Sticky information across the wealth distribution *

Tommaso Di Francesco
t.difrancesco@uva.nl

November 2, 2024

MOST RECENT VERSION HERE

Abstract

This paper investigates the role of wealth-dependent information stickiness in the transmission of monetary policy in a Heterogeneous Agent New Keynesian (HANK) model. Using survey data, I provide empirical evidence that households do not form expectations according to the full-information rational expectations (FIRE) hypothesis but instead exhibit stickiness in updating their information, with wealthier households updating more frequently. I evaluate the effect of this evidence on macroeconomic dynamics using a quantitative model. My findings reveal that inequality significantly affects the aggregate responses to monetary shocks. Specifically, models that neglect heterogeneity in information updating underestimate both the magnitude and the delay of the peak response to monetary policy shocks. Estimating the model by matching simulated impulse response functions (IRFs) to empirical ones shows that stickiness is crucial for accurately capturing the dynamics observed in the data.

1 Introduction

A recently expanding literature in macroeconomics focuses on building models that are consistent not only with macro aggregates but also with micro evidence. Empirically, the impulse response to identified monetary shocks displays a *hump shape* (Romer and Romer, 2004). On the micro side, the literature has instead documented elevated and peaked on impact marginal propensities to consume (MPCs) out of transitory income shocks (Fagereng, Holm and Natvik, 2021).

Models that focus on matching aggregate evidence have typically used New Keynesian models with a representative agent (RANK). To account for the delayed response of macro aggregates to shocks, these models are usually enriched with habits in the consumption behavior of households and frictions, such as sticky prices, on the firm side as in Christiano, Eichenbaum and Evans (2005) or Smets and Wouters (2007). An alternative approach has been to introduce deviations from the full-information rational expectations (FIRE) assumption. The main motif used by the literature has been that of rational inattention, as in Carroll (2003), which argues that agents may face an information processing constraint and therefore not be able to process all available information immediately. Maćkowiak and Wiederholt (2009), Maćkowiak and Wiederholt (2010), and Zorn (2021) have shown that RANK models with some form of rational inattention can match the hump shape of the impulse response function.

*This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956107, "Economic Policy in Complex Environments (EPOC)".

To deal with micro-side evidence, the literature has instead focused on models with agent heterogeneity, incomplete markets, and idiosyncratic productivity risk, giving rise to wealth heterogeneity. The incorporation of these elements with New Keynesian models has given rise to the so-called Heterogeneous Agent New Keynesian (HANK) models. While these models, as in McKay, Nakamura and Steinsson (2016) and Kaplan, Moll and Violante (2018), have added realism and a richer household side, their implications for monetary policy at the aggregate level are remarkably similar to the RANK models.

Some recent papers have started to merge elements from the two traditions but have done so in isolation from each other, assuming that the expectation formation process, even if not fully rational, is uniform across the population and orthogonal to the state space. This paper links the expectation formation process to wealth and studies the implications of this assumption for monetary policy transmission in a HANK model. The type of bounded rationality I use is the sticky information framework, as in Mankiw and Reis (2002) and Carroll et al. (2020). I use survey data, following the approach of Coibion and Gorodnichenko (2015), to show that households do not form their beliefs according to the FIRE assumption but display *stickiness* in incorporating new information. I then provide evidence that this stickiness is lower in households with higher wealth.

I use the HANK model with sticky expectations to revisit the mechanisms through which monetary policy is transmitted to the economy. First, by using a simple model, I show that inequality now matters for the transmission of monetary policy even at the aggregate level, and I show that using a RANK model with sticky expectations misses not only the size but also the timing of the peak response. Second, I quantify the effect of this departure from rationality by estimating a quantitative model by matching simulated impulse responses to empirical ones for an identified Aruoba and Drechsel (2022) monetary policy shock. I show that stickiness is necessary to match the empirical estimates and that the FIRE version of the same model would overestimate the IRF by 50% on impact and 70% cumulatively.

Layout. The rest of the paper is structued as follows. Section 2 presents the empirical evidence of wealth dependent stickiness. Section 3 presents the simple model and highlights the main mechanisms at play. Section 4 presents the quantitative model and the estimation exercise. Section 5 concludes.

2 From uniform to wealth dependent stickiness

I begin by laying out the classical sticky information framework. The economy is populated by a continuum of agents on the unit interval. Their forecast regarding a macroeconomic variable is given by

$$F_t(x_{t+h}) = \begin{cases} \mathbb{E}_t(x_{t+h}) & \text{with p} = (1-\theta) \\ F_{t-1}(x_{t+h}) & \text{with p} = \theta \end{cases} \tag{1}$$

Equation (1) says that in each period an agent updates their information set with probability $(1-\theta)$, in which case they have rational expectations. With probability θ, they do not update their information set and stick with their last forecast. The economy as a whole has expectations that, on average, are given by

$$F_t(x_{t+h}) = \int_0^1 (1-\theta)\mathbb{E}_t(x_{t+h})di + \int_0^1 \theta F_{t-1}(x_{t+h})di = (1-\theta)\mathbb{E}_t(x_{t+h}) + \theta F_{t-1}(x_{t+h}). \quad (2)$$

Note that this holds because θ and hence $F_t(x_{t+h})$ are independent of i and so constant across the population. To test whether FIRE are met in survey data one can then use the fact that a rational forecast needs to satisfy

$$\mathbb{E}_t(x_{t+h}) = x_{t+h} - e_{t+h,t}, \quad (3)$$

with $e_{t+h,t}$ being the forecasting error and uncorrelated with information dated t or earlier. Combining equations (2) and (3) one gets

$$F_t(x_{t+h}) = (1-\theta)(x_{t+h} - e_{t+h,t}) + \theta F_{t-1}(x_{t+h}), \quad (4)$$

and rearranging

$$\frac{F_t(x_{t+h})}{1-\theta} = x_{t+h} - e_{t+h,t} + \frac{\theta}{1-\theta}F_{t-1}(x_{t+h}), \quad (5)$$

that is

$$F_t(x_{t+h}) + \frac{\theta}{1-\theta}F_t(x_{t+h}) = x_{t+h} - e_{t+h,t} + \frac{\theta}{1-\theta}F_{t-1}(x_{t+h}), \quad (6)$$

and finally

$$x_{t+h} - F_t(x_{t+h}) = \frac{\theta}{1-\theta}(F_t(x_{t+h}) - F_{t-1}(x_{t+h})) + e_{t+h,t}. \quad (7)$$

Equation (7) is testable on data since one can regress the forecasting error on the forecast revision $(F_t(x_{t+h}) - F_{t-1}(x_{t+h}))$. This can be done using Ordinary Least Squares (OLS) if the survey elicits forecasts at different moments for the same variable. The null hypothesis of FIRE is met when the coefficient associated with the forecast revision is 0, which would imply $\theta = 0$ and no degree of stickiness.

I proceed by first replicating the baseline estimation in Coibion and Gorodnichenko (2015) using data from the Survey of Professional Forecasters (SPF), the Livingston Survey (LS), and the Michigan Survey of Consumers (MSC). I report the specific questions asked in each survey in appendix A. In the first, the target variable is GDP/GNP inflation forecast for horizons starting at the current quarter to four quarters ahead. For the LS, agents are asked to forecast the CPI at six and twelve months ahead. The LS elicits forecasts at 6 and 12 months ahead, therefore t is the current semester. The SPF elicits forecasts for the next 4 quarters, therefore t is the current quarter. For both datatsets equation (7) can be estimated directly. The MSC has monthly waves, but it is not possible to directly obtain forecast revisions, since individuals are only asked to forecast the percentage change in prices for the next 12 months. Specifically, the regression that one could run is

$$x_{t+h} - F_t(x_{t+h}) = c + \beta(F_t(x_{t+h}) - F_{t-1}(x_{t+h-1})) + \text{error}_t, \tag{8}$$

but now notice that error_t includes not only the rational forecast error but also the term $\beta(F_{t-1}(x_t) - F_{t-1}(x_{t+h}))$ and is thus not orthogonal to the regressor. As a result, this model can not be estimated by OLS. In this case Coibion and Gorodnichenko (2015) suggests using oil price innovations as an instrument for the forecast revision. Oil price innovations are computed log-differences in the oil price. The estimates for the MSC are then obtained by two-stage least squares (2SLS) with quarterly frequency. To compute the forecast error for all surveys I use the first release of the target variable from the Federal Reserve Bank of Philadelphia Real-Time Data Set, as suggested by Croushore and Stark (2003). Table 1 reports the results.

