Estimation methods of Adaptive Learning*

Ajay Kirpekar †

November 6, 2021

Abstract

The paper examines the role of adaptive learning in a DSGE framework and compares the results between two popular approaches of estimating the Likelihood function. The first approach uses conditional linearity and employs a Kalman filter to estimate the State Space model. The Second Approach take such non-linearities in the stochastic dynamics into account and estimates the State Space Model via the Particle Filter with re-sampling. After Bayesian Estimation via Metropolis Algorithm, both models are compared after taking into account E-stability of beliefs assumed by model and compared through posterior inference.

*William Branch has given structure and guidance
†Fabio Milani has offered helpful comments and suggestions

1 Introduction

In the last few decades, much of the cyclical macroeconomic phenomena have been modeled through forward looking agents who form beliefs on macroeconomic variables. Very often, the popular convention is to solve these models with the assumption that agents are perfectly aware of all the relevant data at the current time period before forming expectations of the future[1]. This assumption has been widely problematic with empirical literature (Lovell 1986,Evans Honkapohja 2001) and often cannot generate the persistence seen in the data.

In order to address a lack of persistence, researchers often employ habit formation [2] as well as lagged pricing rules [3] to fit moments of output and inflation observed in the data. Milani(2004) a standard NK model with persistence and compares an adaptive learning framework to that of rational expectations and finds learning offers a higher Marginal Likelihood [4].

However, when Adaptive learning via (Evans and Honkapohja 2001) is used to model expectations, traditional Lineal Models methods miss-specify the Likelihood function that is used in posterior sampling. In order to account for non-linearity in the Likelihood function, this paper uses a particle filter via re sampling (Villaverde and Rubio-Ramirez 2006). After accounting for both Likelihood functions, I evaluate differences in: Posterior Distribution, State dynamics of beliefs, and Overall Model Accuracy.

By assuming agents follow a model via adaptive learning, agents form beliefs that are updated every period based on their accuracy to the data formed by the true data generating process. One question of interest is whether this process of formation of beliefs converges to the true data generating process (See Evans Hankapohja 2001), this convergence is coined the E-stability principle. A given set of parameters in a model are called "E-stable" if they lead agents to eventually form beliefs that exactly fit the inherent process of the underlying

[1] See: ((Sims 1995)
[2] See: (Fuhrer 2000) and (Christiano, Eichenbaum and Evans 2005)
[3] See: (Woodford 2000) and (Christiano,Eichenbaum and Evans 2005)
[4] Geweke's Modified Harmonic Mean was used in order to address the variation of this statistic

state dynamics. The paper compares the results between the parameters estimated from the Likelihood function via particle filter and that of the conditionally Linear Kalman filter.

2 Prelimary Model

This paper first investigates the standard differences highlighted in the baseline NK model. Presented below are the dynamic equations which are augmented with an adaptive learning framework. The reason for the initial specification is to exploit the nature of its properties as expressed by Bullard and Mitra (2002).

Given the *actual law of motion*:

(1)$x_t = E_t[x_{t+1}] - \sigma^{-1}(i_t - r_t^n - E_t[\pi_{t+1}])$

(2) $\pi_t = kx_t + \beta E_t[\pi_{t+1}]$

(3)$i_t = \phi_x x_t + \phi_\pi \pi_t$

(4) $\mathrm{r}_t^n = p_r r_{nt-1} + \epsilon_t \sim N(0, \sigma_r^2)$

where: x_t, π_t, i_t are Output gap, Inflation, and the Federal funds rate respectively.

In the following model, agents hold the following *perceived law of motion*:

(5) $\mathrm{Z}_t = \Lambda_{0,t} + \Lambda_{1,t} Z_{t-1} + \Lambda_{2,t} r_t^n + \eta_t$, where: $\eta_t \sim N(0,1)$

hence: (7)$E_t[Z_{t+1}] = \Lambda_{0,t} + \Lambda_{1,t}(\Lambda_{0,t} + \Lambda_{1,t} Z_{t-1} + \Lambda_{2,t} r_t^n) + \Lambda_{2,t} r_t^n$

where: $\mathrm{Z}_t = [x_t^*, \pi_t^*, i_t^*]'$ and $\Lambda_{i,t}$ represents the beliefs that are revised over time

(8)$\phi_t = vec(\Lambda_{0,t}, \Lambda_{1,t}, \Lambda_{2,t})$

(9) $\phi_t = \phi_{t-1} + gR_t^{-1} X_t (y_t - X_t \phi_t)$

(10)$R_t = R_{t-1} + g(R_t^{-1} - X_t' X_t)$

Where: $\mathrm{y}_t = [x_t, \pi_t, i_t]'$ and: $X_t \phi_t = Z_t$

In the following standard setup, notice in equation (7) how agents, when forecasting variables in the future, observe shocks at time t, but do not observe the macroeconomic variables at time t. This aspect of observation demonstrates a crucial component of persistence generated in the standard NK model. Further more, Equations (9) and (10) are the recursive least squares procedure agents are assumed to follow when the update beliefs when forming expectations in time t+2.Equations (1)-(3) are labeled the "actual law of motion" as to indicate the true data generating process. While Equation (4) is considered the "perceived law of motion", so as to indicate agents perception of the how the Economy behaves. The system of Equations (1)-(7) is considered to be E-stable if:

$\lim_{t\to\infty} \phi_t = A^*$ Where A^* Represents the parameters that govern Equations(1)-(3) in matrix form.

In Models with Rational Expectations, it is often assumed that the parameters of interest are those that lead to a unique and stable equilibrium, otherwise known as a determinate solution. The notion of E-stability, like that of Determinacy, is also dependent upon the parameters of the model. In their seminal work, Bullard and Mitra(2002) derived the property that in the standard NK model, conditions for E-stability and determinacy yield are one in the same.

The parameters that govern equations (1) -(3) are **both** Determinate and E-stable if:

$\phi_\pi + \phi_x \frac{(1-\beta)}{k} > 1$ Note: In models with more parameters or a different specification, it is more the case that E-stability and Determinacy conditions are not only different, but do not jointly apply.[5]

In order to test plausibility of learning in a standard NK model, I estimate the posterior distribution of θ , the structural parameters in equations (1)-(5). Using quarterly data from 1954-2019 for output gap, inflation, and the federal funds rate, I estimate θ using a Random Walk Metropolis Hastings Algorithm.

3 Constant Gain Learning:

In traditional models with learning, agents update beliefs in such a way that over time the changes in beliefs converges to 0. This means that agents dramatically update beliefs in

[5] Mclung 2020 offers an alternative specification which demonstrates this.

earlier periods, and as time goes on, the changes to their belief parameters become incrementally smaller. In light of evidence by Branch and Evans(2005) as well as Orphanides and Williams (2004), Takes more seriously, the view that agents update beliefs each period with equal weight. This constant gain parameter g, determines the degree to which agents update beliefs based on forecast errors. Throughout the paper, this parameter is the subject of key interest in the proceeding sections.

