{q_i}$ is fully observed, but data augmentation by simulating this target distribution as in (3.2) is taken since $\{z_{q_i-2}, z_{q_i-1}\}$ is not observed. Note that $\{z_{q_i-2}, z_{q_i-1}\}$ jointly sampled every quarter and can only be sampled marginally from $\tilde{\xi}_{q_i}$ because DSGE models typically have many predetermined state variables and thus have variance singular which prevents from jointly sampling with missing observations. Thus, the distribution from which missing observations $\{z_{q_i-2}, z_{q_i-1}\}$ are drawn is the marginal normal distribution. Define this distribution as f_{q_i}

$$f_{q_i} \equiv N(z_{q_i-2}, z_{q_i-1} | \tilde{\eta}^{q_i-1}, z_{q_i}, w_{q_i-2}, w_{q_i-1}, w_{q_i})$$

This sampling scheme for data augmentation is similar to Elerian et al. (2001) and Eraker (2001) in a sense that observations adjacent to missing observations are used. w_{q_i-2}, w_{q_i-1} are contemporaneous observations coming from relationships of DSGE models between endogenous variables and $\{\tilde{\eta}^{q_i-1}, w_{q_i}, z_{q_i}\}$ are two periods observations adjacent to missing observations. Then, all the missing observations are drawn sequentially quarter by quarter and thus this constitutes the first stage of the sampling scheme.

$$f \left(\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \hat{\theta}^{(m)} \right) = \prod_{i=1}^Q f_{q_i} (z_{q_i-2}, z_{q_i-1} | \tilde{\eta}^{q_i-1}, z_{q_i}, w_{q_i-2}, w_{q_i-1}, w_{q_i})$$

¹⁶ Derivation is shown in appendix.

Obviously in principle the most efficient sampling scheme would be using the whole dataset as the conditional information, but there is a numerical issue that has to be confronted. The distribution for missing observations for each quarter can also be derived from smoothing Kalman Filter and thus incorporates more information from future observations. However, the variance of the distribution for missing observations, i.e. the analogue of variance f_{q_i} , involves inverting a covariance matrix of state variables which are normally singular due to presence of predetermined variables. Computational trick is to use a generalized inverse¹⁷ of this matrix, but this yields numerically unstable matrix to use it as variance of the target distribution. Because of this computational obstacle with using smoothing Kalman Filter which is potentially most efficient, the sampling scheme of this paper resorts to using information adjacent to missing observations which is also reasonably gaining efficiency compared to alternative estimation strategies as will be shown below¹⁸. The data augmenting Kalman Filter can be summarized by following.

1. Starting with given $\tilde{\xi}_{q_i|q_{i-1}}$ and $\tilde{\Sigma}_{q_i|q_{i-1}}$
2. $\tilde{\Omega}_{q_i|q_{i-1}} = \tilde{H}'\tilde{\Sigma}_{q_i|q_{i-1}}\tilde{H} + \tilde{R}$
3. Simulate $\{\hat{z}_{q_i-2}, \hat{z}_{q_i-1}\} \sim f_{q_i}(z_{q_i-2}, z_{q_i-1}|\tilde{\eta}^{q_i-1}, w_{q_i-2}, w_{q_i-1}, w_{q_i}, z_{q_i})$
4. $\tilde{\eta}_{q_i|q_{i-1}} = \hat{H}'\tilde{\xi}_{q_i|q_{i-1}}$
5. $K_t = \tilde{\Sigma}_{q_i|q_{i-1}}\tilde{H} \left(\tilde{H}'\tilde{\Sigma}_{q_i|q_{i-1}}\tilde{H} + \tilde{R} \right)^{-1}$
6. $\tilde{\Sigma}_{q_i|q_i} = \tilde{\Sigma}_{q_i|q_{i-1}} - \tilde{K}_t\tilde{H}'\tilde{\Sigma}_{q_i|q_{i-1}}$

¹⁷ In Matlab, "pinv.m" is used for generalized inverse.

¹⁸ Sampling scheme for choice of conditional information can further be relaxed for a case of randomly missing observations. See Kim(2009).

7. $\tilde{\xi}_{q_i|q_i} = \tilde{\xi}_{q_i|q_{i-1}} + \tilde{K}_t (\hat{\eta}_t - \tilde{\eta}_{q_i|q_{i-1}})$ where $\hat{\eta}_t = \begin{bmatrix} \hat{z}_{q_i-2} \\ \hat{z}_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix}$
8. $\tilde{\Sigma}_{q_{i+1}|q_i} = \tilde{F}\tilde{\Sigma}_{q_i|q_i}\tilde{F}' + \tilde{Q}$
9. $\tilde{\xi}_{q_{i+1}|q_i} = \tilde{F}\tilde{\xi}_{q_i|q_i}$

Multi-Block Gibbs Sampler Algorithm

After the missing observations are sampled sequentially from distributions conditioning on parameters, sampling parameters of a model in second stage takes place conditioning on this augmented dataset. The second stage is no different than the standard Metropolis-Hasting algorithm for sampling parameters conditioning on this complete dataset. Pseudo-algorithm is summarized in the following.

Pseudo-Algorithm

1. Initialize $\theta^{(0)}$
2. Draw $\{\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2}\}^{(m)} \sim f\left(\{\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2}\} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \hat{\theta}^{(m)}\right)$
3. Evaluate $p_1\left(\theta^{(m)} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \{\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2}\}^{(m)}\right)$
4. Draw θ^*
5. Evaluate $p_2\left(\theta^* | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \{\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2}\}^{(m)}\right)$
6. *posterior odds* $\sim Unif(0, 1)$
7. If accept, record $\theta^{(m+1)} = \theta^*$ else $\theta^{(m+1)} = \theta^{(m)}$
8. Repeat step 2~6 for $m = 1, \dots, M$

Gibbs sampling stage is in step 2. It is important to save the augmented dataset in this stage to be used in both step 3 and step 5. In short, this algorithm explores the shape of the joint density function

$$\tilde{p}(\theta, \tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \mathbf{z}^{q_Q}) = \ell(\theta, \tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) \pi(\theta)$$

and the likelihood function is evaluated via Kalman Filter with augmented dataset

$$\ell(\theta, \hat{\mathbf{z}}^{q_Q-1}, \hat{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) \equiv \prod_{t=1}^T \left[\ell(\theta, \hat{z}_t | \mathbf{w}^t, \mathbf{z}^{t-1})^{1-I(t \in \{q_i\}_{i=1}^Q)} \ell(\theta | \mathbf{w}^t, z_t, \mathbf{z}^{t-1})^{I(t \in \{q_i\}_{i=1}^Q)} \right]$$

where the missing observations are sequentially drawn from

$$\{\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2}\}^{(m)} \sim f\left(\{\tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2}\} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}, \hat{\theta}^{(m)}\right)$$

and $I(t \in \{q_i\}_{i=1}^Q)$ denotes an indicator function which is one if period belongs to last month of each quarter. Finally, the samples of $\{\theta^{(m)}\}_{m=1}^M$ are considered as the posterior distribution of parameters, and this is the marginal density function with integrating out the missing observations so that $p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) = \int_Z \ell(\theta, \tilde{\mathbf{z}}^{q_Q-1}, \tilde{\mathbf{z}}^{q_Q-2} | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) \pi(\theta) dz$.

3.3.3 Stacking Method

The stacking method is simply redefining the state space representation so that the observation variable is fully observed by stacking three months of observations into a one vector. Hence, the observables are transformed into

$$\tilde{\eta}_{q_i} \equiv \begin{bmatrix} z_{q_i} \\ w_{q_i} \\ w_{q_i-1} \\ w_{q_i-2} \end{bmatrix}$$

for $\forall \{q_i\}_{i=1}^Q$ so that it is always observed without missing observations. Accordingly, the state vector can be expressed as

$$\tilde{\xi}_{q_i} \equiv \begin{bmatrix} \xi_{q_i} \\ \xi_{q_i-1} \\ \xi_{q_i-2} \end{bmatrix}$$

Then the observation equations becomes

$$\tilde{\eta}_{q_i} = \tilde{H}\tilde{\xi}_{q_i} + \tilde{u}_{q_i}$$

where

$$\tilde{H} \equiv \begin{bmatrix} H_z & 0 & 0 \\ H_w & 0 & 0 \\ 0 & H_w & 0 \\ 0 & 0 & H_w \end{bmatrix}, \quad \tilde{u}_{q_i} \equiv \begin{bmatrix} u_{q_i}^z \\ u_{q_i}^w \\ u_{q_i-1}^w \\ u_{q_i-2}^w \end{bmatrix}$$

and state equation is

$$\tilde{\xi}_{q_i} = \tilde{F}\tilde{\xi}_{q_i-1} + \tilde{v}_{q_i}$$

where

$$\tilde{F} \equiv \begin{bmatrix} F^3 & 0 & 0 \\ F^2 & 0 & 0 \\ F & 0 & 0 \end{bmatrix}, \quad \tilde{v}_{q_i} \equiv \begin{bmatrix} I_{n_x} & F & F^2 \\ 0 & I_{n_x} & F \\ 0 & 0 & I_{n_x} \end{bmatrix} \begin{bmatrix} v_{q_i} \\ v_{q_i-1} \\ v_{q_i-2} \end{bmatrix}$$

So the posterior density in this method effectively is evaluated by assuming

$$p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) = \prod_{i=1}^Q \ell(\theta | \mathbf{w}^{q_i}, \mathbf{w}^{q_i-1}, \mathbf{w}^{q_i-2}, \tilde{\mathbf{z}}^{q_i}) \pi(\theta)$$

However, notice that the time interval for state space equations is quarterly which implies Kalman Filter gain for optimal estimates of state variables will be updated at quarterly frequency. ξ_{q_i-2} will be updated conditional on the history of observations up to q_{i-1} ($= q_i - 3$) which is still efficient, while ξ_{q_i-1} and ξ_{q_i} are not updated with the new observation at q_{i-2} and at q_{i-1} , respectively. Thus, if a monthly model is to be estimated in which the monthly observations are heavily influenced by the latent variables of the same months, this method will suffer from losing efficiency of state variable's estimates and will potentially lead to biases of parameters' estimates of the model. Furthermore, this method in general can only be applied to the case where mixed frequency data set has consistent frequency of missing observations within the same time series, i.e. it cannot be applied to the randomly missing observation case. For example, due to possibly the less sophisticated method of data collection in earlier

years of a sample which is common with emerging markets, one time series can have multiple mixed frequency observations. So if an econometrician is to estimate using this type of dataset with this method, one either has to synchronize the frequency of that particular time series by aggregating into coarser frequency or has to curtail the earlier part of the sample.

3.3.4 Durbin-Koopman Method

Durbin-Koopman method in this paper is an extension of an example with missing observation originally shown in Durbin and Koopman (2001). They showed whenever η_t is all missing for that particular period as opposed to only observing partially in the case of mixed frequency dataset, they simply estimate the state variable using the optimal estimate from the previous period which was updated up to using available observations. So $\xi_{t+1|t} = \xi_{t+1|t-1} = F\xi_{t|t-1}$ instead of $\xi_{t+1|t} = F\xi_{t|t}$. In this case Kalman Filter gain is zero in period t since there is no extra information to be exploited to estimate state variables. However, in the example in which at least some observations are partially available, Kalman Filter gain can be constructed with this available information at period t . Hence, in a standard case

$$K_t = \Sigma_{t|t-1}H (H'\Sigma_{t|t-1}H + R)^{-1}$$

and the state variable is updated with η_t by

$$\xi_{t|t} = \xi_{t|t-1} + K_t (\eta_t - H'\xi_{t|t-1})$$

Hence when only w_t is available, K_t can be a partitioned accordingly to be consistent with mixed frequency observations so that

$$K_t = \begin{bmatrix} K_t^z & K_t^w \end{bmatrix}$$

then the state variable can be updated by using this submatrix K_t^w ,

$$\xi_{t|t} = \xi_{t|t-1} + K_t^w (w_t - H'_w \xi_{t|t-1})$$

and when η_t is fully observed at the last month of each quarter, Kalman Filter gain is back to the standard one with a full dimension. Hence the posterior density is evaluated with

$$p(\theta | \mathbf{w}^T, \tilde{\mathbf{z}}^{q_Q}) = \prod_{t=1}^T \left[\begin{array}{c} \ell(\theta | \mathbf{w}^t, \tilde{\mathbf{z}}^{t-1})^{1-I(t \in \{q_i\}_{i=1}^Q)} \\ \ell(\theta | \mathbf{w}^t, \tilde{\mathbf{z}}^t)^{I(t \in \{q_i\}_{i=1}^Q)} \end{array} \right] \pi(\theta)$$

So the period likelihood is evaluated based on full observations when $t \in \{q_i\}_{i=1}^Q$ while it is based on only partial observations when $t \notin \{q_i\}_{i=1}^Q$. This method still retains the original state space representation at monthly frequency and thus updates state variables monthly. However, there is still a limitation of gaining efficiency since periods in which only w_t are available suffers lack of information from missing observations on z_t . In contrast to the stacking method, this method in principle is not restricted to monthly and quarterly frequency dataset but can also be applied to randomly missing observations within time series and also possibly the dimension of those time series observed can be time varying.