Table 1: Information Rigidity for Households and Professionals

	LS	**SPF**	**MSC**
Constant	0.25	-0.06	-1.04
	(0.16)	(0.76)	(0.00)
Forecast Revision	0.92	0.95	1.21
	(0.00)	(0.01)	(0.00)
Sample	1969:6 - 2020:12	1969:1 - 2022:4	1982:2 - 2023:3
Observations	103	203	166
Instrument	-	-	Oil Price Innovations
F stat first stage	-	-	87.31

Note: *The table reports estimates of equation (7) using OLS for the Livingston Survey (LS), the Survey of Professional Forecasters (SPF), and the 2SLS for the Michigan Survey of Consumers (MSC). P-values obtained from Newye-West robust standard errors are in parentheses.*

From equation (7), one can directly retrieve the degree of stickiness from the estimated regression coefficient as $\theta = \frac{\hat{\beta}}{1+\hat{\beta}}$, resulting in $\theta = 0.48$ in the LS, $\theta = 0.49$ in the SPF, and $\theta = 0.50$ in the MSC. As discussed this result is derived under the assumption of homoscedasticity of the updating probability across the population.

In what follows I am going to provide evidence of heterogeneity in the updating probability across the wealth distribution. The MSC provides demographic information on multiple dimensions, including education level, age of the respondents, and stock holdings. Although the latter is not a direct measure of wealth, it is likely highly correlated with it[1]. My first step is to partition the dataset into different demographic groups and then estimate equation (7) for each group. For this analysis I move to a monthly frequency, exploiting the monthly waves of the MSC to obtain a larger sample and more powerful tests.

Table 2 reports the results for the whole sample and for the top 10% and bottom 90%

[1]See for example Nord (2022)

of the wealth distribution. The first observation is that the estimated coefficient on the whole sample is higher than in the quarterly estimation. The theoretical model implies that, defining θ_m as the stickiness parameter at the monthly level, at the quarterly level one has $\theta_q = \theta_m^3$. Indeed the 99% confidence interval at the monthly level is $\theta_m \in [0.52, 0.83]$ which would give a quarterly stickiness parameter of $\theta_q \in [0.14, 0.58]$, that includes the quarterly point estimate obtained by the OLS coefficient in table 1 of $\theta_q = 0.55$. Second, the estimated coefficient for the top 10% is both lower and statistically less significant than for the whole sample and the bottom 90%. This is in line with the hypothesis that wealthier individuals are more likely to update their information set. Third, I check whether the coefficients between the top 10% and the bottom 90% are statistically different, by reporting the value of the z-statistic of the coefficient difference.

Table 2: Information Rigidity for different Wealth Levels

	All	**Top 10%**	**Bottom 90%**
Constant	-0.78	-0.33	-0.99
	(0.00)	(0.00)	(0.00)
Forecast Change	3.02	1.08	2.74
	(0.00)	(0.06)	(0.00)
Sample	1986:03 - 2023:06	1990:01 - 2023:06	1990:01 - 2023:06
Observations	306	306	306
Instrument	Oil Price Innovations	Oil Price Innovations	Oil Price Innovations
F stat first stage	39.02	24.68	29.98
p-val difference	-	-	0.09

Note: *The table reports estimates of equation (7) by 2SLS for the Michigan Survey of Consumers (MSC) at monthly frequency. P-values obtained from Newey-West robust standard errors are in parentheses.*

The effect of Age and Education It is well known that wealth is highly correlated with both age and education. Older individuals might update their information set more often as they have lived through more business cycles and could have a better understanding of the economy. Similarly, more educated households might be better able to process relevant information and, therefore, update their information set more often. One might therefore conjecture that the result obtained in table 2 is driven by these two variables. To test for this hypothesis I repeat the analysis by constructing different subgroups. For both variables, I need to consider large bins to have enough observations in each group. For age, I pool individuals into three age groups: under 35 years old, between 35 and 65, and above 65. For education, I consider four groups: individuals with a high school diploma, individuals who attended college but did not graduate, individuals with an undergraduate degree, and individuals with postgraduate education. Figure 1 reports the point estimate and the 90% confidence interval for each group. The results confirm the role of wealth as the main driver of heterogeneity. First, all the estimates of wealth groups in the top 10% are not statistically different from zero, suggesting no degree of stickiness for these groups. Second the differences among coefficients for the different age and education groups are not statistically significant. The only exception is between the group of individuals younger than 35 and older than 65 in the top 10% of the wealth distribution. However this result is more the byproduct of a negative estimated point estimate for the younger group, which is in contrast with the underlying theory. In fact, there are only 295 observations in this subgroup, which might not be enough to obtain a reliable estimate.

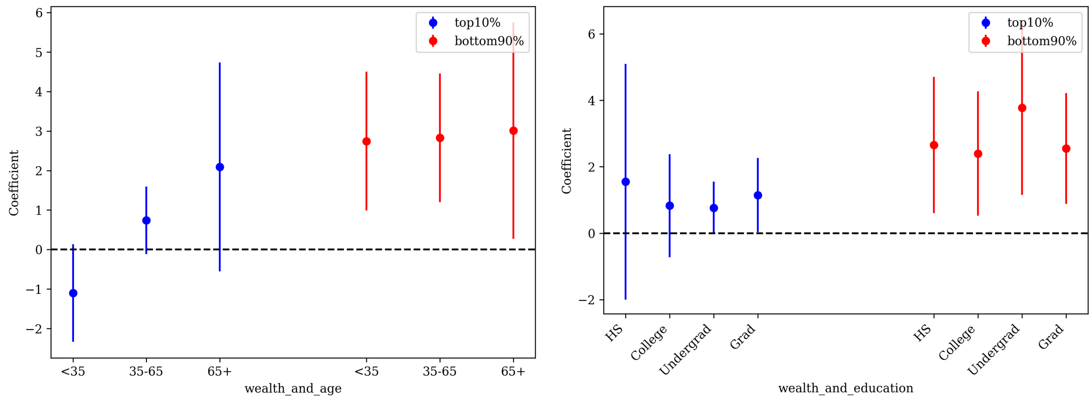

Figure 1: The role of age and education

3 Model

I begin by laying out the simple model which I use to explore the implications of wealth related stickiness. The model is purposely stylized to highlight the main mechanisms at play.

3.1 Households

The economy is populated by a continuum of households that is heterogeneous along two dimensions: wealth and idiosyncratic productivity. Household $i \in [0, 1]$ solves

$$\begin{aligned} &\max_{c_{it}, a_{it}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u(c_{it}) - v(n_{it}) \right) \\ \text{s.t.} \quad & c_{it} + a_{it} \leq (1 + r_t) a_{it-1} + z_{it} \\ & a_{it} \geq \underline{a} \end{aligned} \tag{9}$$

Utility from consumption and disutility from labor are given by $u(c) = \frac{c^{1-\sigma^{-1}}}{1-\sigma^{-1}}$ and $v(n) = \zeta \frac{n^{1+\phi^{-1}}}{1+\phi^{-1}}$. σ and ϕ are the elasticity of intertemporal substitution and the Frisch elasticity and ζ is the disutility of labor. $z_{it} = (1 - \tau_t) y_{it}$ is post tax real income, and $y_{it} = \frac{W_t}{P_t} e_{it} n_{it}$ is labor income. The idiosyncratic productivity e_{it} follows a Markov chain with transition matrix Π^e and is normalized such that $\mathbb{E}_i(e_{it}) = 1$. Households face a borrowing constraint $\underline{a}$ and the real interest rate is r_t.

3.2 Firms

A representative and perfectly competitive firm operates using a linear production technology:

$$Y_t = X_t N_t, \tag{10}$$

where Y_t represents aggregate output, N_t is the level of aggregate labor, and X_t captures productivity. Prices are flexible, and real wage is given by:

$$w_t = \frac{W_t}{P_t} = X_t. \tag{11}$$

Then inflation, $\pi_t = \log(P_t/P_{t-1})$ is given by:

$$\pi_t = \pi_t^w - \log(X_t/X_{t-1}), \tag{12}$$

where $\pi_t^w = \log(W_t/W_{t-1})$ is wage inflation.