4 Estimation Procedure

In most models with bayesian learning, most researchers take the critical stance of assuming Gaussian normality[6]. That is, they assume that conditional on the previous states, the expected value of the next state follows a normal distribution. By utilizing Gaussian Normality, Kalman filter methods can be used to estimate the state space and therefore the Likelihood function. Such assumptions take the critical stance that the learning parameters are fixed at each time period and do not follow a distribution of their own.In the preliminary NK model, this appears as:

$(1a) Y_{obs,t} = \Gamma_0 S_t + \Gamma_1 S_{t-1} + \epsilon_t^m, \epsilon_t^m \sim N(0, I_3)$

$\mathrm{S}_t = A_0 + A_1 S_t + A_2 E_t[S_{t+1}] + \nu_t$

$S_t = C_0 + C_1 E_t[S_{t+1}] + \eta_t, \sim N(0, V_s)$

After substituting for the expectation (7) we get:

(1b) $\mathrm{S}_t = \Lambda_{0,t} + C_1 P(\Lambda_{0,t} + \Lambda_{1,t} S_{t-1} + \Lambda_{2,t} r_t^n) + \eta_t$,where: $\Lambda_{i,t}$ are given by (9)-(10)

After some manipulation, the P matrix which performs T: $\mathbb{R}^3 \rightarrow \mathbb{R}^5$ and S_t becomes:

(1c) $\mathrm{S}_t = L_{0,t} + L_{1,t} S_{t-1} + L_{2,t} n_t, n_t \sim N(0, V_s)$

Where: $\mathrm{S}_t = [x_t, \pi_t, i_t, r_t]'$, and $Y_{obs,t} = [GDP_t - GDP_{Potential,t}, Inflation_t, FFR_t]$.

After obtaining (1a),(1c), and given a specified Prior distribution. I am now able to estimate

[6] See: Milani 2002, Hommes,MavroMatris,Ozden(2018)

the posterior distribution Where: $f(\theta|S_t, Y_t) \propto f(\theta)f(Y_t, S_t|\theta)$

and the Likelihood is derived from evaluating:

$f(Y_t, S_t|\theta) = f(Y_t|S_t, \theta)f(S_t|S_{t-1}, \theta)\ldots f(S_1|S_0, \theta)$. Using the following specification, I compare the Posterior distribution and Marginal Likelihood to the traditional New Keynesian model using the Likelihood generated through solving the Linear Rational Expectations Equilibrium of (1)-(4) via Gensys.[7]

5 Likelihood Miss-specification :

As expressed earlier, When estimating models with Bayesian Learning, it is commonplace to treat the learning parameters as constants. In reality, this assumption comes with Miss-specification issues. We know from the model, that the parameters, since they are endogenous to the states of the economy, should, in reality, be considered states themselves. Given this is the appropriate treatment, the traditional Kalman filter estimation can no longer be used. Referring back to the learning parameters, this becomes more clear. $\phi_t = \phi_{t-1} + gR_t^{-1}X_t(y_t - X_t\phi_t)$, $R_t = R_{t-1} + g(R_t^{-1} - X_t'X_t)$ and, $X_t = h(S_t, S_{t-1})$.

So, $\phi_t = h_1(S_t, S_{t-1})$ and, $R_t = h_2(S_t, S_{t-1})$where h_1 and h_2 are **nonlinear matrix functions**.

We know the Likelihood is derived from:

(11) $f(Y_t, S_t|\theta) = f(Y_t|S_t, \theta)f(S_t|S_{t-1}, \theta)\ldots f(S_1|S_0, \theta)$.

(12) $f(Y_t, S_t|\theta) = f(Y_t|S_t, \theta,)f(S_t|S_{t-1}, \phi_t, R_t, \theta)\ldots f(S_1|S_0, \phi_1, R_1, \theta)$.

(13)$f(Y_t, S_t|\theta) = f(Y_t|S_t, \theta,)f(S_t|S_{t-1}, h_1(S_t, S_{t-1}), h_2(S_t, S_{t-1}), \theta)\ldots f(S_1|S_0, h_1(S_1, S_0), h_2(S_1, S_0), \theta,)$.

In the Linear representation presented in Section 2, (11) is equivalent to (12). After substituting the learning process, (12) is indeed equivalent to (13). Since (13) is intractable, this representation is flawed and leads to erroneous values of S_t, meaning that the model ultimately yields a miss-specified Likelihood function. In order to properly characterize the

[7]This exercise is also done in Milani(2002).

New Keynesian model with an Adaptive Learning framework, we must depart from a conditionally Gaussian framework, append the state variable S_t with the learning parameters, and estimate using particle filter methods.

6 Particle Filter:

With similar motivations as *Rubio-Ramirez* & *Villaverde*(2007) and *Herbst* & *Schorfheide*(2017), I seek to estimate a nonlinear state space system via a particle filter with a re-sampling procedure. In essence, I express the learning parameters ϕ_t and R_t as state variables(rather than take them as given) in order to obtain the "correct" estimate of the Likelihood function. The Economic meaning behind estimating the learning parameters as states can interpreted in the following way: When agents form expectations of the future, the stochastic processes which drive macroeconomic variables are characterized by distributions that follow a mean and variance. Though the shocks are independent of the macro-economy, the distribution macroeconomic variables as well as the learning parameters are endogenous to one another. By adapting the learning parameters as part of the state variable, the researcher is in effect taking the stance that the distribution of macro variables is dependent on the distribution of beliefs and vis versa. Notice However, in the conditionally linear model, since we hold the learning parameters ϕ_t and R_t fixed, we assume that beliefs are driven by macro variables but not the reverse. The Particle Filter in essence simulates the stochastic error terms with a sufficient number of draws, places weights on the realized number states,and calculates the a weighted Likelihood. Often this procedure places high weight on few particles and 0 weight on most particles which in essence induces a high variance of the Likelihood function.[8] In order to avoid this issue, I follow Rubio-Ramirez Villaverde(2007) with the following algorithm.[9]

[8] Kitigawa 1996

[9] For Convergence properties of the Likelihood, please refer to Rubio-Ramirez Villaverde(2007) and Kitigawa....

(14) $\mathrm{Y}_t = M_1 S_t + M_2 S_{t-1} + \epsilon_t^m$.

(15)$S_t = f(S_t, \epsilon_t^s)$.

Step 1(Initialize): Set $e_t^j \sim \epsilon_t^s$.

Step 2(Propagate): $\mathrm{S}_t^j = f(S_{t-1}^j, \epsilon_t^j)$.

Step 3(Evaluate): $w_t^j = \frac{L(Y_t, S_t^j)}{\Sigma w_t^i}$. Where $L(Y_t, S_t^j)$ is the Likelihood function obtained from (14).

Step 4(Re-sample): $\mathrm{q}_i \sim w_{t\,0}^{j\,J}$ and set: $S_{t|t-1}^j = S_t^i$,for all $\{\mathrm{q}_i\}_{i=0}^J$

Set t = t+1; and repeat till t = T.

Step 5(Calculate Likelihood): $\mathrm{P}(\mathrm{Y}^T, \theta) \approx \frac{1}{J}(\prod_{t=1}^{T}(\sum_{j=1}^{J} p(Y_t | w_t^j, S_{t|t-1}^j, Y^{t-1}; \theta)))$.

Where $\mathrm{P}(\mathrm{Y}^T, \theta)$ is obtained from measurement equation(14).

With the stated procedure, as the number of particles, J becomes greater, the Likelihood becomes converges to the true distribution. However, because one must keep track of the states and their relative values across the sample, there exists a trade-off between and accuracy and computational time of the algorithm. In this paper, I find that the results of estimation do not change much between 10 - 100 thousand particles. Hence I assume J to be 100 thousand, a sufficient approximation to the true Likelihood function of interest in the proceeding results. Additionally, because the Likelihood function in the conditionally linear case is indeed subject to miss-specification, as seen in (13). We know traditional measures of fit via marginal Likelihood are not an appropriate comparison of model fit with finite sample size. Because of this reason, other standard measures of model fit may be more appropriate such as MSE or RMSE.

7 Standard NK Results:

As mentioned earlier, in this paper, I append S_t to include ϕ_t, R_t as well as the macro-variables from (1)-(4). With the following equations, I am now able to estimate the Likelihood function with non-linearity taken into account and proceed with a Particle Filter

Metropolis-Hastings procedure. the following procedure, follows the same RWMH algorithm, but instead calculates the log-Likelihood using the particle filter method expressed in the previous section.

