# Departing from full information, rational expectations (FIRE)

**Failings of the representative-agent model.** The representative-agent model has very sharp predictions for the transmission of real interest rates to consumption. Recall that if $r_t^{ante}$ is the ex-ante interest rate (between $t$ and $t+1$), then we have the Euler equation

$$C_t^{-\sigma} = \beta(1 + r_t^{ante})C_{t+1}^{-\sigma} \tag{1}$$

which can be log-linearized to get

$$\frac{dC_t}{C} = \sigma^{-1}\frac{dr_t^{ante}}{1+r} + \frac{dC_{t+1}}{C} \tag{2}$$

Substituting (2) back into itself $n$ times, it becomes

$$\frac{dC_t}{C} = \frac{dC_{t+n+1}}{C} + \sigma^{-1}\sum_{s=0}^{n}\frac{dr_{t+s}^{ante}}{1+r} \tag{3}$$

and then assuming that $dC_{t+n+1} \to 0$ as $n \to \infty$ in general equilibrium,[1] in the limit this becomes just

$$\frac{dC_t}{C} = \sigma^{-1}\sum_{s=0}^{\infty}\frac{dr_{t+s}^{ante}}{1+r} \tag{4}$$

In our matrix notation we wrote this as $d\mathbf{C} = \sigma^{-1}C\mathbf{U}d\mathbf{r}$, where $d\mathbf{r} \equiv (dr_0^{ante}/(1+r), dr_1^{ante}/(1+r), \ldots)$, and $\mathbf{U}$ is the matrix is 1s on and above the main diagonal and 0s below.

**Two big problems.** There are two aspects of (4) that strike many people as unrealistic. First is that it is extremely *forward-looking*: every current and future shock $dr_{t+s}^{ante}$ to interest rates affects today's equilibrium consumption (and, holding other sources of demand equal, output) by the same amount, whether $s = 0$ or $s = 1000$. This is part of the so-called *forward guidance puzzle* in the New Keynesian model, which is that "forward guidance" in the model—commitments to future interest rate policy—seems to many macroeconomists to have too large an effect on consumption and output.[2]

Second, the response happens *immediately*, without any lags. This is in contrast to the long-held view in macroeconomics that monetary policy works with "long and variable lags". More concretely, it is not consistent with estimated impulse responses for the response to a monetary policy shock, which tend to feature hump-shaped impulse responses that start near zero, then become larger, and finally return to zero.[3]

**Quantitative implications.** The combined effect of these two forces—that the far future matters, and that the response happens immediately—leads to quantitative predictions that are remarkably far from reality. For instance, the 10-year TIPS yield, which is a decent proxy for the average real interest rate expected over

[1] This is not guaranteed to happen in general equilibrium, since with exogenous real interest rates there is a multiplicity of equilibria: we can always increase every $dC_t$ by some constant and still satisfy (3). For instance, if $r_t^{ante}$ is constant, any constant level of $C_t$ is in principle an equilibrium consistent with the Euler equation (1). Out of this multiplicity of equilibria, however, in this class we always focus on the equilibrium where $dC_t \to 0$ as $t \to \infty$. There are several justifications for this. One is comparability with heterogeneous-agent models, where the unique equilibrium will generally feature $dC_t \to 0$ in response to a non-permanent shock. Another is sensibility—it doesn't seem like a temporary shock should cause a permanent general equilibrium shift in output. Perhaps the best justification is that in the very long run, we think that in the absence of shocks monetary policy will ensure that we revert to the "natural" steady state, which it could do by implementing a Taylor rule even, say, one million periods in the future.

[2] See, for instance, McKay, Nakamura and Steinsson (2016) (but not their claim that a heterogeneous-agent model solves the problem, which as we've seen is not always true, and turns out to depend on some idiosyncrasies of their model and calibration).

[3] See, for instance, Christiano, Eichenbaum and Evans (2005), who match the hump in consumption by adding habits.

the next 10 years, was approximately -1% at the beginning of 2022. Now, in May 2024, it is around 2%. Most of this shift has come from tightening monetary policy (and anticipated monetary policy) as markets and the Fed have realized that higher rates are necessary to combat inflation.

From the standpoint of the model, this turns out to be a very large shift. If we suppose that the shift in expected rates was in only the decade 2024–2032, then the cumulative change in future $r$ has been approximately $10 \times (2\% - (-1\%)) = 30\%$. Given that steady-state $r$ is near zero, this implies a proportional change in consumption from (4) of about $-\sigma^{-1} \times 30\%$. With the common choice of $\sigma = 1$, we should have seen (all else equal) a 30% decline in consumption! Needless to say, we didn't see that at all—indeed, there is a debate about whether monetary policy had much effect at all.

This is only meant to be a rough, illustrative calculation. One can quibble in many ways—for instance, by saying that perhaps all else was not equal, or that the true elasticity of intertemporal substitution $\sigma^{-1}$ is much lower. But the gap between model and reality is so enormous that it's very hard to rescue the Euler equation (1): it predicts a far larger effect than we see in practice.

**Heterogeneous-agent models on their own not a fix, but seem complementary to deviations from rational expectations.** It seems intuitive to some people that heterogeneous-agent ("HANK") models, like the ones we've seen, might fix these problems, or at least the forward guidance puzzle. After all, in partial equilibrium, households do not respond as much to future real interest rates: many of them are hand to mouth, and the rest have endogenously shorter horizons (because they anticipate that they might hit constraints in the future, and because they don't want to deviate too much from their wealth targets).