3.3 Labor market

Following the approach of Auclert, Rognlie and Straub (2024) and early works by Schmitt-Grohé and Uribe (2005) and Erceg, Henderson and Levin (2019) I model the labor market with sticky wages[2]. A continuum of unions indexed by $k \in [0, 1]$ hires households and aggregates the efficient hours of work supplied to the union into a specific task $N_k t = \int e_{it} n_{ikt} di$. The different union supplies are then aggregated into the total labor supply by a competitive labor market-packer with the constant elasticity of substitution (CES) aggregator

$$N_t = \left(\int_0^1 N_{kt}^{\frac{\varepsilon-1}{\varepsilon}} dk \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{13}$$

and the total supply is sold to the representative firm at the wage W_t. I introduce nominal rigidities by assuming that unions face quadratic adjustment costs in changing the nominal wage

$$\frac{\psi}{2} \int_0^1 \left(\frac{W_{kt}}{W_{kt-1}} \right)^2 dk$$

In each period unions set a uniform wage and allocate hours uniformly across their workers, so that the efficient amount of worked hours $\int e_{it} n_{it} di$ is equal to the aggregate N_t. All unions are symmetric so in equilibrium they set the same wage W_t. Under these assumptions, the linearized wage Phillips curve can be derived as in Appendix C and given by

$$\pi_t^w = \kappa_w \left(\sigma^{-1} \hat{c}_t + \phi^{-1} \hat{n}_t - (\hat{y}_t - \hat{\tau}_t - \hat{n}_t) \right) + \beta \mathbb{E}_t(\pi_{t+1}), \tag{14}$$

[2]Using sicky prices and flexible wages, as common in the New Keynesian literature, leads to an implausible result when embedding heterogenous agents into the model as it would imply countercyclical profits. I refer to Broer et al. (2020) and Auclert, Rognlie and Straub (2024) for a broader discussion

where $\hat{c}_t$, $\hat{n}_t$ and $\hat{y}_t$ are the log-deviations of consumption, labor, and output, from their steady state values, and $\hat{\tau}_t = d\tau_t/(1-\tau)$ and κ_w is the slope of the Phillips curve.

3.4 Government

The government chooses the path of government spending G_t, debt B_t and taxes τ_t to satisfy the government budget constraint

$$B_t = (1+r_t)B_{t-1} + G_t - T_t, \tag{15}$$

with $T_t = \tau_t Y_t$ and where for now I work under the assumption that spending is set exogenously, bonds are fixed at the level B and taxes adjust to satisfy the budget constraint.

3.5 Monetary authority

The monetary authority sets the nominal interest rate according to a Taylor rule

$$i_t = r + \phi_\pi \pi_t + \epsilon_t, \tag{16}$$

in which r is the natural real interest rate, ϕ_π is the response of the nominal interest rate to inflation, and ϵ_t is a monetary policy shock.

3.6 Equilibrium

Aggregate quantities are given by

$$C_t = \int_0^1 c_{it} di, \quad A_t = \int_0^1 a_{it} di. \tag{17}$$

Then given initial conditions for nominal wage, price level, government debt, and the initial distribution of households over assets and productivity, and the exogenous path of government spending, the equilibrium is a sequence of prices, quantities, and allocations such that households, firms, and unions optimize, the government budget constraint and the monetary policy rule are satisfied, and markets clear:

$$\begin{aligned} C_t + G_t &= Y_t, \\ A_t &= B. \end{aligned} \tag{18}$$

3.7 Beliefs

The model layout so far leverages the assumption that beliefs satisfy FIRE. Following the empirical evidence in section 2 I model households with sticky expectations. With respect to the general framework which I laid down so far I make the following additional assumptions. First, I assume that all households are aware of the steady state values of all variables in the economy[3]. Second, I assume that households have sticky information only with respect to aggregate variables and not idiosyncratic ones. Specifically, given that households know the steady states, they have sticky information with respect to *shocks* to aggregate variables. This is in line with the literature and eases the comparison with existing work. Moreover, while I tested directly for stickiness to aggregate variables, it is not straightforward to test for stickiness to idiosyncratic shocks. Third, I follow Carroll et al. (2020) and Auclert, Rognlie and Straub (2020) in assuming that once a shock hits the economy all agents become aware of it and update their information sets accordingly. The first two assumptions together imply that the steady state of the model with sticky information (Sticky HANK model hereafter), coincides with the FIRE one. Finally, and the main novelty of this paper, I assume that the degree of stickiness is endogenously determined by the wealth distribution. Since the survey data evidence is not enough to provide the full functional form for the relationship between wealth and stickiness, I assume that they are related by a power function

$$\theta(a) = \left(\frac{a}{\overline{a}}\right)^{\gamma}. \tag{19}$$

This functional form is convenient as it is determined by the parameter γ only which can then be directly calibrated. Specifically, I will set γ to match the average stickiness, on a quarterly basis, of the top 10% of the distribution.

3.8 Solution Method

Solving the Sticky HANK model is challenging for two reasons. First, one has to deal with household heterogeneity and incomplete markets. For this I rely on the contribution by Auclert, Rognlie and Straub (2021) by using the Sequence-Space representation of the model. Second, one has to take into account the deviation from FIRE imposed by sticky expectations. For this I rely on the methodology developed in Auclert, Rognlie and Straub (2020), Guerreiro (2023), and Bardoczy and Guerreiro (2024). However, a common assumption of these works is that the deviations from FIRE are orthogonal to the state variables of the model. As expectations in the Sticky HANK model are endogenously determined by the wealth distribution, this assumption is not met. A core contribution of this paper is a method to solve the model in this case. To build sequentially towards this solution I proceed as follows. First, I show how to compute the steady state of the economy and then solve for first-order impulse responses to a shock in sequence-space thanks to the methodology in Auclert, Rognlie and Straub (2021). As already remarked the steady state of the Sticky HANK model and the FIRE one coincides. So without having to recompute the steady state, I show how to compute imuplse-responses in the Stikcy HANK by leveraging the way in which the FIRE equivalents are computed.

Steady State. In a steady state aggregate variables are constant over time. Since the steady state is common knowledge, both under FIRE and sticky expectations, agents forecast per-

[3]Note that in the general model given by equation (1) this is true if the economy starts at the steady state and $F_{-1}(x_{t+h}) = x_{-1}$

fectly aggregate variables. The challenging part with respect to a representative agent economy is to solve the household side of the model. The dynamic programming problem associated with the steady state is

$$
\begin{aligned}
V(a,e) = \max_{c,a'} & u(c) - v(n) + \beta \mathbb{E}[V(a',e')|e], \\
\text{s.t.} \quad & c + a' = (1+r)a + z, \\
& a' \geq \underline{a}.
\end{aligned}
\tag{20}
$$

The solution method proceeds as follows. First, the state space is discretized. In this specific case, I use a grid with 500 points for the asset level and 7 possible values for the idiosyncratic productivity. The second step is to obtain the individual policies for consumption and assets by backward iterating on the value function. The specific algorithm used is the endogenous grid method by Carroll (2006), which iterates on the derivative of the value function with respect to assets rather than on the value function itself. A key step which I highlight because it will be important later on, is that in general, the asset policy function is going to point houholds on an asset level that is not in the grid. To deal with this, the solution method uses a lottery that assigns the households to either the closest grid point below or above the optimal asset level. For an agent with desired asset a lying between two grid points a_j and a_{j+1}, the probability of being assigned to a_j is given by

$$
\pi_j = \frac{a_{j+1} - a}{a_{j+1} - a_j}. \tag{21}
$$

Finally given the steady state asset policy function and the exogenous Markov process for idiosyncratic productivity, the steady state distribution can be computed by iterating forward the distribution. Finally, aggregate steady state quantities can be computed as

$$
C = \int c(a,e) dD(a,e), \quad A = \int a'(a,e) dD(a,e), \tag{22}
$$

where $D(a,e)$ is the distribution of households over assets and productivity. In steady state the market clearing conditions are

$$
C + G = Y, \quad A = B, \tag{23}
$$

and a stationary equilibrium is then a sequence of prices, quantities, and allocations such that households, firms, and unions optimize, the government budget constraint and the monetary policy rule are satisfied, markets clear and the steady state constraint is satisfied.

FIRE dynamics away from the steady state. The study of dynamics in response to shocks hitting the economy leverages the Sequence-Space representation of the model. The object of interest is the *sequence* of states of an aggregate endogenous variable, for example, consumption in response to a change in an exogenous variable, for example, the real interest rate. To denote infinite sequences I will use the notation $\{S_t\} \equiv (S_t, S_{t+1}, S_{t+2} \ldots)$. The shocks considered will be MIT shocks, which are shocks that come as a surprise at time 0 in an economy that is in a steady state. For sequences that start at time $t = 0$, I will also use exchangbly the notation $\mathbf{S} \equiv (S_0, S_0, S_0 \ldots) \equiv \{S_0\}$. Then to first-order, the output $\mathbf{o}$ change in response to a change of an input $\mathbf{i}$ is given by

$$d\mathbf{o} = \mathcal{J}^{o,i} d\mathbf{i}, \tag{24}$$

where $\mathcal{J}^{o,i}$ is the Jacobian summarizing the partial derivatives

$$\mathcal{J}^{o,i} = \begin{pmatrix} \frac{\partial o_0}{\partial i_0} & \frac{\partial o_0}{\partial i_1} & \cdots \\ \frac{\partial o_1}{\partial i_0} & \frac{\partial o_1}{\partial i_1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{25}$$