NK Model via Kalman Filter							
Para	Description	Prior Mean	Prior Std.	Posterior Mean	Posterior Std.	Dist.	Bound
p_r	Autoreg. Dem.	.5	.083	.527	.08	Uniform	0<x 1
σ^{-1}	IES	.125	.09	.18	.44	gamma	0<x
ϕ_π	Feedback Infl	1.6	.15	1.61	.019	gamma	0<x < 4
ϕ_x	Feedback Gap	.5	.2	1.16	.014	gamma	0<x<2
σ_r	Demand Shock	.2	.1	.181	.023	invgamma	0<x
k	infl. slope	.015	.011	.1	.0001	gamma	0<x <.8
p_u	Autoreg. Sup.	.5	.08	.52	.068	uniform	0<x < 1
σ_u	Supply Shock	.2	.1	.146	.0014	invgamma	0<x < 1
g	constant gain	.031	.022	.026	.0002	gamma	0<x < 1

NK Model via Particle Filter							
Para	Description	Prior Mean	Prior Std.	Posterior Mean	Posterior Std.	Dist.	Bound
p_r	Autoreg. Dem.	.5	.083	.571	.006	Uniform	0<x 1
σ^{-1}	IES	.125	.09	.1	.2	gamma	0<x
ϕ_π	Feedback Infl	1.6	.15	1.63	.023	gamma	0<x < 4
ϕ_x	Feedback Gap	.5	.2	1.15	.0016	gamma	0<x<2
σ_r	Demand Shock	.2	.1	.258	.1219	invgamma	0<x
k	infl slope	1.2	.15	.1	$9e^{-5}$	gamma	1<x
p_u	Autoreg. Sup.	.5	.08	.457	.0075	uniform	0<x < 1
σ_u	Supply Shock	.2	.1	.203	.033	invgamma	0<x < 1
g	constant gain	.031	.022	.029	.0004	gamma	0<x < 1

8 Hybrid NK Model:

Though the preliminary model offers interesting considerations regarding the the role of learning, it is also important to compare differences in estimation in a more robust New Keynesian model.That is, one that accounts for habits, inflation indexation, and a lagged interest rate rule.[10] This paper appends the preliminary with a more comprehensive model that attributes persistence in inflation and output to price indexation on the part of firms and habit formation of households respectively. On the part of firms, follow Gali(2008) , and attribute their marginal cost proportional to parameters of the production and labor demand.

Households:

First, I assume households have the choice to consume from a basket of goods, offered by a continuum of firms and that their preferences between differing goods follows Dixit-Stiglitz

[10]See: CEE 2005

Preferences. [11]

$C_t = \int_0^1 \left(c_t(i)^{\frac{\epsilon}{\epsilon-1}} di \right)^{\frac{\epsilon-1}{\epsilon}}$. It can be shown that the price of the optimal consumption bundle can be expressed s.t.

$P_t C_t = \int_0^1 p_t(i) c_t(i) di$ where, $P_t \equiv (\int_0^1 p_t(i)^{\frac{\epsilon}{\epsilon-1}})^{\frac{\epsilon-1}{\epsilon}} di$

$U(C_t, L_t) = \frac{(C_t - hC_{t-1})^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\psi}}{1+\psi}$ while the household budget constraint is:

$P_t C_t + B_t = L_t w_t + (1 + i_{t-1}) B_{t-1}$

after taking the first order conditions,log-linearizing near the steady state, and using market clearing conditions ($c_t = y_t$), the following holds:

$y_t = w_1 E_t y_{t+1} + w_2 (E_t \pi_{t+1} - i_t) + w_3 E_t y_{t+2} + w_4 y_{t-1}$.

After plugging for the natural rate of output (where y_t^n is the output when $\pi_t = 0$):

(1c) $x_t = w_1 E_t x_{t+1} + w_2 (E_t \pi_{t+1} - i_t - r_t^n) + w_3 E_t x_{t+2} + w_4 x_{t-1}$.

$r_t^n = p_r r_{t-1}^n + \epsilon_t^r$, $\epsilon_t^r \sim N(0, \sigma_r^2)$

where x_t is the deviation of output to the natural rate of output, and r_t^n is considered the natural rate of interest.[12]

Firms:

On the firm side, I assume that firms operate in a monopolisitically competitive environment and seek to maximize:

$\max_{p_t^*(j)} \; E_t[\Sigma_{k=t}^T \theta^k Q_{t,t+k} \Pi_{t+k}]$

$\Pi_t \equiv y_{t+k}(j) p_t^*(j) - f^{-1}(y_{t+k})$ and $Q_{t,t+k} = \beta^k \frac{P_t}{P_{t+k}} (\frac{\lambda_{t+k}}{\lambda_t})$

In the firm's optimization problem, they seek to maximize via standard Calvo pricing framework where they take into account, $Q_{t,t+k}$, their stochastic discount factor which is expressed as the ratio of marginal utility that comes from the household's Lagrange multiplier via the inter-temporal margin of consumption. Following CEE[13] and Gali(2008)[14], I assume agents

[11] See Dixit Stiglitz

[12] for simplicity in the estimation, I generalize r_t^n as an AR(1) process, similar to Milani(2002)

[13] Christiano,Eichenbam,Evans(2005)

[14] Monetary Policy, Inflation, and the Business Cycle(2008)

who are unable to optimize next period, index their prices partially(by γ) to observed inflation in the previous time period. By doing so, I am able to express Π_{t+k} as:

$\Pi_t \equiv Y_{t+k} p_t^*(j)(\frac{p_t^*(j)\pi_{t+k-1}^{\gamma}}{P_t})^{-\epsilon} \; - w_{t+k} f^{-1}((\frac{p_t^*(j)\pi_{t+k-1}^{\gamma}}{P_t})^{-\epsilon} Y_{t+k})$

Furthermore, I assume that firms produce with the same production function:

$Y_t(i) = f(L_t(i)) \equiv A_t L_t(i)^{1-\alpha}$ (Note: $f^{-1}(x) \equiv x^{\frac{1}{1-\alpha}}$ represents the labor demand function)

After taking the first order condition, substituting for the wage, and log-linearizing around a 0 inflation steady state, I obtain the following Phillips curve:

(2c) $\pi_t = \eta_1 x_t + \eta_2 x_{t-1} + \eta_3 E_t[x_{t+1}] + \eta_4 E_t[\pi_{t+1}] + \eta_5 \pi_{t-1} + u_t u_t = p_u u_{t-1} + \epsilon_t^u;\ \epsilon_t^u \sim N(0, (\sigma_t^u)^2)$

Details of this derivation can be found in the Appendix.

Monetary Policy:

I assume Monetary Policy targets the FFR using the following rule:

(3c) $\mathrm{i}_t = \rho i_{t-1} + (1-\rho)(\phi_x x_t + \phi_\pi \pi_t) + \epsilon_t^i \epsilon_t^i \sim N(0, (\sigma_t^e)^2)$

Taken together, the full model(1c-3c) now depicts a New Keynesian model with persistence driven by both real economic variables as well as expectations driven persistence. Hence, I am able to express the dynamics of the stated model in Matrix form:

(4c) $\mathrm{Z}_t = AZ_{t-1} + BU_t + CE_t[Z_{t+1}] + DE_t[Z_{t+2}] + E\nu_t$

where: $\nu_t \sim N(0, \Sigma_t)$ $Z_t = [x_t, \pi_t, i_t]$ $U_{t-1} = [u_{t-1}, r_{t-1}^n]$

and Σ_t is endogenous to the variances from the cost push, natural rate, and monetary policy shocks($\sigma_r^2, \sigma_e^2, \sigma_u^2$). Unlike the preliminary model, the Hybrid NK model accounts for two period ahead forecasts on the part of agents and thus further propagates non-linearity with cubed terms in the learning parameters agents form for macro variables. This can be seen if one substitutes (5c) and (6c) into (4c). As non-linearity becomes more pronounced, the differences in estimation also become more apparent.

