Level-k DSGE and Monetary Policy

Zhesheng Qiu *

Friday 31st May, 2019

Abstract

This paper develops a new framework of level-k DSGE for monetary policy analysis. Incomplete markets are introduced to guarantee the eductive stability of the equilibrium. k=1.334 is estimated using growth and inflation forecasts from the Michigan Survey of Consumers, capturing the missing indirect channels and the weakened direct channels in households' forecast rules, as well as the wedge between forecasts and realizations. The model produces inflation inertia under Taylor Rule. In pre-Volcker era, more active GDP targeting generates more output mean reversion both in forecasts and in realizations. In Great Recession, the model can explain the missing drop of both inflation and inflation expectations, as well as the stagnant recovery expectations that leads to slow recovery. The model also implies both dampening and accumulation effects of forward guidance. When $k \to +\infty$, the level-k DSGE reduces to a basic three equation New Keynesian DSGE model as in Galí (2015).

*Department of Economics and Finance, City University of Hong Kong, Email: zheshqiu@cityu.edu.hk.
For useful comments, I thank Jośe-Víctor Ríos-Rull, Dirk Krueger, Harold Cole, Enrique G. Mendoza, Jesús Fernández-Villaverde, Joseph Stiglitz, Eric Young, Jian Wang, Yong Wang, Xiang Fang, and seminar participants at University of Pennsylvania, City University of Hong Kong, Chinese University of Hong Kong (shenzhen), Midwest Macroeconomic Meetings, Royal Economic Society Annual Conference and Asian Meeting of Econometric Society (scheduled). A new draft with substantial revision is coming soon.

1 Introduction

The prevalent DSGE models for monetary policy analysis[1] usually impose two assumptions on expectations. First, agents know the aggregate states. Second, agents know the aggregate law of motion. Although these assumptions are often rejected by data, there is less consensus on what alternatives[2] we should make, and in what circumstances it is necessary. In this paper, I relax the second assumption, and develop a new framework of level-k DSGE, based on the idea proposed by Farhi and Werning (2017). The essence of level-k is to turn off the subtle general equilibrium effects in expectations that arise from more than $k-1$ layers of feedbacks. This framework is appealing because it has straight forward setup, transparent mechanisms, sharp empirical support, and reasonable performance under multiple monetary policies.

My first contribution is to lay the foundation for level-k DSGE models. The standard setup of level-k in games (Crawford, Costa-Gomes, and Iriberri, 2013) that level-k players best reply to level-(k-1) is no long sufficient in a DSGE environment for two reasons. First, the ex post budget balance requires agents to observe the prices when making decisions, so that a temporal equilibrium (Grandmont, 1977) structure needs to be imposed as in Farhi and Werning (2017). Second, there can potentially be endogenous state variables[3]. As a result, states determine expectations, expectations drive decisions and decisions affect states. This loop needs to be addressed using a recursive structure. Perceiving all others as one level below is formalized as taking the actual equilibrium objects one level below as the perceived equilibrium objects for decision making. All forecasts are made based on rules as functions of the aggregate states. In addition, the model also allows for non-integer levels by assuming a level-1.3 agent perceiving 30% of the others as level-1 and the rest as level-0[4]. Therefore, the level-k DSGE I propose is more general than those in García-Schmidt and Woodford (2016); Farhi and Werning (2017);

[1] Galí (2015) provides the benchmark for small-scale DSGE, while Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007) provie that for medium-scale DSGE.

[2] Limited attention relaxes the first assumption (Mankiw and Reis, 2002; Woodford, 2003; Angeletos and La'O, 2013; Gabaix, 2014; Afrouzi, 2017), while least-squared learning relaxes the second (Marcet and Sargent, 1989; Evans and Honkapohja, 1998; Milani, 2007; Eusepi and Preston, 2011).

[3] For instance, the interest rate responses to inflation and output gaps is inertial, and level-0 is specified to anchor their decisions to the last period.

[4] 30% percentage of the others are level-1 is different from for 30% probability that all others are level-1, due to recursive structure of the model.

Iovino and Sergeyev (2017).

Another issue is "Eductive Stability". The recursive level-k equilibrium is well defined only if it converges to the recursive competitive equilibrium when $k \to +\infty$. This property does not hold in real business cycles (Evans, Guesnerie, and McGough, 2017), because households respond too aggressively to interest rate expectations. As a result, the interest rate implied by households' decisions would exceed the initial range of interest rate expectations. In my model, I introduce incomplete markets to make the planning horizons of households shorter so that they would respond less aggressively. The extent of market incompleteness is disciplined by the transition probability from non hand-to-mouth consumers to hand-to-mouth consumers in Kaplan and Violante (2014).

The equilibrium conditions can no longer be formulated as intertemporal conditions between current period and next period as in recursive competitive equilibrium. Yet, once solved, the model still has a state space representation. Multiple steps ahead forecasts can be obtained only by iterating on aggregate states using the perceived aggregate law of motion. This feature resembles "Infinite-horizion Learning" as in Eusepi and Preston (2011). However, due to the incomplete markets, the model can have a simple purely forward looking representation only in a special case in which the steady state real interest rate is zero. All theoretical results in this paper are obtained under this condition. In the appendix, a more general algorithm is provided for the computation of the full model.

The specification of level-0 and timing is also worth mentioning. Level-0 households and firms fully anchor their spending the pricing decisions to the last period. This is consistent with the principle in level-k games that level-0 agents should be as "dumb" as possible. In order to circumvent multiple equilibria generated by the simultaneity between decisions making and expectation formation. I specify the timing such that agents do not update their expectations until they finish making decisions at the end each period, even if they have previously observed some new information before. This timing arrangement is neutral under rational expectations, because no observations are informative to the agents if they have already anticipated them using the correct aggregate law of motion. This setup implicitly assumes away other forms of learning, and isolates the eductive learning.

My second contribution is to unravel the essence of level-k DSGE. Level-k DSGE is not the only model to allow for a wedge between the perceived aggregate law of motion and the actual aggregate law of motion. Gabaix (2017) also explicitly specifies a pair of perceived and actual objects. Level-k DSGE is sharper in the sense that it predicts in which dimension the wedge is larger. The model implies that in one-step ahead forecasts, when $k \in [1, 2]$, indirect channels are missing and direct channels are weaker. In multiple-step ahead forecasts, this result will still hold approximately if k is close to 1.

Let's use simple notations to demonstrate these channels. Denote the real GDP as Y, nominal interest rate as R, and inflation as Π. Then, $R \to Y$ and $Y \to \Pi$ are direct channels that are at least partially understandable by agents with $k > 1$, while $Y \to \Pi \to Y$ and $R \to Y \to \Pi$ are indirect ones, and partially understandable only by agents with $k > 2$.

These results only exist in one-step ahead forecasts, because in multiple-step ahead forecasts, the perceived law of motion starts to play a role. Take $Y \to \Pi \to Y$ for example. Households do not understand the effect of inflation expectation of the others, but as their own inflation expectations go up, they anticipate interest rate movement, and hence understand $Y \to \Pi \to Y$ through $R \to Y$. Quantitatively, this channel is much weaker than the previous one.

Another interesting feature of level-k is that the forecasts of the forecasts of others are not identical to the direct forecasts on the same objects. As the forecast horizon becomes longer, this difference becomes larger. It implies that the infinite horizon assumption is not innocuous. This feature is not unique in level-k. Expectations modeled as forecast rules anchored to the past is likely to be incompatible to the common knowledge of individual rationality which can be used to achieve the irrelevance of planning horizons. Angeletos and Lian (2018) and Farhi and Werning (2017) have also mentioned the importance of planning horizons for firms and households respectively, while my paper uncovers the essence of it more generally.

My third contribution is to provide sharp disciplines for level-k DSGE. Despite various empirical works in level-k games to identify the parameter k (Camerer, Ho, and Chong, 2004), there are no empirical counterparts in macroeconomics. There may be two reasons for this. First, it seems that models in which households keep on learning from the past is more plausible for business cycle related issues in normal times. Second, it not clear how to identify k.

I argue that level-k is still relevant given historical data for learning, as expectations data show that households' forecast rules are systematically biased. There are two possible reasons why learning does not make households more rational as players are in dynamic experiments. First, the payoffs of decisions are much less clear along business cycles. Second, recalling and analyzing historical data for business cycles is much more costly.

I show that k can be identified by exploiting the co-movements between macroeconomic data and forecast data. Hence, the DSGE structure actually helps identify level-k by providing dynamics of multiple macroeconomic variables. The missing co-movements for indirect channels identify $k \in [1, 2]$, while the weakened direct channels help identify the exact value of k.

In the estimation, I adopt Bayesian approach, and use five time series including quarterly GDP growth rate, quarterly inflation, quarterly federal fund rate, one-year ahead growth forecasts, and one-ahead inflation forecasts. k is jointly estimated with other parameters, and also fits the macroeconomics data. The prior of k is set to be very dispersed and cover the interval of $[1, +\infty)$. Yet, the posterior has a tight 95% confidence interval $[1.212, 1.467]$.

My forth contribution is to evaluate the performance of the model under different monetary policies. The rational expectations approach was introduced into macroeconomic modeling to overcome "The Lucas Critique". However, agents' endogenous responses to policies does not have to be captured by rational expectations. Level-k DSGE provides an alternative way to model expectations. It can be a useful tool if it performs well under different policy regimes.

Under inflation targeting Taylor Rule during the Great Moderation, the model can produce impulse responses of output, inflation and interest rate to monetary shocks aligned with the data. In particular, the model captures both output and inflation inertia without relying on other frictions. A rational expectations model can have the identical impulse responses if it has external habit, working capital loans, and very sticky marginal cost of production. Still, such a model cannot get even close to the dynamics of forecasts. Level-k DSGE can produce inertia responses because level-1 expectations are anchored, and the anchoring is strong if the value of k is not large. Current inflation is determined by the discounted sum of output forecasts, which is again anchored to the past.

Under Taylor Rule in the pre-Volcker era with more active output targeting, the model can generate output forecasts more aligned with its backcasts. This result comes from the property of the model that inflation targeting stabilizes the economy only by stabilizing households' expenditures, while output targeting goes though households' expectations additionally.

In the Great Recession when interest rate is trapped at zero, the model can explain the missing deflation, stagnant recovery expectations, and slow recovery in data. The drop of inflation realizations and inflation expectations is small in long-lasting recessions because inflation in level-k model does not explode under permanent output gaps, and inflation expectations only captures part of the movement in inflation. The recovery expectations is weak because long-lasting stimulus of low interest rate only has small accumulated effects when households planning horizon is short. The recovery is weak because the expected recovery of output in far future does not help the recovery of current output when planning horizon is short.

The effect of forward guidance is also unique in quantitative results. Unlike the theoretical work of McKay, Nakamura, and Steinsson (2016); Angeletos and Lian (2018); Farhi and Werning (2017) which show how forward guidance is dampened, it is accumulative in empirically relevant level-k models, and can ultimately have a sizable aggregate effect, if the monetary authority has full commitment power in forward guidance.

The rest of the paper is organized as the following. Section 2 documents a set of facts from Michigan Survey of Consumers that are hardly captured by other existing models but consistent with level-k model. Section 3 develops the level-k DSGE model. Section 4 highlights the special role of planning horizon. Section 5 connects the model to data with special emphasis on the identification of parameter k using consumer forecasts. Section 6 evaluates the model performance under different policy regimes. Section 7 concludes.

2 The Missing Forecasts

This section demonstrates a unique set of stylized facts on the consumer forecasts of business condition changes and inflation rate that can hardly be rationalized without a level-k model. The data of interest is from the Michigan Survey of Consumers. I explore the disconnection between forecasts and backcasts of business condition changes to falsify a few theories other than level-k. I also explore forecast errors conditional on federal fund rate to provide supported evidence for the level-k model.

2.1 Facts against Other Theories

Large systematic forecast errors in business condition changes during 1985-2007 can be found in Figure 1. The one year backcasts comove almost perfectly with the ex post output growth rates, indicating that households can observe the output growth rates at least in aggregate. In contrast, the one year forecasts comove very little with either the backcasts or the ex post output growth rates. The discrepancy is large and very persistent, rejecting full information rational expectations.

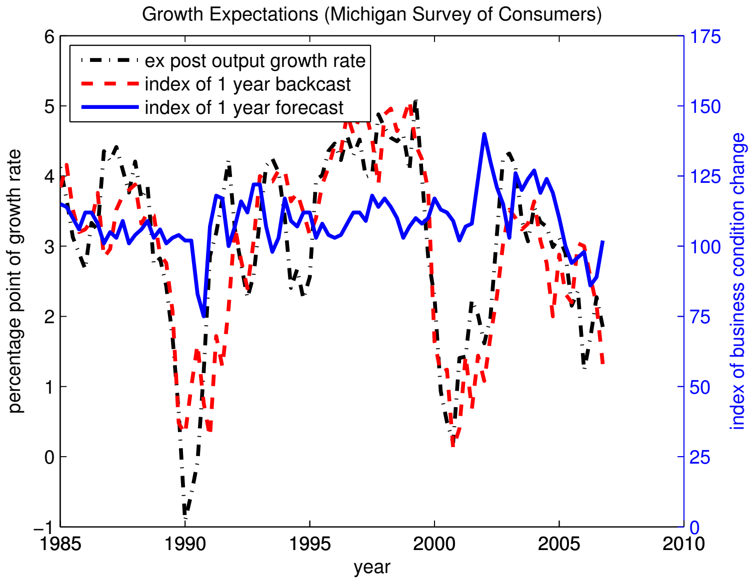

Figure 1: Forecast and Backcast Dynamics

Figure 2 shows that information friction as in Angeletos and La'O (2013) can hardly capture these forecast errors. According to the theory, agents understand the aggregate law of motion, but are confused about the aggregate state. However, Figure 2 indicates that the within sample predicted business condition changes from the autoregression of backcasts is much more precise than the corresponding forecasts. This indicates that consumers are not using the aggregate law of motion of backcasts to make forecasts.

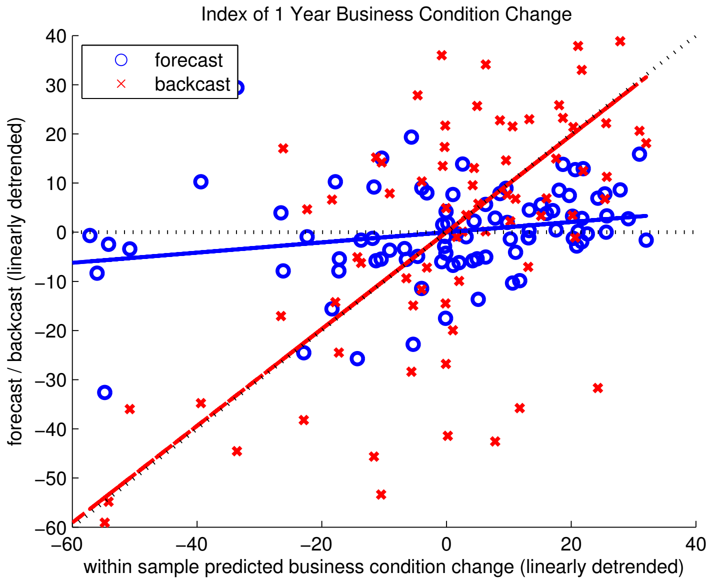

Figure 2: Within Sample Prediction

Figure 3 shows that imperfect knowledge as in Eusepi and Preston (2011) can hardly capture these forecast errors either. According to the theory, agents do not understand the aggregate law of motion, but can gradually learn it from historical data. In order to test this assumption, I replace the within sample prediction in Figure 2 with out of sample prediction in Figure 3. The slop of fitted backcasts is still close to the 45 degree line, while that of the fitted forecasts is still nearly horizontal. This indicates that the aggregate historical data of backcasts is not what consumers can easily learn from. It is also possible that consumers' individual backcasts are subject to large noise (measure errors), so that they greatly underestimate the predictive power of historical backcast data. This possibility still implies that consumers learn very little from historical macroeconomic data.

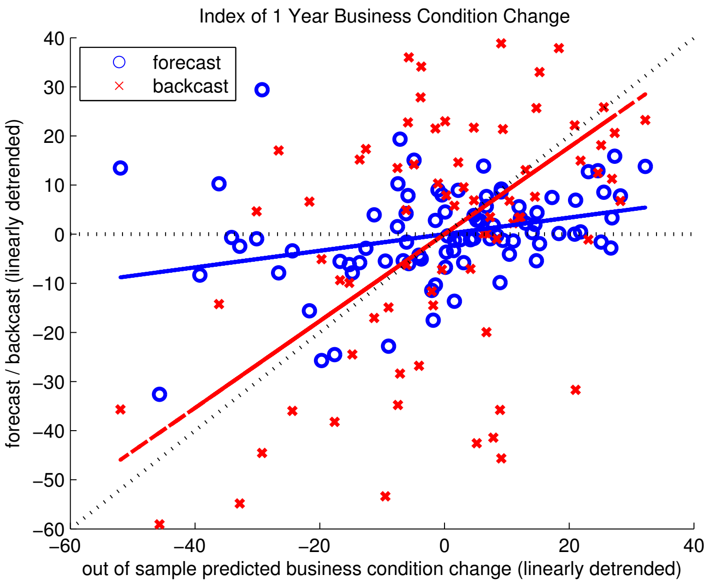

Figure 3: Out of Sample Prediction

2.2 Facts Supporting Level-k Model

The lack of variation in forecasts conditional on backcasts as in Figure 2 and 3 is not conflicting with a level-k model. In the model, level-k agents only understand the comovements between macroeconomic variables that result from no more than k-1 layers of feedback effects. When k is no larger than 2, it will be difficult for consumers to understand how todays' output level feeds back into output growth in the future through any self-stabilizing mechanisms.

In addition to that, level-k model also predicts that inflation rate rise (fall) following interest rate fall (rise) is more difficult to understand than business condition change index rise (fall) following interest rate fall (rise). Figure 4 and 5 confirm these model properties. The slops of one year ahead business condition change to federal fund rate in forecasts and backcasts are very close, while the slops of one year ahead inflation rate to ex ante real interest rate (federal fund rate subtracting inflation expectations) in forecasts and in reality have opposite signs. VAR impulse responses functions are not used here because the main focus of these facts are the biased perceived law of motion instead of the propagation of shocks.