3.3.5 Efficiency

All of above estimation strategies are equivalent in a sense that the estimates from the markov chains of $\theta^{(m)}$ are consistent. However, in reality any estimation strategy will be influenced by the potential biases due to finite sample and thus an efficiency of estimation methods is significant from the methodological point of view. Data augmentation literature has emphasized the advantage of efficiency gain from both theoretical and empirical perspectives. Gelfand and Smith (1990) and Liu et al. (1994) have theoretically shown the smaller variance of sampled estimates with data augmentations. As such, numerous empirical works have shown the efficiency gain of data augmentation estimates by presenting root mean squared errors¹⁹. Hence, in the following section a comparison of estimation methods is based on root mean squared errors of parameters of interest in a New Keynesian model.

¹⁹ $RMSE = \sqrt{Bias(\hat{\theta}, \theta)^2 + var(\hat{\theta})}$

3.4 Medium Scale New Keynesian Model

This model closely follows Fernández-Villaverde et al. (2010) which is similar to Christiano and Eichenbaum (2005) and Smets and Wouters (2005). I adopt this model for both Monte Carlo experiment and estimation with data since it is well known and widely studied. Following paragraphs summarize this medium scale New Keynesian model and the details can be found in technical appendix of this paper.

There is a continuum of households who consumes final good, supplies differentiated labor to labor packer in monopolistic competitive labor market, invests on capital good, saves by purchasing risk free bonds, and also has access to a complete set of Arrow securities. Calvo wage setting with partial indexation is applied in intermediate labor market. Labor packer integrates the intermediate labor supply into homogenous final labor and supply it to the intermediate good producers. While differentiated labor supply induces heterogeneity of households, the complete asset market equalizes the lagrangian multipliers of households and thus yields symmetric equilibrium conditions with respect to all household's decision variables except labor supply. The utility of household is the standard separable utility between consumption and labor hours and exogenously influenced by two preference shocks that influences the wedges in the intertemporal condition and intratemporal condition. Households also earn rental income from capital management with capacity utilization cost incurred. Another source of uncertainty is coming from the marginal efficiency of investment which creates the inverse of relative price of investment good to fluctuate over time.

Intermediate good producers use rental capital and homogenous labor to produce differentiated goods with Cobb-Douglas technology and earn profits facing monopolistic competitive market with Calvo pricing. Production technology faces total factor productivity shocks. Final good producer transforms intermediate goods into a homogenous final good to be demanded by households. Government follows Taylor rule in which risk free interest rate is set to respond to inflation gap and to deviation of growth rate of output from trend with its own persistence. Monetary policy

shock is incorporated in Taylor rule. Aggregate demand is consumption, investment and capacity utilization cost while the aggregate supply is dictated by industry wide Cobb-Douglas production which is implicitly derived from aggregating Cobb-Douglas production of intermediate good producers. And due to Calvo pricing, price dispersion across intermediate good sector creates wedge between these aggregate demand and aggregate supply. Same applies to the labor market due to wage dispersion. In summary, there are five exogenous processes, namely two preference shocks, investment technology shock, total factor productivity shock, and monetary shocks. And risk free interest rates, wage, inflation and consumption are used as monthly observations while output and investment series as quarterly observations. Also growth rates of wage, consumption, output and investment are used as observables which is the standard practice in this literature. Thus the observables vector²⁰ is

$$obs_t = \begin{bmatrix} \log R_t - \log R \\ \log \pi_t - \log \Pi \\ \Delta \log w_t \\ \Delta \log c_t \\ (1 - L^3) \log \tilde{y}_t^Q \\ (1 - L^3) \log \tilde{i}_t^Q \end{bmatrix}$$

Note that the quarterly aggregates for output and investment are taken into account that corresponds to the available data source and this aggregation scheme follows NIPA convention as shown earlier with Simple New Keynesian model.

²⁰ The variables are in terms of real valued levels.

3.5 Estimation Exercise

First, the economy under this model is simulated with a set of calibrated parameters, and then estimate parameters of interest in the model across alternative estimation methods under mixed frequency data. Following subsection "Monte-Carlo Experiment" shows the results of this exercise. Second, raw data are imported from NIPA and BLS, and time series in real terms are constructed following Whelan (2002), and the model is brought to this actual dataset to be estimated which is shown in "Estimation Results". Throughout the estimations in the following, I fix a small set of parameters and set priors for parameters of interest in estimation to get reasonable identification.

Table 3.5: Fixed Parameters

δ	ε	η	ϕ	Φ_2
0.025/3	10	10	0	0.001

δ is depreciation rate of capital and fixed the one third of 0.025 which is standard in quarterly model. Elasticity of substitutions for differentiated labor supply and intermediate goods are fixed to be 10. ϕ is the fixed cost parameter of production technology and Φ_2 is the parameter for capacity utilization cost function which pins down the rental rate of capital in equilibrium condition.

3.5.1 Monte-Carlo Experiment

The medium scale New Keynesian model is simulated over 40 times and with sample size of 100 each. Given from these original datasets, some of observations such as GDP and investment are deleted to construct the mixed frequency dataset. Only subset of parameters of this model is brought to estimation because the convergence properties for some parameters, mostly preference parameters, generally were not desirable for this exercise. Those parameters of choice for estimates are calibrated for the true model as follows. σ_s denotes standard deviations of all the exogenous

shocks.

Table 3.6: True Parameters

θ_p	χ	θ_w	χ_w	γ_R	γ_y	γ_{Π}	ρ_{φ}	ρ_d	$exp(\sigma_s)$
0.85	0.5	0.8	0.5	0.85	0.25	1.5	0.75	0.75	0.01

Priors

Priors are set for the estimation to have a reasonable acceptance rate but as loose as possible.

Table 3.7: MC Priors

θ_p	χ	θ_w	χ_w
<i>Unif</i> (0, 1)	<i>Be</i> (0.5, 0.4)	<i>Unif</i> (0, 1)	<i>Be</i> (0.5, 0.4)

γ_R	γ_y	γ_{Π}
<i>Unif</i> (0, 1)	<i>N</i> (0.25, 0.1)	<i>N</i> (1.5, 0.25)

ρ_{φ}	ρ_d	$exp(\sigma_s)$
<i>N</i> (0.75, 0.15)	<i>N</i> (0.75, 0.15)	<i>InvGamma</i> (0.01, 1)

Simulation Results

For each estimation of one dataset, parameters are drawn 500,000 times and the posterior estimates are posterior modes based on the second half of these draws, i.e. 250,000 draws. Efficiency comparison results across estimation strategies are reported below.

Table 3.8: Simulation Results I

<i>statistics</i>	$\theta_p = 0.85$		$\chi = 0.5$		$\theta_w = 0.8$		$\chi_w = 0.5$	
	$\bar{\theta}_p$	<i>RMSE</i>	$\bar{\chi}$	<i>RMSE</i>	$\bar{\theta}_w$	<i>RMSE</i>	$\bar{\chi}_w$	<i>RMSE</i>
<i>M0</i>	0.8500 0.0078	0.0078	0.4843 0.0636	0.0656	0.7998 0.0094	0.0094	0.4952 0.0557	0.0559
<i>Augment</i>	0.8518 0.0104	0.0106	0.4763 0.0775	0.0810	0.8012 0.0101	0.0102	0.4860 0.0680	0.0695
<i>D – K</i>	0.8511 0.0126	0.0127	0.4935 0.0831	0.0833	0.7996 0.0105	0.0105	0.5050 0.0730	0.0732
<i>Stack</i>	0.8506 0.0118	0.0118	0.4763 0.0796	0.0831	0.8010 0.0106	0.0107	0.4898 0.0777	0.0777

Top row has the parameters of interest with true values. The third row "M0" is the estimation with original simulated dataset, i.e. no missing observations so that all of observables are monthly and thus estimated with the standard procedure. This will serve as a benchmark estimation for the comparison across three methodologies. "Augment" is the estimation with the data augmentation, "D – K" is Durbin-Koopman method and "Stack" is the stacking method. Each parameter has two statistics that are mean of point estimates and root mean squared errors of these point estimates. Numbers below the mean of point estimates are standard deviations of these point estimates. Hence, lower RMSE represents more efficient estimates compared to alternative methods. In Calvo price parameter, θ_p , and the indexation parameter, χ , show a clear advantage with the data augmentation since it brings down RMSE closer to M0. Calvo wage parameter, θ_w , shows only a small difference while the indexation to wage, χ_w , shows more improvement for data augmentation method. Next table shows the Taylor rule parameters.

Table 3.9: Simulation Results II

<i>statistics</i>	$\gamma_R = 0.85$		$\gamma_y = 0.25$		$\gamma_{\Pi} = 1.5$	
	$\bar{\gamma}_R$	<i>RMSE</i>	$\bar{\gamma}_y$	<i>RMSE</i>	$\bar{\gamma}_{\Pi}$	<i>RMSE</i>
<i>M0</i>	0.8499 0.0067	0.0067	0.2456 0.0575	0.0577	1.5136 0.0906	0.0916
<i>Augment</i>	0.8491 0.0080	0.0081	0.2458 0.0509	0.0511	1.5442 0.1084	0.1170
<i>D – K</i>	0.8500 0.0085	0.0085	0.2496 0.0512	0.0512	1.5239 0.1257	0.1279
<i>Stack</i>	0.8505 0.0083	0.0083	0.2652 0.0501	0.0524	1.5552 0.1225	0.1344

The first two parameters, γ_R and γ_y , show some mixed evidence. RMSE for the smoothing parameter are close. γ_y seems to show no difference either but *M0* results show higher RMSE. This was rather one of rare parameters that show inefficiency with the benchmark estimation²¹. The efficiency ranking for γ_{π} is consistent with most of parameters' results and this implies data augmentation is preferable and closer to the benchmark estimation in terms of RMSE. The rest of parameters that are of less interest are reported in the Appendix and similar conclusions can be drawn.

3.5.2 Data

US data covers from 1984:Q1 to 2010:Q2 for quarterly estimation and from 1984:M1 to 2010:M6 for monthly estimation. Interest rate is the effective Federal Funds rates, and quarterly interest rate is simply compounded over three months interest rates. In case of growth rate of wage, the average wage rate for nonfarm business sector is used for quarterly estimation. But the monthly frequency wage rate was available only for total private sector which is the major subcategory of nonfarm business sector. Since using the average wage rate for total private sector at quarterly frequency instead of nonfarm business didn't show different results and thus it can be safely deduced that

²¹ 10 more parameters related to exogenous processes of the medium scale New Keynesian model are estimated. Except for two parameters of these, the efficiency with the benchmark estimation was overall better.

the wage from different scope of a sector do not play much role in monthly frequency as well. Also, this wage rates are adjusted by the ratio between employment rate²² for the corresponding sectors and population rate so that the wage data is consistent with what is implied by the model in which there is no unemployment.

As for GDP components, the consumption is assumed to be the sum of nondurable consumptions and services and the investment to be the sum of durable consumptions and gross domestic private investments following Fernández-Villaverde et al. (2010) and output is the sum of consumption and investment. Since those series are constructed aggregates from GDP components in NIPA tables and thus do not have corresponding aggregate real variables and price indices, I follow Whelan (2002)²³ to derive real terms and price indices of those series. And the price index for this constructed consumption series is used for price level of the model by assuming consumption good as numeraire. Inflation rate is the growth rate of this CPI deflator. Since consumption and output in the model are in terms of same units, the output is normalized by this CPI deflator. Quarterly output series used for monthly estimation is normalized by CPI in the last month of each quarter. Due to the marginal efficiency of investment, investment good is in terms of its own unit in the model and thus investment series are deflated by its own deflator and this helps to identify this investment specific technological progress.

3.5.3 Estimation Results

Since the primary focus is to compare the temporal aggregation bias on Calvo parameters without attributing the bias to the priors, except the Calvo parameters, the priors of the rest of parameters are set equivalently for both monthly and quarterly model. Below is the standard prior specifications following closely to Fernández-Villaverde et al. (2010).

²² This adjustment has been also made in Smets and Wouters (2005) and Chang et al. (2002). US data shows that there is higher growth rates of employment rates than the population growth and thus the raw data on the growth rate of wage has a lower trend than those of per capita GDP components.

²³ Whelan (2002) discusses how Fisher's chain-aggregated data in NIPA are computed and potential pitfalls with simply adding and subtracting real series from those chain aggregates.