$E_t Z_t = \Lambda_{0,t} + \Lambda_{1,t} Z_{t-1} + \Lambda_{2,t} U_t$

$\mathrm{E}_t Z_{t+1} = \Lambda_{0,t} + \Lambda_{1,t} E_t Z_t + \Lambda_{2,t} RU_t$

(5c) $\mathrm{E}_t Z_{t+1} = \Lambda_{0,t} + \Lambda_{1,t}(\Lambda_{0,t} + \Lambda_{1,t} Z_{t-1} + \Lambda_{2,t} RU_t) + \Lambda_{2,t} RU_t$

(6c) $\mathrm{E}_t Z_{t+2} = \Lambda_{0,t} + \Lambda_{1,t} E_t Z_{t+1} + \Lambda_{2,t} R^2 U_t$

Using the same methodology discussed earlier, (4c) is then cast into state space form and becomes: $(7c) Z_t = L_{0,t} + L_{1,t} Z_{t-1} + \epsilon_t^s$, $\epsilon_t \sim N(0, A_t)$
Thus when comparing the Particle filter to the Conditionally linear Kalman filter, The inherent difference comes from whether or not to append Z_t with ϕ_t and R_t, the adaptive learning parameters. The next section discusses the results post estimation.

9 Hybrid NK results:

In the Estimation phase, This paper takes the state space form of (7c) and compares the posterior distribution of the parameters that govern the law of motion. For the purposes of tractability and clarity, I only consider parameters that yield a unique and stable Rational expectations equilibrium for equation (4c). If a draw from the proposal distribution offers an indeterminate solution, it is dropped and the MCMC algorithm continues to draw from the proposal distribution till there is a REE.

A note on Unstable Beliefs: It is well known in the Literature that often initial conditions or a set of realized shocks can lead to Unstable/Non-stationary process.(Marcet and Sergeant 1989) In the context of the Hybrid NK model, this becomes a point of concern when lags from the actual law of motion can induce unstable beliefs. In order to avoid these phenomena, I assume agents stop updating believes whenever they have a PLM that is non-stationary. Furthermore, within the estimation, I also record the "dropout rates" which count the number of times the agent no longer employs Recursive Least Squares. Thus, the dropout rate records what percent of the time agents must revise their beliefs due to miss-specifying a non-stationary process. I take the stance that agents recenter their beliefs somewhere near the REE of the Eigenvalues inherent in the Actual Law of Motion: $\phi_t \sim N(\phi^*, I)$ if $|eig(L_{1,t})| > 1$. Furthermore, within the estimation, I also record the "dropout rates" which count the number of times the projection facility is used. For example, for a given set of parameters, the max number of times the projection can be used is the length of the sample of data, 260. Thus, the dropout rate records what percent of the time agents must revise their beliefs due to miss-specifying a non-stationary process.

Hybrid NK Model via Kalman Filter							
Para	Description	Prior Mean	Prior Std.	Posterior Mean	Posterior Std.	Dist.	Bound
p_r	Autoreg. Dem.	.5	.083	.569	.06	Uniform	0<x 1
σ^{-1}	IES	.125	.09	.419	.065	gamma	0<x
ϕ_π	Feedback Infl	1.6	.15	1.67	.024	gamma	0<x < 4
ϕ_x	Feedback Gap	.5	.2	.188	.018	gamma	0<x<2
ν	Frisch Labor	.3	.2	.331	.029	gamma	0<x
θ	Calvo Stick.	.015	.011	.0016	.0001	gamma	0<x <.8
σ_r	Demand Shock	.2	.1	.165	.014	invgamma	0<x
α	Capital Share	.3	.05	.298	$2.3e^{-6}$	Normal	.2<x < .3
ψ	Markup	1.2	.15	1.175	.01	gamma	1<x
h	Habit	.5	.083	.447	.073	uniform	0<x < 1
γ	Infl. Indexation	.5	.083	.429	.04	uniform	0<x < 1
ρ	MP Persistence	.5	.2	.47	.02	gamma	0<x < 1
p_u	Autoreg. Sup.	.5	.08	.493	.06	uniform	0<x < 1
σ_e	Demand Shock	.2	.1	.265	.037	invgamma	0<x < 1
σ_u	Supply Shock	.2	.1	.263	.067	invgamma	0<x < 1
g	constant gain	.031	.022	.0279	.0036	gamma	0<x < 1

Hybrid NK Model via Particle Filter							
Para	Description	Prior Mean	Prior Std.	Posterior Mean	Posterior Std.	Dist.	Bound
p_r	Autoreg. Dem.	.5	.083	.649	.005	Uniform	0<x < 1
σ^{-1}	IES	.125	.09	0.28	.38	gamma	0<x
ϕ_π	Feedback Infl	1.6	.15	1.63	.02	gamma	0<x < 4
ϕ_x	Feedback Gap	.5	.2	.18	.017	gamma	0<x < 2
ν	Frisch Labor	.3	.2	0.636	.13	gamma	0<x
θ	Calvo Stick.	.015	.015	.0193	.0001	gamma	0<x < 1
σ_r	Demand Shock	.2	1.00	1.56	.5	invgamma	0<x < 1
α	Capital Share	.3	.05	.298	2.3e^{-6}	Normal	0.2<x < .3
ψ	Markup	1.2	.15	1.25	.019	gamma	1<x
h	Habit	.5	.083	.445	.043	uniform	0<x < 1
γ	Infl. Indexation	.5	.083	.389	.068	uniform	0<x < 1
ρ	MP Persistence	.5	.2	.544	.032	gamma	0<x < 1
p_u	Autoreg. Sup.	.5	.08	.457	.041	uniform	0<x < 1
σ_e	Demand Shock	.2	.1	.488	.14	invgamma	0<x < 1
σ_u	Supply Shock	.2	.1	.195	.02	invgamma	0<x < 1
g	gain coef.	.031	.022	.0316	.0004	gamma	0<x < 1

KF Mean Dropout rate: 14.51% Variance of Dropout: 11.05

PF Mean Dropout rate: 25.36 % Variance of Dropout: 7.77

A Note on Marginal Likelihood: Although at first glance it may appear as though the Kalman Filter outperforms its counterpart. Accurate Marginal Likelihood estimates