Clearly, the most challenging Jacobians to compute pertains to the household side of the model. A brute force approach to compute them is to use the same method described to obtain the steady state for each entry of the matrix. As an example assume that one wishes to compute the partial derivative of consumption at date t to the interest rate at date s. Then one would do a backward iteration starting from a period T at which it is assumed that the economy is back in the steady state. The backward iteration can handle time-varying inputs and one can carry it out by setting $\mathbf{r}$ at the steady state except at date s, in which it is shocked by an infinitesimal amount h After obtaining the respective policies one would iterate forward the distribution and aggregate to compute C_t. Then the partial derivative would be given by

$$\frac{\partial C_t}{\partial r_s} = \frac{C_t - C}{h}. \tag{26}$$

This method is computationally expensive as it requires $T \times T$ backward and forward iterations for each Jacobian. One of the central contributions of Auclert, Rognlie and Straub (2021) is a method to compute these Jacobians in a more efficient way. This method relies on the assumption that agents have perfect foresight, but Auclert, Rognlie and Straub (2020) shows how to compute the Jacobians in the model with deviations from FIRE at almost no computational cost starting from the FIRE one. However, their method relies on the assumption that the deviations from FIRE are orthogonal to the state variables of the model. Moving forward I will then show how to compute Jacobians for the households side in the Sticky HANK model in two ways. First by brute force, and then by leveraging the method in Auclert, Rognlie and Straub (2020) and adapting it to the case in which the deviations from FIRE are not orthogonal to the state variables.

Sticky dynamics away from the steady state. The brute force solution leverages the following idea. Consider an individual i with wealth a at time 0. Assume that this individual is going to keep their wealth constant over time. Then when a shock is announced to hit the economy at time s, the probability that the individual will learn about it at a time $\tau < s$ is $(1 - \theta(a))\theta(a)^\tau$. Moreover, as long as the population of agents with the same asset level is large enough, then this probability coincides with the fraction of agents with the same asset level that updated their information set. In this case, across that asset level, at time τ the economy reacts as if the shock hitting at time s has size $(1-\theta(a))\theta(a)^{\tau}$ of the original shock. The brute force algorithm then consists of computing the Jacobian as described in the previous part but rescaling the size of the shock to the correct level at each asset level: $h \to h(1-\theta(a))\theta(a)^{\tau}$. This method relies on the approximation that while agents are moving in the distribution, they move in such a way that

$$\theta(a_k) \approx \frac{\sum_{a_j} m_{a_j,a_k}\theta(a_j)}{\sum_{a_j} m_{a_j,a_k}}, \tag{27}$$

where m_{a_j,a_k} is the mass of agents that have wealth a_j at time $\tau-1$ and a_k at time τ. In other words, the probability of updating at time τ for an agent with wealth a_k is the weighted average of the updating probability of agents with wealth a_j at time $\tau-1$. I discuss the validity of this approximation in appendix B. In any case this method is computationally expensive and I propose a different way of obtaining the Jacobians in the Sticky HANK model. The brute force method will then serve as a benchmark to test the validity of the new method.

The methodology builds on the one in Auclert, Rognlie and Straub (2020) to compute Jacobians with deviations from FIRE. It needs however to be adapted to the case in which the degree of stickiness θ is endogenous and determined by the distribution. I begin by reassuming the method used in Auclert, Rognlie and Straub (2020). This follows two assumptions. The first one is that it is possible to partition the population at any time t in two groups: those who have updated their information set at least once in a period $\tau \leq t$, thus learning about the shock and those who have not. This assumption is met also in this case. The second one is that the updating probability is orthogonal to idiosyncratic shocks and constant over the population and time. This assumption is not met in this case. If it were, then it would be possible to derive the sticky expectations Jacobian as

$$\mathcal{J}^{o,i} = (1-\theta)\sum_{\tau=0}^{\infty}\theta^{\tau}\mathcal{J}^{o,i,\tau}. \tag{28}$$

In the context of this paper, a similar relationship can be derived, but this involved an additional step to deal with the endogenous evolving stickyness as I explain below.

FIRE Jacobian manipulation

Start by considering a representative household with time-varying stickiness parameter θ_t. I define the probability that this agent has not updated their information set in any period up to τ as

$$P_\tau = \begin{cases} 1 & \text{if } \tau = -1 \\ \prod_{k=0}^{\tau}\theta_k & \text{if } \tau \geq 0 \end{cases} \tag{29}$$

Then the probability of learning about about a shock at time $\tau < t$ for the households can be derived as:

$$\ell_\tau = (1-\theta_\tau)P_{\tau-1}.$$

Then, consider an economy populated by a continuum of representative households, this can be divided at any point in time into two groups: those who have updated their information set at least once in a period $\tau \leq t$, thus learning about the shock and those who have not.

Then the aggregate Jacobian describing the output sequence o response to input sequence i can be written as

$$\mathcal{J}^{o,i} = \sum_{\tau=0}^{\infty} \ell_\tau \mathcal{J}^{o,i,\tau}. \tag{30}$$

The insights provided in appendix D.3 of Auclert, Rognlie and Straub (2020) still apply and can be used to derive the following relationship between the sticky information Jacobian and the Jacobian of the model with perfect foresight. Specifically, it is still true that a household learning at date τ about a shock at time s to input i will have the same response as a household learning at time 0 about the same shock at time $s-\tau$ to input i, shifted by τ periods. That is

$$\mathcal{J}^{o,i,\tau}_{t,s} = \mathcal{J}^{o,i,\tau-1}_{t-1,s-1} = \cdots = \mathcal{J}^{o,i,0}_{t-\tau,s-\tau}. \tag{31}$$

I also maintain the assumption that all households are aware of the shock when it hits at time s so that if $\tau > s$ then $\mathcal{J}^{o,i,\tau}_{t,s} = \mathcal{J}^{o,i,s}_{t,s}$. With this I can rewrite equation (30) as

$$\mathcal{J}^{o,i}_{t,s} = \sum_{\tau=0}^{s-1} \ell_\tau \mathcal{J}^{o,i,\tau}_{t,s} + \sum_{\tau=s}^{\infty} \ell_\tau \mathcal{J}^{o,i,s}_{t,s}. \tag{32}$$

The term $\sum_{\tau=s}^{\infty} \ell_\tau$ can be simplified as follows. First, note that

$$\ell_\tau = P_{\tau-1} - P_\tau, \tag{33}$$

since

$$P_{\tau-1} - P_\tau = P_{\tau-1} - (P_{\tau-1}\cdot\theta_\tau) = P_{\tau-1}(1-\theta_\tau) \tag{34}$$

Then, the sum can be written as

$$S = \sum_{\tau=s}^{\infty}(P_{\tau-1} - P_\tau) = \left(\sum_{\tau=s}^{\infty} P_{\tau-1}\right) - \left(\sum_{\tau=s}^{\infty} P_\tau\right). \tag{35}$$

Then noticing that the first sum is equal to the second sum shifted by one period I have

$$S = P_{s-1} + \left(\sum_{n=s}^{\infty} P_n\right) - \left(\sum_{n=s}^{\infty} P_n\right) = P_{s-1}. \tag{36}$$

Therefore

$$S = P_{s-1} = \sum_{\tau=s}^{\infty} \ell_\tau \tag{37}$$

Then (32) becomes

$$\mathcal{J}^{o,i}_{t,s} = \sum_{\tau=0}^{s-1} \ell_\tau \mathcal{J}^{o,i,\tau}_{t,s} + \mathcal{J}^{o,i,s}_{t,s} P_{s-1}. \tag{38}$$

Now one can write

$$\mathcal{J}^{o,i}_{t,s} = \mathcal{J}^{o,i,s}_{t,s} P_{s-1} + \sum_{\tau=1}^{s-1} \ell_\tau \mathcal{J}^{o,i,\tau}_{t,s} + \ell_0 \mathcal{J}^{o,i,0}_{t,s}, \tag{39}$$

and applying (**??**) I have

$$\mathcal{J}^{o,i}_{t,s} = \mathcal{J}^{o,i,s-1}_{t-1,s-1} P_{s-1} + \sum_{\tau=0}^{s-2} \ell_{\tau+1} \mathcal{J}^{o,i,\tau}_{t-1,s-1} + \ell_0 \mathcal{J}^{o,i,0}_{t,s}. \tag{40}$$