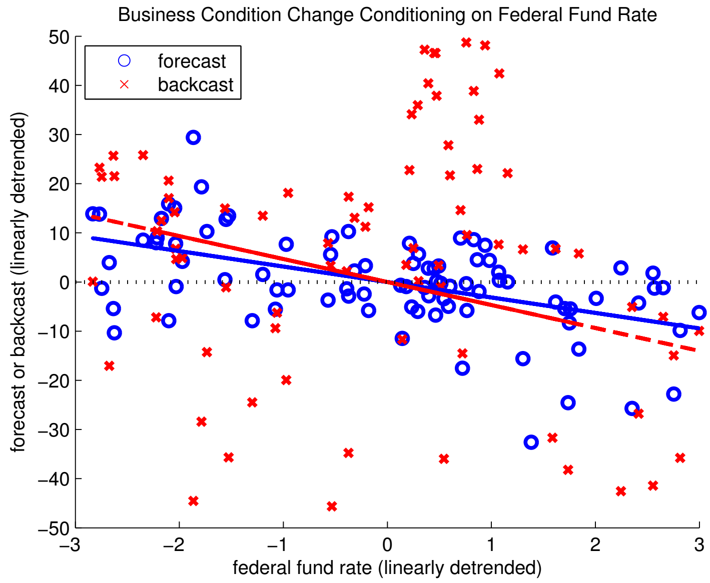

Figure 4: Interest Rate and 1 Year Ahead Business Condition Change

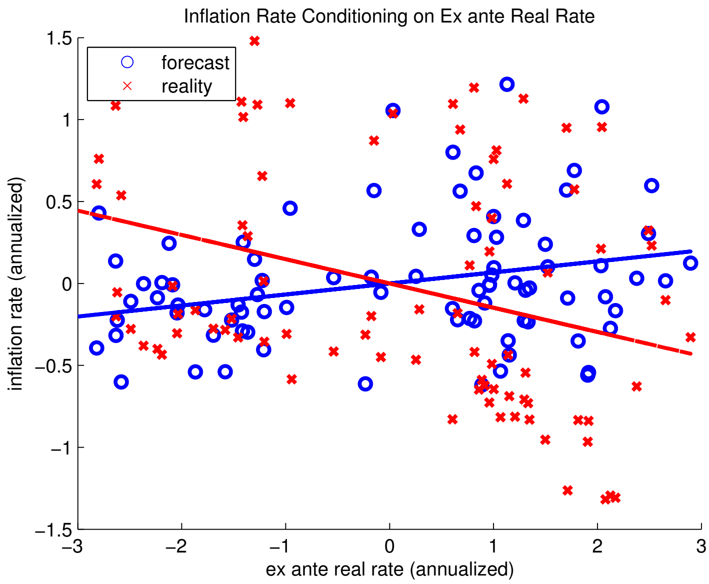

Figure 5: Interest Rate and 1 Year Ahead Inflation Rate

3 Recursive Level-k Equilibrium

This section demonstrates how to put level-k friction into a basic New Keynesian DSGE model as in Galí (2015). The log-linearized form of this basic model is

$$\begin{aligned}
\hat{y}_t =& -\omega(\mathbb{E}_{t-1}\hat{r}_t - \mathbb{E}_{t-1}\hat{\pi}_{t+1} + \eta_t^d) + \mathbb{E}_{t-1}\hat{y}_{t+1}, \\
\hat{\pi}_t =& \frac{(1-\theta)^2}{\theta}(\kappa_y \hat{y}_t - \kappa_z \eta_t^z) + \mathbb{E}_{t-1}\hat{\pi}_{t+1}, \\
\hat{r}_t =& \rho_r \hat{r}_{t-1} + (1-\rho_r)(\phi_\pi \hat{\pi}_t + \phi_y \hat{y}_t) + \sigma_r \epsilon_t^r, \\
\eta_{t+1}^d =& \rho_d \eta_t^d + \sigma_d \epsilon_t^d, \\
\eta_{t+1}^z =& \rho_z \eta_t^z + \sigma_z \epsilon_t^z.
\end{aligned}$$

The first equation represents the IS curve that connects the current output $\hat{y}_t$ to the expected nominal interest $\mathbb{E}_{t-1}\hat{r}_t$, expected next period inflation rate $\mathbb{E}_{t-1}\hat{\pi}_{t+1}$, expected next period output $\mathbb{E}_{t-1}\hat{y}_{t+1}$, and demand wedge η_t^d. The second equation represents the Phillips Curve that connects the current inflation rate $\hat{\pi}_t$ to the current output, next period inflation expectation, and efficiency wedge η_t^z. The third equation represents the Taylor Rule that describes the law of motion for nominal interest rate $\hat{r}_t$. Both the demand wedge η_t^d and the efficiency wedge η_t^z follow an AR(1) process. The timing of shocks satisfies

$$\mathbb{E}_{t-1}(\epsilon_t^r, \epsilon_t^d, \epsilon_t^z) = (0, 0, 0).$$

This timing assumption resembles Christiano, Eichenbaum, and Trabandt (2016) in which the current period nominal interest rate is known only after all decisions have been made.

In order to introduce level-k into this model, we need to introduce recursive level-k equilibrium as a solution concept to deal with endogenous state variables in expectation formation, assume that expectations are updated only after decisions are made to avoid the feedback effects from current decisions to expectations on future, and impose market incompleteness to guarantee "eductive stability". For quantitative purposes, the space of k is expanded from $\mathbb{N}$ to $\{0\} \cap [1, +\infty)$, and the labor input margin in production is replaced by material input margin. With all these modifications, the level-k DSGE model still collapses to the basic model as $k \to +\infty$.

3.1 Households

There are a measure one of infinitely-lived households with a constant discount factor β, and an aggregate time-varying demand wedge $\exp(\eta^d)$ multiplied to it.

Timing. Within each period, events happen in the following order:
(1) Households inherit observations and expectations from the end of last period.
(2) Households observe the current gross inflation rate Π. The real net wealth a is determined by the last period real bond position b_-, the last period nominal gross interest rate R_-, and the current gross inflation rate Π through $a = b_- R_- / \Pi$.
(3) Households are hit by idiosyncratic preference shocks $\zeta \in \{1, \overline{\zeta}\}$, with $\overline{\zeta} \geq 1$ and transition probability $\Pr(\zeta|\zeta_-) = \lambda_{\zeta|\zeta_-}$. This yields an unconditional probability $\Pr(\zeta) = \lambda_\zeta$.
(4) Households observe the real wage rate W and receive a lump sum transfer of real dividend D. They consume c and supply labor $\ell \leq 1$ to get utility $\zeta u(c)$. The rest of budget is saved in real bond $b = -c + W + D + a$, with borrowing constraint $b \geq 0$ only for $\zeta = \overline{\zeta}$.
(5) Households observe the aggregate output Y as well as all aggregate shocks $\epsilon = (\epsilon^r, \epsilon^d, \epsilon^z)$ that drives the current gross nominal interest rate R, and the next period wedges $(\eta^{d\prime}, \eta^{z\prime})$.
(6) Expectations are updated via level-k reasoning and reported to surveys[5].

States and Equilibrium Objects. Denote the vector of aggregate states as S. Individual states include the real net wealth a and idiosyncratic preference shocks ζ. Level-k households take as given the following equilibrium objects:
(1) perceived and actual real wage rate $\{W^{e,(k)}(S), W^{(k)}(S)\}$,
(2) perceived and actual real dividend $\{D^{e,(k)}(S), D^{(k)}(S)\}$,
(3) perceived gross nominal interest rate $R^{e,(k)}(S, \epsilon^r)$,
(4) perceived gross inflation rate $\Pi^{e,(k)}(S)$,
(5) perceived aggregate law of motion $H^{e,(k)}(S, \epsilon)$.

[5] I assume that households do not update expectations until the end of each period using the newly arrived information of $\{\Pi, W, D\}$ to aviod a twoway feedback between current equilibrium outcomes and expectations on future. As a result, expectations are inferred exclusively from the aggregate states. Under full information rational expectations, expectations inferred from the aggregate states are consistent with equlibrium outcomes, hence $\{\Pi, W, D\}$ provide no additional information. In my specification of level-k reasoning, there are ex post forecast errors in $\{\Pi, W, D\}$ but households do not learn from them.

Households' Problems. Households have perceptions on their equilibrium policy and value functions $\{c^{e,(k)}, b^{e,(k)}, V^{h,e,(k)}\}$ on (ζ, a, S) for the future. These functions solve

$$
\begin{aligned}
V^{h,e,(k)}(\zeta, a, S) = & \max_{\{c,b\}} \{\zeta u(c) + \beta \exp(\eta^d) \cdot \mathbb{E}[V^{h,e,(k)}(\zeta', a', S')|(\zeta, a, S)]\} \\
& s.t.\ b = -c + W^{e,(k)}(S) + D^{e,(k)}(S) + a, \\
& \quad b \geqslant 0 \text{ when } \zeta = \overline{\zeta}, \\
& \quad a' = b \cdot R^{e,(k)}(S, \epsilon^r)/\Pi^{e,(k)}(S'), \\
& \quad S' = H^{e,(k)}(S, \epsilon).
\end{aligned}
$$

In the current period, households observe the actual real wage and dividend $\{W^{(k)}(S), D^{(k)}(S)\}$, and have perceived continuation value function $V^{h,e,(k)}$. The actual equilibrium policy and value functions $\{c^{(k)}, b^{(k)}, V^{h,(k)}\}$ on (ζ, a, S) solve

$$
\begin{aligned}
V^{h,(k)}(\zeta, a, S) = & \max_{\{c,b\}} \{\zeta u(c) + \beta \exp(\eta^d) \cdot \mathbb{E}[V^{h,e,(k)}(\zeta', a', S')|(\zeta, a, S)]\} \\
& s.t.\ b = -c + W^{(k)}(S) + D^{(k)}(S) + a, \\
& \quad b \geqslant 0 \text{ when } \zeta = \overline{\zeta}, \\
& \quad a' = b \cdot R^{e,(k)}(S, \epsilon^r)/\Pi^{e,(k)}(S'), \\
& \quad S' = H^{e,(k)}(S, \epsilon).
\end{aligned}
$$

Labor supply does not show up in this optimization problem because it does not induce any disutility and will always have conner solution $\ell = 1$.

Aggregation. With the assumption that bond b is in zero net supply, and $b \geq 0$ binds only when $\zeta = \overline{\zeta}$, an initially degenerate wealth distribution will always induce equilibrium with degenerate wealth distribution. Start with a degenerate wealth distribution with $a = 0$ for all households. Those with $\zeta = \overline{\zeta}$ would like to consume more, but are borrowing constrained. For others with $\zeta = 1$, the equilibrium wage and dividend[6] will clear the goods market such that they would like to choose $b = 0$. This can be formalized in the following.

Assumption 1. $u(\cdot)$ *is twice continuously differentiable, increasing and strictly concave.*

[6] Interest rate will not help clear any market because by the timing specification, it is not known ex ante.

Assumption 2. $B^{(k)}(S) = \lambda_1 b^{(k)}(1,0,S) + (1-\lambda_1) b^{(k)}(\overline{\zeta},0,S) = 0.$

Proposition 1. *Starting from a degenerate wealth distribution in which $a = 0$ for all households, the aggregate equilibrium objects $\{B^{(k)}, C^{(k)}\}$ satisfy*

$$
\begin{aligned}
B^{(k)}(S) =& b^{(k)}(1,0,S) = b^{(k)}(\overline{\zeta},0,S) = 0, \\
C^{(k)}(S) =& c^{(k)}(1,0,S) = c^{(k)}(\overline{\zeta},0,S) = W^{(k)}(S) + D^{(k)}(S).
\end{aligned}
$$

3.2 Firms

Production Technology. There are a measure one of infinitely-lived firms. Firm $j \in [0,1]$ uses material inputs m_j and labor n_j to produce variety

$$
x_j = \exp(\eta^z) m_j^{\alpha} n_j^{1-\alpha},
$$

where $\exp(\eta^z)$ denotes the efficiency wedge. X is the gross final goods through Dixit-Stiglitz aggregator, and we can also have the variety demand and price aggregator in the following.

$$
X = \left(\int_0^1 x_j^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}}, \qquad x_j = \left(\frac{p_j}{P} \right)^{-\varepsilon} X, \qquad P = \left(\int_0^1 p_j^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}.
$$

With $M = \int_0^1 m_j dj$, the net aggregate supply is

$$
Y = X - M.
$$

Timing. Within each period, events happen in the following order:
(1) A random fraction θ of the firms are drawn to keep their previous prices unchanged.
(2) Each other Firm j sets price p_j before observing $\{\Pi, W, x_j\}$.
(3) Each Firm $j \in [0,1]$ produces with cost minimization after observing $\{W, x_j\}$.
(4) Real wage W and dividend D are paid to households evenly.
(5) Firms observe aggregate output Y and all aggregate shocks $\epsilon = (\epsilon^r, \epsilon^d, \epsilon^z)$.
(6) Expectations are updated via level-k reasoning and reported to surveys.

Cost Minimization. Given real wage W, the cost minimization problem of Firm j yields a material input m_j proportional to labor input n_j

$$m_j = \frac{\alpha}{1-\alpha} W n_j.$$

Given $N = \int_0^1 n_j dj = 1$ and $M = \int_0^1 m_j dj$, we have

$$W = \frac{1-\alpha}{\alpha} M = \frac{1-\alpha}{\alpha} \int_0^1 \left(\frac{x_j}{\exp(\eta^z)} \right)^{\frac{1}{\alpha}} dj.$$

The real cost of producing one unit of x_j is

$$Z = \frac{W^{1-\alpha}}{\alpha^\alpha (1-\alpha)^{1-\alpha} \exp(\eta^z)} \equiv \frac{W^{1-\alpha}}{\alpha_z \exp(\eta^z)}.$$

The profit of firm j becomes

$$D_j = \left(\frac{p_j}{P} - Z \right) \left(\frac{p_j}{P} \right)^{-\epsilon} X.$$

States and Equilibrium Objects. The firms that reset prices choose $p^a = p/P_-$ as the new price over the previous aggregate price. Other firms have individual state $p_-^n = p_-/P_-$.

The Dixit-Stiglitz aggregator and cost minimization problem in production are both static, and hence do not induce any inconsistency between equilibrium objects and perceived objects. In the equilibrium, we have individual objects $\{x^{(k)}, m^{(k)}, n^{(k)}\}$ for firms as functions of (p_-^n, S). The aggregate counterpart is $\{X^{(k)}, M^{(k)}, N^{(k)}\}$.

In the price setting stage, level-k firms take as given the following equilibrium objects:
(1) perceived gross inflation rate $\Pi^{e,(k)}(S)$,
(2) perceived real wage rate $W^{e,(k)}(S)$,
(3) perceived aggregate output $Y^{e,(k)}(S)$,
(4) perceived nominal gross interest rate $R^{e,(k)}(S, \epsilon^r)$,
(5) perceived aggregate law of motion $H^{e,(k)}(S, \epsilon)$.

Firms' Problems. Functions $\{p^{a,(k)}, V^{a,e,(k)}, V^{n,e,(k)}(p^n_-, \cdot)\}$ on S solve

$$
\begin{aligned}
V^{a,e,(k)}(S) = \max_{p^a} & \underbrace{\left(\frac{p^a}{\Pi^{e,(k)}(S)} - \frac{W^{e,(k)}(S)^{1-\alpha}}{\alpha_z \exp(\eta^z)}\right)}_{\text{relative price } - \text{ marginal cost}} \underbrace{\left(\frac{p^a}{\Pi^{e,(k)}(S)}\right)^{-\varepsilon} X^{e,(k)}(S)}_{\text{variety demand}} \\
& + \mathbb{E}\left[\frac{\Pi^{e,(k)}(S')}{R^{e,(k)}(S,\epsilon^r)}\left(\theta V^{n,e,(k)}\left(\frac{p^a}{\Pi^{e,(k)}(S)}, S'\right) + (1-\theta)V^{a,e,(k)}(S')\right)\Bigg| S\right], \\
V^{n,e,(k)}(p^n_-, S) = & \left(\frac{p^n_-}{\Pi^{e,(k)}(S)} - \frac{W^{e,(k)}(S)^{1-\alpha}}{\alpha_z \exp(\eta^z)}\right)\left(\frac{p^n_-}{\Pi^{e,(k)}(S)}\right)^{-\varepsilon} X^{e,(k)}(S) \\
& + \mathbb{E}\left[\frac{\Pi^{e,(k)}(S')}{R^{e,(k)}(S,\epsilon^r)}\left(\theta V^{n,e,(k)}\left(\frac{p^n_-}{\Pi^{e,(k)}(S)}, S'\right) + (1-\theta)V^{a,e,(k)}(S')\right)\Bigg| (p^n_-, S)\right]. \\
s.t.\ S' = & H^{e,(k)}(S,\epsilon).
\end{aligned}
$$

Profits are discounted using the real interest rate instead of the stochastic discount factor of households. It does not make much difference up to first order approximation.

Aggregation. The aggregate inflation and dividend are determined by

$$
\begin{aligned}
\Pi^{(k)}(S) &= \left(\theta + (1-\theta)p^{a,(k)}(S)^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}, \\
D^{(k)}(S) &= Y^{(k)}(S) - \frac{W^{(k)}(S)Y^{(k)}(S)}{\exp(\eta^z)}.
\end{aligned}
$$

3.3 Monetary Policy and Market Clearing

Monetary Policy. Denote steady state nominal gross interest rate, gross inflation rate and output as $\{R_{SS}, \Pi_{SS}, Y_{SS}\}$. Assume these are common knowledge to households.

In normal times, the monetary authority follows a conventional Taylor Rule

$$
\frac{R^{(k)}(S,\epsilon^r)}{R_{SS}} = \left(\frac{R_-}{R_{SS}}\right)^{\rho_r}\left(\frac{\Pi^{(k)}(S)}{\Pi_{SS}}\right)^{(1-\rho_r)\phi_\pi}\left(\frac{Y^{(k)}(S)}{Y_{SS}}\right)^{(1-\rho_r)\phi_y}\exp(\sigma_r\epsilon^r),
$$

where ρ_r denotes the level of interest rate smooth, (ϕ_π, ϕ_y) denotes the response coefficients of nominal interest rate to inflation and output, ϵ^r denotes the i.i.d. federal fund rate shock, and σ_r denotes its standard deviation. If not specified, monetary policy follows Taylor Rule. When discussing liquidity trap related issues, monetary policy rule must satisfy $R^{(k)}(S, \epsilon^r) \geq 0$.