Table 3.10: Priors

B	γ	ψ	κ	α
$Be(0.75, 0.13)$	$N(1, 0.25)$	$N(9, 3)$	$N(4, 1.5)$	$N(0.3, 0.0125)$

χ	χ_w	γ_R	γ_y	γ_π
$Be(0.5, 0.142)$	$Be(0.5, 0.102)$	$Be(0.75, 0.13)$	$N(0.15, 0.05)$	$N(1.5, 0.1)$

ρ_d	ρ_φ	σ_s
$Be(0.5, 0.142)$	$Be(0.5, 0.142)$	$InvGamma(0.1, 2)$

The prior for Calvo parameters in each frequency is set to imply equivalent average price durations. Quarterly model's prior for Calvo parameter is set with mean ,0.5, implying 6 months price duration. Thus monthly model's Calvo parameters are set with a mode of the prior being 0.833.

Table 3.11: Priors on Calvo parameters

θ_p, θ_w	
Q	M
$Be(0.5, 0.28)$	$Be(0.833, 0.25)$

The following table is the estimation results on Calvo parameters with two different frequencies.

Table 3.12: Results on Price Durations

<i>methods</i>	θ_p	<i>PriceDuration</i>	θ_w	<i>WageDuration</i>
<i>Q</i>	0.8549 0.0163	20.67	0.7233 0.0518	10.84
<i>Augment</i>	0.7984 0.0259	4.96	0.8240 0.0308	5.68

As consistent with the simulation exercises with a parsimonious New Keynesian model, the calvo parameters from quarterly model is 0.855 implying approximately 20.7 months of average price duration while the monthly model when estimated with data augmentation implies approximately 5 months price duration. The gap with Calvo wage is relatively smaller than Calvo price. Wage duration for quarterly model implies 11.8 months while 5.7 months with the monthly model.

3.6 Conclusions

This chapter investigates the temporal aggregation issue with a New Keynesian model and finds that Calvo parameter is upward biased in the sense that the quarterly model has a stronger degree of price stickiness. Monte Carlo simulation result suggests that a frequency misspecification of a New Keynesian model generates this upward bias and the estimation with data consistently confirms this finding. This paper also examines three estimation strategies to accommodate mixed frequency dataset in DSGE model's estimations and shows methodological improvements with the data augmentation method borrowed from Bayesian statistics literature.

The results and the method provided in this paper can potentially lead to another research agenda since it can address various interesting questions in macroeconomic studies. For example, this data augmentation method can naturally conduct inferences on unobserved movements of GDP at a monthly frequency and thus potentially can tune the forecasts.

Appendix A

Appendix for Chapter 2

A.1 Standard Kalman Filter

1. Starting with given $\xi_{t|t-1}$ and $\Sigma_{t|t-1}$
2. $\Omega_{t|t-1} = H'\Sigma_{t|t-1}H + R$
3. $\eta_{t|t-1} = H'\xi_{t|t-1}$
4. $K_t = \Sigma_{t|t-1}H (H'\Sigma_{t|t-1}H + R)^{-1}$
5. $\Sigma_{t|t} = \Sigma_{t|t-1} - K_t H'\Sigma_{t|t-1}$
6. $\xi_{t|t} = \xi_{t|t-1} + K_t (\eta_t - \eta_{t|t-1})$
7. $\Sigma_{t+1|t} = F\Sigma_{t|t}F' + Q$
8. $\xi_{t+1|t} = F\xi_{t|t}$

A.2 Derivation of Target Distribution

Derivation of the distribution of $f\left(\begin{bmatrix} z_t \\ \xi_t \end{bmatrix} \middle| \xi_{t|t-1}^{(m)}, P_{t|t-1}^{(m)}, w_t, \theta^{(m)}\right)$ is not trivial. In this section, we present the derivation from a general state space form of loglinearized DSGE model:

$$\xi_{t+1} = F\xi_t + v_{t+1}, v_t \sim N(0, Q)$$

$$\eta_t = H'\xi_t + u_t, u_t \sim N(0, R)$$

Given normality of errors, the joint distribution of states and data is normal with the following mean and variance,

$$\begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix} \middle| \eta^{t-1} \sim N\left(\begin{bmatrix} \xi_{t|t-1} \\ H'\xi_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}H \\ H'P_{t|t-1} & H'P_{t|t-1}H + R \end{bmatrix}\right)$$

Suppose η_t has some missing observations. In this case, it turns useful to partition η_t into two components

$$\eta_t \equiv \begin{bmatrix} z_t \\ w_t \end{bmatrix} = \begin{bmatrix} H'_z \\ H'_w \end{bmatrix} \xi_t + \begin{bmatrix} u_t^z \\ u_t^w \end{bmatrix}$$

$$\begin{bmatrix} u_t^z \\ u_t^w \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R_z & 0 \\ 0 & R_w \end{bmatrix}\right)$$

Where, z_t is assumed to be formed by series with missing observations and w_t is fully observed.

In this case, under normality assumptions, we can rewrite

$$\begin{aligned} \begin{bmatrix} \xi_t \\ z_t \\ w_t \end{bmatrix} \middle| \begin{bmatrix} z^{t-1} \\ w^{t-1} \end{bmatrix} &\sim N\left(\begin{bmatrix} \xi_{t|t-1} \\ H'_z \xi_{t|t-1} \\ H'_w \xi_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}H_z & P_{t|t-1}H_w \\ H'_z P_{t|t-1} & H'_z P_{t|t-1}H_z + R_z & H'_z P_{t|t-1}H_w \\ H'_w P_{t|t-1} & H'_w P_{t|t-1}H_z & H'_w P_{t|t-1}H_w + R_w \end{bmatrix}\right) \\ &\equiv N\left(\begin{bmatrix} \mu_{\xi z} \\ \mu_w \end{bmatrix}, \begin{bmatrix} \Sigma_{\xi z} & \Sigma_{\xi z, w} \\ \Sigma'_{\xi z, w} & \Sigma_w \end{bmatrix}\right) \end{aligned}$$

where

$$\hat{z}^{t-1} = \{z_1, \hat{z}_2, \hat{z}_3, z_4, \dots, \hat{z}^{t-1}\}$$

Note that Kalman-Filter with data augmentation is to exploit the following normality to draw $z_t^{(m)} \sim f(z_t | \hat{z}^{t-1}, w^{t-1}, w_t, \theta^{(m)})$. So the desired normality with using w_t as a new information is the following.

$$\begin{bmatrix} \xi_t \\ z_t \end{bmatrix} \Big| \begin{bmatrix} \hat{z}^{t-1} \\ w^{t-1} \\ w_t \end{bmatrix} \sim N([\mu_{\xi z} + \Sigma_{\xi z, w} \Sigma_w^{-1} (w_t - \mu_w)], [\Sigma_{\xi z} - \Sigma_{\xi z, w} \Sigma_w^{-1} \Sigma'_{\xi z, w}]) \quad (\text{A.1})$$

A.3 Kalman Filter with Data Augmentation

Initialize $\xi_{1|0}, P_{1|0}$

$$\xi_{2|1}^{(m)}, P_{2|1}^{(m)} \sim KF(\xi_2, P_2 | \xi_{1|0}, P_{1|0}, z_1, w_1, \theta^{(m)})$$

$$\begin{bmatrix} \hat{z}_2^{(m)} \\ \hat{\xi}_2^{(m)} \end{bmatrix} \sim f\left(\begin{bmatrix} z_2 \\ \xi_2 \end{bmatrix} \Big| \xi_{2|1}^{(m)}, P_{2|1}^{(m)}, w_2, \theta^{(m)}\right)$$

$$\xi_{3|2}^{(m)}, P_{3|2}^{(m)} \sim KF(\xi_3, P_3 | \xi_{2|1}^{(m)}, P_{2|1}^{(m)}, \hat{z}_2^{(m)}, w_2, \theta^{(m)})$$

$$\begin{bmatrix} \hat{z}_3^{(m)} \\ \hat{\xi}_3^{(m)} \end{bmatrix} \sim f\left(\begin{bmatrix} z_3 \\ \xi_3 \end{bmatrix} \Big| \xi_{3|2}^{(m)}, P_{3|2}^{(m)}, w_3, \theta^{(m)}\right)$$

$$\xi_{4|3}^{(m)}, P_{4|3}^{(m)} \sim KF(\xi_4, P_4 | \xi_{3|2}^{(m)}, P_{3|2}^{(m)}, \hat{z}_3^{(m)}, w_3, \theta^{(m)})$$

$$\xi_{5|4}^{(m)}, P_{5|4}^{(m)} \sim KF(\xi_5, P_5 | \xi_{4|3}^{(m)}, P_{4|3}^{(m)}, z_4, w_4, \theta^{(m)})$$

¹ In general, if X and Y conditional on w are jointly normal

$$[X|w \quad Y|w]' \sim N\left(\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}\right)$$

then $X|y, w$ is also jointly normally distributed with the following distribution

$$X|y, w \sim N(\bar{x} + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \bar{y}), \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})$$

$$\begin{aligned} \begin{bmatrix} \hat{z}_5^{(m)} \\ \hat{\xi}_5^{(m)} \end{bmatrix} &\sim f \left(\begin{bmatrix} z_5 \\ \xi_5 \end{bmatrix} \mid \widehat{\xi}_{5|4}^{(m)}, P_{5|4}^{(m)}, w_5, \theta^{(m)} \right) \\ &\vdots \\ \begin{bmatrix} \hat{z}_T^{(m)} \\ \hat{\xi}_T^{(m)} \end{bmatrix} &\sim f \left(\begin{bmatrix} z_T \\ \xi_T \end{bmatrix} \mid \widehat{\xi}_{T|T-1}^{(m)}, P_{T|T-1}^{(m)}, w_T, \theta^{(m)} \right) \end{aligned}$$

A.4 Alternative Algorithm

Note that $\hat{\xi}^{(m)}$'s that were jointly generated with missing data are discarded but replaced by Kalman Filter. In principle, it is still valid to use the generated $\hat{\xi}^{(m)}$'s to augment next periods' missing data. So alternatively, between step 1 and step 2 of the pseudo-algorithm we can introduce

$$\begin{aligned} &\text{Initiate } \xi_{1|0}, P_{1|0} \\ \hat{\xi}_1^{(m)} &\sim g(\xi_1 \mid \xi_{1|0}, P_{1|0}, z_1, w_1, \theta^{(m)}) \\ \begin{bmatrix} z_2^{(m)} \\ \hat{\xi}_2^{(m)} \end{bmatrix} &\sim f \left(\begin{bmatrix} z_2 \\ \xi_2 \end{bmatrix} \mid \hat{\xi}_1^{(m)}, P_{2|1}^{(m)}, w_2, \theta^{(m)} \right) \\ \begin{bmatrix} z_3^{(m)} \\ \hat{\xi}_3^{(m)} \end{bmatrix} &\sim f \left(\begin{bmatrix} z_3 \\ \xi_3 \end{bmatrix} \mid \hat{\xi}_2^{(m)}, P_{3|2}^{(m)}, w_3, \theta^{(m)} \right) \\ \hat{\xi}_4^{(m)} &\sim g(\xi_4 \mid \hat{\xi}_3^{(m)}, P_{4|3}^{(m)}, z_4, y_4, \theta^{(m)}) \\ \begin{bmatrix} z_5^{(m)} \\ \hat{\xi}_5^{(m)} \end{bmatrix} &\sim f \left(\begin{bmatrix} z_5 \\ \xi_5 \end{bmatrix} \mid \hat{\xi}_4^{(m)}, P_{5|4}^{(m)}, w_5, \theta^{(m)} \right) \\ &\vdots \\ \begin{bmatrix} z_T^{(m)} \\ \hat{\xi}_T^{(m)} \end{bmatrix} &\sim f \left(\begin{bmatrix} z_T \\ \xi_T \end{bmatrix} \mid \hat{\xi}_T^{(m)}, P_{T|T-1}^{(m)}, w_T, \theta^{(m)} \right) \end{aligned}$$

where $g(\cdot)$ denotes the distribution from which we Gibbs sample as in the third step of 3.2. And the step 2 will just compute $p_1\left(\theta^{(m)}; \hat{z}_{-q}^{(m)}, z_q, w\right)$ by standard Kalman Filter using $\hat{z}_{-q}^{(m)}$ as was generated by above. But there are at least two shortcomings of this algorithm. First, the convergence property of this algorithm was not promising based on observation of MCMC chains. The convergence was slower than the first alternative algorithm. Second, when we enter second stage of drawing parameter with Metropolis Hastings algorithm, $\hat{\xi}^{(m)}$'s should be discarded and $\xi_{t|t-1}^{(m)}$ have to be computed in order to evaluate $p_1\left(\theta^{(m)}; \hat{z}_{-q}^{(m)}, z_q, w\right)$ with standard Kalman Filter in step 2. So this algorithm will be computationally inefficient since there is redundancy of computing ξ 's. Thus, instead of using $\hat{\xi}_{t-1}^{(m)}$ to simulate missing data z_t , we simply use $\xi_{t|t-1}^{(m)}$ and $P_{t|t-1}^{(m)}$ that are backed out by standard Kalman filter. In this way, the convergence property of MCMC chain is better and the efficiency of computation time is achieved.