Notice now that shifting (38) by 1 period in both t and s I have

$$\mathcal{J}^{o,i}_{t-1,s-1} = \sum_{\tau=0}^{s-2} \ell_\tau \mathcal{J}^{o,i,\tau}_{t-1,s-1} + \mathcal{J}^{o,i,s-1}_{t-1,s-1} P_{s-2}. \tag{41}$$

which is almost the same expression as the first two terms in (40), except for $\ell_{\tau+1}$ in place of ℓ_τ and P_{s-2} in place of P_{s-1}. It is easy to apply the tranformation $P_{s-1} = P_{s-2}\theta_{s-1}$, but the relationship betwen $\ell_{\tau+1}$ and ℓ_τ can be derived as follows

$$\ell_{\tau+1} = (1-\theta_{\tau+1})P_\tau = (1-\theta_{\tau+1})\theta_\tau P_{\tau-1} = \ell_\tau \frac{(1-\theta_{\tau+1})\theta_\tau}{1-\theta_\tau}. \tag{42}$$

The term $\frac{1-\theta_{\tau+1}}{1-\theta_\tau}$ is time dependent and prevents me from directly expressing $\mathcal{J}^{o,i}_{t,s}$ in terms of $\mathcal{J}^{o,i}_{t-1,s-1}$. To overcome this, I approximate the term as

$$\frac{(1-\theta_{\tau+1})\theta_\tau}{1-\theta_\tau} \approx \theta_{s-1}. \tag{43}$$

With this I can write (40) as

$$\mathcal{J}^{o,i}_{t,s} = \mathcal{J}^{o,i}_{t-1,s-1}\theta_{s-1} + \ell_0 \mathcal{J}^{o,i,0}_{t,s}. \tag{44}$$

Now, noitce that for $s = 0$ then (38) becomes $\mathcal{J}^{o,i}_{t,0} = \mathcal{J}^{o,i,0}_{t,0}$. For $t = 0$ and $s > 0$ then households react only if $\tau = 0$, so $\mathcal{J}^{o,i}_{0,s} = \ell_0 \mathcal{J}^{o,i,0}_{0,s}$. Combining these insights, and realizing that $\mathcal{J}^{o,i,0}_{t,s} = \mathcal{J}^{o,i,\text{FI}}_{t,s}$ since it just the full-information Jacobian, I can write

$$\mathcal{J}^{o,i}_{t,s} = \begin{cases} \mathcal{J}^{o,i}_{t-1,s-1}\theta_{s-1} + \ell_0 \mathcal{J}^{o,i,\text{FI}}_{t,s} & t > 0, s > 0 \\ \mathcal{J}^{o,i,\text{FI}}_{t,s} & s = 0 \\ \ell_0 \mathcal{J}^{o,i,\text{FI}}_{t,s} & t = 0, s > 0 \end{cases} \tag{45}$$

which is a nice formulation since it relates the sticky information Jacobian to the full information Jacobian and the time-varying stickyness parameter and can thus be computed efficiently as long as the path of θ_t is known. However, the path of θ_t is endogenously determined by the distribution of wealth in the economy and in essence this is a fixed point problem. To solve it I can just use the following insight. At the aggregate level, the stickiness of the economy in steady state is given by

$$\theta = \int \theta(a) dD(a). \tag{46}$$

I can then treat the path of θ_t as any aggregate output and compute the Jacobian mapping any shock to input i to the path of θ_t, $\mathcal{J}^{\theta,i}$ under FIRE assumption. To compute the effect of a shock to input i on the path of θ_t I can then use the following algorithm.

1. Guess an initial path for $\{\theta_t\}^0$.
2. Starting from $\mathcal{J}_{t,s}^{\theta,i,\text{FIRE}}$, compute $\mathcal{J}_{t,s}^{\theta,i}$ using (45).
3. Apply (24) to obtain the change in θ_t and add it to the steady state value to get a new path $\{\theta_t\}^1$.
4. Check if $||\{\theta_t\}^0 - \{\theta_t\}^1|| < \epsilon$, for an exogenously set tolerance level ϵ.
5. Run until convergence.

The resulting path of $\{\theta_t\}$ can then be used to obtain the sticky information Jacobian relating any output o to input i by applying (45) to the appropriate FIRE Jacobian.

3.9 Calibration

I calibrate the model to a quarterly frequency. The steady state of interest is the one with 0 inflation which implies that nominal and real interest rates coincide. Table 3 reports the value of the parameters used in the calibration.

Table 3: Calibration

Parameter	Description	Value
Households		
σ	Elasticiy of intertemporal substition	0.5
ϕ	Firsch elasticity	0.5
ζ	Disitulity of labor	1.0
ρ_e	Persistence of idiosyncratic productivity shocks	0.92
sd_e	Standard deviation of idiosyncratic productivity shocks	0.92
n_e	Number of productivity states	11
$\underline{a}$	Minimum asset level	0
$\bar{a}$	Maximum asset level	4000.0
n_a	Number of asset grid points	500
γ	Sensitivity of stickiness to wealth	0.08
Firms		
X	Steady state TFP level	1.0
κ_w	Wage rigidity	0.16
Monetary		
ϕ_π	Taylor rule coefficient	1.5
r	Real interest rate	0.5%
Fiscal		
G/Y	Spending to GDP ratio	0.16
B	Bond supply	5.6
Targeted		
β	Discount factor	0.95

3.10 Monetary policy

I now consider the implications of the model for monetary policy. As a useful reference model, I also consider a representative agent (RANK) model, in which the only difference is the household side of the model. Specifically, the representative household solves

$$\begin{aligned} &\max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma}}{1-\sigma} \\ \text{s.t.} \quad &C_t + A_t \leq (1+r_t)A_{t-1} + Z_t. \end{aligned} \tag{47}$$

The calibration for this model is also the same as in the HANK model, except for the value of the discount factor β which is now set to $\beta = 1/(1+r)$. For consistency to the remainder of the paper, I assume that the monetary authority actually shocks the beginning of the period interest rate r^{ante}. This will allow to handle valuation effects when adding dividends. This can be accommodated into the model by introducing a simple rule mapping the ex-ante to the ex-post interest rate

$$\begin{cases} r_t = r_{t-1}^{\text{ante}} & \text{if } t \geq 1 \\ r_t = r & \text{if } t = 0 \end{cases} \tag{48}$$

I start in the FIRE framework and consider the response of aggregate output to a shock to the real interest rate which deviates 1% from its steady state on impact and goes back to the steady state with an exponentially declining weight of 0.7. As remarked by Auclert, Rognlie and Straub (2021) it is useful to visualize the macro model as a Directed Acyclic Graph (DAG). Since B and G are constant, given the fiscal rule T will simply adapt to match the fiscal authority budget constraint. Also for the moment, I can focus on the real side of the economy, given that steady state inflation is 0. Under these assumptions, the DAG, which is common for both HANK and RANK is shown in figure 2. Given that

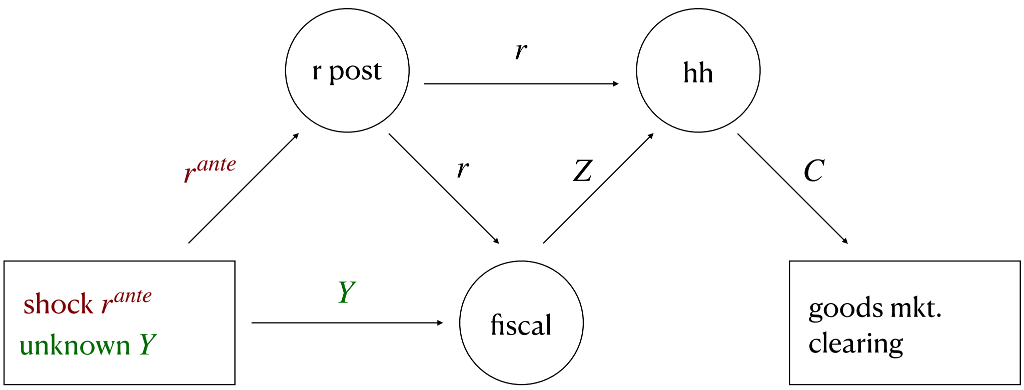

Figure 2: DAG of the model for monetary shock

market clearing is 0, the DAG represents the following relationship in sequence space

$$\mathbf{Y} = \mathbf{G} + \mathcal{C}(\mathbf{Z}, \mathbf{r}), \tag{49}$$

where the notation $\mathcal{C}$ stresses that aggregate consumption and aggregate net of taxes income are functions of sequences. Indeed also the ex-post interest rate and $\mathbf{Z}$ are functions. Then in response to a shock to the real interest rate, differentiating (49) yields (since $\mathbf{G}$ is constant)

$$d\mathbf{Y} = \mathcal{J}^{C,r} d\mathbf{r} + \mathcal{J}^{C,Z} d\mathbf{Z}, \tag{50}$$

and by explicitly considering the dependence of the fiscal and ex-post interest rates

$$d\mathbf{Y} = \mathcal{J}^{C,r} d\mathbf{r} + \mathcal{J}^{C,Z} \left(\mathcal{J}^{Z,Y} d\mathbf{Y} + \mathcal{J}^{Z,r} d\mathbf{r} \right), \tag{51}$$

whit $d\mathbf{r} = \left(\mathcal{J}^{r,r^{\text{ante}}} d\mathbf{r}^{\text{ante}} \right)$. This Jacobians in (50) are what Auclert, Rognlie and Straub (2024) call intertemporal marginal propensities to consume (iMPCs) and are of particular interest since they are sufficient statistics for the response of aggregate output to a monetary policy shock. This formulation is also useful to decompose the effect of the shock into direct (via intertemporal substitution) and indirect (via income) effects. The left panel of figure 3 shows the IRFs of the HANK and RANK models. Within the proposed calibration,