The perceived monetary policy rule is

$$\frac{R^{e,(k)}(S, \epsilon^r)}{R_{SS}} = \left(\frac{R_-}{R_{SS}}\right)^{\rho_r} \left(\frac{\Pi^{e,(k)}(S)}{\Pi_{SS}}\right)^{(1-\rho_r)\phi_\pi} \left(\frac{Y^{e,(k)}(S)}{Y_{SS}}\right)^{(1-\rho_r)\phi_y} \exp(\sigma_r \epsilon^r).$$

Market Clearing. In each period when agents are making decisions, demand must be equal to supply in goods market, labor market and bond market.

$$\begin{aligned} C^{(k)}(S) =& Y^{(k)}(S), \\ N^{(k)}(S) =& 1, \\ B^{(k)}(S) =& 0. \end{aligned}$$

3.4 Recursive Level-k Equilibrium

This subsection establishes a Recursive Level-k Equilibrium. Both households and firms are level-k. Perceiving others as one level below is equivalent to using the actual equilibrium objects of this level as perceived equilibrium objects. Therefore, the recursive level-k equilibrium can be established by iterating on the equilibrium objects. The Recursive Level-k Equilibrium nests the definition of level-k equilibrium in Farhi and Werning (2017) as a special case, and allows for a state space representation of the model.

Level-0 Initialization. Level-0 needs to be specify to initialize the iteration on equilibrium objects. It is natural to assume that level-0 agents' expenditure and level-0 firms' price choice are fully anchored to the last period[7].

[7] Fehr and Tyran (2008) and Gill and Prowse (2016) have provided experimental evidence showing that in a dynamic setting, players' decisions are indeed anchored to the past.

Assumption 3. *The level-1 expectations are given by the following statements.*
(1) Level-0 households and firms do not change decisions.

$$(Y^{(0)}(S), \Pi^{(0)}(S)) = (\underbrace{Y_{-}, 1}_{\textit{behavioral type}}).$$

(2) $\{W^{(0)}, D^{(0)}, M^{(0)}, X^{(0)}\}$ *satisfies*

$$\begin{aligned}
W^{(0)}(S) =& \frac{1-\alpha}{\alpha} M^{(0)}(S), \\
D^{(0)}(S) =& Y^{(0)}(S) - W^{(0)}(S), \\
M^{(0)}(S) =& \left(\frac{X^{(0)}(S)}{\exp(\eta^z)}\right)^{\frac{1}{\alpha}} \int_0^1 (p^n_{j,-})^{-\frac{\epsilon}{\alpha}} dj, \\
X^{(0)}(S) =& Y^{(0)}(S) + M^{(0)}(S).
\end{aligned}$$

(3) Given $\{Y^{(0)}, \Pi^{(0)}\}$, $R^{(0)}$ *satisfies Taylor Rule.*
(4) $H^{(0)}$ *is consistent with equilibrium objects.*

Level-k Updating. For $\forall j \geq 1$ and $j \in \mathbb{N}_{+}$, expectations are updated according to

$$(Y^{e,(j+1)}, \Pi^{e,(j+1)}, W^{e,(j+1)}, D^{e,(j+1)}, R^{e,(j+1)}, H^{e,(j+1)}) = (Y^{(j)}, \Pi^{(j)}, W^{(j)}, D^{(j)}, R^{(j)}, H^{(j)}).$$

For $\forall k \geq 1$, and $j \leq k \leq j+1$, expectations are defined[8] as

$$(Y^{e,(k)}, \cdots) = (j+1-k)(Y^{e,(j)}, \cdots) + (k-j)(Y^{e,(j+1)}, \cdots).$$

The solution to the temporary equilibrium[9] with given expectations yields the mapping

$$T : (Y^{e,(k)}, \cdots) \longrightarrow (Y^{(k)}, \cdots).$$

This specification of level-k updating expands the space of k from $\mathbb{N}$ to $\{0\} \cap [1, +\infty)$ for empirical purposes.

[8]It extends integer k in most level-k models in a way different from García-Schmidt and Woodford (2016), but has more transparent implications in more complex models and when $k \in [1, 2]$.
[9]The equilibrium for each level-k is a temporary equilibrium (Grandmont, 1977).

Aggregate Shocks and States. Aggregate shocks $\epsilon = (\epsilon^r, \epsilon^d, \epsilon^z)$ and wedge (η^d, η^z) satisfy

$$\begin{aligned}
\epsilon^r \sim& \mathcal{N}(0,1), \\
\eta^{d\prime} =& \rho_d \eta^d + \sigma_d \epsilon^d, \quad \epsilon^d \sim \mathcal{N}(0,1), \\
\eta^{z\prime} =& \rho_z \eta^z + \sigma_z \epsilon^z, \quad \epsilon^z \sim \mathcal{N}(0,1).
\end{aligned}$$

In the Taylor Rule monetary regimes, the aggregate state is $S = \{Y_-, R_-, \int_0^1 (p^n_{j,-})^{-\frac{\epsilon}{\alpha}} dj, \eta^d, \eta^z\}$[10], while in other regimes, it may also include time index.

Recursive Level-k Equilibrium.

Definition 1. *The Recursive Level-k Equilibrium consists of a set of*
(1) $\{c^{(k)}, b^{(k)}, V^{h,(k)}\}$ *and* $\{c^{e,(k)}, b^{e,(k)}, V^{h,e,(k)}\}$ *on* (ζ, a, S) *for households,*
(2) $\{x^{(k)}, m^{(k)}, n^{(k)}, V^{n,e,(k)}\}$ *on* (p^n_-, S)*, and* $\{p^{a,(k)}, V^{a,e,(k)}\}$ *on* S *for firms,*
(3) $\{C^{(k)}, B^{(k)}, X^{(k)}, M^{(k)}, N^{(k)}, Y^{(k)}, \Pi^{(k)}, W^{(k)}, D^{(k)}, R^{(k)}, H^{(k)}(\cdot, \epsilon)\}$ *and*
$\{X^{e,(k)}, Y^{e,(k)}, \Pi^{e,(k)}, W^{e,(k)}, D^{e,(k)}, R^{e,(k)}, H^{e,(k)}(\cdot, \epsilon)\}$ *on* S*, such that*
1. individual policy and value functions solve the corresponding problems,
2. individual policy functions are consistent with the law of motion of individual and aggregate states, as well as the aggregate objects,
3. monetary policy follows Taylor Rule,
4. goods, labor and bond markets clear,
5. perceived aggregate objects are determined by level-k updating.

Definition 2. *Replacing the level-k updating with consistency between the actual and perceived objects yields the Recursive Competitive Equilibrium.*

By definition, when $k \to +\infty$, if the Recursive Level-k Equilibrium converges, it must converge to the Recursive Competitive Equilibrium. The convergence enables us to nest rational expectations as a special case in estimation. This can be viewed as a way to compare level-k with rational expectations in empirical performance.

[10] The distribution of variety prices has real aggregate effects.

3.5 Equilibrium Representation

In models with rational expectations, the equilibrium can usually be represented by a set of non-linear difference equations. However, it is no longer true in level-k DSGE. Still, the model has a state space representation in general, and an infinite horizon representation under perfect foresight or certainty equivalence, but the law of iterated expectations no longer applies.

State Space Representation. Definition 1 naturally allows for a transition equation

$$S_{t+1} = H^{(k)}(S_t, \epsilon_t),$$

and a set of observation equations

$$\begin{aligned}
\ln(Y_{t+1}) &= \ln(Y^{(k)}(S_{t+1})), \\
\ln(\Pi_{t+1}) &= \ln(\Pi^{(k)}(S_{t+1})), \\
\ln(R_t) &= \ln(R^{(k)}(S_{t+1})), \\
\ln(Y^e_{t+4|t}) &= \ln(Y^{e,(k)}(H^{e,(k)})^3(S_{t+1})) + \sigma_y \epsilon^y_t, \\
\ln(\Pi^e_{t+4|t}) &= \ln(\Pi^{e,(k)}(H^{e,(k)})^3(S_{t+1})) + \sigma_p \epsilon^p_t.
\end{aligned}$$

Here, the multiple step ahead forecasts are calculated assuming all shocks are absent. $\{\epsilon^y_t, \epsilon^p_t\}$ denote the measurement errors of forecasts.

Infinite Horizon Representation. When all aggregate shocks are turned off, the s period ahead forecast of the aggregate states becomes[11]

$$S^{e,(k)}_{t+s|t-1} = (H^{e,(k)})^s(S_t).$$

The long horizon expectations are determined by

$$Y^{e,(k)}_{t+s|t-1} = Y^{(k)}(S^{e,(k)}_{t+s|t-1}).$$

[11] Eusepi and Preston (2011) also iterate foreward on the perceived aggregate law of motion to obtain long horizon expectations.

The whole sequence of forecasts at all horizons, although biased, pin down the optimal decisions of households and firms, and hence the equilibrium. Similar results can be obtained if shocks are not turned off but the solution is certainty equivalent or linear.

Iterated Expectations. When there is no aggregate uncertainty, agents would still revise their forecasts after they have new observations that are inconsistent with what they expected ex ante. When they try to forecast the forecasts of the others, the law of motion that is used will be downgraded by one level. This will not be equivalent to the direct forecasts, and the deviation will be amplified by the time horizon.

The accumulation of biases in expectations can be illustrated in the following example.

Example 1. *Consider $k \in \mathbb{N} \cap [2, +\infty)$ and $\epsilon = 0$, the long horizon forecast of the aggregate state can be obtained by iterating on the aggregate state using the aggregate law of motion at one level below, i.e.*

$$S^{e,(k)}_{t+1+s|t-1} = (H^{(k-1)})^{s+1}(S_t).$$

while the forecast of the forecast of others is

$$S^{e,(k)}_{(t+1+s|t)|t-1} = (H^{(k-2)})^{s} H^{(k-1)}(S_t),$$

Therefore, in each iteration of higher order beliefs, small deviations in short horizon forecasts can accumulate into large deviations in long horizons.

Accumulative Forecast Errors. The accumulative forecast errors along forecast horizons are not unique in level-k DSGE. In Eusepi and Preston (2011), the accumulation of forecast errors allows the model to have quantitatively sizable effects from learning, but the biases in short horizon forecast rules are small due to learning. In Angeletos and Lian (2018), forecasts errors are also accumulative but do not come from biased forecast rules.

4 Planning Horizon Non-Neutrality

This section explores the crucial role of "planning horizon" in level-k DSGE. First, the model can be viewed as a pair of dynamic beauty contests connecting the discounted sum of output and inflation forecasts to their current values. Second, "planning horizon" affects the extent of forward looking due to the violation of "Law of Iterated Expectations" across agents. Third, short "planning horizon" is necessary for "Eductive Stability".

4.1 Beauty Contests

A few natural assumptions can make the characterization of level-k DSGE more transparent.

Assumption 4. *Constant intertemporal elasticity of substitution ω:*

$$u(c) = \frac{c^{1-\frac{1}{\omega}}}{1-\frac{1}{\omega}}.$$

This assumption allows us to linearize the decision rules of households, and expressed it as a function of discounted present value of income and substitution effects. The absence of labor supply in utility function removes the wealth effect on labor supply from ω, while Proposition 1 implies that ω does not affect precautionary saving.

Assumption 5. *The steady state net real interest rate is zero, i.e.* $R_{SS} = \Pi_{SS}$.

This assumption simplifies the connection between the frequency of borrowing constraints and the effective discount factor in households' decision rules.

Assumption 6. $\lambda_{\bar{\zeta}|1}\bar{\zeta}$ *is large enough, so that assuming* $b^{e,(k)}(\bar{\zeta}, a, S) = 0$ *is innocuous.*

This assumption requires that the borrowing constraint always binds under $\zeta = \bar{\zeta}$ also in the off-equilibrium belief. The wealth dynamics of each individual household is degenerate also in expectations so that we do not need to keep track of it when solving households' optimization problem.

Lemma 1. *After log-linearization, the aggregate state for decisions at period* t *becomes* $\hat{s}_t = (\hat{y}_{t-1}, \hat{r}_{t-1}, \eta_t^d, \eta_t^z)$. *This information set is indexed by* $t-1$.

Proposition 2. *The recursive level-k equilibrium must satisfy*

$$\hat{y}_t^{(k)} = (1-\gamma)\sum_{s=0}^{+\infty}\gamma^s\left(\varphi_y \hat{y}^{e,(k)}_{t+1+s|t-1} + \varphi_\pi \hat{\pi}^{e,(k)}_{t+1+s|t-1}\right) - \frac{\omega}{1-\gamma\rho_r}\hat{r}^{e,(k)}_{t|t-1} - \frac{\omega}{1-\gamma\rho_d}\eta^d_t,$$
$$\hat{\pi}_t^{(k)} = (1-\theta)\sum_{s=0}^{+\infty}\theta^s\left((1-\theta)\kappa_y \hat{y}^{e,(k)}_{t+s|t-1} + \hat{\pi}^{e,(k)}_{t+s|t-1}\right) - \kappa_z\frac{1-\theta}{1-\theta\rho_z}\eta^z_t,$$

where $\gamma = 1 - \sqrt{\frac{\lambda_{\bar{\zeta}|1}\bar{\zeta}}{\lambda_{\bar{\zeta}|1}\bar{\zeta}+1-\lambda_{\bar{\zeta}|1}}}$,
$(\varphi_y, \varphi_\pi) = \left(1 - \frac{\omega\gamma(1-\rho_r)}{(1-\gamma)(1-\gamma\rho_r)}\phi_y, \frac{\omega}{1-\gamma} - \frac{\omega\gamma(1-\rho_r)}{(1-\gamma)(1-\gamma\rho_r)}\phi_\pi\right)$,
$(\kappa_y, \kappa_z) = \left(\frac{1-\alpha}{\alpha}\varepsilon - (1-\alpha)(\varepsilon-1), \frac{1-\alpha}{\alpha}\varepsilon + 1\right)$.

This representation of the model can be viewed as a dynamic beauty contest in which the current equilibrium variable is a vector of output and inflation gaps depending on the discounted sequence of its future expected values.

The determinant of the parameter γ indicates that when borrowing constraints are more likely to bind, households have shorter planning horizons. Farhi and Werning (2017) uses positive mortality rate to capture this effect in their theoretical framework. In their model, households react less to long horizon events because higher return rate makes the present value of future incomes smaller. This assumption is no longer satisfactory in quantitative work because an unrealistically high mortality rate is needed, which has no clear interpretation. In my model, a 1% quarterly binding rate is sufficient to generate 10% quarterly discount rate on responses to future events, due to the incentive of the households to keep a certain level of precautionary saving, while the present value effect is much less important.

The price rigidity parameter θ plays two roles. The first one is similar to γ, which determines firms' planning horizon. The second one is the slop of Phillips Curve.

From this representation, it is obvious that higher expectations on future output and inflation always raise the current inflation rate, but not necessarily the current output. The reason is that the interest rate response to higher output level and inflation rate would possibly reverse the incentive of spending, and hence result in even lower current output level.

Example 2. *when $\gamma \to 0$,*

$$\hat{y}_t^{(k)} = \hat{y}^{e,(k)}_{t+1|t-1} - \omega(\hat{r}^{e,(k)}_{t|t-1} - \hat{\pi}^{e,(k)}_{t+1|t-1} + \eta^d_t).$$

This example shows that when $\gamma \to 0$, the model with arbitrary subjective expectations has the identical households block as the approach which first derives equilibrium conditions under rational expectations, and then replace rational expectations with subjective expectations as in Milani (2007).

Example 3. *If the model converges when $k \to +\infty$, then when $k \to +\infty$,*

$$\begin{aligned}
\hat{y}_t^{(k)} =& \hat{y}^{e,(k)}_{t+1|t-1} - \omega(\hat{r}^{e,(k)}_{t|t-1} - \hat{\pi}^{e,(k)}_{t+1|t-1} + \eta^d_t), \\
\hat{\pi}_t^{(k)} =& \hat{\pi}^{e,(k)}_{t+1|t-1} + \frac{(1-\theta)^2}{\theta}(\kappa_y \hat{y}^{e,(k)}_{t|t-1} - \kappa_z \eta^z_t).
\end{aligned}$$

Since expectations are unbiased when $k \to +\infty$, this indicates that level-k DSGE nests the rational expectations DSGE as a special case.

4.2 Planning Horizon

The following corollary can be used to demonstrate what makes planning horizon relevant in level-k DSGE.

Corollary 1. *The level-k IS curve can be reformulated in the following*

$$\begin{aligned}
\hat{y}_t^{(k)} = \hat{y}^{e,(k)}_{t+1|t-1} - \omega(\hat{r}^{e,(k)}_{t|t-1} - \hat{\pi}^{e,(k)}_{t+1|t-1} + \eta^d_t) + (1-\gamma)\sum_{s=1}^{+\infty}\gamma^s(\hat{y}^{e,(k)}_{t+1+s|t-1} - \hat{y}^{e,(k)}_{(t+1+s|t)|t-1}) \\
- \omega\sum_{s=1}^{+\infty}\gamma^s[(\hat{r}^{e,(k)}_{t+s|t-1} - \hat{r}^{e,(k)}_{(t+s|t)|t-1}) - (\hat{\pi}^{e,(k)}_{t+1+s|t-1} - \hat{\pi}^{e,(k)}_{(t+1+s|t)|t-1})].
\end{aligned}$$

When the Law of Iterated Expectations holds across agents, this equation reduces to

$$\hat{y}_t^{(k)} = \hat{y}^{e,(k)}_{t+1|t-1} - \omega(\hat{r}^{e,(k)}_{t|t-1} - \hat{\pi}^{e,(k)}_{t+1|t-1} + \eta^d_t).$$

According to Corollary 1, planning horizon matters because the “Law of Iterated Expectation” no longer holds across agents. For example, a level-2 agent form expectations as if others are level-1, while they forecast the forecasts of others as if they arise from level-0.

This issue not only exists in level-k models. The derivation of Corollary 1 implicitly assumes that agents understand “individual rationality”. If we relax the level-k assumptions expectations but instead impose “Law of Iterated Expectation” on Corollary 1, then the admissible set of subjective expectations will be very restricted. See the following example for illustration.