A.5 State Space Representation of Coarse Estimation

The true model was

$$\begin{aligned}\xi_{t+1} &= F\xi_t + Gv_{t+1} \\ \eta_t &= H'\xi_t + u_t \\ \forall t &= 1, 2, \dots, T\end{aligned}$$

Then the coarse estimation would be

$$\begin{aligned}\tilde{\xi}_{t+3} &= \tilde{F}\xi_t + \tilde{G}\tilde{v}_{t+3} \\ \eta_t &= \tilde{H}'\tilde{\xi}_t + u_t \\ \forall t &= 1, 4, 7, \dots, T\end{aligned}$$

So

$$\begin{bmatrix} \xi_{t+3} \\ \xi_{t+2} \\ \xi_{t+1} \end{bmatrix} = \begin{bmatrix} F^3 & 0 & 0 \\ F^2 & 0 & 0 \\ F & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_t \\ \xi_{t-1} \\ \xi_{t-2} \end{bmatrix} + \begin{bmatrix} G & F & F^2 \\ 0 & G & F \\ 0 & 0 & G \end{bmatrix} \begin{bmatrix} v_{t+3} \\ v_{t+2} \\ v_{t+1} \end{bmatrix}$$

$$\eta_t = \begin{bmatrix} H' & 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_t \\ \xi_{t-1} \\ \xi_{t-2} \end{bmatrix} + u_t$$

A.6 State Space Representation of BMF Estimation

In practice, we usually find that data shown at quarterly frequency is aggregated.

$$w_t = \{r_t, \text{wag}_t, l_t\} \quad \forall t = 1, 2, 3, 4, \dots, T$$

$$\tilde{z}_t = \{\tilde{y}_t, \tilde{c}_t, \tilde{i}_t\} \quad \forall t = 1, 4, 7, 10, \dots, T$$

where tildes denote quarterly aggregated data. The observation equation in this case can be rewritten as,

$$\begin{aligned} \tilde{\eta}_t &\equiv \begin{bmatrix} w_t \\ \tilde{z}_t \end{bmatrix} = \begin{bmatrix} w_t \\ \frac{1}{3}(z_t + z_{t-1} + z_{t-2}) \end{bmatrix} = B\eta_t + A\eta_{t-1} + A\eta_{t-2} \\ &= BH'\xi_t + AH'\xi_{t-1} + AH'\xi_{t-2} + Bu_t + Au_{t-1} + Au_{t-2} \\ &= \tilde{H}'\tilde{\xi}_t + \tilde{u}_t \end{aligned}$$

where

$$\begin{aligned} A &\equiv \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}, B \equiv \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \\ \tilde{H}' &\equiv [BH' \quad AH' \quad AH'] \\ \tilde{u}_t &\sim N(0, \tilde{R}), \tilde{R} \equiv BRB' + 2ARA' \end{aligned}$$

The state equation is

$$\tilde{\xi}_{t+1} = \tilde{F}\tilde{\xi}_t + \tilde{G}\tilde{v}_{t+1}$$

where

$$\begin{aligned} \tilde{\xi}_t &\equiv \begin{bmatrix} \xi_t \\ \xi_{t-1} \\ \xi_{t-2} \end{bmatrix}, \tilde{v}_{t+1} \equiv \begin{bmatrix} v_{t+1} \\ 0 \\ 0 \end{bmatrix} \\ \tilde{F} &\equiv \begin{bmatrix} F & 0 & 0 \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix}, \tilde{G} \equiv \begin{bmatrix} G & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Appendix B

Appendix for Chapter 3

B.1 Temporal Aggregation of Autoregressive Process

The true monthly model is

$$a_t = \rho_m a_{t-1} + \sigma \varepsilon_t, \forall t = 1, 2, \dots, T$$

$$\varepsilon_t \sim iid N(0, 1)$$

Then

$$a_t = \rho_m a_{t-1} + \sigma \varepsilon_t$$

$$a_{t-1} = \rho_m a_{t-2} + \sigma \varepsilon_{t-1}$$

$$a_{t-2} = \rho_m a_{t-3} + \sigma \varepsilon_{t-2}$$

Define

$$\tilde{a}_t \equiv a_t + a_{t-1} + a_{t-2}$$

$$\tilde{a}_t = \rho_m (a_{t-1} + a_{t-2} + a_{t-3}) + \sigma (\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2})$$

$$= \rho_m^2 (a_{t-2} + a_{t-3} + a_{t-4}) + \rho_m \sigma (\varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3}) + \sigma (\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2})$$

$$= \rho_m^3 (a_{t-3} + a_{t-4} + a_{t-5}) + \rho_m^2 \sigma (\varepsilon_{t-2} + \varepsilon_{t-3} + \varepsilon_{t-4}) + \rho_m \sigma (\varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3}) + \sigma (\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2})$$

Also define

$$\tilde{\varepsilon}_t \equiv \rho_m^2 \sigma (\varepsilon_{t-2} + \varepsilon_{t-3} + \varepsilon_{t-4}) + \rho_m \sigma (\varepsilon_{t-1} + \varepsilon_{t-2} + \varepsilon_{t-3}) + \sigma (\varepsilon_t + \varepsilon_{t-1} + \varepsilon_{t-2})$$

Then

$$\tilde{a}_t = \rho_m^3 \tilde{a}_{t-3} + \tilde{\varepsilon}_t, \forall t = 3, 6, \dots, T$$

However,

$$Cov(\tilde{\varepsilon}_t, \tilde{\varepsilon}_{t-3}) = Cov(\rho_m^2 \sigma (\varepsilon_{t-3} + \varepsilon_{t-4}), \sigma (\varepsilon_{t-3} + \varepsilon_{t-4}) + \rho_m \sigma \varepsilon_{t-4}) \neq 0$$

B.2 Derivation of Target Distribution

Derivation of the distribution of $f_{q_i}(z_{q_i-2}, z_{q_i-1} | \tilde{\eta}^{q_i-1}, z_{q_i}, w_{q_i-2}, w_{q_i-1}, w_{q_i})$ is not trivial. This section presents the derivation from a general state space form of loglinearized DSGE model:

$$\xi_{t+1} = F\xi_t + v_{t+1}, v_t \sim N(0, Q)$$

$$\eta_t = H'\xi_t + u_t, u_t \sim N(0, R)$$

Suppose η_t has some missing observations. In this case, it turns useful to partition η_t into two components

$$\eta_t \equiv \begin{bmatrix} z_t \\ w_t \end{bmatrix} = \begin{bmatrix} H'_z \\ H'_w \end{bmatrix} \xi_t + \begin{bmatrix} u_t^z \\ u_t^w \end{bmatrix}$$

Now, stacking three months of variables into one vector transforms those equations to

$$\tilde{\eta}_{q_i} = \tilde{H}\tilde{\xi}_{q_i} + \tilde{u}_{q_i}$$

$$\tilde{\xi}_{q_i} = \tilde{F}\tilde{\xi}_{q_i-1} + \tilde{v}_{q_i}$$

where

$$\tilde{\eta}_{q_i} \equiv \begin{bmatrix} z_{q_i-2} \\ z_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix}, \tilde{u}_{q_i} \equiv \begin{bmatrix} u_{q_i-2}^z \\ u_{q_i-1}^z \\ u_{q_i}^z \\ u_{q_i-2}^w \\ u_{q_i-1}^w \\ u_{q_i}^w \end{bmatrix}$$

$$\tilde{\xi}_{q_i} \equiv \begin{bmatrix} \xi_{q_i} \\ \xi_{q_i-1} \\ \xi_{q_i-2} \end{bmatrix}, \tilde{v}_{q_i} = \begin{bmatrix} \nu_{q_i} + F\nu_{q_i-1} + F^2\nu_{q_i-2} \\ \nu_{q_i-1} + F\nu_{q_i-2} \\ \nu_{q_i-2} \end{bmatrix}$$

with

$$\tilde{v}_{q_i} \sim N(0, \tilde{Q})$$

$$\tilde{u}_{q_i} \sim N(0, \tilde{R})$$

and

$$\tilde{H} \equiv \begin{bmatrix} 0 & 0 & H'_w \\ 0 & H'_w & 0 \\ H'_w & 0 & 0 \\ 0 & 0 & H'_z \\ 0 & H'_z & 0 \\ H'_z & 0 & 0 \end{bmatrix}$$

$$\tilde{F} \equiv \begin{bmatrix} F^3 & 0 & 0 \\ F^2 & 0 & 0 \\ F & 0 & 0 \end{bmatrix}$$

Given normality of errors, the joint distribution of states and data is normal with the following mean and variance,

$$\begin{bmatrix} \tilde{\xi}_{q_i} \\ \tilde{\eta}_{q_i} \end{bmatrix} | \tilde{\eta}^{q_i-1} \sim N \left(\begin{bmatrix} \tilde{\xi}_{q_i|q_i-1} \\ \tilde{H}'\xi_{q_i|q_i-1} \end{bmatrix}, \begin{bmatrix} \tilde{P}_{q_i|q_i-1} & \tilde{P}_{q_i|q_i-1}\tilde{H} \\ \tilde{H}'\tilde{P}_{q_i|q_i-1} & \tilde{H}'\tilde{P}_{q_i|q_i-1}\tilde{H} + \tilde{R} \end{bmatrix} \right)$$

Define

$$\tilde{H}_1 \equiv \begin{bmatrix} 0 & 0 & H'_w \\ 0 & H'_w & 0 \end{bmatrix}$$

$$\tilde{H}_2 \equiv \begin{bmatrix} H'_w & 0 & 0 \\ 0 & 0 & H'_z \\ 0 & H'_z & 0 \\ H'_z & 0 & 0 \end{bmatrix}$$

$$\tilde{R}_1 \equiv Var \left(\begin{bmatrix} u_{q_i-2}^z \\ u_{q_i-1}^z \end{bmatrix} \right)$$

$$\tilde{R}_2 \equiv Var \left(\begin{bmatrix} u_{q_i}^z \\ u_{q_i-2}^w \\ u_{q_i-1}^w \\ u_{q_i}^w \end{bmatrix} \right)$$

Rewriting with partitioned matrices

$$\begin{aligned} \begin{bmatrix} \tilde{\xi}_{q_i} \\ z_{q_i-2} \\ z_{q_i-1} \\ z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} | \tilde{\eta}^{q_i-1} &\sim N \left(\begin{bmatrix} \tilde{\xi}_{q_i|q_i-1} \\ \tilde{H}'_1 \tilde{\xi}_{q_i|q_i-1} \\ \tilde{H}'_2 \tilde{\xi}_{q_i|q_i-1} \end{bmatrix}, \begin{bmatrix} \tilde{P}_{q_i|q_i-1} & \tilde{P}_{q_i|q_i-1} \tilde{H}_1 & \tilde{P}_{q_i|q_i-1} \tilde{H}_2 \\ \tilde{H}'_1 \tilde{P}_{q_i|q_i-1} & \tilde{H}'_1 \tilde{P}_{q_i|q_i-1} H_1 + \tilde{R}_1 & \tilde{H}'_1 \tilde{P}_{q_i|q_i-1} H_2 \\ \tilde{H}'_2 \tilde{P}_{q_i|q_i-1} & \tilde{H}'_2 \tilde{P}_{q_i|q_i-1} H_1 & \tilde{H}'_2 \tilde{P}_{q_i|q_i-1} H_2 + \tilde{R}_2 \end{bmatrix} \right) \\ &\equiv N \left(\begin{bmatrix} \mu_\xi \\ \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_\xi & \Sigma_{\xi,1} & \Sigma_{\xi,2} \\ \Sigma'_{\xi,1} & \Sigma_1 & \Sigma_{1,2} \\ \Sigma'_{\xi,2} & \Sigma'_{1,2} & \Sigma_2 \end{bmatrix} \right) \end{aligned}$$

So the desired normality with updated information is the following.

$$\begin{aligned} \begin{pmatrix} \tilde{\eta}^{q_i-1} \\ \tilde{\xi}_{q_i} \\ z_{q_i-2} \\ z_{q_i-1} \end{pmatrix} \left| \begin{matrix} z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{matrix} \right. &\sim N(\tilde{\mu}, \tilde{V}) \\ \tilde{\mu} &\equiv \begin{bmatrix} \mu_\xi \\ \mu_1 \end{bmatrix} + \begin{bmatrix} \Sigma_{\xi,2} \\ \Sigma_{1,2} \end{bmatrix} \Sigma_2^{-1} \left(\begin{bmatrix} z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} - \mu_2 \right) \\ \tilde{V} &\equiv \begin{bmatrix} \Sigma_\xi & \Sigma_{\xi,1} \\ \Sigma'_{\xi,1} & \Sigma_1 \end{bmatrix} - \begin{bmatrix} \Sigma_{\xi,2} \\ \Sigma_{1,2} \end{bmatrix} \Sigma_2^{-1} \begin{bmatrix} \Sigma'_{\xi,2} & \Sigma'_{1,2} \end{bmatrix} \end{aligned}$$

Since $f_{q_i}(z_{q_i-2}, z_{q_i-1} | \tilde{\eta}^{q_i-1}, z_{q_i}, w_{q_i-2}, w_{q_i-1}, w_{q_i})$ is the marginal distribution of the