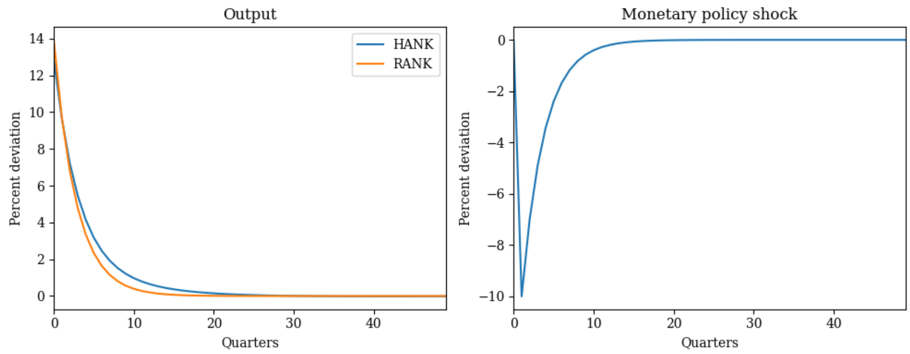

Figure 3: Monetary policy in the HANK and RANK

the finding is that under FIRE there difference in terms of aggregate response from RANK to HANK is minimal. Indeed this is a well-known result in the literature, (Ahn et al., 2018). While the aggregate response is similar, the decomposition into direct and indirect effects in figure 4reveals that in HANK indirect effects, via income are much more important. Indeed in the RANK model, marginal propensities to consume are zeros and therefore the only channel is the direct one. Again, however, for a policymaker, interested only in the aggregate response, the difference is minimal.

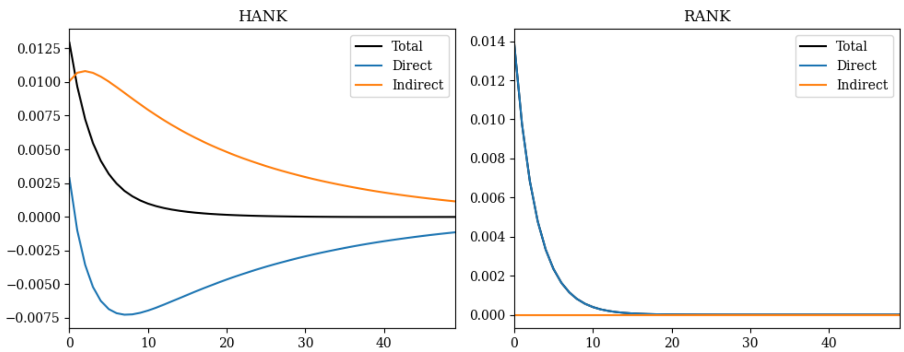

Figure 4: Decomposition

Now I consider the same shock in the case in which agents have sticky information. In order to maintain a fair comparison I also consider a sticky version of the RANK model. Absent assets heterogeneity, however, I set the stickiness in RANK to the steady state value

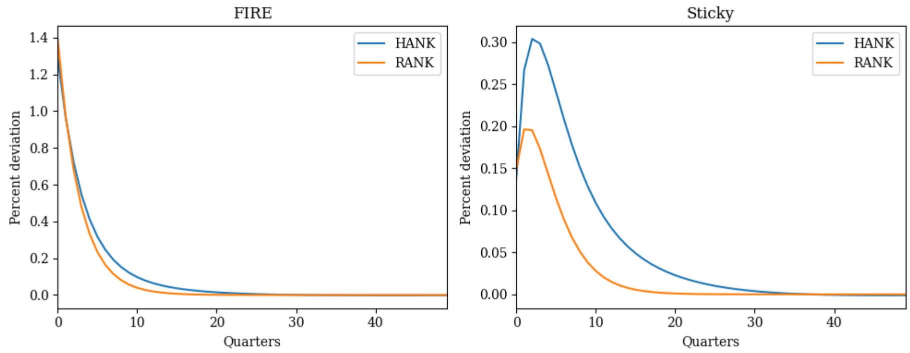

Figure 5: Monetary policy in the HANK and RANK under sticky information

of the HANK model and constant for all agents. Then I apply the algorithm described in section 3.8 to the iMPCs of the equation (50) to obtain the sticky information Jacobians.[4] The right panel of figure 5 shows the IRFs of the HANK and RANK models under sticky information. In both cases, the response displays the hump shape typical of empirical impulse responses. However now by neglecting heterogeneity, the response is underestimated and also the peak of the response is anticipated. When individuals have sticky information, inequality matters.

4 Estimation

In this section I move to the estimation of the model. In order to do so I consider a slightly different model, including a more realistic supply side, that could accomodate the importance of investements in the transmission of monetary policy, as pointed out in Auclert, Rognlie and Straub (2024).

4.1 Model

Households

The household side is almost the same, but now it is assumed that households also receive dividends from ownership of firms. Morevoer taxes and dividends follow an incidence rule. Specifically I will assume that they are both proportional to the household productivity. Their optimization problem is then

$$\begin{aligned} &\max_{c_{it},a_{it}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(u(c_{it}) - v(n_{it})\right) \\ \text{s.t.} \quad & c_{it} + a_{it} \leq (1+r_t)a_{it-1} + y_{it} - \tau_t \bar{\tau}(e_{it}) + d_t \bar{d}(e_{it}) \\ & a_{it} \geq \underline{a} \end{aligned} \tag{52}$$

[4]In the RANK case, since the stickiness is constant, there is no need to iterate, and the transformation in (45) can be directly applied.

Firms The main difference is the introduction of capital into the economy. The firm side is populated by a final good producer and a continuum of intermediate goods producers, indexed by j. Intermediate producers have Cobb-Douglas production function

$$y_{jt} = F(n_{jt}, k_{jt-1}) = n_{jt}^{1-\alpha} k_{jt-1}^{\alpha}, \tag{53}$$

and face adjustment costs when choosing their capital stock. I follow Auclert, Rognlie and Straub (2021) in setting quadratic adjustment costs

$$\Psi_t^k (k_{jt}, k_{jt-1}) = \frac{k_{jt}}{k_{jt-1}} - (1-\delta) + \frac{1}{2\delta\epsilon_I} \left(\frac{k_{jt} - k_{jt-1}}{k_{jt-1}} \right)^2, \tag{54}$$

with δ the depreciation rate and ϵ_I the capital adjustment cost parameter, both positive. Aggregate investment then evolves by

$$I_t = K_t - (1-\delta)K_{t-1} + \Psi_t^k (K_t, K_{t-1}) K_{t-1}. \tag{55}$$

Price setting of intermediate firms is also subject to quadratic adjustment costs,

$$\Psi_t^p (p_{jt}, p_{jt-1}) = \frac{\mu}{\mu - 1} \frac{1}{2\kappa} - [\log(p_{jt}) - \log(p_{jt-1})]^2, \tag{56}$$

where μ is the constant elasticity of substition of the final good producer and κ is the price adjustment cost parameter. The Phillips curve for aggregate inflation is then

$$\log(1+\pi_t) = \kappa \left(\frac{w_t}{F'_N(N_t, K_{t-1})} - \frac{1}{\mu} \right) + \frac{1}{1+r_{t+1}} \frac{Y_{t+1}}{Y_t} \log(1+\pi_{t+1}). \tag{57}$$

Dividends are given by output net of investements, the remuneration of labor and the price adjustment costs, so that $d_t = Y_t - I_t - w_t N_t - \Psi_t^p(P_t, P_{t-1})Y_t$. The evolution of capital is determined jointly with the Tobin's Q

$$Q_t = 1 + \frac{1}{\delta\epsilon_I} \frac{K_t - K_{t-1}}{K_{t-1}} \tag{58}$$

$$\begin{aligned} (1+r_{t+1})Q_t = \alpha_t \frac{Y_{t+1}}{K_t} \frac{w_{t+1}}{F'_N(N_{t+1}, K_t)} - \\ \left[\frac{K_{t+1}}{K_t} - (1-\delta) + \frac{1}{2\delta\epsilon_I} \left(\frac{K_{t+1} - K_t}{K_t} \right)^2 \right] + \frac{K_{t+1}}{K_t} Q_{t+1}. \end{aligned} \tag{59}$$

Fiscal and Monetary authority

The fiscal authority follows a similar rule as before, but now taxes are proportional to labor so $\tau_t w_t N_t = r_t B_t + G_t$. The taylor rule now includes also reaction to the output gap $i_t = r + \phi_\pi \pi_t + \phi_y (Y_t - Y_{ss})$, and the Fisher equation reads $r_t = (1 + i_{t-1})/(1 + \pi_t)$.