Example 4. *Consider an example in which*
(1) level-k is relaxed,
(2) $1-\theta=\epsilon^r=\eta^d=0$, $\phi_y\geq 0$, *and*
(3) the perceived Taylor Rule is $\hat{r}^e_{t+s|t-1}=\phi_y\hat{y}^e_{t+s|t-1}$, *then*

$$\hat{y}^e_{t+1|t-1}=(1-\gamma)\sum_{s=0}^{+\infty}\gamma^s\left(\hat{y}^e_{t+2+s|t-1}-\frac{\omega\phi_y}{1-\gamma}\hat{y}^e_{t+1+s|t-1}\right).$$

Now, consider a perceived law of motion $\hat{y}^e_{t|t-1}=\rho^e_y\hat{y}^e_{t-1}$. *Assume that the “Law of Iterated Expectation” also holds across time, which is a natural assumption when individual rationality is not violated, we can iterate on the perceived law of motion for output to obtain long horizon expectations. As a result, we must have either* $\rho^e_y=0$ *or* $\rho^e_y=1+\omega\phi_y\geq 1$. *In another word, anchored expectations are not admissible in this example.*

There are two cases in which planning horizon does not play a role. The first case is $k\to+\infty$, which corresponds to rational expectations as is standard in most DSGE models. In this case, the “Law of Iterated Expectations” holds across both time and agents, but expectations will no long be anchored by historical data. The second case is $\gamma\to 0$, which corresponds to the “Euler Equation Learning” approach as in Milani (2007) and Milani (2011). In this case, long run expectations are assumed not to play a role. As in Example 4, anchored expectations are not likely to be compatible with the common knowledge of individual rationality. In summary, planning horizon is likely to play a key role in determining equilibrium output, if we would like to have expectations anchored to the past. This issue exists even if the anchoring does not arise from non-rational expectations (Angeletos and Lian, 2018).

The following proposition summarizes the real effects of planning horizon γ.

Proposition 3. *When $\eta^d = \eta^z = 0$ and $k \in [1,2]$, we have*

$$\hat{y}_t^{(k)} = -\frac{\omega}{1-\gamma\rho_r}\hat{r}^{e,(k)}_{t|t-1} + (k-1)(1-\gamma)(1-\theta)\kappa_y\varphi_\pi\hat{y}^{e,(k)}_{t|t-1} + \delta(1-\gamma)\sum_{s=0}^{+\infty}\gamma^s\hat{y}^{e,(k)}_{t+1+s|t-1},$$

where $\delta = 1 - \frac{\omega\gamma}{1-\gamma}\left[(k-1)(1-\theta)\kappa_y\left(\frac{1-\rho_r}{1-\gamma\rho_r}\gamma\phi_\pi - 1\right) + \frac{1-\rho_r}{1-\gamma\rho_r}\phi_y\right]$.

This proposition indicates that $(1-\gamma)\sum_{s=0}^{+\infty}\gamma^s\hat{y}^{e,(k)}_{t+1+s|t-1}$ is crucial in determining the current equilibrium output, and δ captures the size of this effect. Here, γ plays two roles. First, smaller γ implies that households care more about the near future than the far future. Second, smaller γ leads to larger δ and hence makes households more responsive to expectations. Taking stock, when level-k output expectations anchors the past, smaller γ leads to stronger anchoring.

In addition to γ, $\{k, \theta, \phi_\pi, \phi_y\}$ all affect δ. $\{k, \theta\}$ affect δ in the same way, as lower level reasoning and price flexibility both dampen the self-stabilizing channel in expectations through making inflation expectations less responsive. $\{\phi_\pi, \phi_y\}$ both dampens the effects of expectations, but the relative role of ϕ_π is affected by $k-1$ because it has to operate through inflation expectations, and affects the equilibrium output only indirectly.

4.3 Eductive Stability

The convergence to rational expectations when $k \to +\infty$ is difficult to characterize in the full model. Still, some transparent results can be obtained when $\rho_r = \phi_\pi = 0$. The goal is to show why γ plays a crucial role in "Eductive Stability". The role of γ in the full model is similar.

Lemma 2. *When $\rho_r = \phi_\pi = \eta^d = \eta^z = 0$ and $k \in \mathbb{N}_{++}$, the law of motion for output satisfies $\hat{y}_t^{(k)} = \rho_y^{(k)}\hat{y}_{t-1}$, and $\rho_y^{(k)}$ satisfies*

$$\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}} = -\omega\phi_y + \left(1 - \frac{\omega\gamma}{1-\gamma}\phi_y\right)\frac{(1-\gamma)\rho_y^{(k-1)}}{1-\gamma\rho_y^{(k-1)}}.$$

Definition 3. *Under the same environment, "Eductive Stability" is defined as the following:* $\exists M \in [0, 1)$*, such that for* $\forall \rho_y^{(k-1)}$ *with* $|\rho_y^{(k-1)}| \leq 1$*,* $|\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}}| \leq M$*.*

The requirement that the convergence speed must be uniformly fast than some lower bound does not lose any generality because the mapping from $\rho_y^{(k-1)}$ to $\rho_y^{(k)}$ itself is uniform.

If we allow $\rho_y^{(0)}$ to be specified in an arbitrary way, then "Eductive Stability" will become a sufficient and necessary condition to guarantee that the convergence to rational expectations when $\kappa \to +\infty$ is monotonic in absolute values. Due to the monotonicity of $\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}}$ in $\rho_y^{(k-1)}$ for $|\rho_y^{(k-1)}| \leq 1$, it is easy to prove the following proposition.

Proposition 4. *When* $\rho_r = \phi_\pi = \eta^d = \eta^z = 0$ *and* $k \in [1, +\infty) \cap \mathbb{N}$*, the sufficient and necesary condition for "Eductive Stability" is* $\gamma \in (\frac{1}{2}\omega\phi_y, 1 - \frac{1}{2}\omega\phi_y)$*.*

In standard parameterization, the $\gamma > \frac{1}{2}\omega\phi_y$ is always satisfied, while $\gamma < 1 - \frac{1}{2}\omega\phi_y$ is usually not. For instance, when $(\omega, \phi_y) = (0.5, 0.2)$, "Eductive Stability" requires $\gamma \in (0.05, 0.95)$. Complete market models are observationally equivalent to $\gamma = \beta > 0.95$, while incomplete market models can have $\gamma < 0.95$. This result resembles the findings in Evans et al. (2017) that the long planning horizon destroys the "Eductive Stability".

The proof of "Eductive Stability" is complex and less illuminating in the full model. Yet, the main insight is similar. Too strong self-stabilizing feedback from Taylor Rule and consumption response would make the equilibrium output responding negatively to output expectations, which is strongly contradictory to common sense.

There is a crucial parameter δ we could take a closer look. This parameter describes how the equilibrium output reacts to its expectations along all horizons. "Eductive Stability" requires this parameter to be large enough. From the expression of δ as in Proposition 3, we can see that (1) γ plays a similar role as in Lemma 2; (2) smaller k and $1 - \theta$ both increase the likelihood of "Eductive Stability" directly, and in addition makes inflation targeting less likely to destroy "Eductive Stability" indirectly; (3) ρ_r does not play a large role. Therefore, we can conclude that level-k model is more likely to be a useful tool for business cycle questions in the presence of incomplete markets and nominal rigidities.

5 Asymmetric Reasoning

The major feature of level-DSGE is the "asymmetric reasoning" in forecast rules. Households can understand the comovements between macroeconomic variables that have direct connections, but not those that must be inferred from feedback effects. This section formalizes this feature of the model, takes the model to take, and evaluates its quantitative impact.

5.1 Testable Forecast Rules

One Step Ahead Forecasts. When $k \in (1, 2]$, the macroeconomic comovements that arise from feedback effects are not understandable by agents, so that part of the state variables are missing in their forecast rules.

Proposition 5. *Level-k expectations for* $k \in [1, 2]$ *are given by*

$$\begin{aligned}
\hat{y}^{e,(k)}_{t|t-1} - \hat{y}_{t-1} =& (k-1)\omega \left(-\frac{1-\rho_r}{1-\gamma\rho_r} \frac{\phi_y}{1-\gamma} \hat{y}_{t-1} - \frac{\rho_r}{1-\gamma\rho_r} \hat{r}_{t-1} - \frac{1}{1-\gamma\rho_d} \eta^d_t \right), \\
\hat{\pi}^{e,(k)}_{t|t-1} =& (k-1)(1-\theta) \left(\kappa_y \hat{y}_{t-1} - \frac{\kappa_z}{1-\theta\rho_z} \eta^z_t \right).
\end{aligned}$$

This proposition has the following implications. First, the monetary policy rule is only partly understood, in the sense that only the interest rate response to output fluctuations is incorporated into the one-quarter ahead forecast rules. Second, if interest rate does not respond to output, then the forecast rules exhibits "asymmetric reasoning", in the sense that interest rate is only used to forecast output growth, while output level is only used to forecast inflation in the next quarter. Third, agents' reasoning is asymmetric in a way that only the direct effects shows up in forecast rules, while the feedback effects are absent. Forth, the coefficients on all state variables in forecast rules are proportional to $k-1$, which looks as if $k-1$ represents households' awareness of direct effects.

Mutiple Step Ahead Forecasts. The nice results in one quarter ahead forecasts do not carry on directly to multiple step ahead forecasts.

Corollary 2. *When* $\phi_y = \eta^d = \eta^z = 0$ *and* $k \in [1, 2]$, *we have*

$$\frac{\partial \hat{y}^{e,(k)}_{t+3|t-1}}{\partial \hat{y}_{t-1}} - 1 = -\psi\phi_\pi(1 + \rho_r^2 - \psi\phi_\pi),$$

$$\frac{\partial \hat{\pi}^{e,(k)}_{t+3|t-1}}{\partial \hat{r}_{t-1}} = -\frac{\psi(1 + \rho_r + \rho_r^2 - \psi\phi_\pi)}{1 - \rho_r},$$

where $\psi = (k-1)^2(1-\theta)\kappa_y\omega\frac{\rho_r(1-\rho_r)}{1-\gamma\rho_r}$,

The missing channels in Proposition 5 are no longer missing here. Yet, their sizes are in the same order of magnitude as the small ψ. Agents are aware of the self-stabilizing forces because they have non-trivial inflation expectations, and understand that interest rate responds to it. These two channels will be used to compute the perceived law of motion. However, both of them are much weaker than reality due to the dampening effects of level-k when k is close to 1. In addition, growth and inflation expectations are formed without using the expectations of the others, which also makes expectations less responsive than reality.

As a result, as long as k is close to 1, the "asymmetric reasoning" is still a distinct feature of the one-year ahead forecast rules. Therefore, we can still use the data implied forecast rules to test whether expectations are correctly specified in the level-k model.

5.2 Taking Model to Data

Calibrate Planning Horizon γ. Recall the equation that determines the value of γ

$$\gamma = 1 - \sqrt{\frac{\lambda_{\bar{\zeta}|1}\bar{\zeta}}{\lambda_{\bar{\zeta}|1}\bar{\zeta} + 1 - \lambda_{\bar{\zeta}|1}}}.$$

Two parameters $\{\lambda_{\bar{\zeta}|1}, \bar{\zeta}\}$ need to be calibrated.

According to Kaplan, Violante, and Weidner (2014), the probability that a non-hand-to-mouth consumer remains non-hand-to-mouth after one quarter is 0.967, and hence $\lambda_{\bar{\zeta}|1} = 0.967$. Set $\bar{\zeta} = 1.2$ or 1.5 or 2.0 to get $\gamma = 0.802$ or 0.779 or 0.747. $\gamma = 0.8$ is taken as the benchmark.

State Space Representations. The state space representation has a transition equation

$$\hat{s}_{t+1} = \begin{bmatrix} \hat{y}_t \\ \hat{r}_t \\ \eta^d_{t+1} \\ \eta^z_{t+1} \end{bmatrix} = \Gamma^{ss,(k)} \begin{bmatrix} \hat{y}_{t-1} \\ \hat{r}_{t-1} \\ \eta^d_t \\ \eta^z_t \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma_r \epsilon^r_t \\ \sigma_d \epsilon^d_t \\ \sigma_z \epsilon^z_t \end{bmatrix},$$

and a measurement equation

$$\begin{bmatrix} \hat{y}_{t+1} - \hat{y}_t \\ \hat{p}_{t+1} - \hat{p}_t \\ \hat{r}_t \\ \hat{y}^e_{t+4|t} - \hat{y}_t \\ \hat{p}^e_{t+4|t} - \hat{p}_t \end{bmatrix} = \begin{bmatrix} \Gamma^{ys,(k)} - \Gamma^y \\ \Gamma^{\pi s,(k)} \\ \Gamma^r \\ \Gamma^{ys,e,(k)} (\Gamma^{ss,e,(k)})^3 - \Gamma^y \\ \Gamma^{\pi s,e,(k)} \sum_{\tau=0}^{3} (\Gamma^{ss,e,(k)})^\tau \end{bmatrix} \hat{s}_{t+1} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \sigma_y \epsilon^y_t \\ \sigma_p \epsilon^p_t \end{bmatrix},$$

where $\Gamma^{ss,(k)}$ and $\Gamma^{ss,e,(k)}$ denote the actual and perceived law of motion for the recursive level-k equilibrium, $\Gamma^{ys,(k)}$ and $\Gamma^{ps,(k)}$ denote the actual function of output and inflation, $\Gamma^{ys,e,(k)}$ and $\Gamma^{ps,e,(k)}$ denote the perceived functions or the forecast rules of output and inflation, Γ^y and Γ^r extract output and interest rate from the vector of aggregate states. $\{\epsilon^r_t, \epsilon^d_t, \epsilon^z_t, \epsilon^y_t, \epsilon^p_t\}$ are i.i.d. standard normally distributed, in which $\{\epsilon^r_t, \epsilon^d_t, \epsilon^z_t\}$ stand for exogenous shocks driving aggregate states, while $\{\epsilon^y_t, \epsilon^p_t\}$ are measurement errors. $\{\sigma^r_t, \sigma^d_t, \sigma^z_t, \sigma^y_t, \sigma^p_t\}$ denote the standard deviation of all these shocks.

Time Series Observables. The macroeconomic data range from the first quarter of 1985 to the last quarter of 2007, covering the period of the Great Moderation. The quarterly real GDP Y_t is used to construct output growth rate, quarter GDP deflator P_t is used to construct inflation, and annualized federal fund rate r_t is used to construct nominal interest rate.

$$\begin{aligned} \hat{y}_{t+1} - \hat{y}_t &= detrend(\ln Y_{t+1} - \ln Y_t), \\ \hat{p}_{t+1} - \hat{p}_t &= detrend(\ln P_{t+1} - \ln P_t), \\ \hat{r}_t &= detrend(r_t)/4, \end{aligned}$$

where $detrend(\cdot)$ represents for the linear detrending operator.

The forecast data is taken from Michigan Survey of Consumers in the same time range. It includes the backcast index of yearly business condition change $bago_t$, forecast index of yearly business condition change $bexp_t$, and forecast of yearly inflation rate $inflmedian12_t$. The one year ahead output growth rate forecast is constructed assuming perfect recall and the one year ahead inflation rate forecast is taken from the median value of the survey response directly.

$$\begin{aligned}
\hat{y}^e_{t+4|t} - \hat{y}_t =& detrend\left[\frac{\sum bago_t(\ln Y_t - \ln Y_{t-4})}{\sum bago_t^2} \cdot bexp_t\right],\\
\hat{p}^e_{t+4|t} - \hat{p}_t =& detrend(inflmedian12_t).
\end{aligned}$$

Baynesian Estimation. Set $\kappa_y = 1$ to capture the 50% intermediate input share as in Bils, Klenow, and Malin (2018). The prior and posterior distributions of other parameters are summarized in Table 1.

Table 1: Prior and Posterior Distribution of Parameters

Parameters	Prior shape	Prior Mean	Prior S.D.	Post. Mean	95% Band
k	Uniform 1/k			1.334	[1.212,1.467]
ω	Normal	1.000	0.500	0.075	[0.036,0.114]
θ	Uniform			0.925	[0.896,0.945]
ϕ_π	Normal	1.500	0.500	1.636	[0.767,2.844]
ϕ_y	Normal	0.200	0.100	0.102	[-0.001,0.214]
ρ_r	Uniform			0.915	[0.869,0.957]
ρ_d	Uniform			0.237	[0.076,0.440]
ρ_z	Uniform			0.878	[0.844,0.914]
$400\sigma_r$	InvGamma	0.500	4.000	0.435	[0.384,0.507]
$100\sigma_d$	InvGamma	1.000	4.000	4.160	[2.178,8.996]
$100\kappa_z\sigma_z$	InvGamma	5.000	4.000	5.562	[3.397,8.794]
$100\sigma_y$	InvGamma	0.300	4.000	0.307	[0.262,0.358]
$100\sigma_p$	InvGamma	0.500	4.000	0.373	[0.323,0.431]

The prior of k is chosen such that $k = +\infty$ can be drawn with the same probability as $k = 1$ without data. For parameters $\{\omega, \phi_\pi, \phi_y\}$ that can take the values in $\mathbb{R}$, the prior distribution is assumed to be normal. For parameters $\{\theta, \rho_r, \rho_d, \rho_z\}$ that can only take values in $[0, 1]$, the prior distribution is assumed to be uniform.

$k = 1.334$ is close to the results in experimental games as in Camerer et al. (2004). Although ω is much smaller than 1, it is consistent with macro level estimates as summarized by Havranek (2015). θ is very close to 1 due to the absence of wage rigidity. $\phi_\pi = 1.636$ is quite standard. ϕ_y is close to 0 because higher values will lead to very strong self-stabilizing force in expectations, which is absent in the data. $\rho_r = 0.915$ is standard in the literature. ρ_d much smaller than 1 indicates that the model has strong internal propagation mechanism for aggregate demand. ρ_z not close to 1 indicates that the model is unlikely to be misspecified in terms of stationarity. σ_y and σ_p are both smaller than the standard deviations of growth and inflation forecasts, which are very difficult to get if the expectations formation process is not properly specified.

5.3 Quantitative Forecast Wedges

Forecast Dynamics. Figure 6 and 7 compare forecasts in the model, with forecasts in the data and reality in the data for both output growth and inflation.

The volatility of output growth forecasts in both model and data is much lower than reality. The forecast errors are large, persistent and countercyclical, which indicates that households expectations' are driven by endogenous waves of optimism and pessimism. The model implied forecasts generally fits the data counterpart well. The rise of growth forecasts during the early 2000s and the fall of them during the late 2000s in the data that seems difficult to capture by the model are aligned with news shocks as in Barsky and Sims (2012).

Inflation forecasts are also much less volatile than the reality. The model generally captures the overpredict of inflation during the IT boom in the late 1990s, and the underpredict of inflation during the housing booms in the middle of 2000s. Yet, it does not fully capture the rise of inflation expectations during the late 1980s. This discrepancy is very difficult to clean up because k is exogenous in the model. Once I raise k to capture the inflation expectations during the late 1980s, the fit of inflation expectations during other periods will be much worse.

It would be ideal if we have the data of interest rate expectations. Unfortunately, the data provided by Michigan Survey of Consumers is only qualitative, and there is no easy to transform it into quantitative measures.