¹ In general, if X and Y conditional on w are jointly normal

$$[X'|w \quad Y'|w]' \sim N \left(\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix} \right)$$

then $X'|y, w$ is also jointly normally distributed with the following distribution

$$X|y, w \sim N(\bar{x} + \Sigma_{xy} \Sigma_{yy}^{-1} (y - \bar{y}), \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{yx})$$

above normal distribution

$$f_{q_i} \sim N \left(\mu_1 + \Sigma_{1,2} \Sigma_2^{-1} \left(\begin{bmatrix} z_{q_i} \\ w_{q_i-2} \\ w_{q_i-1} \\ w_{q_i} \end{bmatrix} - \mu_2 \right), \Sigma_1 - \Sigma_{1,2} \Sigma_2^{-1} \Sigma'_{1,2} \right)$$

B.3 More Monte Carlo Results

To be updated

B.4 Medium Scale New Keynesian Model

Households Problem

There is a continuum of households in the economy index by i which maximizes the lifetime utility function.

$$E_0 \sum_{t=0}^{\infty} \beta^t d_t \left\{ \log(c_{it} - bc_{it-1}) + \nu \log\left(\frac{m_{it}}{p_t}\right) - \varphi_t \psi \frac{l_{it}^{1+\gamma}}{1+\gamma} \right\}$$

where b is the parameter that controls habit persistence, d_t , is an intertemporal preference shock and φ_t is a labor supply (intra-temporal) shock :

$$\log d_t = \rho_d \log d_{t-1} + \sigma_d \varepsilon_{d,t} \text{ where } \varepsilon_{d,t} \sim N(0, 1)$$

$$\log \varphi_t = \rho_\varphi \log \varphi_{t-1} + \sigma_\varphi \varepsilon_{\varphi,t} \text{ where } \varepsilon_{\varphi,t} \sim N(0, 1)$$

The i^{th} household's budget constraint is given by :

$$\begin{aligned} c_{it} + i_{it} + \frac{m_{it}}{p_t} + \frac{B_{it+1}}{p_t} + \int q_{t+1,t} a_{it+1} d\omega_{i,t+1|t} \\ = w_{it} l_{it} + (r_t u_{i,t} - q_t a[u_{it}]) \bar{k}_{i,t} + \frac{m_{it-1}}{p_t} + R_{t-1} \frac{B_{it}}{p_t} + a_{it} + T_t + F_t \end{aligned}$$

where p_t is price level of final good, w_{jt} is the real wage, r_t is the rental price of capital, $u_{jt} > 0$ is the intensity of use of capital, $q_t a[u_{jt}]$ is the physical cost of use of capital in resource terms where

$$a[u] = \gamma_1 (u - 1) + \gamma_2 (u - 1)^2$$

Here, we assume the household has technology that transforms the final good into investment good that faces this exogenous process. Thus the investment good is

$$I_{it} = \xi_t i_{it}$$

ξ_t is an investment-specific technology shock or also its inverse is interpreted as the relative price of investment good in final good unit. Its exogenous process is

$$\xi_t = \xi_{t-1} \exp(\Lambda_\xi + \sigma_\xi \varepsilon_{\xi,t})$$

where

$$\varepsilon_{\xi,t} \sim N(0, 1)$$

Later, I substitute with stationary variable $\mu_{\xi,t} \equiv \frac{\xi_t}{\xi_{t-1}}$ so that

$$\log \mu_{\xi,t} = \Lambda_\xi + \sigma_\xi \varepsilon_{\xi,t}$$

And this investment good is newly installed to capital stock and thus the capital stock² evolves with

$$\bar{k}_{it+1} = (1 - \delta) \bar{k}_{it} + \left(1 - S\left(\frac{I_{it}}{I_{it-1}}\right)\right) I_{it}$$

where

$$S\left(\frac{I_{it}}{I_{it-1}}\right) = \frac{\kappa}{2} \left(\frac{I_{it}}{I_{it-1}} - \Lambda_I\right)^2$$

is the investment adjustment cost. For ease of notation, define

$$F(I_{it}, I_{it-1}) \equiv \left(1 - S\left(\frac{I_{it}}{I_{it-1}}\right)\right) I_{it}$$

and

$$F_{1t} = 1 - S\left(\frac{I_{it}}{I_{it-1}}\right) - S'\left(\frac{I_{it}}{I_{it-1}}\right) \frac{I_{it}}{I_{it-1}}$$

$$F_{2t+1} = S'\left(\frac{I_{it+1}}{I_{it}}\right) \left(\frac{I_{it+1}}{I_{it}}\right)^2$$

² Here, we denote \bar{k}_t as installed capital stock and k_t as capital service.

Our lagrangian problem is summarized by choosing $c_{it}, B_{it}, u_{it}, \bar{k}_{it+1}, i_{it}, I_{it}, a_{it+1,t}, w_{it}, l_{it}$ to maximize

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \left[-\lambda_{it} \left\{ \begin{array}{l} d_t \left\{ \log(c_{it} - bc_{it-1}) + v \log\left(\frac{m_{it}}{p_t}\right) - \varphi_t \psi_{1+\gamma}^{\frac{l_{it}}{1+\gamma}} \right\} \\ c_{it} + i_{it} + \frac{m_{it}}{p_t} + \frac{B_{it+1}}{p_t} + \int q_{t+1,t} a_{it+1,t} d\omega_{i,t+1|t} \\ -w_{it} l_{it} - (r_t u_{i,t} - q_t a[u_{it}]) k_{i,t} - \frac{m_{it-1}}{p_t} - R_{t-1} \frac{B_{it}}{p_t} - a_{it} - T_t - F_t \\ -q_{it} [\bar{k}_{it+1} - (1-\delta)\bar{k}_{it} - F(I_{it}, I_{it-1})] \\ -\zeta_{it} [I_{it} - \xi_t i_{it}] \end{array} \right\} \right]$$

And HH will determine w_{it} and l_{it} by maximizing relevant part of the lagrangian under Calvo wage setting which will be characterized separately.

Household Conditions

FOCs of the above problem with respect to $c_{it}, B_{it}, u_{it}, \bar{k}_{it+1}, i_{it}, I_{it}, a_{it+1|t}$ are

$$d_t (c_{it} - bc_{it-1})^{-1} - b\beta E_t d_{t+1} (c_{it+1} - bc_{it})^{-1} = \lambda_{it}$$

$$\lambda_{it} = \beta E_t \lambda_{it+1} \frac{R_t}{\Pi_{t+1}}$$

$$r_t = q_t a' [u_{it}]$$

$$\lambda_{it} q_{it} = \beta E_t \{ \lambda_{it+1} [(1-\delta) q_{it+1} + r_{t+1} u_{it+1} - q_{t+1} a(u_{it+1})] \}$$

$$1 = \zeta_{it} \xi_t$$

$$\lambda_{it} \zeta_{it} = \lambda_{it} q_{it} F_{1,t} + \beta E_t \lambda_{it+1} q_{it+1} F_{2,t+1}$$

$$\lambda_{it+1} q_{t+1,t} = \lambda_{it}$$

Symmetric Equilibrium

Since we consider a symmetric equilibrium due to complete asset market (the complete set of state contingent Arrow securities and perfect risk sharing) so that $c_{it} = c_t, B_{it} = B_t, \lambda_{it} = \lambda_t, u_{it} = u_t, q_{it} = q_t, \zeta_{it} = \zeta_t, i_{it} = i_t, I_{it} = I_t, \bar{k}_{it} = \bar{k}_t, a_{it+1|t} = a_{t+1|t}$. After substituting $\zeta_t = \frac{1}{\xi_t}$ and rearranging,

$$d_t (c_t - bc_{t-1})^{-1} - b\beta E_t d_{t+1} (c_{t+1} - bc_t)^{-1} = \lambda_t$$

$$\lambda_t = \beta E_t \lambda_{t+1} \frac{R_t}{\Pi_{t+1}}$$

$$r_t = q_t a' [u_t]$$

$$\lambda_t q_t = \beta E_t \{ \lambda_{t+1} [(1 - \delta) q_{t+1} + r_{t+1} u_{t+1} - q_{t+1} a(u_{t+1})] \}$$

$$1 = q_t \xi_t F_{1,t} + \beta E_t \frac{\lambda_{t+1}}{\lambda_t} q_{t+1} \xi_{t+1} \frac{1}{\mu_{\xi,t+1}} F_{2,t+1}$$

Household labor problem Calvo wage problem for household

$$\max_{w_{jt}} E_t \sum_{\tau=0}^{\infty} (\beta \theta_w)^\tau \left\{ -d_t \varphi_t \psi \frac{l_{jt+1}^{1+\gamma}}{1+\gamma} + \lambda_{jt+1} \prod_{s=1}^{\tau} \frac{\Pi_{t+s}^{\chi_w}}{\Pi_{t+s}} w_{jt} l_{jt+\tau} \right\}$$

subject to

$$l_{jt+\tau} = \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s}^{\chi_w}}{\Pi_{t+s}} \frac{w_{jt}}{w_{t+\tau}} \right)^{-\eta} l_{t+\tau}^d \quad \forall j$$

This gives the law of motion

$$f_t = \frac{\eta - 1}{\eta} (w_t^*)^{1-\eta} \lambda_t w_t^\eta l_t^d + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta-1} f_{t+1}$$

$$f_t = \psi d_t \varphi_t \left(\frac{w_t}{w_t^*} \right)^{\eta(1+\gamma)} (l_t^d)^{1+\gamma} + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{-\eta(1+\gamma)} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta(1+\gamma)} f_{t+1}$$

The real wage index evolves :

$$w_t^{1-\eta} = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{1-\eta} w_{t-1}^{1-\eta} + (1 - \theta_w) w_t^{*1-\eta}$$

which can be rewritten

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{1-\eta} \left(\frac{w_{t-1}}{w_t} \right)^{1-\eta} + (1 - \theta_w) (\Pi_t^{*w})^{1-\eta}$$

where

$$\Pi_t^{*w} = \frac{w_t^*}{w_t}$$

Firms

Final Good Producer Final good producer produces one final good in perfectly competitive market using intermediate good with following technology.

$$y_t^d = \left(\int_0^1 y_{jt}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where ε controls the elasticity of substitution between intermediated goods. And thus the intermediate good producers' markup is $\frac{\varepsilon}{\varepsilon-1}$. The problem of final good producer is

$$\max_{y_{jt}} p_t y_t^d - \int_0^1 p_{jt} y_{jt} dj$$

subject to

$$y_t^d = \left(\int_0^1 y_{jt}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

gives input demand function

$$y_{jt} = \left(\frac{p_{jt}}{p_t} \right)^{-\varepsilon} y_t^d \quad \forall j$$

where the aggregate price level is

$$p_t = \left(\int_0^1 p_{jt}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}$$

Intermediate Good Producer Intermediate good producer's technology is

$$y_{jt} = A_t k_{jt}^\alpha l_{jt}^{1-\alpha} - \phi z_t$$

where k_{jt} and l_{jt} are capital services and homogenous labor and A_t follows

$$A_t = A_{t-1} \exp(\Lambda_A + \sigma_A \varepsilon_{A,t}) \quad \varepsilon_{A,t} \sim N(0, 1)$$

or define $\mu_{A,t} \equiv \frac{A_t}{A_{t-1}}$

$$\log \mu_{A,t} = \Lambda_A + \sigma_A \varepsilon_{A,t}$$

also

$$z_t = A_t^{\frac{1}{1-\alpha}} \xi_t^{\frac{\alpha}{1-\alpha}}$$

or define $\mu_{z,t} \equiv \frac{z_t}{z_{t-1}}$

$$\log \frac{\mu_{z,t}}{\Lambda_z} = \frac{1}{1-\alpha} \log \frac{\mu_{A,t}}{\Lambda_A} + \frac{\alpha}{1-\alpha} \log \frac{\mu_{\xi,t}}{\Lambda_{\xi}}$$

and

$$\Lambda_z \equiv \Lambda_A^{\frac{1}{1-\alpha}} \Lambda_{\xi}^{\frac{\alpha}{1-\alpha}}$$

ϕ is fixed cost parameter and usually calibrated either to zero or to guarantee zero profits in the economy at steady state.