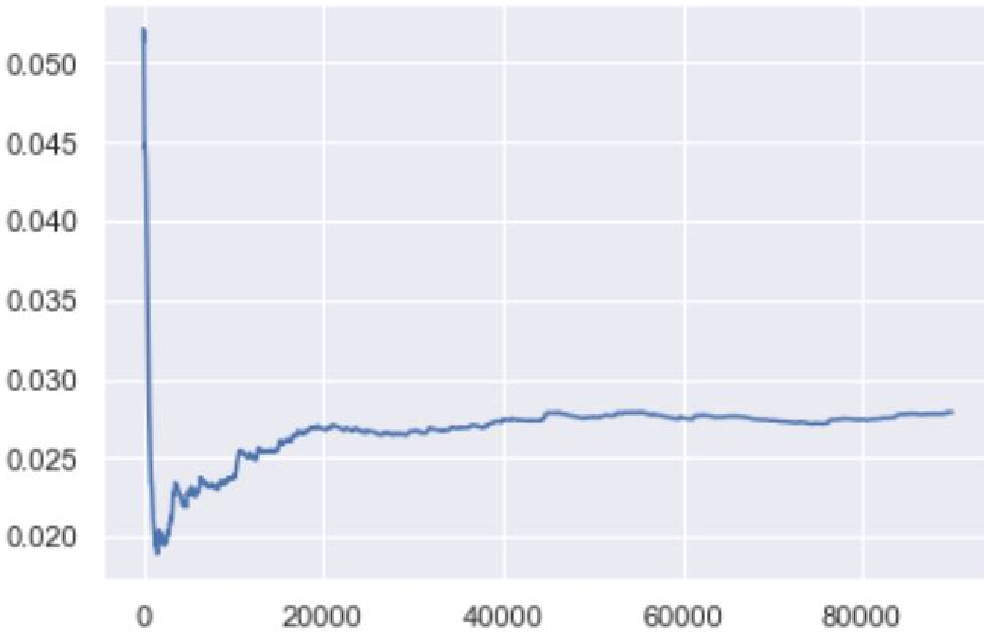

Figure 1: Moving Avg for g in KF Estimation

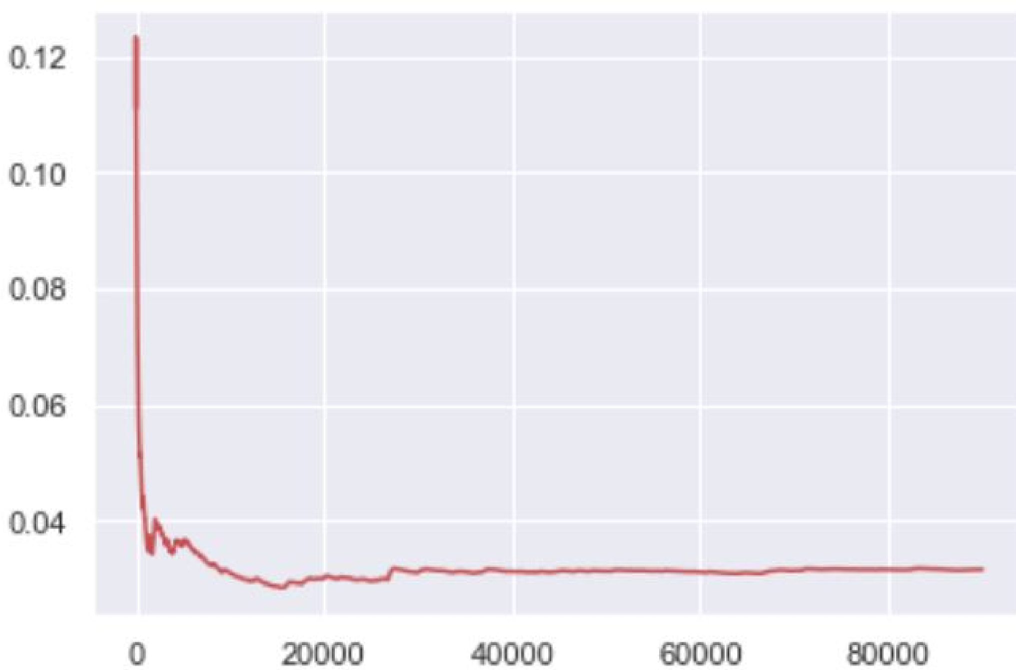

Figure 2: Moving Avg for g in PF Estimation

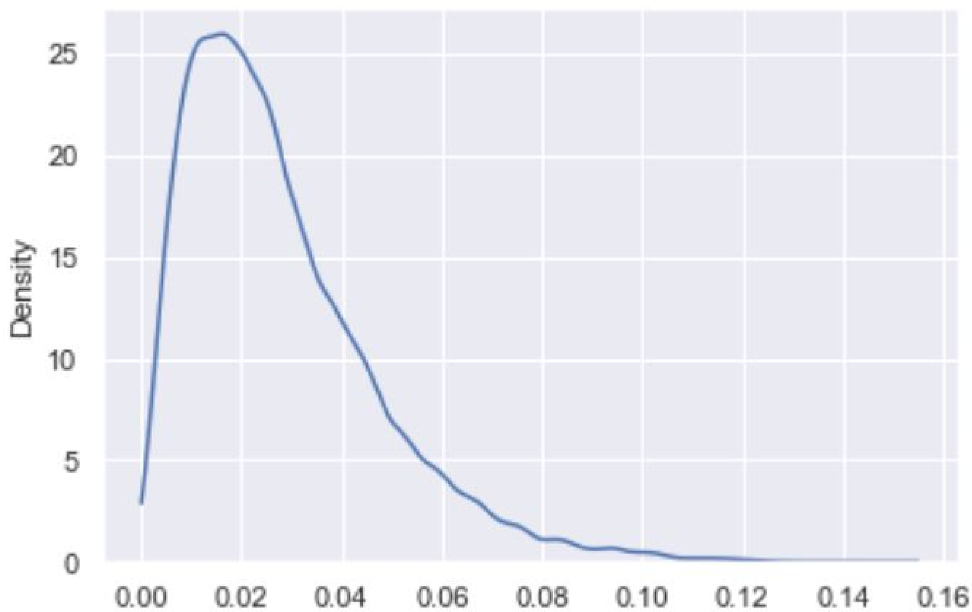

Figure 3: Posterior Distribution of g in KF Estimation

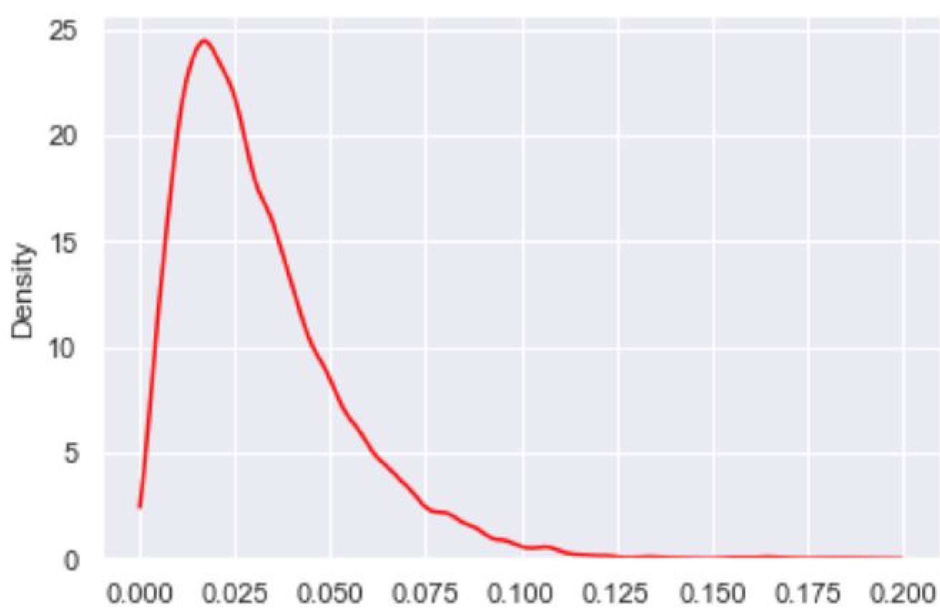

Figure 4: Posterior Distribution of g in PF Estimation

inherently depend on correct likelihood specification, Hence Kalman filter approximations yield noticeable upward bias. One proof of concept to show this is to track the Kalman filter's Mean Squared error to the Particle Filter's Mean Squared error given a set of parameters. Below, I report the corresponding SSE as well as plots of the generated data to the actual data for the output gap.[15]. For the Sum of Squared errors computation, I measure the difference between the model generated output and inflation to the actual output gap and inflation.[16]

I calculate the Particle Filter portion using the following method:

$$\text{SSE}(\theta)_{PF} = \Sigma_{t=0}^{T}\left(Y_t^{obs} - M_1\hat{E}_t[S_t] + M_2\hat{E}_t[S_{t-1}]\right) = \Sigma_{t=0}^{T}\left(Y_t^{obs} - \frac{M_1}{J}\left(\Sigma_{j=1}^{J} S_{t|t}^{j}\right) + \frac{M_2}{J}\left(\Sigma_{j=1}^{J} S_{t-1|t}^{j}\right)\right)$$

Here, the conditional expectation denotes the expected value of the state after re-sampling.