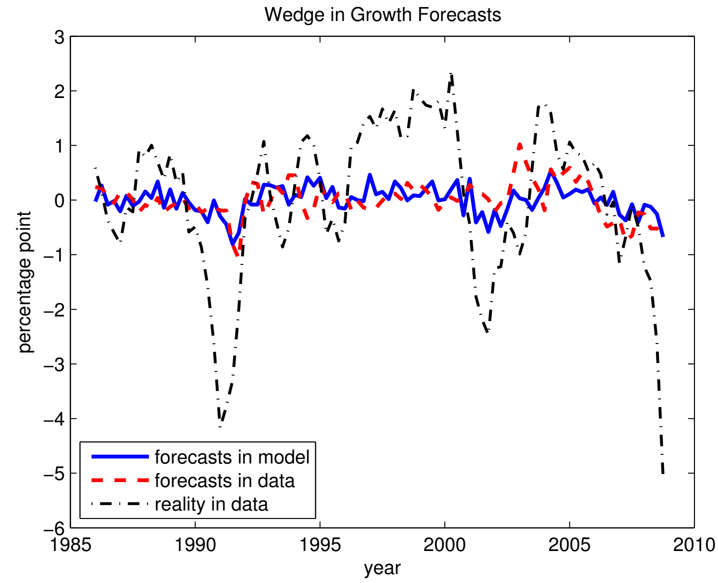

Figure 6: Forecast Dynamics of Output Growth

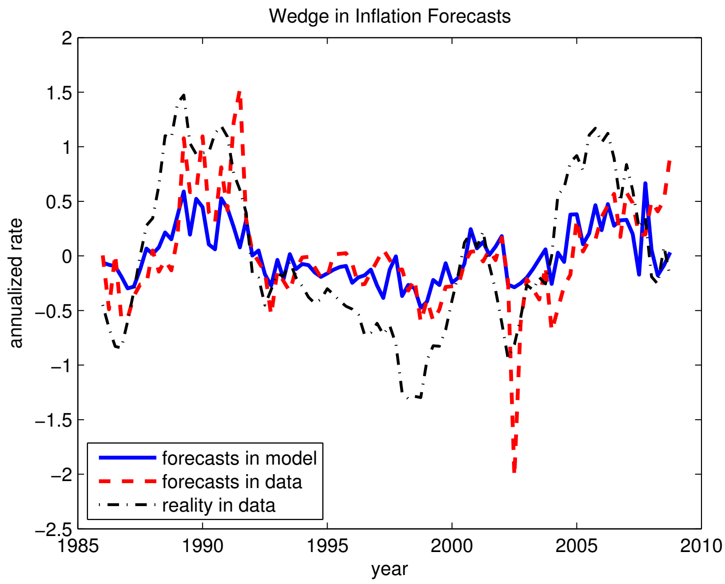

Figure 7: Forecast Dynamics of Inflation Rate

Wedges in Forecast Rules. A distinct feature of level-k DSGE model is the wedge between forecasts and reality. Table 2 summarizes the wedges by comparing the corresponding law of motion for output growth and inflation. The indirect channels that are hard to understand by households are marked by shaded areas.

Table 2: Wedges in Forecast Rules

	$\hat{y}^e_{t+4\|t} - \hat{y}_t$		$\mathbb{E}_t\hat{y}_{t+4} - \hat{y}_t$	
	Model	Data	Model	Data
$\hat{y}_t$	-0.022	0.023 [-0.009,0.055]	-0.249	-0.126 [-0.248,-0.005]
$\hat{r}_t$	-0.298	-0.531 [-0.683,-0.379]	-1.945	-1.352 [-1.930,-0.773]
η^d_{t+1}	-0.040	-0.009 [-0.020,0.001]	-0.141	-0.136 [-0.176,-0.096]
η^z_{t+1}	0.001	0.008 [-0.001,0.016]	0.016	0.030 [-0.002,0.062]
R^2		0.402		0.526
	$\hat{p}^e_{t+4\|t} - \hat{p}_t$		$\mathbb{E}_t\hat{p}_{t+4} - \hat{p}_t$	
	Model	Data	Model	Data
$\hat{y}_t$	0.100	0.068 [0.028,0.107]	0.320	0.279 [0.233,0.326]
$\hat{r}_t$	-0.012	0.457 [0.271,0.643]	-0.463	-0.316 [-0.537,-0.094]
η^d_{t+1}	-0.003	0.009 [-0.004,0.022]	-0.043	-0.013 [-0.028,0.003]
η^z_{t+1}	-0.034	-0.040 [-0.050,-0.030]	-0.110	-0.077 [-0.089,-0.065]
R^2		0.552		0.719

In Table 2, state variables are used as the regressors in the forecast rule regression to check model fit. $\{\hat{\eta}^d, \hat{\eta}^z\}$ are model implied structural shocks. $\hat{y}^e_{t+4|t} - \hat{y}_t$ and $\hat{p}^e_{t+4|t} - \hat{p}_t$ are one year ahead growth and inflation forecasts. $\mathbb{E}_t\hat{y}_{t+4} - \hat{y}_t$ and $\mathbb{E}_t\hat{p}_{t+4} - \hat{p}_t$ are the unbiased one year ahead growth and inflation expectations. If forecast rules in the model are correctly specified, then the model implied measurement errors should be i.i.d., hence the estimated forecast rules should be identical for model and data. If the equilibrium problem in the model are correctly specified, then structural shocks should be i.i.d., and the estimated unbiased expectations rules should also be identical for model and data. The regression results indicate that the misspecification is only mild.

The regressions in Table 2 indicate that the model generally captures the wedges between the forecast rules and the unbiased expectations rules. The indirect connection between output level and output growth forecasts, and that between interest rate and inflation rate forecasts are missing in both model and data. The direct connection between interest rate and output growth forecasts, and that between output level and inflation rate forecasts are there but much weaker in both model and data.

Identification of Level-k. The main identification of level-k comes from the forecast rules. Figure 8 plots the coefficient differences between the model implied forecast rules and the data implied forecast rules, normalized by the standard errors of the later.

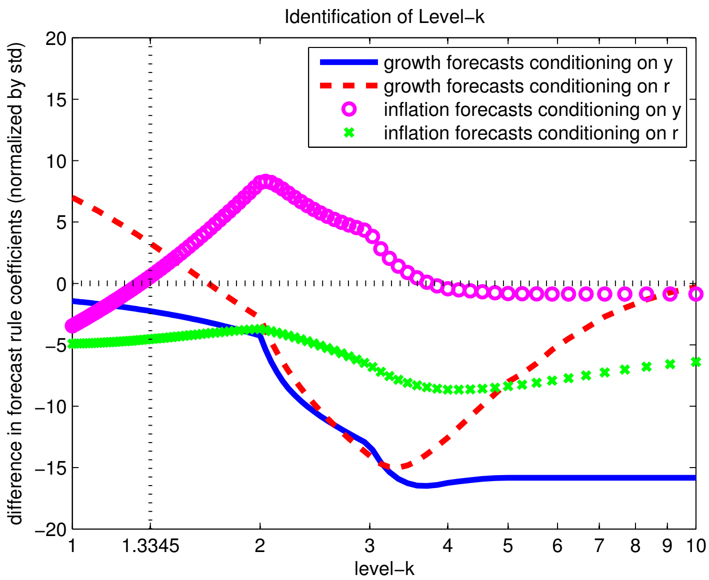

Figure 8: Identification of Level-k

The two indirect channels, growth forecasts conditioning on output level and inflation forecast conditioning on interesting rate, can identify $k \in [1, 2]$, because they obtain best fit only in this range; while the two direct channels, growth forecasts conditioning on interest rate and inflation forecast conditioning on output level, help identify k given $k \in [1, 2]$, because they have additional best fit points outside $[1, 2]$ but are more sensitive to k within $[1, 2]$.

6 Alternative Policies

This section evaluates the model performance under four alternative monetary policies. The robustness of model performance can provide additional support for the usefulness of level-k DSGE. The four polices includes Taylor Rule during the Great Moderation, Taylor Rule during in Pre-Volcker era, Liquidity Trap in the Great Recession, and Forward Guidance.

6.1 Great Moderation

Impulse Reponses. Figure 9 plots the impulse response functions to -1 standard deviation of federal fund rate shocks. The model generated responses are well aligned with those produced by a structural VAR.

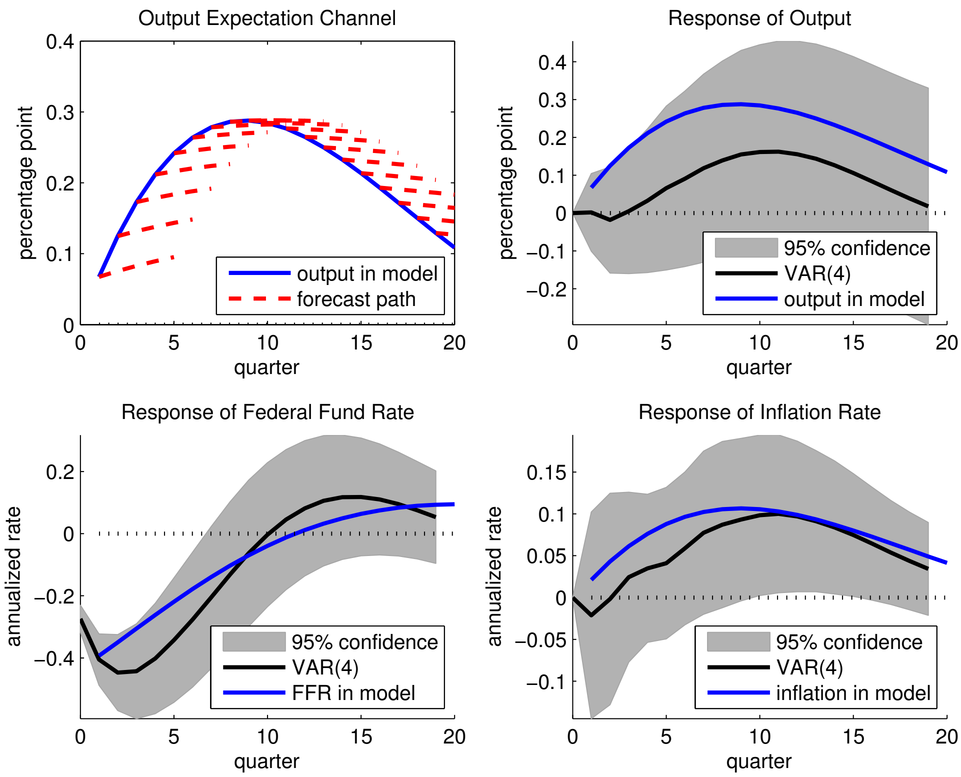

Figure 9: Impulse Reponses to Monetary Shocks

Inflation Inertia. According to the top-left panel, the response of output is hump-shaped. One potential reason is the highly anchored forecasts of future output dynamics. Corollary 2 has provided the intuition of it.

More interesting is the inflation inertia from the right-bottom panel. Under rational expectations, it is not easy to have such results without inflation indexation or working capital loans. The following proposition demonstrates how level-k DSGE produces inflation inertia.

Proposition 6. *When $k \in [1, 2]$ and $\eta_t^z = 0$,*

$$\hat{\pi}_t^{(k)} = (1-\theta)\kappa_y \left\{ (k-1)(1-\theta)\hat{y}_{t-1} + [1+\theta(k-1)](1-\theta)\sum_{s=0}^{+\infty} \theta^s \hat{y}_{t+s|t-1}^{e,(k)} \right\}.$$

This proposition shows that the current inflation is driven by output forecasts, and we have known previously that the forecasts of output paths are highly anchored to the past.

Observational Equivalence. Under rational expectations, other frictions are needed to produce the same impulse responses.

Proposition 7. *The following model with full information rational expectations, external habit and within period working capital loans*

$$\begin{aligned}
\hat{y}_t - \hat{y}_{t-1} =& -\frac{1-\widetilde{h}}{\widetilde{h}}\widetilde{\omega}(\hat{r}_t - \hat{\pi}_{t+1}) + \frac{1-wy_{ss}(1-\widetilde{\omega})}{\widetilde{h}}(\hat{y}_{t+1} - \hat{y}_t),\\
\hat{\pi}_t =& \frac{(1-\widetilde{\theta})^2}{\widetilde{\theta}}\left[\widetilde{\kappa}_y\left(\frac{1}{1-\widetilde{h}}\hat{y}_t - \frac{\widetilde{h}}{1-\widetilde{h}}\hat{y}_{t-1}\right) + \hat{r}_t\right] + \hat{\pi}_{t+1},\\
\hat{r}_t =& \rho_r\hat{r}_{t-1} + (1-\rho_r)(\phi_\pi\hat{\pi}_t + \phi_y\hat{y}_t),
\end{aligned}$$

where $(\widetilde{h}, \widetilde{\omega}, \widetilde{\kappa}_y, \widetilde{\theta}) = (0.863, 0.730, 0.068, 0.749)$ has the identical impulse response functions to FFR shocks as the level-k DSGE. Here $wy_{ss} = 2/3$ denotes the steady state labor share.

Compared with this model, level-k DSGE has its own merit because it has the same number of parameters but captures additionally the forecast dynamics. $\widetilde{\kappa}_y = 0.068$ also explains why very sticky wage is needed in Christiano et al. (2005) to amplify the role of working capital.

6.2 Pre-Volcker Era

Growth Forecasts Across Regimes. Extending the sample range from 1985q1-2007q4 to 1965q1-2007q4 for Figure 1 to get Figure 10, we see that the backcasts of business condition changes are not improved, but the forecasts are much more aligned with both the backcasts and the reality.

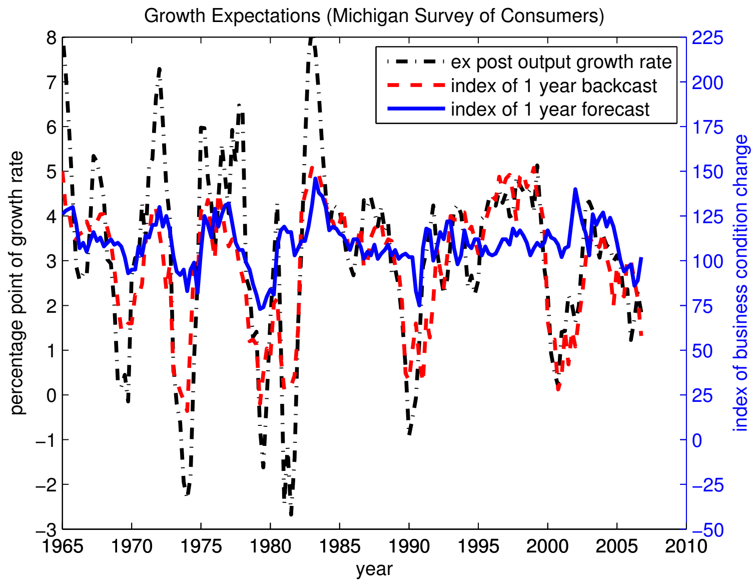

Figure 10: Growth Forecasts Across Regimes

In the empirical literature of estimating monetary policy rules, there is one line of research led by Orphanides (2004) arguing that in the pre-Volcker regime, monetary policy is too active to poorly measured output gaps. Boivin (2006) conducts more careful analysis and confirms the more active response to output gaps in the Pre-Volcker era.

Ideally, we should estimate the whole model in the pre-Volcker era to quantify the contribution of alternative policy rules in explaining the forecasts during the pre-Volcker era. However, this cannot be achieved because the current version of level-k DSGE does not handle stochastic trend very well. A much less ambitious approach is to check whether the alternative forecast rule under alternative policy rule is also aligned with data.

Inflation vs GDP Targeting. The coefficients ϕ_π and ϕ_y reflect the dual goals of inflation targeting and GDP targeting in monetary policy rules. Despite the disagreement in empirical results, it is generally believed that ϕ_π is lower (Clarida, Galí, and Gerlter, 2000) and ϕ_π is higher (Orphanides, 2004) in the pre-Volcker era. Boivin (2006) adopts more sophisticated methods, and gets $(\phi_\pi, \phi_y) = (1.10, 0.47)$ before 1979q3, and $(\phi_\pi, \phi_y) = (1.50, 0.00)$ after that.

Table 3 summarizes the role of policy rules in affecting the mean reversion of output. The 4 columns correspond to the model based forecast rule, data based forecast regression, model based unbiased expectation rule, data based auto-regression. The upper half of corresponds to the Great Moderation period, while the lower half corresponds to the pre-Volcker era.

Table 3: Mean Reversion Across Policy Regimes

	$\hat{y}^e_{t+4\|t} - \hat{y}_t$		$\mathbb{E}_t\hat{y}_{t+4} - \hat{y}_t$	
	$(1.64, 0.10)$	1985q1-2007q4	$(1.64, 0.10)$	1985q1-2007q4
$\hat{y}_t$	-0.022	-0.008 $[-0.050, 0.034]$	-0.250	-0.290 $[-0.447, -0.133]$
Obs.		88		88
R^2		0.002		0.135
	$(1.10, 0.47)$	1965q1-1979q3	$(1.10, 0.47)$	1965q1-1979q3
$\hat{y}_t$	-0.092	-0.176 $[-0.272, -0.079]$	-0.682	-0.775 $[-1.034, -0.516]$
Obs.		55		55
R^2		0.201		0.405

The results indicate that GDP targeting does enlarge the mean reversion of output both in forecast and in reality. The intuition has been highlighted in Proposition 5.

Stablization Mechanism. As in Proposition 3, the current output is determined by the nowcast of interest rate, the nowcast of output, and the forecast of output. When the federal fund rate does not respond much to output gaps, lower interest rate raises the output growth not through the nowcast or forecast of output, but through the nowcast of interest rate. As a result, Taylor Rule stabilizes the economy not through making expectations mean reverting, but through making the economic outcomes much more mean reverting than expectations.

6.3 Great Recession

Missing Deflation. Under rational expectations, permanent negative demand gaps can make deflation explode. This implies that strong deflation would be observed if the demand gap is very persistent. In level-k DSGE, this is no longer the case.

Proposition 8. *For hypothetical permanent output gap $\hat{y}_{perm}$ without aggregate shocks,*

$$\begin{aligned}\pi^{e,(k)}_{perm} =&(k-1)\cdot(1-\theta)\kappa_y\hat{y}_{perm},\\ \pi^{(k)}_{perm} =&k\cdot(1-\theta)\kappa_y\hat{y}_{perm}.\end{aligned}$$

This proposition indicates that permanent output gap is compatible stable inflation. However, it does not mean that the monetary authority can permanently raise the output level. Suppose it does, then the original inflation target can no longer be maintained, and it is no longer innocuous to assume that households are unaware of it.