Firms are competitive in factor markets where they confront rents, w_t and r_t^k , from l_{jt}^d and k_{jt}^d . Thus, the firm solves the static cost minimization problem,

$$\min_{l_{jt}^d, k_{jt}^d} w_t l_{jt}^d + r_t k_{jt}^d$$

subject to the production

$$y_{jt} = A_t k_{jt}^{\alpha} l_{jt}^{1-\alpha} - \phi z_t$$

Assuming interior solution, FOCs are

$$w_t = \varrho (1-\alpha) A_t k_{jt}^{\alpha} (l_{jt}^d)^{-\alpha}$$

$$r_t = \varrho \alpha A_t k_{jt}^{1-\alpha} (l_{jt}^d)^{1-\alpha}$$

where ϱ is the Lagrangian multiplier. Then we can find real marginal cost mc_t by

setting $A_t k_{jt}^\alpha (l_{jt}^d)^{1-\alpha} = 1$. This implies

$$\begin{aligned}
1 &= A_t K_{jt}^\alpha (l_{jt}^d)^{1-\alpha} = A_t \left(\frac{K_{jt}}{l_{jt}^d} \right)^\alpha l_{jt}^d \\
&= A_t \left(\frac{w_t}{r_t} \frac{\alpha}{1-\alpha} \right)^\alpha l_{jt}^d \\
l_{jt}^d &= \frac{\left(\frac{\alpha}{1-\alpha} \frac{w_t}{r_t} \right)^{-\alpha}}{A_t} \\
mc_t &= \left(\frac{1}{1-\alpha} \right) w_t l_{jt}^d \\
&= \left(\frac{1}{1-\alpha} \right) w_t \frac{\left(\frac{\alpha}{1-\alpha} \frac{w_t}{r_t} \right)^{-\alpha}}{A_t} \\
&= \left(\frac{1}{1-\alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^\alpha \frac{w_t^{1-\alpha} (r_t)^\alpha}{A_t}
\end{aligned}$$

Intermediate good producer price decision Calvo Pricing decision

$$\max_{p_{it}} E_t \sum_{\tau=0}^{\infty} (\beta \theta_p)^\tau \frac{\lambda_{t+\tau}}{\lambda_t} \left\{ \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^X \frac{p_{it}}{p_{t+\tau}} - mc_{t+\tau} \right) y_{it+\tau} \right\}$$

subject to

$$y_{it+\tau} = \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^X \frac{p_{it}}{p_{t+\tau}} \right)^{-\varepsilon} y_{t+\tau}^d \quad \forall i$$

The law of motion

$$\begin{aligned}
g_t^1 &= \lambda_t mc_t y_t^d + \beta \theta_p E_t \left(\frac{\Pi_t^X}{\Pi_{t+1}} \right)^{-\varepsilon} g_{t+1}^1 \\
g_t^2 &= \lambda_t \Pi_t^* y_t^d + \beta \theta_p E_t \left(\frac{\Pi_t^X}{\Pi_{t+1}} \right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*} \right) g_{t+1}^2 \\
\varepsilon g_t^1 &= (\varepsilon - 1) g_t^2
\end{aligned}$$

Price level evolves

$$1 = \theta_p \left(\frac{\Pi_{t-1}^x}{\Pi_t} \right)^{1-\varepsilon} + (1 - \theta_p) \Pi_t^{*1-\varepsilon}$$

Government

Government sets the nominal interest rates according to the Taylor rule :

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi} \right)^{\gamma_\Pi} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp(m_t)$$

through open market operations that are financed through lump-sum transfers T_t such that the deficit are equal to zero :

$$T_t = \frac{\int_0^1 m_{it} di}{p_t} - \frac{\int_0^1 m_{it-1} di}{p_t} + \frac{\int_0^1 B_{it+1} di}{p_t} - R_{t-1} \frac{\int_0^1 B_{it} di}{p_t}$$

Π represents the target levels of inflation (equal to inflation in the steady state), R steady state gross return of capital, and Λ_{y^d} the steady state gross growth rate of y_t^d . The term m_t is a random shock to monetary policy that follows $m_t = \sigma_m \varepsilon_{mt}$ where ε_{mt} is distributed according to $N(0, 1)$. Consequently, the HH aggregate budget constraint is reduced to

$$c_t + \frac{1}{\xi_t} I_t = w_t l_t + (r_t u_t - q_t a [u_t]) \bar{k}_t + \Omega_t$$

Aggregation

The aggregate demand is

$$y_t^d = c_t + \frac{1}{\xi_t} I_t + q_t a [u_t] \bar{k}_t$$

Calvo pricing produces price dispersion in the economy, thus

$$y_t^d \int_0^1 \left(\frac{p_{jt}}{p_t} \right)^{-\varepsilon} dj = A_t K_t^\alpha (l_t^d)^{1-\alpha} - \phi z_t$$

By defining $v_t^p \equiv \int_0^1 \left(\frac{p_{jt}}{p_t}\right)^{-\varepsilon} dj$, with the properties of indexation under Calvo pricing,

$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^x}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + (1 - \theta_p) \Pi_t^{*-\varepsilon}$$

and we have

$$y_t^d = \frac{A_t k_t^\alpha (l_t^d)^{1-\alpha} - \phi z_t}{v_t^p}$$

Similarly define $v_t^w \equiv \int_0^1 \left(\frac{w_{it}}{w_t}\right)^{-\eta} di$, then

$$l_t^d = \frac{1}{v_t^w} l_t^s$$

Also

$$v_t^w = \theta_w \left(\frac{w_{t-1}}{w_t} \frac{\Pi_{t-1}^{xw}}{\Pi_t}\right)^{-\eta} v_{t-1}^w + (1 - \theta_w) (\Pi_t^{w*})^{-\eta}$$

Also in capital market

$$\int_0^1 k_{jt} dj = u_t \bar{k}_t$$

and thus

$$k_{jt} = u_t \bar{k}_t$$

Also capital stock evolves

$$\bar{k}_{t+1} = (1 - \delta) \bar{k}_t + \left(1 - S\left(\frac{I_t}{I_{t-1}}\right)\right) I_t$$

Equilibrium Conditions

- Intermediate good producer

$$\frac{u_t \bar{k}_t}{l_t^d} = \frac{\alpha}{1 - \alpha} \frac{w_t}{r_t}$$

$$mc_t = \left(\frac{1}{1-\alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^\alpha \frac{w_t^{1-\alpha} (r_t)^\alpha}{A_t}$$

$$g_t^1 = \lambda_t mc_t y_t^d + \beta \theta_p E_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}} \right)^{-\varepsilon} g_{t+1}^1$$

$$g_t^2 = \lambda_t \Pi_t^* y_t^d + \beta \theta_p E_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}} \right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*} \right) g_{t+1}^2$$

$$\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$$

$$1 = \theta_p \left(\frac{\Pi_{t-1}^\chi}{\Pi_t} \right)^{1-\varepsilon} + (1 - \theta_p) \Pi_t^{*1-\varepsilon}$$

- Households

$$d_t (c_t - bc_{t-1})^{-1} - b\beta E_t d_{t+1} (c_{t+1} - bc_t)^{-1} = \lambda_t$$

$$\lambda_t = \beta E_t \lambda_{t+1} \frac{R_t}{\Pi_{t+1}}$$

$$r_t = q_t a' [u_t]$$

$$\lambda_t q_t = \beta E_t \{ \lambda_{t+1} [(1 - \delta) q_{t+1} + r_{t+1} u_{t+1} - q_{t+1} a(u_{t+1})] \}$$

$$1 = q_t \xi_t F_{1,t} + \beta E_t \frac{\lambda_{t+1}}{\lambda_t} q_{t+1} \xi_{t+1} \frac{1}{\mu_{\xi,t+1}} F_{2,t+1}$$

$$f_t = \frac{\eta - 1}{\eta} (w_t^*)^{1-\eta} \lambda_t w_t^\eta l_t^d + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi w}}{\Pi_{t+1}} \right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta-1} f_{t+1}$$

$$f_t = \psi d_t \varphi_t \left(\frac{w_t}{w_t^*} \right)^{\eta(1+\gamma)} (l_t^d)^{1+\gamma} + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi w}}{\Pi_{t+1}} \right)^{-\eta(1+\gamma)} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta(1+\gamma)} f_{t+1}$$

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi w}}{\Pi_t} \right)^{1-\eta} \left(\frac{w_{t-1}}{w_t} \right)^{1-\eta} + (1 - \theta_w) (\Pi_t^{*w})^{1-\eta}$$

$$\Pi_t^{*w} = \frac{w_t^*}{w_t}$$

- Government

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi} \right)^{\gamma_\Pi} \left(\frac{y_t^d}{\Lambda_{y^d}} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp(m_t)$$

- Aggregation

$$y_t^d = c_t + \frac{1}{\xi_t} I_t + q_t a [u_t] \bar{k}_t$$

$$y_t^d = \frac{A_t k_t^\alpha (l_t^d)^{1-\alpha} - \phi z_t}{v_t^p}$$

$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^X}{\Pi_t} \right)^{-\varepsilon} v_{t-1}^p + (1 - \theta_p) \Pi_t^{*-\varepsilon}$$

$$v_t^w = \theta_w \left(\frac{w_{t-1} \Pi_{t-1}^{X_w}}{w_t \Pi_t} \right)^{-\eta} v_{t-1}^w + (1 - \theta_w) (\Pi_t^{w*})^{-\eta}$$

$$\bar{k}_{t+1} = (1 - \delta) \bar{k}_t + \left(1 - S \left(\frac{I_t}{I_{t-1}} \right) \right) I_t$$

- Exogenous Process

$$A_t = A_{t-1} \exp(\Lambda_A + \sigma_A \varepsilon_{A,t})$$

$$\xi_t = \xi_{t-1} \exp(\Lambda_\xi + \sigma_\xi \varepsilon_{\xi,t})$$

$$\log d_t = \rho_d \log d_{t-1} + \sigma_d \varepsilon_{d,t}$$

$$\log \varphi_t = \rho_\varphi \log \varphi_{t-1} + \sigma_\varphi \varepsilon_{\varphi,t}$$

$$m_t = \sigma_m \varepsilon_{m,t}$$

- Definition for growth term

$$z_t = A_t^{\frac{1}{1-\alpha}} \xi_t^{\frac{\alpha}{1-\alpha}}$$

Stationary Equilibrium Conditions

Preliminaries

- Variables $\bar{k}_t, l_t^d, w_t, r_t^k, mc_t, g_t^1, g_t^2, \Pi_t, \Pi_t^*, \lambda_t, y_t^d, I_t, q_t, u_t, R_t, c_t, f_t, w_t^*, \Pi_t^{w*}, v_t^p, v_t^w, z_t$
- Stationary variables.

$$\begin{aligned}
 - \tilde{c}_t &= \frac{c_t}{z_t}, \tilde{w}_t = \frac{w_t}{z_t}, \tilde{w}_t^* = \frac{w_t^*}{z_t}, \tilde{l}_t = \frac{l_t}{z_t \xi_t}, \tilde{y}_t^d = \frac{y_t^d}{z_t} \\
 - \tilde{r}_t &= r_t \xi_t, \tilde{k}_{t+1} = \frac{\bar{k}_{t+1}}{z_t \xi_t}, \tilde{q}_t = q_t \xi_t, \tilde{\lambda}_t = \lambda_t z_t,
 \end{aligned}$$

Stationarize I

- Intermediate good producer

$$\begin{aligned}
 \frac{u_t \bar{k}_t}{l_t^d z_{t-1} \xi_{t-1}} \frac{z_{t-1} \xi_{t-1}}{z_t \xi_t} &= \frac{\alpha}{1-\alpha} \frac{w_t}{r_t} \frac{1}{z_t \xi_t} \\
 mc_t &= \left(\frac{1}{1-\alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^\alpha \frac{\left(\frac{w_t}{z_t} \right)^{1-\alpha} (r_t \xi_t)^\alpha}{A_t} \frac{z_t^{1-\alpha}}{\xi_t^\alpha} \\
 g_t^1 &= \lambda_t z_t mc_t \frac{y_t^d}{z_t} + \beta \theta_p E_t \left(\frac{\Pi_t^X}{\Pi_{t+1}} \right)^{-\varepsilon} g_{t+1}^1 \\
 g_t^2 &= \lambda_t z_t \Pi_t^* \frac{y_t^d}{z_t} + \beta \theta_p E_t \left(\frac{\Pi_t^X}{\Pi_{t+1}} \right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*} \right) g_{t+1}^2 \\
 \varepsilon g_t^1 &= (\varepsilon - 1) g_t^2 \\
 1 &= \theta_p \left(\frac{\Pi_{t-1}^X}{\Pi_t} \right)^{1-\varepsilon} + (1 - \theta_p) \Pi_t^{*1-\varepsilon}
 \end{aligned}$$