In contrast, the Kalman filter SSE is computed via Conditional Gaussian Linearity:

$$\text{SSE}_{KF}(\theta) = \Sigma_{t=0}^{T}\left(Y_t^{obs} - M_1\hat{E}_t[S_t] + M_2\hat{E}_t[S_{t-1}]\right)$$

$$\hat{E}_t[S_t] = S_{t|t-1} + L_t D_t^{-1}(Y_t^{obs} - \hat{Y}_t)$$

$$\text{L}_t = Cov(Y_t^{obs}, S_t|I_t)$$

$$\text{D}_t = Var(Y_t^{obs}|I_t)$$

I_t = Information Set[17]

$SSE_{KF}(\theta^*_{KF}) = .0169$

$SSE_{PF}(\theta^*_{KF}) = .0116$

$SSE_{PF}(\theta^*_{KF}) = 0.02049977$

$SSE_{PF}(\theta^*_{PF}) = 0.01163876$

The Value of importance is:

$\text{SSE}_{PF}(\theta^*_{PF}) < SSE_{KF}(\theta^*_{KF})$ Which offers some evidence that the Particle filter is superior in data fit.

[15] The Output gap has been set through a band pass filter in order to avoid Steady State parameters in the estimation phase...See Christiano and Fitzgerald(1999)

[16] I omit the calculation of the Federal Funds rate because both estimate the federal funds rate with noticeable inaccuracy. This is most likely due to issues with ZLB...See Wu and Zhang(2019)

[17] for more details, refer to Durban and Koopman(2001)

10 Model Fit

Because Sum of Squared errors is subject to potential variance. It is indeed plausible that the Particle filter is subject to over-fitting. In order adress this issue, I employ a Bayes Factor calculation to determine whether the KF or PF performs with better model accuracy relative to the true Data. In this context, when the Bayes Factor is computed via likelihood comparison, the Particle filter's likelihood is penalized for particle size.(See Step 5 via Particle Filter Algorithm.)

P(M$_i|Y_{data}) = \frac{P(M_i)P(Y_{data}|M_i)}{P(M_i)}$

By setting the prior wieghts for both Models equal, I obtain:

BF $= \frac{P(Y_{data|M_i})}{P(Y_{data}|M_j)}$

where: P(M$_i|Y_{data}) = \int P(Y_{data}|\theta_i)$ dθ_i

P(M$_i|Y_{data}) \approx \sum_{\tau=1,N} LL(Y_{data}|\theta^i_\tau) = \Sigma^N_{\tau=1}\Sigma^T_{t=1}LL(y_t|\theta_\tau)$

We can evaluate which Model offers better data fit across the parameters obtained from the Metropolis algorithm.Hence, if model i is the PF and j is the KF, if the BF term has a value greater than 1, this implies the PF corresponds with the true data set of Macroeconomic time series. The results of this for both the baselines and habit formation NK model are shown below.

11 E-stability:

In the preliminary model, conditions for E-stability and Determinacy are one in the same. However, in the Hybrid NK model, conditions for E-stability and Determinacy are different. Thus there are a variety of cases where parameters can be: 1.Estable/determinate 2.Unstable/determinate 3.Unstable/indeterminate 4. Estable/indeterminate. This section describes the conditions as provided by Evans and Honkapohja(2012) in greater detail, and seeks to highlight the Economic significance of these principles. In a general framework this can be expressed as:

ALM: $\mathrm{x}_t = Ax_{t-1} + BE_t[x_{t+1}] + CU_t$

PLM: $\mathrm{x}_t = L_0 x_{t-1} + L_1 U_t + \eta_t$

Stationary Shocks: $\mathrm{U}_t = RU_{t-1} + \epsilon_t$

After plugging in the Perceived law of motion(PLM) into the Actual law of motion(ALM) and simplifying:

$\mathrm{x}_t = C_1 x_{t-1} + C_2 U_t$, where: $C_1 \equiv (A + BL_0 L_0)$ and $C_2 \equiv B(L_0 L_1 + L_1 R) + C$.

A model where agents hold beliefs such that: $\mathrm{L}_0^* = C_1$ and $L_1^* = C_2$ are labeled a rational expectations equilibrium because the beliefs they form exactly match the equilibrium ALM. However, if we relax this assumption and allow agents divert off the equilibrium path with recursive learning, we cannot guarantee that the parameters that make up the model will always lead to $\lim_{t\to\infty} L_{0,t} = C_1 = L_0^*$ and $\lim_{t\to\infty} L_{1,t} = C_2 = L_1^*$. As per Evans and Honkapohja, the E-stability condition is then satisfied if: $Dh(\theta)$ contains negative eigenvalues. Where $h(\theta) = \lim_{t\to\infty} E[Q(t,\theta,Z_t(\theta))]$ and $\theta_t = \theta_{t-1} + gQ(t,\theta_{t-1},Z_t)$. Here, θ_t represents the belief parameters of L_0 and L_1. While Z_t represents the corresponding information used to form the PLM[18] and $Q(t,\theta,Z_t(\theta))$ is the forecast error of the agent. With respect to the above model, this condition is satisfied when:

$\mathrm{h}_1(\theta) = C_1$

$\mathrm{h}_2(\theta) = C_2$

$\mathrm{Dh}(\theta) = \begin{bmatrix} \frac{\partial h_1}{\partial L_0} & \frac{\partial h_1}{\partial L_1} \\ \frac{\partial h_2}{\partial L_0} & \frac{\partial h_2}{\partial L_1} \end{bmatrix}_{L_0=L_0^*\ ,\ L_1=L_1^*}$ contains negative real eigenvalues.[19]

In the next section, the paper takes the posterior parameter draws from the previous estimation and filters them for those which are E-stable. Thus the results show the parameters

[18]For more details on the stability of beliefs, see Evans and Honkapohja(2012)

[19]For another implementation of E-stability, refer to McClung(2020)

which are both Determinate and E-stable in the Hybrid NK model.[20]

12 Results for E-stability:

	E-Stable Particle Filter				
Para	Descr	Estable KF	Estable PF	KF	PF
p_r	Autoreg. Dem.	.76	.83	.569	.649
σ^{-1}	IES	1.10	1.26	.419	.281
ϕ_π	Feedback Infl.	1.58	1.67	1.57	1.63
ϕ_x	Feedback Gap	.93	.827	.188	.184
ν	Frisch Labor	1.36	1.386	.331	.636
θ	Calvo Stick.	.007	.024	.0016	.015
σ_r	Demand Shock	1.61	1.166	.243	1.01
α	Capital Share	.298	.29	.298	.298
ψ	Markup	1.04	1.005	1.175	1.25
h	Habit	.047	.112	.447	.445
γ	Infl. Indexation	.109	.048	.428	.389
ρ	MP Persistence	.392	.521	.47	.544
p_u	Autoreg. Sup.	.9	.848	.493	.457
σ_e	Demand Shock	1.26	1.16	.265	.488
σ_u	Supply Shock	.91	.96	.263	.195
g	constant gain	.0217	.029	.0279	.0316

[20] See Appendix for E-stability conditions of the Hybrid NK model.