With estimated parameters, 1.0% permanent downturn generates 0.1% disinflation expectations and 0.4% disinflation. The lack of drop in inflation expectations is a fact that has been explored by Negro, Giannoni, and Schorfheide (2015). Level-k DSGE interprets it as the lack of deep reasoning, instead of anything from surprise.

The observational equivalent model in Proposition 7 will produce the same size of deflation if the output gap lasts for 4 years. Yet, that model will produce a drop of inflation expectations to the same size of the drop of inflation, which does not happen during the Great Recession.

Stagnant Recovery Expectations. Consider an economy initially with a negative output gap, and no aggregate shocks. The natural rate of interest has changed permanently by $\Delta\hat{r}^n$, and the federal fund rate is binded at zero until the output gap becomes non-negative.

Proposition 9. *Denote $\hat{\pi}_{SS}$ as the inflation target. When there are no exogenous shocks, for $k\in[1,2]$, during the liquidity trap, we have*

$$\hat{y}^{e,(k)}_{t+s|t-1}=\hat{y}_{t-1}+(s+1)(k-1)\frac{\omega}{1-\gamma}(\hat{\pi}_{SS}+\Delta\hat{r}^n).$$

With estimated parameters, $(k-1)\frac{\omega}{1-\gamma} = 0.125$, which indicates very slow expected recovery. This is confirmed by Figure 11. In the figure, the average index of business condition change after the end of 2009 is nearly the same as that before, which supports the model implication that households do not expect strong recovery during the Great Recession.

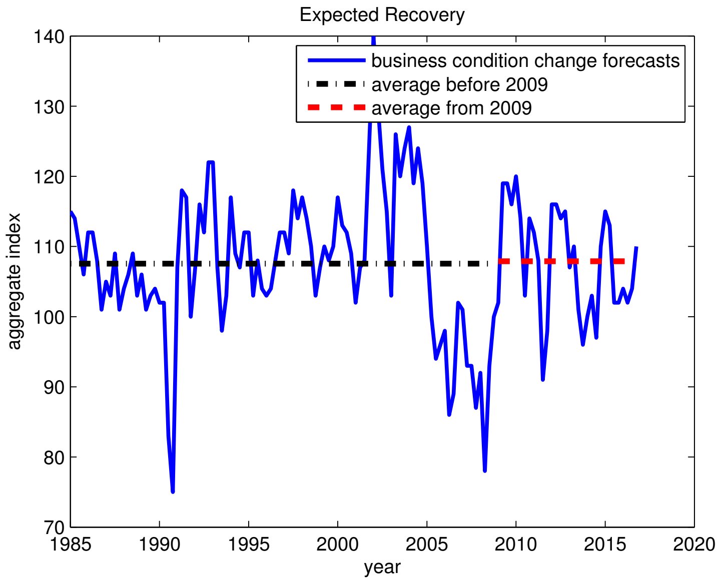

Figure 11: Stagnant Recovery Expectations

This figure overturns the common belief that the economy would recover very fast after the recession. If their forecasts are relevant for economic decisions, then what we need to explain in the model is not why recovery is surprisingly slow, but whether the slow recovery can be the consequence of pessimistic expectations.

Slow Recovery. Now consider a numerical example, in which the neutral rate of interest in terms of discount factor drops by 1.5% permanently, and the economy starts with some negative output gaps. Nominal interest rate is trapped as zero until output is fully recovered. All shocks are turned off.

As in Figure 12, the output recovery can be very slow if the natural rate of interest is low as well, and the recovery is slower if the initial output gap is larger.

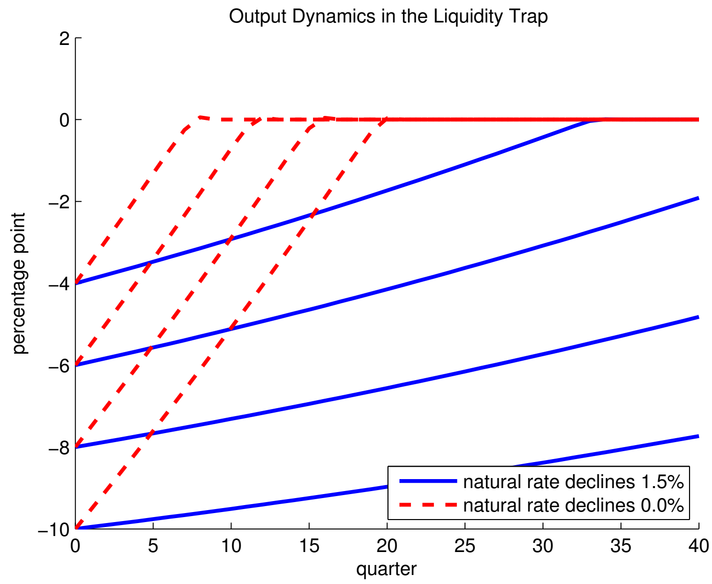

Figure 12: Slow Recovery

Proposition 3 and 9 jointly provide the intuition why the recovery can be so lower. In both propositions, planning horizon plays a very important role. According to Proposition 9, given a long period of lower interest rate stimulus, households would expect very strong recovery if γ is close to one. According to Proposition 3, given some expectations of recovery, large negative output gap cannot be sustained if γ is close to one.

6.4 Forward Guidance

Amplifying or Dampening. Similar to Farhi and Werning (2017), the initial response of output to an interest rate shock in future will be dampened by level-k. However, due to the different specification of level-0, the effects of monetary shocks are accumulated across time. The ultimate effect of forward guidance may not be small. It just takes a while to fully realize.

Policy Experiment. Now consider a thought experiment that the $t+\tau$ period interest rate expectation is shocked by -1 percentage point. All other shocks are turned off. Interest rates are pegged before the shock, and follow the Taylor Rule otherwise.

Proposition 10. *Under this environment, for $k \in [1,2]$ and $s \in [0, \tau - 1] \cap \mathbb{N}$,*

$$
\begin{aligned}
\hat{y}^{e,(k)}_{t+1+s|t-1} =& \hat{y}_{t-1} - (k-1)\omega \frac{\gamma^{\tau-s-1} - \gamma^{\tau+1}}{1-\gamma} \hat{r}^{e,(k)}_{t+\tau|t-1}, \\
\hat{\pi}^{e,(k)}_{t+1+s|t-1} =& (k-1)\kappa \hat{y}^{e,(k)}_{t+s|t-1}.
\end{aligned}
$$

Figure 13 plots the dynamic effects of forward guidance with a -1% shock in interest rate at 1-10 years horizons. The results indicate that forward guidance at shorter horizons have larger initial responses, but those at longer horizons can have larger cumulative effects.

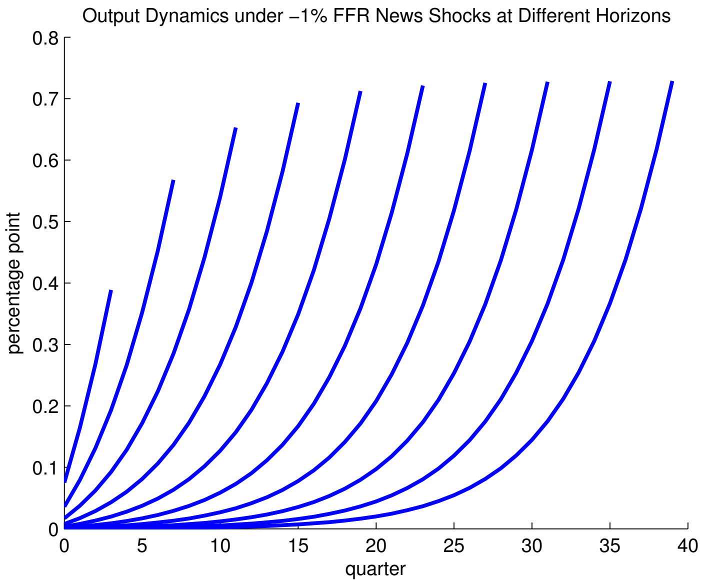

Figure 13: Horizon Effect of Forward Guidance

This result implies that forward guidance can be effective only if the central bank keeps the announcement for a while and if the announcement is too far away in the future, it will be very difficult to detect its effect empirically.

7 Conclusion

This paper establishes a level-k DSGE framework for monetary policy analysis. A recursive level-k equilibrium is established to handle endogenous state variables, and incomplete markets are introduce to discipline households' planning horizons, as well as guarantee eductive stability. The framework is easy to use and has transparent mechanisms.

The model structure and expectation data help identify parameter k. The interaction between households and firms allows us to use output growth and inflation expectations to separate direct and indirect transmission channels. The expectation data support the model implication that indirect channels are missing, while direct channels are weak in households' forecast rules. The formal evidence identifies $k \in [1, 2]$, while the later evidence identifies the exact value of k. Level-k also has data supported implications different from limited attention.

The model performance is evaluated under four alternative monetary policies. Under the Taylor Rule during the Great Moderation, the model well captures the impulse response functions to federal fund rate shocks, and is observationally equivalent to a rational expectations model with external habit, working capital loans and very sticky marginal cost production in terms of output and inflation dynamics. In the pre-Volcker era, the more active GDP targeting rule can partially explain why growth forecasts are more aligned with the backcasts. In the Great Recession, when interest rate is trapped at zero, the model can explain the both the missing deflation, and the missing drop in inflation expectations. It also well captures the stagnant recovery expectations as in the data, and produce very slow recovery under such expectations. Under forward guidance shock, the model is consistent with the consensus that the initial response should be small, but also indicates that forward guidance can be effective if the announcement is kept for a few quarters.

There are still a few related research questions worth further exploring. First, planning horizon has larger effects on equilibrium dynamics under non-rational expectations, but it is still not clear under what general conditions, this part should be explicitly specified. Second, the dual dynamic beauty contests in this paper only provides an example to separate direct and indirect effects in expectation data. Wage stickiness will add one more beauty contest, and provides sharper views for inflation. Third, reasoning in terms of real variables actually involves some level of rationality. The level-k model can be used to deal with this issue by initializing level-0 with nominal anchors. Forth, the level-k framework could be applied to finance related topics in which long horizon expectations play a role, such as bubbles and private money. I leave all these valuable questions for future research.

References

Afrouzi, H. (2017): "Strategic Inattention, Inflation Dynamics and the Non-Neutrality of Money," *Working Paper*.

Angeletos, G.-M. and J. La'O (2013): "Sentiments," *Econometrica*, 81, 739–779.

Angeletos, G.-M. and C. Lian (2018): "Forward Guidance without Common Knowledge," *American Economic Review forthcoming*.

Barsky, R. B. and E. R. Sims (2012): "Information, Animal Spirits, and the Meaning of Innovation in Consumer Confidence," *American Economic Review*, 102, 1343–1377.

Bils, M., P. J. Klenow, and B. A. Malin (2018): "Resurrecting the Role of the Product Market Wedge in Recessions," *American Economic Review*, 108, 1118–1146.

Boivin, J. (2006): "Has U.s. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," *Journal of Money, Credit and Banking*, 38, 1149–1173.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004): "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics*, 119, 861–898.

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

Christiano, L. J., M. S. Eichenbaum, and M. Trabandt (2016): "Unemployment and Business Cycles," *Econometrica*, 84, 1523–1569.

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

Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013): "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications," *Journal of Economic Literature*, 51.

Eusepi, S. and B. Preston (2011): "Expectations, Learning, and Business Cycle Fluctuations," *American Economic Review*, 101, 2844–2872.

Evans, G. W., R. Guesnerie, and B. McGough (2017): "Eductive Stability in Real Business Cycle Models," *Working Paper*.

Evans, G. W. and S. Honkapohja (1998): "Economic Dynamics with Learning: New Stability Results," *Review of Economic Studies*, 65.

Farhi, E. and I. Werning (2017): "Monetary Policy, Bounded Rationality, and Incomplete Markets," *Working Paper*.

Fehr, E. and J.-R. Tyran (2008): "Limitied Rationality and Strategic Interaction: The Impact of the strategic Environment on Nominal Inertia," *Econometrica*, 76.

Gabaix, X. (2014): "A Sparsity-Based Model of Bounded Rationality," *Quarterly Journal of Economics*, 129, 1661–1710.

——— (2017): "Behavioral Macroeconomics Via Sparse Dynamic Programming," *Working Paper*.

Galí, J. (2015): *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press, 2 ed.

García-Schmidt, M. and M. Woodford (2016): "Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis," *Working Paper*.

Gill, D. and V. Prowse (2016): "Cognitive Ability, Charater Skills, and Learning to Play Equilibrium: A Level-k Analysis," *Journal of Political Economy*, 124.

Grandmont, J. M. (1977): "Temporary General Equilibrium Theory," *Econometrica*, 45.

Havranek, T. (2015): "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting," *Journal of European Economic Association*, 13, 1180–1204.

Iovino, L. and D. Sergeyev (2017): "Quantitative Easing without Rational Expectations," *Working Paper*.

Kaplan, G. and G. L. Violante (2014): "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82, 1199–1239.

Kaplan, G., G. L. Violante, and J. Weidner (2014): "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 48, 77–138.

Mankiw, N. G. and R. Reis (2002): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 117, 1295–1328.

Marcet, A. and T. J. Sargent (1989): "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models," *Journal of Economic Theory*, 48.

McKay, A., E. Nakamura, and J. Steinsson (2016): "The Power of Forward Guidance Revisited," *American Economic Review*, forthcoming.

Milani, F. (2007): "Expectations, Learning and Macroeconomic Persistence," *Journal of Monetary Economics*, 54, 2065–2082.

——— (2011): "Expectation Shocks and Learning as Drivers of the Business Cycle," *Economic Journal*, 121, 379–401.

Negro, M. D., M. P. Giannoni, and F. Schorfheide (2015): "Inflation in the Great Recession and New Keynesian Models," *American Economic Journal: Macroeconomics*, 7, 168–196.

Orphanides, A. (2004): "Monetary Policy Rules, Macroeconomic Stability, and Inflation: A View from the Trenches," *Journal of Money, Credit and Banking*, 36, 151–175.

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

Woodford, M. (2003): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps.*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, Princeton: Princeton University Press.

Appendix A: Proofs

Proof of Proposition 1

Suppose $b^{(k)}(\overline{\zeta}, 0, S) > 0$, Assumption 2 implies that $b^{(k)}(1, 0, S) < 0$.
Consider households' budget $c^{(k)}(\zeta, 0, S) = W^{(k)}(S) + D^{(k)}(S) - b^{(k)}(\zeta, 0, S)$.
We must have $c^{(k)}(\overline{\zeta}, 0, S) < c^{(k)}(1, 0, S)$.
This is contradictory to the concavity of perceived and actual value functions.

Proof of Lemma 3

Lemma 3. *In the recursive level-k equilibrium, output satisfies*

$$\hat{y}_t^{(k)} = (1-\gamma)\sum_{s=0}^{+\infty}\gamma^s \hat{y}^{e,(k)}_{t+1+s|t-1} - \omega \sum_{s=0}^{+\infty}\gamma^s(\hat{r}^{e,(k)}_{t+s|t-1} - \hat{\pi}^{e,(k)}_{t+1+s|t-1} + \mathbb{E}_{t-1}\eta^d_{t+s}), \tag{1}$$

where $\gamma = 1 - \sqrt{\frac{\lambda_{\overline{\zeta}|1}\overline{\zeta}}{\lambda_{\overline{\zeta}|1}\overline{\zeta}+1-\lambda_{\overline{\zeta}|1}}}$.

Consider a perfect foresighted optimization problem. With common income $\hat{y}_{t+1}$, the linearized consumption of a just constrained household i in period $t+1$ satisfies

$$\hat{c}^{con}_{i,t+1} = \hat{a}^{unc}_{i,t+1} + \hat{y}_{t+1}. \tag{2}$$

Denote $\lambda = \lambda_{\overline{\zeta}|1}\overline{\zeta}$. The consumption of an unconstrained household i in period t satisfies

$$\hat{c}^{unc}_{i,t} = (1-\lambda)\hat{c}^{unc}_{i,t+1} + \lambda\hat{c}^{con}_{i,t+1} - \omega(\eta^d_t + \hat{r}_t - \hat{\pi}_{t+1}). \tag{3}$$

$\hat{a}^{unc}_{i,t+1}$ for unconstrained household i satisfies

$$\hat{a}^{unc}_{i,t+1} = \hat{a}_{i,t} - \hat{c}^{unc}_{i,t} + \hat{y}_t. \tag{4}$$

Combining equation (2)(3)(4) yields

$$(1+\lambda)\hat{c}^{unc}_{i,t} = (1-\lambda)\hat{c}^{unc}_{i,t+1} + \lambda(\hat{a}_{i,t} + \hat{y}_t + \hat{y}_{t+1}) - \omega(\eta^d_t + \hat{r}_t - \hat{\pi}_{t+1}). \tag{5}$$

Use guess and verify approach to find the expression of $\hat{c}^{unc}_{i,t}$

$$\hat{c}^{unc}_{i,t} = \iota_a \hat{a}_{i,t} + \iota_0 \hat{y}_t + \sum_{s=0}^{+\infty} \gamma^s [\iota_y \hat{y}_{t+s} - \iota_r \omega(\eta^d_{t+s} + \hat{r}_{t+s} - \hat{\pi}_{t+1+s})].$$

$\hat{c}^{unc}_{i,t+1}$ for households unconstrained in period t satisfies

$$\hat{c}^{unc}_{i,t+1} = \iota_a(\hat{a}_{i,t} + \hat{y}_t - \hat{c}^{unc}_{i,t}) + \iota_0 \hat{y}_{t+1} + \sum_{s=0}^{+\infty} \gamma^s [\iota_y \hat{y}_{t+1+s} - \iota_r \omega(\eta^d_{t+1+s} + \hat{r}_{t+1+s} - \hat{\pi}_{t+2+s})].$$

Compare the coefficient of $\hat{a}_t$ in equation (5) to get ι_a

$$(1+\lambda)\iota_a = (1-\lambda)\iota_a(1 - \iota_\ell \iota_a) + \lambda \implies \iota_a = \frac{\sqrt{\lambda}}{1+\sqrt{\lambda}}.$$

Compare the coefficient of $\{\hat{y}_t, \hat{y}_{t+2}, \hat{y}_{t+1}\}$ in equation (5) to get $\iota_0 + \iota_y$, γ and ι_y

$$\begin{aligned}
&(1+\lambda)(\iota_0 + \iota_y) = (1-\lambda)\iota_a(1 - \iota_0 - \iota_y) + \lambda \implies \iota_0 + \iota_y = \iota_a. \\
&(1+\lambda)\gamma^2 \iota_y = (1-\lambda)(-\iota_a \gamma^2 \iota_y + \gamma \iota_y) \implies \gamma = 1 - \sqrt{\lambda}. \\
&(1+\lambda)\gamma \iota_y = (1-\lambda)(-\iota_a \gamma \iota_y + \iota_0 + \iota_y) + \lambda \implies \iota_y = \gamma^{-1}(\iota_0 + \iota_y).
\end{aligned}$$

Compare coefficients on $\eta^d_t + \hat{r}_t - \hat{\pi}_{t+1}$ in equation (5) to get ι_r

$$-(1+\lambda)\iota_r \omega = (1-\lambda)\iota_a \iota_r \omega - \omega \implies \iota_r = 1 - (\iota_0 + \iota_y).$$

In the equilibrium, we have $\hat{y}_t = \hat{c}_{i,t}^{unc}$ and $\hat{a}_{i,t} = 0$ and then

$$\hat{y}_t = (1-\gamma)\sum_{s=0}^{+\infty}\gamma^s \hat{y}_{t+1+s} - \omega \sum_{s=0}^{+\infty}\gamma^s(\eta_{t+s}^d + \hat{r}_{t+s} - \hat{\pi}_{t+1+s}).$$

Now, impose level-k on it and obtain

$$\hat{y}_t^{(k)} = (1-\gamma)\sum_{s=0}^{+\infty}\gamma^s \hat{y}_{t+1+s|t-1}^{e,(k)} - \omega \sum_{s=0}^{+\infty}\gamma^s(\hat{r}_{t+s|t-1}^{e,(k)} - \hat{\pi}_{t+1+s|t-1}^{e,(k)} + \mathbb{E}_{t-1}\eta_{t+s}^d).$$

Proof of Lemma 4

Lemma 4. *In the recursive level-k equilibrium, inflation satisfies*

$$\hat{\pi}_t^{(k)} = (1-\theta)\sum_{s=0}^{+\infty}\theta^s((1-\theta)\kappa_y \hat{y}_{t+s|t-1}^{e,(k)} + \hat{\pi}_{t+s|t-1}^{e,(k)}) - \kappa_z \frac{1-\theta}{1-\theta\rho_z}\eta_t^z,$$

where $(\kappa_y, \kappa_z) = \left(\frac{1-\alpha}{\alpha}\varepsilon - (1-\alpha)(\varepsilon-1), \frac{1-\alpha}{\alpha}\varepsilon + 1\right)$.