- Households

$$d_t \left(\frac{c_t}{z_t} - b \frac{c_{t-1}}{z_{t-1}} \frac{z_{t-1}}{z_t} \right)^{-1} - b \beta E_t d_{t+1} \left(\frac{c_{t+1}}{z_{t+1}} \frac{z_{t+1}}{z_t} - b \frac{c_t}{z_t} \right)^{-1} = \lambda_t z_t$$

$$\lambda_t z_t = \beta E_t \lambda_{t+1} z_{t+1} \frac{z_t}{z_{t+1}} \frac{R_t}{\Pi_{t+1}}$$

$$r_t \xi_t = q_t \xi_t a' [u_t]$$

$$\lambda_t z_t q_t \xi_t = \beta E_t \left\{ \lambda_{t+1} z_{t+1} \frac{z_t}{z_{t+1}} \left[(1 - \delta) q_{t+1} \xi_{t+1} \frac{\xi_t}{\xi_{t+1}} + r_{t+1} \xi_{t+1} \frac{\xi_t}{\xi_{t+1}} u_{t+1} \right] - q_{t+1} \xi_{t+1} \frac{\xi_t}{\xi_{t+1}} a(u_{t+1}) \right\}$$

$$1 = q_t \xi_t F_{1,t} + \beta E_t \frac{\lambda_{t+1} z_{t+1}}{\lambda_t z_t} \frac{z_t}{z_{t+1}} q_{t+1} \xi_{t+1} \frac{1}{\mu_{\xi,t+1}} F_{2,t+1}$$

where

$$F_{1t} = 1 - S \left(\frac{\tilde{v}_t}{\tilde{v}_{t-1}} \frac{z_t \xi_t}{z_{t-1} \xi_{t-1}} \right) - S' \left(\frac{\tilde{v}_t}{\tilde{v}_{t-1}} \frac{z_t \xi_t}{z_{t-1} \xi_{t-1}} \right) \frac{\tilde{v}_t}{\tilde{v}_{t-1}} \frac{z_t \xi_t}{z_{t-1} \xi_{t-1}}$$

$$F_{2t+1} = S' \left(\frac{\tilde{v}_{t+1}}{\tilde{v}_t} \frac{z_{t+1} \xi_{t+1}}{z_t \xi_t} \right) \left(\frac{\tilde{v}_{t+1}}{\tilde{v}_t} \frac{z_{t+1} \xi_{t+1}}{z_t \xi_t} \right)^2$$

$$f_t = \frac{\eta - 1}{\eta} \left(\frac{w_t^*}{z_t} \right)^{1-\eta} \lambda_t z_t \left(\frac{w_t}{z_t} \right)^\eta l_t^d + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{1-\eta} \left(\frac{w_{t+1}^* z_t}{w_t^* z_{t+1}} \frac{z_{t+1}}{z_t} \right)^{\eta-1} f_{t+1}$$

$$f_t = \psi d_t \varphi_t \left(\frac{w_t z_t}{w_t^* z_t} \right)^{\eta(1+\gamma)} (l_t^d)^{1+\gamma} + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{-\eta(1+\gamma)} \left(\frac{w_{t+1}^* z_t}{w_t^* z_{t+1}} \frac{z_{t+1}}{z_t} \right)^{\eta(1+\gamma)} f_{t+1}$$

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{1-\eta} \left(\frac{w_{t-1}}{z_{t-1}} \frac{z_{t-1}}{z_t} \right)^{1-\eta} + (1 - \theta_w) (\Pi_t^{*w})^{1-\eta}$$

$$\Pi_t^{*w} = \frac{w_t^*}{w_t}$$

- Government

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi} \right)^{\gamma_\Pi} \left(\frac{y_t^d}{y_{t-1}^d} \frac{z_{t-1}}{z_t} \frac{z_t}{z_{t-1}} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp(m_t)$$

- Aggregation

$$\frac{y_t^d}{z_t} = \frac{c_t}{z_t} + \frac{1}{z_t \xi_t} I_t + \frac{q_t \xi_t a[u_t] \bar{k}_t}{z_{t-1} \xi_{t-1}} \frac{z_{t-1} \xi_{t-1}}{z_t \xi_t}$$

$$\frac{y_t^d}{z_t} = \frac{\frac{A_t}{z_t} \left(\frac{k_t}{\xi_{t-1} z_{t-1}} \right)^\alpha \left(\frac{\xi_{t-1} z_{t-1}}{\xi_t z_t} \right)^\alpha (\xi_t z_t)^\alpha (l_t^d)^{1-\alpha} - \phi}{v_t^p}$$

$$v_t^p = \theta_p \left(\frac{\Pi_t^X}{\Pi_t} \right)^{-\varepsilon} v_{t-1}^p + (1 - \theta_p) \Pi_t^{*- \varepsilon}$$

$$v_t^w = \theta_w \left(\frac{w_{t-1}}{w_t} \frac{z_t}{z_{t-1}} \frac{z_{t-1}}{z_t} \frac{\Pi_{t-1}^{Xw}}{\Pi_t} \right)^{-\eta} v_{t-1}^w + (1 - \theta_w) (\Pi_t^{w*})^{-\eta}$$

$$\frac{\bar{k}_{t+1}}{z_t \xi_t} = (1 - \delta) \frac{\bar{k}_t}{z_{t-1} \xi_{t-1}} \frac{z_{t-1} \xi_{t-1}}{z_t \xi_t} + \left(1 - S \left(\frac{\tilde{i}_t}{\tilde{i}_{t-1}} \frac{z_t \xi_t}{z_{t-1} \xi_{t-1}} \right) \right) \frac{i_t}{z_t \xi_t}$$

- Exogenous Process

$$A_t = A_{t-1} \exp(\Lambda_A + \sigma_A \varepsilon_{A,t})$$

$$\xi_t = \xi_{t-1} \exp(\Lambda_\xi + \sigma_\xi \varepsilon_{\xi,t})$$

$$\log d_t = \rho_d \log d_{t-1} + \sigma_d \varepsilon_{d,t}$$

$$\log \varphi_t = \rho_\varphi \log \varphi_{t-1} + \sigma_\varphi \varepsilon_{\varphi,t}$$

$$m_t = \sigma_m \varepsilon_{m,t}$$

- Definition for growth term

$$z_t = A_t^{\frac{1}{1-\alpha}} \xi_t^{\frac{\alpha}{1-\alpha}}$$

Stationarize II

- Intermediate good producer

$$\frac{u_t \tilde{k}_t}{l_t^d} \frac{1}{\mu_{z,t} \mu_{\xi,t}} = \frac{\alpha}{1 - \alpha} \frac{\tilde{w}_t}{\tilde{r}_t}$$

$$mc_t = \left(\frac{1}{1 - \alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^\alpha (\tilde{w}_t)^{1-\alpha} (\tilde{r}_t)^\alpha$$

$$\begin{aligned}
g_t^1 &= \tilde{\lambda}_t m c_t \tilde{y}_t^d + \beta \theta_p E_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}} \right)^{-\varepsilon} g_{t+1}^1 \\
g_t^2 &= \tilde{\lambda}_t \Pi_t^* \tilde{y}_t^d + \beta \theta_p E_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}} \right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*} \right) g_{t+1}^2 \\
\varepsilon g_t^1 &= (\varepsilon - 1) g_t^2 \\
1 &= \theta_p \left(\frac{\Pi_{t-1}^\chi}{\Pi_t} \right)^{1-\varepsilon} + (1 - \theta_p) \Pi_t^{*1-\varepsilon}
\end{aligned}$$

- Households

$$\begin{aligned}
d_t \left(\tilde{c}_t - b \tilde{c}_{t-1} \frac{1}{\mu_{z,t}} \right)^{-1} - b \beta E_t d_{t+1} (\tilde{c}_{t+1} \mu_{z,t+1} - b \tilde{c}_t)^{-1} &= \tilde{\lambda}_t \\
\tilde{\lambda}_t &= \beta E_t \tilde{\lambda}_{t+1} \frac{1}{\mu_{z,t+1}} \frac{R_t}{\Pi_{t+1}} \\
\tilde{r}_t &= \tilde{q}_t a' [u_t] \\
\tilde{\lambda}_t \tilde{q}_t &= \beta E_t \left\{ \frac{\tilde{\lambda}_{t+1}}{\mu_{\xi,t+1} \mu_{z,t+1}} \left[\begin{array}{c} (1 - \delta) \tilde{q}_{t+1} + \tilde{r}_{t+1} u_{t+1} \\ - \tilde{q}_{t+1} a(u_{t+1}) \end{array} \right] \right\} \\
1 &= \tilde{q}_t F_{1,t} + \beta E_t \frac{\tilde{\lambda}_{t+1}}{\tilde{\lambda}_t \mu_{z,t+1}} \frac{\tilde{q}_{t+1}}{\mu_{\xi,t+1}} F_{2,t+1}
\end{aligned}$$

where

$$\begin{aligned}
F_{1t} &= 1 - S \left(\frac{\tilde{i}_t}{\tilde{i}_{t-1}} \mu_{z,t} \mu_{\xi,t} \right) - S' \left(\frac{\tilde{i}_t}{\tilde{i}_{t-1}} \mu_{z,t} \mu_{\xi,t} \right) \frac{\tilde{i}_t}{\tilde{i}_{t-1}} \mu_{z,t} \mu_{\xi,t} \\
F_{2t+1} &= S' \left(\frac{\tilde{i}_{t+1}}{\tilde{i}_t} \mu_{z,t+1} \mu_{\xi,t+1} \right) \left(\frac{\tilde{i}_{t+1}}{\tilde{i}_t} \mu_{z,t+1} \mu_{\xi,t+1} \right)^2 \\
f_t &= \frac{\eta - 1}{\eta} (\tilde{w}_t^*)^{1-\eta} \tilde{\lambda}_t (\tilde{w}_t)^\eta l_t^d + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi w}}{\Pi_{t+1}} \frac{\tilde{w}_t^*}{\tilde{w}_{t+1}^*} \frac{1}{\mu_{z,t+1}} \right)^{1-\eta} f_{t+1} \\
f_t &= \psi d_t \varphi_t \left(\frac{\tilde{w}_t}{\tilde{w}_t^*} \right)^{\eta(1+\gamma)} (l_t^d)^{1+\gamma} + \beta \theta_w E_t \left(\frac{\Pi_t^{\chi w}}{\Pi_{t+1}} \frac{\tilde{w}_t^*}{\tilde{w}_{t+1}^*} \frac{1}{\mu_{z,t+1}} \right)^{-\eta(1+\gamma)} f_{t+1}
\end{aligned}$$

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{1-\eta} \left(\frac{\tilde{w}_{t-1}}{\tilde{w}_t} \frac{1}{\mu_{z,t}} \right)^{1-\eta} + (1 - \theta_w) (\Pi_t^{*w})^{1-\eta}$$

$$\Pi_t^{*w} = \frac{\tilde{w}_t^*}{\tilde{w}_t}$$

- Government

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi} \right)^{\gamma_\Pi} \left(\frac{\tilde{y}_t^d}{\tilde{y}_{t-1}^d} \mu_{z,t} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp(m_t)$$

- Aggregation

$$\tilde{y}_t^d = \tilde{c}_t + \tilde{i}_t + \frac{\tilde{q}_t a [u_t] \tilde{k}_t}{\mu_{z,t} \mu_{\xi,t}}$$

$$\tilde{y}_t^d = \frac{\frac{\mu_{A,t}}{\mu_{z,t}} (u_t \tilde{k}_t)^\alpha (l_t^d)^{1-\alpha} - \phi}{v_t^p}$$

$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^\chi}{\Pi_t} \right)^{-\varepsilon} v_{t-1}^p + (1 - \theta_p) \Pi_t^{*- \varepsilon}$$

$$v_t^w = \theta_w \left(\frac{\tilde{w}_{t-1}}{\tilde{w}_t} \frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{-\eta} v_{t-1}^w + (1 - \theta_w) (\Pi_t^{*w})^{-\eta}$$

$$\tilde{k}_{t+1} = (1 - \delta) \frac{\tilde{k}_t}{\mu_{z,t} \mu_{\xi,t}} + \left(1 - S \left(\frac{\tilde{i}_t}{\tilde{i}_{t-1}} \mu_{z,t} \mu_{\xi,t} \right) \right) \tilde{i}_t$$

- Exogenous Process

$$\log \mu_{A,t} = \Lambda_A + \sigma_A \varepsilon_{A,t}$$

$$\log \mu_{\xi,t} = \Lambda_\xi + \sigma_A \varepsilon_{\xi,t}$$

$$\log d_t = \rho_d \log d_{t-1} + \sigma_d \varepsilon_{d,t}$$

$$\log \varphi_t = \rho_\varphi \log \varphi_{t-1} + \sigma_\varphi \varepsilon_{\varphi,t}$$

$$m_t = \sigma_m \varepsilon_{m,t}$$

- Definition for growth term

$$\log \frac{\mu_{z,t}}{\exp(\Lambda_z)} = \frac{1}{1-\alpha} \log \frac{\mu_{A,t}}{\exp(\Lambda_A)} + \frac{\alpha}{1-\alpha} \log \frac{\mu_{\xi,t}}{\exp(\Lambda_{\xi})}$$

Steady State

Equilibrium Conditions Let $\mu^z \equiv \Lambda_z = \Lambda_A^{\frac{1}{1-\alpha}} \Lambda_k^{\frac{\alpha}{1-\alpha}}$ where $\mu^A \equiv \Lambda_A$ and $\mu^k \equiv \Lambda_k$. Given the definitions, the mean growth rate of the economy is $\Lambda_c = \Lambda_x = \Lambda_w = \Lambda_{w^*} = \Lambda_{y^d} = \Lambda_z$. $\tilde{u} = 1$ at steady state.