Market Clearing

Asset market clearing is the same as before, but goods market clearing also includes investements and price adjustment costs, so that $Y_t = C_t + I_t + G_t + \Psi_t^p(P_t, P_{t-1})Y_t$.

4.2 Empirical IRFs

I estimate the model by matching model implied IRFs to empirical one. The causal shock of intereste is the estimated monetary policy shock of Aruoba and Drechsel (2022), which builds on the Romer and Romer (2004) procedure, but uses also Natural Language Processing techniques, and compute the shock as changes in the Federal Funds Rate orthogonal with respect to all available FED forecasts and text-based time series. To pin down the causal dynamic effect of this shock on the macroeconomic aggregates I use a Bayesian Vector Autoregression (BVAR) model, following Caravello, McKay and Wolf (2024). I use three target outcomes: output, inflation and the nominal interest rate and collect their impulse responses over 25 quarters by stacking them into the vector $\boldsymbol{\Upsilon}$, with covariance matrix $\boldsymbol{\Sigma}$.

I then estimate the following set of parameters: the degree of household inattention γ, the coefficient of the Taylor rule reaction to inflation ϕ_π, the coefficient of the Taylor rule reaction to output gap ϕ_y, the price adjustement cost parameter κ, the wage rigidity parameter κ_w and the investment adjustment cost parameter ϵ_I. Collecting these parameters in the vector $\boldsymbol{\Psi} \equiv (\gamma, \phi_\pi, \phi_y, \kappa, \kappa_w, \epsilon_I)$, an approximate likelihood of the data $\hat{\boldsymbol{\Upsilon}}$, as a function of $\boldsymbol{\Psi}$ is given by:

$$p(\hat{\boldsymbol{\Upsilon}} \mid \boldsymbol{\Psi}) \propto \exp\left[-0.5\left(\hat{\boldsymbol{\Upsilon}} - \theta_v(\boldsymbol{\Psi})\right)' V_{\theta_v}^{-1}\left(\hat{\boldsymbol{\Upsilon}} - \theta_v(\boldsymbol{\Psi})\right)\right]. \tag{17}$$

The posterior for $\boldsymbol{\Psi}$ given the policy shock causal effect data $\hat{\boldsymbol{\Upsilon}}$ is then

$$p(\boldsymbol{\Psi} \mid \hat{\boldsymbol{\Upsilon}}) = \frac{p(\hat{\boldsymbol{\Upsilon}} \mid \boldsymbol{\Psi})p(\boldsymbol{\Psi})}{p(\hat{\boldsymbol{\Upsilon}})},$$

and where

$$p(\hat{\boldsymbol{\Upsilon}}) = \int p(\hat{\boldsymbol{\Upsilon}} \mid \boldsymbol{\Psi})p(\boldsymbol{\Psi})d\boldsymbol{\Psi}.$$

Computation of these objects is standard, relying on the usual random walk Metropolis-Hastings algorithm both to draw from the posterior distribution and to compute the marginal likelihood.

The final step is to recover posterior model probabilities. I have

$$p(\mathcal{M}_j \mid \hat{\boldsymbol{\Upsilon}}) = \frac{p(\hat{\boldsymbol{\Upsilon}})p(\mathcal{M}_j)}{\sum_{i=1}^{M} p(\hat{\boldsymbol{\Upsilon}} \mid \mathcal{M}_i)p(\mathcal{M}_i)}. \tag{18}$$

The priors and posteriors results are displayed in table 4.

Fure 6 shows the empirical IRFs and the model implied ones, evaluated at the posterior mode.

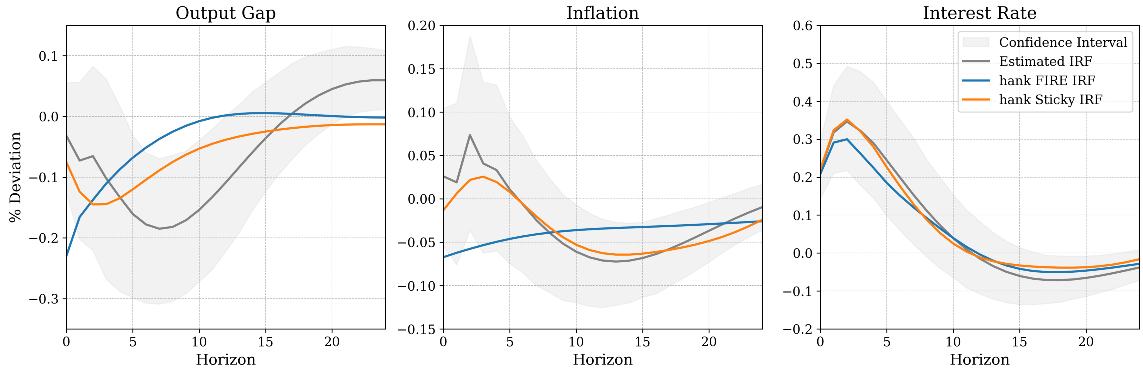

Figure 6: Empirical IRFs

Table 4: Prior and Posterior Distributions

Parameter	Prior Distribution	Mode	Mean	[0.05, 0.95] CI
γ	Gamma(0.08, 0.1)	0.047	0.05	[0.031, 0.072]
ϕ_{pi}	Gamma(1.5, 0.5)	1.715	1.721	[1.231, 1.964]
ϕ_y	Gamma(0.5, 0.5)	0.001	0.001	[0.000, 0.0064]
κ	Gamma(0.1, 0.1)	0.325	0.313	[0.100, 0,631]
κ_w	Gamma(0.1 0.1)	0.293	0.287	[0.072, 0.524]
ϵ_I	Gamma(4, 2)	0.743	0.897	[0.347, 0.986]

5 Conclusion

This paper advances the understanding of monetary policy transmission by introducing wealth-dependent information stickiness into a HANK model. Through empirical analysis of survey data, I provide evidence of wealth-dependent stickiness, which is incorporated into the model using a novel computational methodology that extends the sequence-space methods introduced by Auclert, Rognlie and Straub (2021). My approach modifies the computation of the General Equilibrium Jacobians to account for the endogenous evolution of the wealth distribution and state-dependent updating probabilities.

I use this methodology to efficiently solve the model, simulate the economy's response to monetary policy shocks, and estimate the model by matching impulse responses. The results reveal that wealth-dependent stickiness significantly alters the transmission mechanisms of monetary policy. Specifically, ignoring heterogeneity in information updating—as in representative agent models with uniform stickiness—leads to an underestimation of both the magnitude and the delay of the peak response to monetary shocks.

References

Ahn, Se Hyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf. 2018. "When inequality matters for macro and macro matters for inequality." *NBER Macroeconomics Annual*, 32.

Aruoba, S. Boragan, and Thomas Drechsel. 2022. "Identifying Monetary Policy Shocks: A Natural Language Approach." *SSRN Electronic Journal*.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2020. "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model." *Manuscript*.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2021. "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated Hank Model." *SSRN Electronic Journal*.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub. 2024. "The Intertemporal Keynesian Cross." *Journal of Political Economy*.

Bardoczy, Bence, and Joao Guerreiro. 2024. "Unemployment Insurance in Macroeconomic Stabilization with Imperfect Expectations." *Working Paper*.

Broer, Tobias, Niels Jakob Harbo Hansen, Per Krusell, Erik ÖbergNiels Jakob Harbo Hansen, Per Krusell, and Erik Öberg. 2020. "The new keynesian transmission mechanism: A heterogeneous-agent perspective." *Review of Economic Studies*, 87.

Caravello, Tomas E., Alisdair McKay, and Christian K. Wolf. 2024. "Evaluating Policy Counterfactuals: A "VAR-Plus" Approach." *Working Paper*.

Carroll, Christopher D. 2003. "Macroeconomic expectations of households and professional forecasters." *Quarterly Journal of Economics*, 118.

Carroll, Christopher D. 2006. "The method of endogenous gridpoints for solving dynamic stochastic optimization problems." *Economics Letters*, 91.

Carroll, Christopher, Edmund Crawley, Jiri Slacalek, Kiichi Tokuoka, and Matthew White. 2020. "Sticky Expectations and Consumption Dynamics." *American Economic Journal: Macroeconomics*, 12(3): 40–76.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 2005. "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of Political Economy*, 113.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review*, 105.