The optimal price satisfies

$$\begin{aligned}
\hat{p}_t^a =& (1-\theta)\sum_{s=0}^{+\infty}\theta^s \sum_{\tau=0}^{s}\hat{\pi}_{t+\tau|t-1}^{e,(k)} + (1-\theta)\sum_{s=0}^{+\infty}\theta^s(\kappa_y \hat{y}_{t+s|t-1}^{e,(k)} - \kappa_z \mathbb{E}_{t-1}\eta_{t+s}^z) \\
=& (1-\theta)\sum_{\tau=0}^{+\infty}\sum_{s=\tau}^{+\infty}\theta^s \hat{\pi}_{t+\tau|t-1}^{e,(k)} + (1-\theta)\sum_{s=0}^{+\infty}\theta^s(\kappa_y \hat{y}_{t+s|t-1}^{e,(k)} - \kappa_z \mathbb{E}_{t-1}\eta_{t+s}^z) \\
=& \sum_{s=0}^{+\infty}\theta^s \hat{\pi}_{t+s|t-1}^{e,(k)} + (1-\theta)\sum_{s=0}^{+\infty}\theta^s(\kappa_y \hat{y}_{t+s|t-1}^{e,(k)} - \kappa_z \mathbb{E}_{t-1}\eta_{t+s}^z).
\end{aligned}$$

According to the price aggregator,

$$\hat{\pi}_t = (1-\theta)\hat{p}_t^a.$$

Proof of Lemma 5

Lemma 5. *The expected cumulative effect of interest rate is,*

$$\sum_{s=0}^{+\infty}\gamma^s\hat{r}^{e,(k)}_{t+s|t-1}=\frac{1}{1-\gamma\rho_r}\left[\rho_r\hat{r}_{t-1}+(1-\rho_r)\sum_{s=0}^{+\infty}\gamma^s(\phi_\pi\hat{\pi}^{e,(k)}_{t+s|t-1}+\phi_y\hat{y}^{e,(k)}_{t+s|t-1})\right].$$

The interest rate forecast satisfies

$$\begin{aligned}\hat{r}^{e,(k)}_{t+s|t-1}=&\rho_r\hat{r}^{e,(k)}_{t+s-1|t-1}+(1-\rho_r)(\phi_\pi\hat{\pi}^{e,(k)}_{t+s|t-1}+\phi_y\hat{y}^{e,(k)}_{t+s|t-1})+\sigma_r\mathbb{E}_{t-1}\epsilon^r_{t+s|t-1}\\ =&\rho_r^{s+1}\hat{r}_{t-1}+(1-\rho_r)\sum_{\tau=0}^{s}\rho_r^\tau(\phi_\pi\hat{\pi}^{e,(k)}_{t+s-\tau|t-1}+\phi_y\hat{y}^{e,(k)}_{t+s-\tau|t-1})+\sigma_r\sum_{\tau=0}^{s}\rho_r^\tau\mathbb{E}_{t-1}\epsilon^r_{t+s-\tau}\\ =&\rho_r^{s+1}\hat{r}_{t-1}+(1-\rho_r)\sum_{\tau=0}^{s}\rho_r^{s-\tau}(\phi_\pi\hat{\pi}_{t+\tau|t-1}+\phi_y\hat{y}_{t+\tau|t-1})+\sigma_r\sum_{\tau=0}^{s}\rho_r^{s-\tau}\mathbb{E}_{t-1}\epsilon^r_{t+\tau}.\end{aligned}$$

Use the following identity

$$\begin{aligned}\sum_{s=0}^{+\infty}\gamma^s\sum_{\tau=0}^{s}\rho_r^{s-\tau}=&\sum_{s=0}^{+\infty}\sum_{\tau=0}^{s}\gamma^s\rho_r^{s-\tau}=\sum_{\tau=0}^{+\infty}\sum_{s=\tau}^{+\infty}\gamma^s\rho_r^{s-\tau}\\ =&\frac{1}{1-\gamma\rho_r}\sum_{\tau=0}^{+\infty}\gamma^\tau=\frac{1}{1-\gamma\rho_r}\sum_{s=0}^{+\infty}\gamma^s.\end{aligned}$$

The cumulative effect of interest rate forecasts becomes

$$\begin{aligned}&\sum_{s=0}^{+\infty}\gamma^s\hat{r}^{e,(k)}_{t+s|t-1}\\ =&\sum_{s=0}^{+\infty}\gamma^s\left[\rho_r^{s+1}\hat{r}_{t-1}+(1-\rho_r)\sum_{\tau=0}^{s}\rho_r^{s-\tau}(\phi_\pi\hat{\pi}^{e,(k)}_{t+\tau|t-1}+\phi_y\hat{y}^{e,(k)}_{t+\tau|t-1})+\sigma_r\sum_{\tau=0}^{s}\rho_r^{s-\tau}\mathbb{E}_{t-1}\epsilon^r_{t+\tau}\right]\\ =&\frac{1}{1-\gamma\rho_r}\left[\rho_r\hat{r}_{t-1}+(1-\rho_r)\sum_{s=0}^{+\infty}\gamma^s(\phi_\pi\hat{\pi}^{e,(k)}_{t+s|t-1}+\phi_y\hat{y}^{e,(k)}_{t+s|t-1})+\sigma_r\sum_{s=0}^{+\infty}\gamma^s\mathbb{E}_{t-1}\epsilon^r_{t+s}\right].\end{aligned}$$

Proof of Proposition 2

Proposition 2 can be proved by combining Lemma 3, Lemma 4 and Lemma 5.

Proof of Corollary 1

Change all time index t to $t+1$ in equation (1) and forecast it based on period $t-1$ information.

$$\hat{y}_{t+1}^{(k)} = (1-\gamma)\sum_{s=0}^{+\infty}\gamma^s \hat{y}^{e,(k)}_{(t+2+s|t)|t-1} - \omega\sum_{s=0}^{+\infty}\gamma^s(\hat{r}^{e,(k)}_{(t+1+s|t)|t-1} - \hat{\pi}^{e,(k)}_{(t+2+s|t)t-1} + \mathbb{E}_{t-1}\eta^d_{t+1+s}), \tag{6}$$

Combine equation (1) and (6), and we can get it.

Proof of Proposition 3

Follow Proposition 2 directly.

Proof of Lemma 2

Set $\rho_r = \phi_\pi = \eta^d = \eta^z = 0$ in Proposition 3, and we can get

$$\hat{y}_t^{(k)} = -\omega\phi_y \hat{y}^{e,(k)}_{t|t-1} + \left(1 - \frac{\omega\gamma}{1-\gamma}\phi_y\right)(1-\gamma)\sum_{s=0}^{+\infty}\gamma^s \hat{y}^{e,(k)}_{t+1+s|t-1}. \tag{7}$$

Since the only state variable now is $\hat{y}_{t-1}$, and it must converge the 0 in the long run, the law of motion becomes $\hat{y}_t^{(k)} = \hat{\rho}_y^{(k)}\hat{y}_{t-1}$. When $k \in [1, +\infty) \cap \mathbb{N}$, the perceived law of motion becomes $\hat{y}^{e,(k)}_{t+s|t-1} = (\rho_y^{e,(k)})^{s+1}\hat{y}_{t-1} = (\rho_y^{(k-1)})^{s+1}\hat{y}_{t-1}$.

Substituting the actual and perceived law of motion back to equation (7) yields Lemma 2.

Proof of Proposition 4

When $|\rho_y^{(k-1)}| \leq 1$, $\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}}$ is monotonic in $\rho_y^{(k-1)}$. Hence, we only need to check the bounds.

When $\rho_y^{(k-1)} = 1$, eductive stability requires

$$\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}} = 1 - \frac{\omega}{1-\gamma}\phi_y > -1 \implies \gamma < 1 - \frac{1}{2}\omega\phi_y.$$

When $\rho_y^{(k-1)} = -1$, eductive stability requires

$$\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}} = -\omega\phi_y - \left(1 - \frac{\omega\gamma}{1-\gamma}\phi_y\right)\frac{1-\gamma}{1+\gamma} = -\frac{1-\gamma}{1+\gamma} - \frac{\omega}{1+\gamma}\phi_y > -1 \implies \gamma > \frac{1}{2}\omega\phi_y.$$

Once these conditions are satisfied, $|\frac{\rho_y^{(k)}}{\rho_y^{(k-1)}}| \leq M$ for $M = \max\left\{|1 - \frac{\omega}{1-\gamma}\phi_y|, \frac{1-\gamma}{1+\gamma} + \frac{\omega}{1+\gamma}\phi_y\right\}$.

Proof of Proposition 5

Base on the initialization of level-0, we have

$$\hat{y}^{e,(1)}_{t|t-1} = \hat{y}_{t-1}, \quad \hat{\pi}^{e,(1)}_{t|t-1} = 0.$$

This implies that in $H^{e,(1)}$, output is full anchored. Hence,

$$\hat{y}^{e,(1)}_{t+s|t-1} = \hat{y}_{t-1}.$$

Applied this output forecasts in Proposition 2, we get

$$\begin{aligned}
\hat{y}^{e,(2)}_{t|t-1} - \hat{y}_{t-1} =& \omega\left(-\frac{1-\rho_r}{1-\gamma\rho_r}\frac{\phi_y}{1-\gamma}\hat{y}_{t-1} - \frac{\rho_r}{1-\gamma\rho_r}\hat{r}_{t-1} - \frac{1}{1-\gamma\rho_d}\eta^d_t\right), \\
\hat{\pi}^{e,(2)}_{t|t-1} =& \kappa\left(\hat{y}_{t-1} - \omega\frac{1-\theta}{1-\theta\rho_z}\eta^z_t\right).
\end{aligned}$$

For $k \in [1,2]$, $(\hat{y}^{e,(k)}_{t|t-1}, \hat{\pi}^{e,(k)}_{t|t-1}) = (2-k)(\hat{y}^{e,(1)}_{t|t-1}, \hat{\pi}^{e,(1)}_{t|t-1}) + (k-1)(\hat{y}^{e,(2)}_{t|t-1}, \hat{\pi}^{e,(2)}_{t|t-1})$.

Proof of Corollary 2

When $\eta_t^d = \eta_t^z = 0$ and $k \in [1, 2]$, the perceived aggregate law of motion becomes

$$\begin{bmatrix} \hat{y}^{e,(k)}_{t|t-1} \\ \hat{r}^{e,(k)}_{t|t-1} \end{bmatrix} = \begin{bmatrix} 1 & -(k-1)\frac{\omega\rho_r}{1-\gamma\rho_r} \\ (k-1)(1-\rho_r)[(1-\theta)(\omega^{-1}+\xi^{-1})\phi_\pi + \phi_y] & \rho_r - (k-1)\frac{\omega\rho_r}{1-\gamma\rho_r}(1-\rho_r)\phi_y \end{bmatrix} \begin{bmatrix} \hat{y}_{t-1} \\ \hat{r}_{t-1} \end{bmatrix}.$$

Denote the perceived law of motion as $\hat{h}^{e,(k)}$, then

$$\begin{bmatrix} \hat{y}^{e,(k)}_{t+s|t-1} \\ \hat{r}^{e,(k)}_{t+s|t-1} \end{bmatrix} = (\hat{h}^{e,(k)})^{s+1} \begin{bmatrix} \hat{y}_{t-1} \\ \hat{r}_{t-1} \end{bmatrix}.$$

Hence, $\hat{y}^{e,(k)}_{t+3|t-1}$ can be obtained directly, and $\hat{\pi}^{e,(k)}_{t+3|t-1}$ can be obtained from

$$\hat{\pi}^{e,(k)}_{t+3|t-1} = (k-1)\kappa\hat{y}^{e,(k)}_{t+2|t-1}.$$

Proof of Proposition 6 and Proposition 8

Following Proposition 2 directly.

Proof of Proposition 7

Consider an external habit model with no uncertainty

$$\begin{aligned} \hat{y}_t - \hat{y}_{t-1} =& \frac{1 - wy_{ss}(1-\omega)}{h}(\hat{y}_{t+1} - \hat{y}_t) - \frac{1-h}{h}\omega(\hat{r}_t - \hat{\pi}_{t+1}), \\ \hat{\pi}_t =& \frac{(1-\theta)(1-\beta_f\theta)}{\theta}\left[\psi\left(\frac{1}{1-h}\hat{y}_t - \frac{h}{1-h}\hat{y}_{t-1}\right) + \hat{r}_t\right] + \hat{\pi}_{t+1}, \\ \hat{r}_t =& \rho_r\hat{r}_{t-1} + (1-\rho_r)(\phi_\pi\hat{\pi}_t + \phi_y\hat{y}_t). \end{aligned}$$

Suppose the actual law of motion is identical to the level-k model, then

$$\Gamma^{ys,(k)} - \Gamma^{y} = \frac{1 - wy_{ss}(1-\omega)}{h}(\Gamma^{ys,(k)} - \Gamma^{y})\Gamma^{ss,(k)} - \frac{1-h}{h}\omega(\Gamma^{r} - \Gamma^{\pi s,(k)})\Gamma^{ss,(k)},$$
$$\Gamma^{\pi s,(k)} = \frac{(1-\theta)(1-\beta_f\theta)}{\theta}\left[\psi\left(\frac{1}{1-h}\Gamma^{ys,(k)} - \frac{h}{1-h}\Gamma^{y}\right) + \Gamma^{rs,(k)}\right] + \beta_f\Gamma^{\pi s,(k)}\Gamma^{ss,(k)}.$$

The new set of parameters can be solved as

$$\begin{bmatrix} \frac{1-wy_{ss}(1-\omega)}{h} & \frac{1-h}{h}\omega \end{bmatrix} = (\Gamma^{ys,(k)} - \Gamma^{y})(\Gamma^{ss,(k)})^{-1}\begin{bmatrix} \Gamma^{ys,(k)} - \Gamma^{y} \\ \Gamma^{\pi s,(k)} - \Gamma^{r} \end{bmatrix}^{-1},$$
$$\begin{bmatrix} \frac{(1-\theta)(1-\beta_f\theta)}{\theta}\psi & \frac{(1-\theta)(1-\beta_f\theta)}{\theta} \end{bmatrix} = \Gamma^{\pi s,(k)}(\mathbb{I} - \beta_f\Gamma^{ss,(k)})\begin{bmatrix} \frac{1}{1-h}\Gamma^{ys,(k)} - \frac{h}{1-h}\Gamma^{y} \\ \Gamma^{rs,(k)} \end{bmatrix}^{-1}.$$

Proof of Proposition 9

When nominal interest rate becomes zero, it declines π_{SS} compared to the original steady state. As the natural rate of interest changes by $\hat{r}^n$, the interest rate gap becomes $\hat{r}_t = -(\pi_{SS} + \hat{r}^n)$ in the liquidity trap. Neither level-0 nor level-1 households react to expectations on future.

$$\hat{y}^{e,(2)}_{t+s|t-1} = \hat{y}_{t-1} + (s+1)\frac{\omega}{1-\gamma}(\hat{\pi}_{SS} + \hat{r}^n),$$
$$\hat{\pi}^{e,(2)}_{t+s|t-1} = \kappa\hat{y}^{e,(2)}_{t-1+s|t-1}.$$

For $k \in [1,2]$, $(\hat{y}^{e,(k)}_{t|t-1}, \hat{\pi}^{e,(k)}_{t|t-1}) = (2-k)(\hat{y}^{e,(1)}_{t|t-1}, \hat{\pi}^{e,(1)}_{t|t-1}) + (k-1)(\hat{y}^{e,(2)}_{t|t-1}, \hat{\pi}^{e,(2)}_{t|t-1})$.

Proof of Proposition 10

The proof is identical to Proposition 9 except that

$$\hat{y}^{e,(2)}_{t+1+s|t-1} = \hat{y}_{t-1} - \omega\sum_{\nu=0}^{s+1}\gamma^{\tau-\nu}\hat{r}^{e,(k)}_{t+\tau|t-1} = \hat{y}_{t-1} - \omega\frac{\gamma^{\tau-s-1} - \gamma^{\tau+1}}{1-\gamma}\hat{r}^{e,(k)}_{t+\tau|t-1}.$$

Appendix B: Solving the Model with Endogenous Labor Supply

The standard solution procedure for rational expectations DSGE models cannot be directly applied here. Hence, it is useful to describe how to write the model into a state space form. Use Γ to denote the coefficients in linearized equilibrium objects, and the solution procedure can be briefly described in the following.