- Households satisfy

$$\left(\tilde{c} - \frac{b\tilde{c}}{\Lambda_z} \right)^{-1} - b\beta (\tilde{c}\Lambda_z - b\tilde{c})^{-1} = \tilde{\lambda}$$

$$1 = \frac{\beta}{\Lambda_z \Lambda_k} (\tilde{r} + 1 - \delta)$$

$$\beta \frac{1}{\Lambda_z} \frac{R}{\Pi} = 1$$

$$\tilde{r} = \gamma_1$$

$$1 = \tilde{q}$$

$$f = \frac{\eta - 1}{\eta} (\tilde{w}^*)^{1-\eta} (\Pi^{*w})^{-\eta} \tilde{\lambda} (\tilde{w}^*)^{\eta} l^d + \beta \theta_w \left(\frac{\Pi^{xw}}{\Pi \Lambda_z} \right)^{1-\eta} f$$

$$f = \psi (\Pi^{*w})^{-\eta(1+\gamma)} (l^d)^{1+\gamma} + \beta \theta_w \left(\frac{\Pi^{xw}}{\Pi \Lambda_z} \right)^{-\eta(1+\gamma)} f$$

- Firms that can change prices set them to satisfy (4 eqs)

$$g^1 = \tilde{\lambda} m c \tilde{y}^d + \beta \theta_p \left(\frac{\Pi^x}{\Pi} \right)^{-\varepsilon} g^1$$

$$g^2 = \tilde{\lambda} \Pi^* \tilde{y}^d + \beta \theta_p \left(\frac{\Pi^x}{\Pi} \right)^{1-\varepsilon} g^2$$

$$\varepsilon g^1 = (\varepsilon - 1) g^2$$

where

$$\Pi^{*w} = \frac{\tilde{w}^*}{\tilde{w}}$$

- They rent inputs to satisfy their static minimization problem(2 eqs)

$$\frac{\tilde{k}}{l^d} = \frac{\alpha}{\alpha - 1} \frac{\tilde{w}}{\tilde{r}} \Lambda_z \Lambda_k$$

$$mc = \left(\frac{1}{1 - \alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^\alpha (\tilde{w})^{1-\alpha} \tilde{r}^\alpha$$

- The wages evolve

$$1 = \theta_w (\Pi^{\chi_w - 1})^{1-\eta} (\Lambda_z)^{-1+\eta} + (1 - \theta_w) (\Pi^{*w})^{1-\eta}$$

- The price level evolve

$$1 = \theta_p (\Pi^{\chi - 1})^{1-\varepsilon} + (1 - \theta_p) \Pi^{*1-\varepsilon}$$

- Markets clear

$$\tilde{y}^d = \tilde{c} + \tilde{x}$$

$$v^p \tilde{y}^d = \frac{\Lambda_A}{\Lambda_z} \tilde{k}^\alpha (l^d)^{1-\alpha} - \phi$$

where

$$v^p = \theta_p (\Pi^{\chi - 1})^{-\varepsilon} v^p + (1 - \theta_p) \Pi^{*- \varepsilon}$$

$$v^w = \theta_w \left(\frac{\Pi^{\chi_w - 1}}{\Lambda_z} \right)^{-\eta} v^w + (1 - \theta_w) (\Pi^{*w})^{-\eta}$$

and

$$\tilde{k} = (1 - \delta) \frac{\tilde{k}}{\Lambda_z \Lambda_\xi} + \tilde{i}$$

- Exogenous processes evolve (6 eqs)

$$d = 1$$

$$\varphi = 1$$

$$\mu^k = \Lambda_k$$

$$\mu^A = \Lambda_A$$

$$m = 0$$

where

$$\Lambda_z = \Lambda_A^{\frac{1}{1-\alpha}} \Lambda_k^{\frac{\alpha}{1-\alpha}}$$

Steady State computation

- Fixed Parameters

Table 1 : Fixed Parameters

β	δ	ε	η	ϕ	γ_2
0.9992	0.025/3	10	10	0	0.001

- Estimated parameters :

$$\{b, \gamma, \psi, \kappa, \alpha, \theta_p, \chi_p, \theta_w, \chi_w, \Lambda_k, \Lambda_A, \gamma_R, \gamma_y, \gamma_{\Pi}, \Pi, \rho_d, \rho_{\varphi}, \sigma_K, \sigma_a, \sigma_d, \sigma_{\varphi}, \sigma_M\}$$

- Free parameters : set $u = 1$ so that $\gamma_1 = \tilde{r}$.
- Parameters related to exogenous processes : $d = 1, \varphi = 1, m = 0$.
- Growth terms

$$\Lambda_z = \Lambda_A^{\frac{1}{1-\alpha}} \Lambda_k^{\frac{\alpha}{1-\alpha}}$$

$$\Lambda_y = \Lambda_z$$

$$\Lambda_x = \Lambda_z \Lambda_k$$

- Interest rate

$$\tilde{r} = \frac{\Lambda_z \Lambda_k}{\beta} - 1 + \delta$$

$$\gamma_1 = \tilde{r}$$

$$R = \frac{\Pi \Lambda_z}{\beta}$$

- Prices

$$\Pi^* = \left(\frac{1 - \theta_p \Pi^{-(1-\varepsilon)(1-\chi)}}{1 - \theta_p} \right)^{\frac{1}{1-\varepsilon}}$$

$$mc = \frac{\varepsilon - 1}{\varepsilon} \frac{1 - \beta \theta_p \Pi^{(1-\chi)\varepsilon}}{1 - \beta \theta_p \Pi^{-(1-\chi)(1-\varepsilon)}} \Pi^*$$

$$v^p = \frac{1 - \theta_p}{1 - \theta_p \Pi^{(1-\chi)\varepsilon}} \Pi^{*\varepsilon}$$

- Wages

$$\Pi^{*w} = \left(\frac{1 - \theta_w (\Pi)^{(1-\eta)(\chi_w-1)} (\Lambda_z)^{-1+\eta}}{1 - \theta_w} \right)^{\frac{1}{1-\eta}}$$

$$\tilde{w} = (1 - \alpha) \left(mc \left(\frac{\alpha}{\tilde{r}} \right)^\alpha \right)^{\frac{1}{1-\alpha}}$$

$$\tilde{w}^* = \tilde{w} \Pi^{*w}$$

$$v^w = \frac{1 - \theta_w}{1 - \theta_w \Pi^{(1-\chi_w)\eta} (\Lambda_z)^\eta} (\Pi^{*w})^{-\eta}$$

- Capital/labor ratio

$$\frac{\tilde{k}}{\tilde{l}^d} = \frac{\alpha}{\alpha - 1} \frac{\tilde{w}}{\tilde{r}} \Lambda_z \Lambda_k$$

- Assuming $\phi = 0$ or ϕ satisfying the zero profits at steady state

$$\frac{\tilde{y}^d}{\tilde{l}^d} = \frac{\frac{\Lambda_A}{\Lambda_z} \left(\frac{\tilde{k}}{\tilde{l}^d} \right)^\alpha - \phi}{v^p}$$

$$\frac{\tilde{k}}{\tilde{l}^d} \left(1 - \frac{1 - \delta}{\Lambda_z \Lambda_\xi} \right) = \frac{\tilde{i}}{\tilde{l}^d}$$

$$\frac{\tilde{i}}{\tilde{l}^d} = \frac{\tilde{k}}{\tilde{l}^d} \left(\frac{\Lambda_z \Lambda_\xi - 1 + \delta}{\Lambda_z \Lambda_\xi} \right)$$

$$\frac{\tilde{c}}{\tilde{l}^d} = \frac{\tilde{y}^d}{\tilde{l}^d} - \frac{\tilde{i}}{\tilde{l}^d}$$

- Labor demand

$$\begin{aligned} \tilde{\lambda} l^d &= \left(\frac{\tilde{c}}{\tilde{l}^d} - \frac{b\tilde{c}}{\Lambda_z \tilde{l}^d} \right)^{-1} - b\beta \left(\frac{\tilde{c}}{\tilde{l}^d} \Lambda_z - b \frac{\tilde{c}}{\tilde{l}^d} \right)^{-1} \\ &= \left(\frac{\tilde{c}}{\tilde{l}^d} \right)^{-1} \left(1 - \frac{b}{\Lambda_z} \beta \right) \left(1 - \frac{b}{\Lambda_z} \right)^{-1} \\ &= \left(\frac{\tilde{c}}{\tilde{l}^d} \right)^{-1} \frac{(\Lambda_z - b\beta)}{(\Lambda_z - b)} \end{aligned}$$

$$f = \frac{\frac{\eta-1}{\eta} \tilde{w}^* (\Pi^{*w})^{-\eta} \tilde{\lambda} l^d}{\left(1 - \beta \theta_w \left(\frac{\Pi^{xw}}{\Pi \Lambda_z} \right)^{1-\eta} \right)}$$

$$l^d = \left[\frac{f \left(1 - \beta \theta_w \left(\frac{\Pi^{xw}}{\Pi \Lambda_z} \right)^{-\eta(1+\gamma)} \right)}{\psi (\Pi^{*w})^{-\eta(1+\gamma)}} \right]^{\frac{1}{1+\gamma}}$$

Bibliography

- An, S. and Schorfheide, F. (2007), “Bayesian analysis of DSGE models,” *Econometric Reviews*, 26, 113–172.
- Bils, M. and Klenow, P. (2004), “Some evidence on the importance of sticky prices,” *Journal of political economy*, 112, 947–985.
- Chang, Y., Gomez, J., and Schorfheide, F. (2002), “VLearning’by’Doing as a Propagation Mech’anism,” *V American Economic Review*, 92, 1498.
- Christiano, L. and Eichenbaum, M. (2005), “C. Evans (2005), Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*.
- Del Negro, M. and Schorfheide, F. (2008), “Forming priors for DSGE models (and how it affects the assessment of nominal rigidities),” *Journal of Monetary Economics*, 55, 1191–1208.
- Diebolt, J. and Robert, C. (1994), “Estimation of finite mixture distributions through Bayesian sampling,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 363–375.
- Durbin, J. and Koopman, S. (2001), *Time series analysis by state space methods*, Oxford Univ Pr.
- Elerian, O., Chib, S., and Shephard, N. (2001), “Likelihood inference for discretely observed nonlinear diffusions,” *Econometrica*, 69, 959–993.
- Eraker, B. (2001), “MCMC analysis of diffusion models with application to finance,” *Journal of Business and Economic Statistics*, 19, 177–191.
- Eraker, B., Chiu, C., Foerster, A., Kim, T., and Seoane, H. (2008), “Bayesian Mixed Frequency VARs,” .
- Fernández-Villaverde, J., Guerrón-Quintana, P., and Rubio-Ramirez, J. (2010), “Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in US Data,” *NBER Working Paper*.

- Gelfand, A. and Smith, A. (1990), “Sampling-based approaches to calculating marginal densities,” *Journal of the American statistical association*, 85, 398–409.
- Harvey, A. (1991), *Forecasting, structural time series models and the Kalman filter*, Cambridge Univ Pr.
- Liu, J., Wong, W., and Kong, A. (1994), “Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes,” *Biometrika*, 81, 27.
- Marcellino, M. (1999), “Some consequences of temporal aggregation in empirical analysis,” *Journal of Business & Economic Statistics*, 17, 129–136.
- Nakamura, E. and Steinsson, J. (2008), “Five Facts about Prices: A Reevaluation of Menu Cost Models*,” *Quarterly Journal of Economics*, 123, 1415–1464.
- Schmitt-Grohé, S. and Uribe, M. (2004), “Solving dynamic general equilibrium models using a second-order approximation to the policy function,” *Journal of Economic Dynamics and Control*, 28, 755–775.
- Smets, F. and Wouters, R. (2005), “Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE approach,” *Journal of Applied Econometrics*, 20, 161–183.
- Whelan, K. (2002), “A guide to US chain aggregated NIPA data,” *Review of Income and Wealth*, 48, 217–233.
- Working, H. (1960), “Note on the correlation of first differences of averages in a random chain,” *Econometrica: Journal of the Econometric Society*, 28, 916–918.

Biography

Profile

1. Tae Bong Kim
2. born in Seoul, Korea on May 13, 1978
3. Pd.D. in Economics at Duke University, Durham, NC, U.S.A., May 2011
4. M.A. in Economics at Duke University, Durham, NC, U.S.A., May 2006
5. B.A. in Economics/Business at Yonsei University, Seoul, Korea, Feb 2004

Academic Achievement

1. Duke Graduate School Fellowship, 2006 - 2007, 2009
2. Kwanjeong Scholarship for MA program, 2004 - 2006
3. Yonsei Academic Achievement Scholarship, 1997
4. Temporal Aggregation Bias and Mixed Frequency Estimation of a New Keynesian Model, Fall 2010, Job Market Paper
5. Mixed Frequency Estimation of DSGE Models : Bayesian Approach(2009)
6. Bayesian Mixed Frequency VAR's with B. Eraker, et al. (2008)
7. After graduation, Tae Bong Kim will be a reserach fellow at Korea Development Institute.