Croushore, Dean, and Tom Stark. 2003. "A real-time data set for macroeconomists: Does the data vintage matter?" *Review of Economics and Statistics*, 85: 605–617.

Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin. 2019. "Optimal monetary policy with staggered wage and price contracts." *Credit and Capital Markets*, 52.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik. 2021. "MPC Heterogeneity and Household Balance Sheets." *American Economic Journal: Macroeconomics*, 13.

Guerreiro, Joao. 2023. "Belief Disagreement and Business Cycles." *Working Paper*.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante. 2018. "Monetary policy according to HANK."

Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve." *Quarterly Journal of Economics*, 117.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal sticky prices under rational inattention." *American Economic Review*, 99.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2010. "Business cycle dynamics under rational inattention." *Review of Economic Studies*, 82.

Mckay, Alisdair, and Christian K. Wolf. 2022. "Optimal Policy Rules in HANK." *Working Paper*.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The power of forward guidance revisited." *American Economic Review*, 106.

Nord, Lukas. 2022. "Who Cares about Inflation? Endogenous Expectation Formation of Heterogeneous Households." *SSRN Electronic Journal*.

Romer, Christina D, and David H Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review*, 94.

Schmitt-Grohé, Stephanie, and Martín Uribe. 2005. "Optimal fiscal and monetary policy in a medium-scale macroeconomic model." *NBER Macroeconomics Annual*, 20.

Smets, Frank, and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review*, 97.

Wolf, Christian. 2021. "Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result." *SSRN Electronic Journal*.

Zorn, Peter. 2021. "Investment Under Rational Inattention: Evidence from US Sectoral Data." *SSRN Electronic Journal*.

A Survey Data

I report either a description of the variable or the question asked to the respondents, for the surveys used in the analysis.

Survey of Professional Forecasters: *Querterly expectations of inflation (measured by the GNP/GDP price index and, alternatively, the CPI).*

Michigan Survey of Consumers: *By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?*

Livingston Survey: *Give the current-month (June and December) forecasts [of the CPI] and then base [your] six-month and 12-month forecasts on [your] current-month predictions.*

- Gini Index for the US. SIPOVGINIUSA from the FRED at yearly frequency. To obtain the index for each quarter or semester I linearly interpolate between available observations.

When dealing with individual forecasts I had to deal with missing values. In doing so I followed the procedure of the survey as reported in the technical note. Specifically NaNs and "don't know if up or down" responses were dropped from the sample. However "don't know how much up" and "don't know how much down" were imputed by the mean of the other responses. In the orignial procedure the imputation is done by matching the distribution of the other responses. However since the sample mean would be unaffected and higher order moments are not used in this analysis, I opted for the simpler approach.

B Brute Force Method approximation

What is required for the approximation to be accurate is that agents move across the distribution in such a way that the probability of ending up in wealth level a is almost the same for agents starting at wealth levels $a - \Delta$ and $a + \Delta$, for any Δ. In order for this to be the case, two conditions should be met, the first is that the mapping from current to future wealth levels must be almost linear. The second that the wealth space being unbounded. To assess wheter the first condition is met, in figure 7 I plot net savings as a function of current wealth for different income levels, in the steady state.

As can be seen the relationship is almost linear, except for extremely low level of wealth. Zooming in on the low wealth region one can see that the nonlinearity is really accentuated only for the bottom 0.1% of the distribution. The second condition is clearly not met since the grid is discretized between a minimum and maximum asset level. Then especially at the boundaries the approximation will be less accurate. However given the shape of the function mapping wealth to probability of stickiness, for high values of assets the function is almost flat, so the error will be small. For example at the 95% level one has $\theta = 0.517$ while at the maximum asset level $\theta = 0.515$, a difderence of 0.2 percentage points. At the bottom of the distribution the error will be larger. I can only argue that the

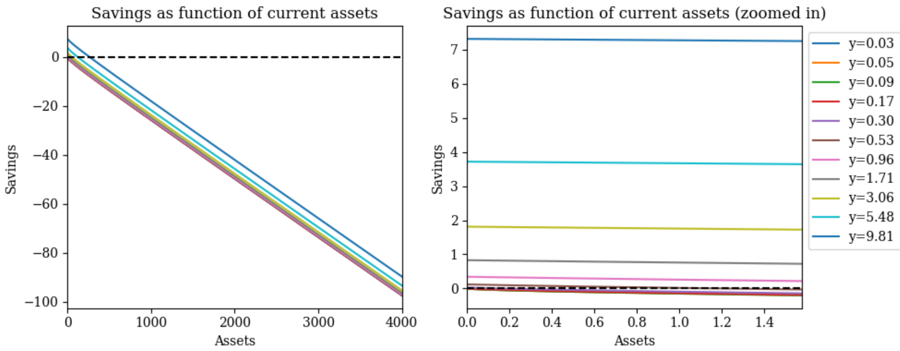

Figure 7: Net Savings as a Function of Assets for different Income Levels

proportion of agents in this region is small, as can be seen in figure 8. The only excpetion being the lower bound in which in steady state there is a mass of 1.4% of agents.

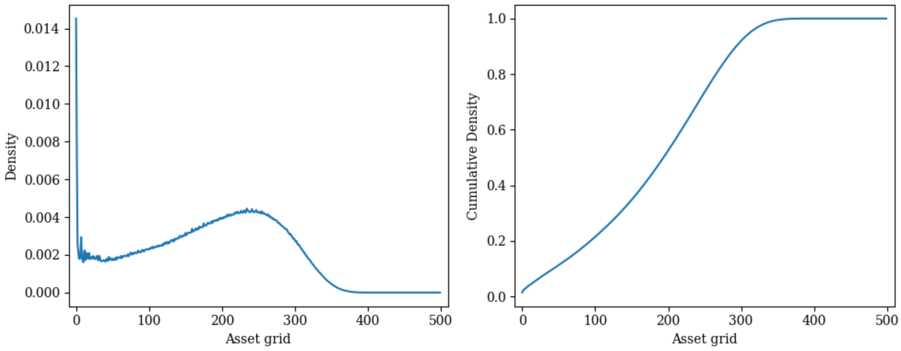

Figure 8: Wealth Distribution

C Phillips Curve

The union k faces the following maximization problem at time t

$$\sum_{h=0}^{\infty} \beta^h \left[u'(C_{t+h})(1-\tau_{t+h}) \frac{W_{u,t+h} N_{u,t+h}}{P_{t+h}} - v'(N_{t+h}) N_{u,t+h} - \frac{1}{2\psi} \left(\frac{W_{u,t+h}}{W_{u,t+h-1}} - 1 \right)^2 \right], \tag{60}$$

which depends on the marginal utililies of aggregate quantities, which ignores the effect of the union decision on the distribution.[5] The union sets wage monopolistically, considering the demand curve of the labor packer

$$N_{kt} = \left(\frac{W_{kt}}{W_t} \right)^{-\varepsilon} N_t. \tag{61}$$

Under these assumptions the first order condition reads

[5]In this I follow Wolf (2021), Mckay and Wolf (2022) and Guerreiro (2023). Using the approach in Auclert, Rognlie and Straub (2024) of considerng an average utility keeping track of these distributional consequences would provide little quantitative difference at a great computational cost.

$$\left(e^{\pi_t^w}-1\right)e^{\pi_t^w}=\psi(\varepsilon-1)\left[-u'(C_t)(1-\tau_t)Y_t+\frac{\varepsilon}{\varepsilon-1}v'(N_t)N_t\right]+\beta_t\left(e^{\pi_{t+1}^w}-1\right)e^{\pi_{t+1}^w}, \quad (62)$$

and can be linearized to

$$\pi_t^w=\kappa_w\left(\sigma^{-1}\hat{c}_t+\phi^{-1}\hat{n}_t-(\hat{y}_t-\hat{\tau}_t-\hat{n}_t)\right)+\beta\pi_{t+1}, \quad (63)$$

where $\kappa_w=\psi\varepsilon v'(N)N$.

D Posterior Distribution

I use a standard Random Walk Metropolis Hastings algorithm, with a multivariate normal for the proposal distribution. The variance-covariance matrix is initially assumed to be equal to the prior variance-covariance matrix, scaled by a constant. I use the first N_a draws to estimate the variance- covariance matrix of the proposal distribution, updating the proposal variance-covariance matrix to the observed variance-covariance matrix of parameters in the first N_a draws. Once updated, I sample another N_b+N_c draws, burn the first N_b and keep the last N_c draws, which I use as our posterior distribution. I set $N_a=N_c=100000, N_b=50000$. Theacceptance rates for all models considered range between 20 and 30 percent.