1. Solve for $\Gamma^{ca,e}$ without using equilibrium objects.
2. Initialize $(\Gamma^{ys,e,(1)}, \Gamma^{\pi s,e,(1)}, \Gamma^{ws,e,(1)}, \Gamma^{\tau s,e,(1)})$ from level-0, and obtain $\Gamma^{ss,e,(1)}$.
3. Solve for $\Gamma^{cs,e,(1)}$ from the perceived households' problem.
4. Solve for $(\Gamma^{ys,(1)}, \Gamma^{\ell s,(1)}, \Gamma^{ws,(1)}, \Gamma^{\tau s,(1)})$ from the temporary equilibrium.
5. Solve for $\Gamma^{\pi s,(1)}$ from the firms' problem, and obtain $\Gamma^{ss,(1)}$.
6. Use $(\Gamma^{ys,e,(j+1)}, \Gamma^{\pi s,e,(j+1)}, \Gamma^{ws,e,(j+1)}, \Gamma^{\tau s,e,(j+1)}) = (\Gamma^{ys,(j)}, \Gamma^{\pi s,(j)}, \Gamma^{ws,(j)}, \Gamma^{\tau s,(j)})$ to update.
7. Obtain the state space representation.

Step 1: Solve for $\Gamma^{ca,e}$

Log-linearizing the optimality conditions for the constrained households yields

$$\begin{aligned}\omega^{-1}\hat{c}^e(\bar{\zeta}) =& \hat{w}^e - \xi^{-1}\hat{\ell}^e(\bar{\zeta}),\\ \hat{c}^e(\bar{\zeta}) =& \hat{a} + wy_{SS}(\hat{w}^e + \hat{\tau}^e + \hat{\ell}^e(\bar{\zeta})).\end{aligned}$$

$(\Gamma^{ca,e}(\bar{\zeta}), \Gamma^{\ell a,e}(\bar{\zeta}))$ can be obtained from

$$\begin{bmatrix}\omega^{-1} & \xi^{-1}\\ 1 & -wy_{SS}\end{bmatrix}\begin{bmatrix}\Gamma^{ca,e}(\bar{\zeta})\\ \Gamma^{\ell a,e}(\bar{\zeta})\end{bmatrix} = \begin{bmatrix}0\\ 1\end{bmatrix}.$$

The solution is $\Gamma^{ca,e}(\bar{\zeta}) = \frac{1}{1+\xi\omega^{-1}wy_{SS}}$.

The unconstrained households have

$$
\begin{aligned}
\hat{c}^e(1) =& -\omega(\eta^d + \hat{r}^e - \hat{\pi}^{e\prime}) + \underline{\lambda}\hat{c}^{e\prime}(1) + (1-\underline{\lambda})\hat{c}^{e\prime}(\overline{\zeta}),\\
\hat{a}^{e\prime}(1) =& (R_{SS}/\Pi_{SS})[\hat{a} + wy_{SS}(\hat{w}^e + \hat{\tau}^e + \hat{\ell}^e(1)) - \hat{c}^e(1)],\\
\hat{\ell}^e(1) =& \xi\hat{w}^e - \xi\omega^{-1}\hat{c}^e(1).
\end{aligned}
$$

$(\Gamma^{ca,e}(1), \Gamma^{\ell a,e}(1), \Gamma^{aa,e}(1))$ satisfy

$$
\begin{aligned}
\Gamma^{ca,e}(1) =& [\underline{\lambda}\Gamma^{ca,e}(1) + (1-\underline{\lambda})\Gamma^{ca,e}(\overline{\zeta})]\Gamma^{aa,e}(1),\\
\Gamma^{aa,e}(1) =& (R_{SS}/\Pi_{SS})(1 + wy_{SS}\Gamma^{\ell a,e}(1) - \Gamma^{ca,e}(1)),\\
\Gamma^{\ell a,e}(1) =& -\xi\omega^{-1}\Gamma^{ca,e}(1).
\end{aligned}
$$

This yields a quadratic function for $\Gamma^{ca,e}(1)$

$$
(\Pi_{SS}/R_{SS})\Gamma^{ca,e}(1) = [\underline{\lambda}\Gamma^{ca,e}(1) + (1-\underline{\lambda})\Gamma^{ca,e}(\overline{\zeta})][1 - (1 + \xi\omega^{-1}wy_{SS})\Gamma^{ca,e}(1)].
$$

Solving $\underline{\lambda}$ from $\Gamma^{ca,e}(1)$ yields

$$
\underline{\lambda} = \frac{\Gamma^{ca,e}(\overline{\zeta}) - \Gamma^{ca,e}(1)\frac{(\Pi_{SS}/R_{SS})}{1-(1+\xi\omega^{-1}wy_{SS})\Gamma^{ca,e}(1)}}{\Gamma^{ca,e}(\overline{\zeta}) - \Gamma^{ca,e}(1)}.
$$

The notation $\underline{\lambda} = \frac{\lambda_1}{\lambda_1 + (1-\lambda_1)\overline{\zeta}}$ yields

$$
\lambda_1 = [1 + (\underline{\lambda}^{-1} - 1)\overline{\zeta}^{-1}]^{-1}.
$$

The fraction of hand-to-mouth households λ_{HtM} satisfies

$$
\begin{aligned}
\lambda_{HtM} =& (1-\lambda_1)/(2-\lambda_1-\lambda_2),\\
\lambda_2 =& 1 - (1-\lambda_1)(1-\lambda_{HtM})/\lambda_{HtM}.
\end{aligned}
$$

Step 2: Initialize $(\Gamma^{ys,e,(1)}, \Gamma^{\pi s,e,(1)}, \Gamma^{ws,e,(1)}, \Gamma^{\tau s,e,(1)})$ and $\Gamma^{ss,e,(1)}$

Specify the level-1 expectations.

$$
\begin{aligned}
\Gamma^{ys,e,(1)}\hat{s} =&\hat{y}_-, \\
\Gamma^{\pi s,e,(1)}\hat{s} =&0, \\
\Gamma^{\ell s,e,(1)}\hat{s} =&\hat{y}_- - \eta^z, \\
\Gamma^{ws,e,(1)} =&\omega^{-1}\Gamma^{ys,e,(1)} + \xi^{-1}\Gamma^{\ell s,e,(1)}, \\
\Gamma^{\tau s,e,(1)} =&wy_{SS}^{-1}\Gamma^{ys,e,(1)} - \Gamma^{\ell s,e,(1)} - \Gamma^{ws,e,(1)}.
\end{aligned}
$$

According to the perceived Taylor Rule,

$$
\Gamma^{rs,e,(1)} = \rho_r\Gamma^r + (1-\rho_r)(\phi_\pi\Gamma^{\pi s,e,(1)} + \phi_y\Gamma^{ys,e,(1)}).
$$

The state variable is $\hat{s} = (\hat{y}_-, \hat{r}_-, \eta^d, \eta^z)$. $\Gamma^{ss,e,(1)}$ can be obtained from $(\Gamma^{ys,e,(1)}, \Gamma^{rs,e,(1)})$ and the exogenous law of motion for (η^d, η^z).

Step 3: Solve for $\Gamma^{cs,e,(1)}$

Recall the optimality conditions of the constrained households

$$
\begin{aligned}
\omega^{-1}\hat{c}^e(\bar{\zeta}) =&\hat{w}^e - \xi^{-1}\hat{\ell}^e(\bar{\zeta}), \\
\hat{c}^e(\bar{\zeta}) =&\hat{a} + wy_{SS}(\hat{w}^e + \hat{\tau}^e + \hat{\ell}^e(\bar{\zeta})).
\end{aligned}
$$

$(\Gamma^{cs,e,(1)}(\bar{\zeta}), \Gamma^{\ell s,e,(1)}(\bar{\zeta}))$ can be obtained from

$$
\begin{bmatrix} \omega^{-1} & \xi^{-1} \\ 1 & -wy_{SS} \end{bmatrix}
\begin{bmatrix} \Gamma^{cs,e,(1)}(\bar{\zeta}) \\ \Gamma^{\ell s,e,(1)}(\bar{\zeta}) \end{bmatrix}
=
\begin{bmatrix} \Gamma^{ws,e,(1)} \\ wy_{SS}(\Gamma^{ws,e,(1)} + \Gamma^{\tau s,e,(1)}) \end{bmatrix}.
$$

Recall the optimality conditions of the unconstrained households

$$\begin{aligned}
\hat{c}^e(1) =& -\omega(\eta^d + \hat{r}^e - \hat{\pi}^{e\prime}) + \underline{\lambda}\hat{c}^{e\prime}(1) + (1-\underline{\lambda})\hat{c}^{e\prime}(\overline{\zeta}),\\
\hat{a}^{e\prime}(1) =&(R_{SS}/\Pi_{SS})[\hat{a} + wy_{SS}(\hat{w}^e + \hat{\tau}^e + \hat{\ell}^e(1)) - \hat{c}^e(1)],\\
\hat{\ell}^e(1) =&\xi\hat{w}^e - \xi\omega^{-1}\hat{c}^e(1).
\end{aligned}$$

$(\Gamma^{cs,e,(1)}(1), \Gamma^{\ell s,e,(1)}(1), \Gamma^{as,e,(1)}(1))$ satisfy

$$\begin{aligned}
\Gamma^{cs,e,(1)}(1) =& -\omega(\Gamma^{ds} + \Gamma^{rs,e,(1)} - \Gamma^{\pi s,e,(1)}\Gamma^{ss,e,(1)})\\
&+ [\underline{\lambda}\Gamma^{ca,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{ca,e,(1)}(\overline{\zeta})]\Gamma^{as,e,(1)}(1),\\
&+ [\underline{\lambda}\Gamma^{cs,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{cs,e,(1)}(\overline{\zeta})]\Gamma^{ss,e,(1)},\\
\Gamma^{as,e,(1)}(1) =&(R_{SS}/\Pi_{SS})[wy_{SS}(\Gamma^{ws,e,(1)} + \Gamma^{\tau s,e,(1)} + \Gamma^{\ell s,e,(1)}(1)) - \Gamma^{cs,e,(1)}(1)],\\
\Gamma^{\ell s,e,(1)}(1) =&\xi\Gamma^{ws,e,(1)} - \xi\omega^{-1}\Gamma^{cs,e,(1)}(1).
\end{aligned}$$

Eliminate $(\Gamma^{\ell s,e,(1)}(1), \Gamma^{as,e,(1)}(1))$ to obtain a single equation of $\Gamma^{cs,e,(1)}(1)$

$$\begin{aligned}
\Gamma^{cs,e,(1)}(1) =& -\omega(\Gamma^{ds} + \Gamma^{rs,e,(1)} - \Gamma^{\pi s,e,(1)}\Gamma^{ss,e,(1)})\\
&+ [\underline{\lambda}\Gamma^{ca,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{ca,e,(1)}(\overline{\zeta})]\\
&\cdot (R_{SS}/\Pi_{SS})\{wy_{SS}[(1+\xi)\Gamma^{ws,e,(1)} + \Gamma^{\tau s,e,(1)}] - (1+\xi\omega^{-1}wy_{SS})\Gamma^{cs,e,(1)}(1)\},\\
&+ [\underline{\lambda}\Gamma^{cs,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{cs,e,(1)}(\overline{\zeta})]\Gamma^{ss,e,(1)}.
\end{aligned}$$

The solution for $\Gamma^{cs,e,(1)}(1)$ is

$$\begin{aligned}
\Gamma^{cs,e,(1)}(1) =&\{(R_{SS}/\Pi_{SS})wy_{SS}[\underline{\lambda}\Gamma^{ca,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{ca,e,(1)}(\overline{\zeta})][(1+\xi)\Gamma^{ws,e,(1)} + \Gamma^{\tau s,e,(1)}]\\
&+ (1-\underline{\lambda})\Gamma^{cs,e,(1)}(\overline{\zeta})\Gamma^{ss,e,(1)} - \omega(\Gamma^{ds} + \Gamma^{rs,e,(1)} - \Gamma^{\pi s,e,(1)}\Gamma^{ss,e,(1)})\}\\
&\{\mathbb{I} + (R_{SS}/\Pi_{SS})(1+\xi\omega^{-1}wy_{SS})[\underline{\lambda}\Gamma^{ca,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{ca,e,(1)}(\overline{\zeta})] * \mathbb{I} - \underline{\lambda}\Gamma^{ss,e,(1)}\}^{-1}.
\end{aligned}$$

Step 4: Solve for $(\Gamma^{ys,(1)}, \Gamma^{\ell s,(1)}, \Gamma^{ws,(1)}, \Gamma^{\tau s,(1)})$

The temporary equilibrium satisfies

$$\begin{aligned}
\hat{y} =& -\omega(\eta^d + \hat{r}^e - \hat{\pi}^{e\prime}) + \underline{\lambda}\hat{c}^{e\prime}(1) + (1-\underline{\lambda})\hat{c}^{e\prime}(\bar{\zeta}),\\
\hat{\ell} =& \hat{y} - \eta^z,\\
\hat{w} =& \xi^{-1}\hat{\ell} + \omega^{-1}\hat{y},\\
\hat{\tau} =& wy_{SS}^{-1}\hat{y} - \hat{\ell} - \hat{w}.
\end{aligned}$$

$(\Gamma^{ys,(1)}, \Gamma^{\ell s,(1)}, \Gamma^{ws,(1)}, \Gamma^{\tau s,(1)})$ can be obtained from

$$\begin{aligned}
\Gamma^{ys,(1)} =& -\omega(\Gamma^d + \Gamma^{rs,e,(1)} - \Gamma^{\pi s,e,(1)}\Gamma^{ss,e,(1)}) + [\underline{\lambda}\Gamma^{cs,e,(1)}(1) + (1-\underline{\lambda})\Gamma^{cs,e,(1)}(\bar{\zeta})]\Gamma^{ss,e,(1)},\\
\Gamma^{\ell s,(1)} =& \Gamma^{ys,(1)} - \Gamma^z,\\
\Gamma^{ws,(1)} =& \xi^{-1}\Gamma^{\ell s,(1)} + \omega^{-1}\Gamma^{ys,(1)},\\
\Gamma^{\tau s,(1)} =& wy_{SS}^{-1}\Gamma^{ys,(1)} - \Gamma^{\ell s,e,(1)} - \Gamma^{ws,(1)}.
\end{aligned}$$

Step 5: Solve for $\Gamma^{\pi s,(1)}$ and Obtain $\Gamma^{ss,(1)}$

Denote $\beta_f = \Pi_{SS}/R_{SS}$ The linearized Phillips Curve with arbitrary expectations $\widetilde{\mathbb{E}}_t$ is

$$\hat{\pi}_t = (1-\theta)(1-\beta_f\theta)\sum_{s=0}^{\infty}(\beta_f\theta)^s\widetilde{\mathbb{E}}_t(\hat{w}_{t+s} - \eta^z_{t+s}) + (1-\theta)\sum_{s=0}^{\infty}(\beta_f\theta)^s\widetilde{\mathbb{E}}_t\hat{\pi}_{t+s}.$$

The matrix representation for $(\Gamma^{\pi s,(1)}, \Gamma^{rs,(1)})$ is

$$\begin{aligned}
\Gamma^{\pi s,(1)} =& (1-\theta)\left\{[(1-\beta_f\theta)\Gamma^{ws,e,(1)} + \Gamma^{\pi s,e,(1)}](\mathbb{I} - \beta_f\theta\Gamma^{ss,e,(1)})^{-1} - \frac{1-\beta_f\theta}{1-\beta_f\theta\rho_z}\Gamma^z\right\},\\
\Gamma^{rs,(1)} =& \rho_r\Gamma^r + (1-\rho_r)(\phi_\pi\Gamma^{\pi s,(1)} + \phi_y\Gamma^{ys,(1)}).
\end{aligned}$$

$\Gamma^{ss,(1)}$ can be obtained from $(\Gamma^{ys,(1)}, \Gamma^{rs,(1)})$ and the exogenous law of motion for (η^d, η^z).

Step 6: Update Expectations

For $\forall k \in [1, +\infty)$, first update expectations to $[k]$ using

$$\left(\Gamma^{ys,e,(j+1)}, \Gamma^{\pi s,e,(j+1)}, \Gamma^{ws,e,(j+1)}, \Gamma^{\tau s,e,(j+1)}\right) = \left(\Gamma^{ys,(j)}, \Gamma^{\pi s,(j)}, \Gamma^{ws,(j)}, \Gamma^{\tau s,(j)}\right).$$

Level-k expectations are defined as

$$\text{level-k} = (1 - k + [k]) \cdot \text{level-[k]} + ([k] - k) \cdot \text{level-[k+1]}.$$

Step 7: State Space Representation

The transition equation is

$$\hat{s}_{t+1} = \begin{bmatrix} \hat{y}_t \\ \hat{r}_t \\ \hat{\eta}^d_{t+1} \\ \hat{\eta}^z_{t+1} \end{bmatrix} = \Gamma^{ss,(k)} \begin{bmatrix} \hat{y}_{t-1} \\ \hat{r}_{t-1} \\ \hat{\eta}^d_t \\ \hat{\eta}^z_t \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma_m \epsilon^r_t \\ \sigma_d \epsilon^d_t \\ \sigma_z \epsilon^z_t \end{bmatrix}.$$

The measurement equation is

$$\begin{bmatrix} \hat{y}_{t+1} - \hat{y}_t \\ \hat{p}_{t+1} - \hat{p}_t \\ \hat{r}_t \end{bmatrix} = \begin{bmatrix} \Gamma^{ys} - \Gamma^y \\ \Gamma^{\pi s} \\ \Gamma^r \end{bmatrix} \hat{s}_{t+1}.$$

The expectation equations and ex post counterparts are

$$\begin{bmatrix} \hat{y}^e_{t+4} - \hat{y}_t \\ \hat{p}^e_{t+4} - \hat{p}_t \\ \hat{y}_{t+4} - \hat{y}_t \\ \hat{p}_{t+4} - \hat{p}_t \end{bmatrix} = \begin{bmatrix} \Gamma^{ys,e,(k)}\left(\Gamma^{ss,e,(k)}\right)^3 - \Gamma^y \\ \Gamma^{\pi s,e,(k)} \sum_{\tau=0}^3 \left(\Gamma^{ss,e,(k)}\right)^\tau \\ \Gamma^{ys,(k)}\left(\Gamma^{ss,(k)}\right)^3 - \Gamma^y \\ \Gamma^{\pi s,(k)} \sum_{\tau=0}^3 \left(\Gamma^{ss,(k)}\right)^\tau \end{bmatrix} \hat{s}_{t+1}.$$