Percent of Post. Estimation Parameters KF: 5.272 %

Percent of Post. Estimation Parameters PF: 51.342 %

Mean Dropout Rate KF: 14.51 %

Mean Dropout Rate PF: 25.36 %

Estable Mean Dropout Rate KF: 2.35 %

Variance of Dropout: 11.41

Estable Mean Dropout Rate PF: 32.93%

Variance of Dropout: 7.56

$\text{SSE}_{PF}(\theta^{E*}_{PF}) = 0.01161553$

$\text{SSE}_{PF}(\theta^{*}_{PF}) = .01163876$

$\text{SSE}_{KF}(\theta^{E*}_{KF}) = .0184101$

$\text{SSE}_{KF}(\theta^{*}_{KF}) = .0169$

$\theta^{E*}_{j} \equiv$ Posterior Mean of E-stable Parameters for j type estimation

$\theta^{*}_{j} \equiv$ Posterior Mean of all Parameters for j type estimation

Posterior estimation shows that in both models, E-stability implies lower constant gains than the without E-stability. Additionally, it is clear after filtering for the E-stable values in KF estimation, the parameters are indeed closer to both the PF as well as the Estable Particle filter parameters. This implies that E-stability attenuates parameter differences produced from KF estimation. However, since the SSE values for the KF are an order of 10 times higher, this offers just cause to apply PF estimation.
One may initially consider the projection facility to be an incorrect account of belief formation. Holding this view would however imply the projection facility should be utilized less overall in both models. However, it is interesting that in the PF model, E-stable parameters yield **higher** dropout ratios. This implies that even when agents are assumed to form expectations that will exactly match the process they estimate, the projection facility is still utilized.Thus offering some credibility that agents, will in fact, periodically update beliefs irrespective of a Bayesian update rule, i.e Recursive Least squares.

13 Conclusion

Using an MCMC method, this paper has shown the significant differences between linear and nonlinear estimation yield in both data fit as well as strucural parameters. Ultimately, we have seen that in order estimate agent based learning, one must adopt a nonlinear estimation technique such as the Particle filter to capture a better estimate of your Likelihood function. Furthermore, this work has examined the performance of E-stability conditions and offers evidence in support of the projection facility. One key drawback that where future work is needed, pertains to capturing nonlinearity in the FFR. More specifically, it is imperative that agent based learning can append knowledge of the zero lower bound in ways that may depart from the traditional Rational Expectations.

14 References

Branch, W., and; Evans, G. (2011). Monetary policy and heterogeneous expectations. Economic Theory.

Christiano, L., and; Evans, C. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. The University of Chicago Press.

Christiano, L., and; Fitzgerald, T. (1999). THE BAND PASS FILTER. NBER.

Durbin, J., and; Koopman, S. J. (2012). Time Series Analysis by State Space Methods: Second Edition (9780199641178, 019964117X). doi:9780199641178, 019964117X

Evans, G. (2012). Learning and Expectations in Macroeconomics. Princeton, NJ: Princeton University Press.

Fernandez-Villaverde, J., and; Rubio-Ramirez, J. (2006). Estimating Macroeconomic Models: A Likelihood Approach. doi:10.3386/t0321 Fuhrer, J. C. (2000). Habit Formation in Consumption and Its Implications for Monetary-Policy Models. American Economic Review, 90(3), 367-390. doi:10.1257/aer.90.3.367

Gali, J. (2015). Monetary policy, inflation, and the business cycle: An introduction to the new Keynesian framework and its applications.

Hommes, C. H., Mavromatis, K., Ozden, T., and; Zhu, M. (2019). Behavioral Learning Equilibria in the New Keynesian Model. SSRN Electronic Journal. doi:10.2139/ssrn.3450570

Kitigawa, G. (n.d.). Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models. Journal of Computational and Graphical Statistics.

Marcet, A., and; Nicolini, J. P. (1998). Recurrent Hyperinflations and Learning. SSRN Electronic Journal. doi:10.2139/ssrn.95989

McCallum, B. (2004). On the Relationship between Determinate and Msv Solutions in Linear Re Models. NBER.

McClung, N. (2020). E-stability vis-‘a-vis Determinacy in Regime-Switching Models. Journal Of Economic Literature.

Milani, F. (n.d.). Expectations, Learning and Macroeconomic Persistence.

Orphanides, A., and; Williams, J. (2003). Imperfect Knowledge, Inflation Expectations, and Monetary Policy. doi:10.3386/w9884

Stiglitz, J., and; Dixit, A. (n.d.). Monopolistic Competition and Optimum Product Diversity. American Economic Review.

Woodford, M. (2000). Monetary Policy in a World Without Money‘. NBER. Wu, J., and;

Zhang, J. (2019). A shadow rate New Keynesian model. Journal of Economic Dynamics and Control.

Appendix A. Investment Spending Curve

Households seek to maximize:

$\mathrm{E}_t[\Sigma_{t+k}^T \beta^k U(C_{t+k}, L_{t+k})]$ s.t $P_{t+k}C_{t+k} + B_t = L_{t+k}w_{t+k} + (1 + R_{t+k-1})B_{t+k-1}$ for k=1,2...,T

This becomes, $\mathcal{L} = \mathrm{E}_t[\Sigma_{t+k}^T \beta^{t+k} U(C_{t+k}, L_{t+k}) - \lambda_{t+k}(P_{t+k}C_{t+k} + B_t - L_{t+k}w_{t+k} - (1+i_{t+k-1})B_{t+k-1})]$

(1) $\dfrac{\partial \mathcal{L}}{\partial C_t} = (C_t - hC_{t-1})^{-\sigma} - h\beta(E_t C_{t+1} - hC_t) - \lambda_t P_t = 0$

(2) $\dfrac{\partial \mathcal{L}}{\partial L_t} = -(L_t)^{\psi} + \lambda_t(w_t) = 0$

(3) $\dfrac{\partial \mathcal{L}}{\partial B_t} = -\lambda_t + E_t[\lambda_{t+1}](1 + R_t) = 0$

After Substitution of (1) in (3):

$\mathrm{E}_t[\frac{(C_{t+1}-hC_t)^{-\sigma} - h\beta(C_{t+2}-hC_{t+1})}{(C_t - hC_{t-1})^{-\sigma} - h\beta(C_{t+1}-hC_t)} \frac{P_t}{P_{t+1}}] = (1 + R_t)^{-1}$

After Log Linearization around the S.S. and setting $c_t = y_t$:

$y_t = w_1 E_t y_{t+1} + w_2(E_t \pi_{t+1} - i_t) + w_3 E_t y_{t+2} + w_4 y_{t-1}$. Where:

$i_t = log(R_t) - log(R)$

$w_1 \equiv [\frac{1+h^2\beta+h\beta}{1+h+h^2\beta}]$

$w_2 \equiv [\frac{(1-h)(1-h\beta)}{\sigma(1+h+h^2\beta)}]$

$w_3 \equiv [\frac{-h\beta}{1+h+h^2\beta}]$

$w_4 \equiv [\frac{h}{1+h+h^2\beta}]$

using the fact that $x_t = y_t - y_t^f$ and $y_t^f = w_1 E_t y_{t+1}^f + w_2(-i_t) + w_3 E_t y_{t+2}^f + w_4 y_{t-1}^f$,

we can express the IS curve as: $x_t = w_1 E_t x_{t+1} + w_2(E_t \pi_{t+1} - i_t - r_t^n) + w_3 E_t x_{t+2} + w_4 x_{t-1}$.

where $r_t^n = \frac{1}{w_2}(-y_t^f + y_{t+1}^f + w_3 y_{t+2}^f + w_4 y_{t-1}^f)$.

Appendix B. Phillips Curve

Firms seek to maximize the following objective function:

$\mathcal{V} = \max_{p_t^*(j)} \; E_t[\Sigma_{k=t}^T \theta^k Q_{t,t+k} \Pi_{t+k}]$

$\Pi_t \equiv y_{t+k}(j)p_t^*(j) - f^{-1}(y_{t+k})$ and $Q_{t,t+k} = \beta^k(\frac{P_t}{P_{t+k}}(\frac{\lambda_{t+k}}{\lambda_t})$

$$\frac{\partial \mathcal{V}}{\partial p_t^*(j)} = E_t[\Sigma_{k=t}^T \theta^k \beta^k (\frac{\lambda_{t+k}}{\lambda_t})(\frac{p_t^*}{p_{t+k}})(\frac{\pi_{t+k-1}^{\gamma} p_t^*}{p_{t+k}})^{-\epsilon}(Y_{t+k}) - f^{-1}((\frac{\pi_{t+k-1}^{\gamma} p_t^*}{p_{t+k}})^{-\epsilon}(Y_{t+k}))(\frac{w_{t+k}}{p_{t+k}})] = 0$$

Note: Using the labor supply condition and the production function of the firms, the real wage can be expressed in terms of λ_t.

$L_t^{\psi} = (\frac{\lambda_t}{P_t})w_t$

$w_t = MPL = \dfrac{\partial(N_t^{1-\alpha})}{\partial N_t}w_t = c_2\lambda_t.$

where, $c_2 \equiv \frac{\psi}{\alpha+\psi}$.

After Log-Linearizing around a 0 Inflation S.S, the FOC becomes :

$p_t^* = \frac{1-\beta\theta}{1+\psi\epsilon}[-c_2\lambda_t - \gamma\epsilon\psi\pi_{t-1} + \psi y_{t+k} + \Sigma(\beta\theta)^k(-c_2\lambda_{t+k} - \gamma\epsilon\psi\pi_{t+k-1} - \psi y_{t+k})] + (1-\beta\theta)\Sigma p_{t+k}(\beta\theta)^k.(*)$

Using the formula for aggregate price as a function of the optimal price by optimizing firms and the previous price:[21] $P_t = [(p_t^*)^{-\epsilon}(1-\theta) + (P_{t-1}\pi_{t-1}^{\gamma})^{1-\epsilon}\theta]$

we get the Phillips curve:

$\pi_t = \eta_1 x_t + \eta_2 x_{t-1} + \eta_3 E_t[x_{t+1}] + \eta_4 E_t[\pi_{t+1}] + \eta_5\pi_{t-1} + u_t$[22]

$\eta_1 \equiv [kc_2c_1 + kh_1^2c_2 - k\psi]$

$\eta_2 \equiv [-khc_2c_1]$

$\eta_3 \equiv [-kh\beta c_1c_2]$

$\eta_4 \equiv \frac{\beta}{1+\gamma\beta\theta}$

$\eta_5 \equiv [\frac{\gamma}{1+\gamma\beta\theta} - \gamma\epsilon\psi k]$

$c_1 \equiv (\frac{-\sigma}{1-h\beta})(\frac{1}{1-h})$

Using flexible prices, we know we are able to express λ_{t+k} in terms of x_t.

Under flexible prices, we know $\theta = 0$ and $P_{t+k} = p_t^*$ Which yields: $c_2\lambda_{t+k}^f = \psi y_{t+k}^f$.

Using the log linearized FOC, we can express (*) in terms of x_t by using the fact that: $\lambda_t = \lambda_t^x - \lambda_t^f$. Where superscript f denotes the flexible price equilibrium output, and superscript x denotes the terms in λ_t as x_t.

[21] for ease of notation, all of the variables in the following lines are "hatted" $\widehat{a_t}$, and are their respective log linearized values.

[22] here, we add cost push shocks to the Phillips curve

Appendix C. E-stability Condition

The Equations in the Hybrid NK model(4c) are considered the ALM(actual law of motion) while the PLM(percieved law of motion) which agents use is:

$Z_t = \Lambda_{t,0} + \Lambda_{t,1} Z_{t-1} + \Lambda_{t,2} U_t$

$E_t[Z_{t+1}] = \Lambda_{t,0} + \Lambda_{t,0}\Lambda_{t,1} + \Lambda_{t,1}^2 Z_{t-1} + (\Lambda_{t,1}\Lambda_{t,2} + \Lambda_{t,2}R)U_t$

$E_t[Z_{t+2}] = (\Lambda_{t,0} + \Lambda_{t,0}\Lambda_{t,1} + \Lambda_{t,1}\Lambda_{t,0}\Lambda_{t,1}) + \Lambda_{t,1}^3 Z_{t-1} + [\Lambda_{t,1}(\Lambda_{t,1}\Lambda_{t,2} + \Lambda_{t,2}R) + \Lambda_{2,t}RR]U_t$
plugging the expectation terms into the ALM from (4c) gives the result:

$Z_t = \Omega_0 + \Omega_1 Z_{t-1} + \Omega_2 U_t + E\nu_t$

$\Omega_0 \equiv (CV_0^1 + DV_0^2)$ $\Omega_1 \equiv (A + CV_1^1 + DV_1^2)$ $\Omega_2 \equiv (B + CV_2^1 + DV_2^2)$.

where $V_1^1 \equiv \Lambda_1^2$, $V_1^2 \equiv \Lambda_1^3$, $V_0^1 \equiv \Lambda_0 + \Lambda_0\Lambda_1$, $V_0^2 \equiv \Lambda_0 + \Lambda_1\Lambda_0 + \Lambda_1\Lambda_0\Lambda_1$

and $V_2^1 = \Lambda_1\Lambda_2 + \Lambda_2 R$, $V_2^2 = \Lambda_1[\Lambda_1\Lambda_2 + \Lambda_2 R] + \Lambda_2 R^2$

Hence , we can express the Mapping between the ALM and the PLM as:

$G(\theta) \equiv (\Omega_0, \Omega_1, \Omega_2) - (\Lambda_0, \Lambda_1, \Lambda_2)$[23]

$h_0(\theta) \equiv \Omega_0$

$h_1(\theta) \equiv \Omega_1$

$h_2(\theta) \equiv \Omega_2$

Conditions for E-stability can be broken down by taking advantage of the fact that

Ω_1 is a function of only Λ_1 :

$$\frac{\partial h_1}{\partial \Lambda_1^*} = (I \otimes C\Lambda_1^*) + (\Lambda_1^{*\prime} \otimes C) + (\Lambda^{*2\prime}_{\ 1} \otimes D) + (\Lambda_1^{*\prime} \otimes D\Lambda_1^*) + (I \otimes D\Lambda^{*2}_{\ 1})$$

If the above Matrix given by the Gradient only contains Real Eigenvalues less than 1, then we can say that $\Lambda_1 \xrightarrow{inf} \Lambda_1^*$.Assuming this is satisfied, the next condition we must have is:

$$\frac{\partial h_0}{\partial \Lambda_0^*} = (I \otimes C) + (\Lambda_1^* \otimes C) + (I \otimes D) + (I \otimes D\Lambda_1^*) + (\Lambda_1^{*\prime} \otimes D\Lambda_1^*)$$

Similarly, If the real parts of the eigenvalues of the above equation are less than unity, we can say: $\Lambda_0 \xrightarrow{inf} \Lambda_0^*$. The Last set of parameters in question are now Λ_2 :

[23] $E[Q(\theta, Z_t, t)]$ is equivalent to the given condition of $h(\theta)$

$$\frac{\partial h_2}{\partial \Lambda_2^*} = (I \otimes C\Lambda_1^*) + (R' \otimes C) + (I \otimes D{\Lambda^*}_1^2) + (R' \otimes D\Lambda_1^*) + (R^{2\prime} \otimes D)$$

If the real parts of the eigenvalues of the above equation are less than unity, we can then say that the set of parameters satisfy the E-stability criterion.