But as we've seen, a basic HANK model actually delivers effects of real interest rates (whether shocked today or in the future) very similar to those in a representative-agent ("RANK") model, and sometimes even larger. The reason is that even though the "direct" effect of future real interest rates on today's consumption is smaller, the "indirect" effects—especially those working through endogenously higher output feeding back into consumption—are larger, because the HANK model has high MPCs. So on its own, this turns out to be a dead end if we want to address the forward guidance puzzle.

But there is a ray of hope. The "indirect" effects that make future real interest rates powerful in HANK models work in part through expectations: people anticipate higher income in the future due to higher output, and consume more today. If we dampen these expectations—perhaps because people don't realize how much output will increase—then we will presumably dampen the effects of forward guidance. Indeed, we saw an example of this at the end of the monetary policy lecture, where we made one ad-hoc change: we shut down all expectations of future output changes, so that households would not consume out of anticipated future income.[4] This one change alone substantially decreased the power of a real interest rate shock, and was especially important for the effect of forward guidance. The same change had very little effect in the RANK and TANK models, neither of which has large indirect effects that operate through expectations.[5]

This suggests that we should further investigate the intersection of HANK models with deviations from full information and rational expectations ("FIRE"). Indeed, that's where we'll go. But first we'll return to the RANK model, to introduce deviations from FIRE in a simpler setting. In RANK, we'll have to also dampen expectations of future real interest rate changes to get anything, since the "indirect" effects in

[4] Importantly, households still anticipated the full path of shocked real interest rates, just not the higher general equilibrium output that would result from them.

[5] The indirect effects in the TANK model are substantial, but operate through *current* income affecting the *current* consumption of the hand-to-mouth, and this does not involve expectations of the future.

RANK are quite small.

## Deviations from FIRE in the RANK model

**Cognitive discounting.** We will consider two main deviations from full information and rational expectations (FIRE). First, we consider so-called *cognitive discounting*, introduced by Gabaix (2020). This is a somewhat ad-hoc deviation from FIRE, which assumes that for all variables $u$ periods in the future that enter their decision problem, agents "cognitively discount" deviations from the steady state by $\alpha^u$, where $\alpha \in [0,1]$ is some parameter. Essentially, if there is some shock to $dX_{t+u}$ that should be known at date $t$, agents will act at date $t$ as if they think the shock is actually $\alpha^u dX_{t+u}$.

You can probably guess that the idea behind cognitive discounting is to try and dampen the forward guidance puzzle: if households act as if future shocks aren't as large as they actually will be, then those shocks will have a smaller effect today.

Let's now think about the effect of a shock $dr_s$ to the real interest rate in the RANK model when there is cognitive discounting. For simplicity, let's assume that aggegate assets are zero (i.e. there are no bonds) and that output equals consumption (i.e. there is no government spending or other sources of demand).

For $t > s$, there are no more shocks, so we should be back at steady state with $dY_t = dC_t = 0$. At $t = s$, the shock $dr_s$ is not cognitively discounted, so we should have the FIRE outcome $\frac{dY_s}{Y} = \frac{dC_s}{C} = \sigma^{-1}\frac{dr_s}{1+r}$. Then, for any $t < s$, we can write (2) as

$$\frac{dC_t}{C} = \alpha \frac{dC_{t+1}}{C}$$

where all future variables enter the household's decision via $dC_{t+1}$, so we can just discount that by $\alpha$ (and $dr_t = 0$). It follows that for $t \leq s$

$$\frac{dY_t}{Y} = \frac{dC_t}{C} = \alpha^{s-t}\sigma^{-1}\frac{dr_s}{1+r} \tag{5}$$

so that relative to the standard RANK model, we get "discounting" by the factor $\alpha^{s-t}$, which grows with the horizon $s - t$ until the real interest rate shock.

**Sticky information / expectations.** *Sticky information*, introduced by Mankiw and Reis (2002), is an alternative deviation from FIRE, which assumes that agents only update their information set about aggregate variables each period with iid probability $1 - \theta$. When they do update, they then have full information and rational expectations.

In response to an MIT shock at date 0 that changes the path of aggregate variables, this implies that at date 0, fraction $1 - \theta$ of agents know about the shock (those who updated right away at date 0), then at date 1, fraction $1 - \theta^2$ of agents know (those who updated at either dates 0 or 1), and so on. In general, at date $t$, a fraction $1 - \theta^{t+1}$ of agents knows about the changes that have occurred. Note the difference vs. cognitive discounting: here, the friction depends on the time that has elapsed since the MIT shock first happened at date 0 (when a FIRE agent would have learned about everything). In contrast, the effect of cognitive discounting depends on the time until a variable will deviate from steady state.

For instance, if real interest rates are expected to deviate at date $s = 100$, then by date 20, if $\theta = 0.8$, sticky-information agents will have nearly full information about $dr_{100}$: $1 - 0.8^{20} \approx 0.988$ of them will know about this change and react to it (and its consequences). But even with $\alpha = 0.9$, cognitively-discounting agents at date 20 will essentially ignore this change in rates, since the discount factor is $0.9^{100-20} \approx 0.0002$.

One can show, although it is a bit more tricky than our deviation above, that sticky information in the

RANK model changes the effect of a future real interest rate shock by approximately the factor $1-\theta^{t+1}$, equal to the fraction of people who are aware of the shock. (This is not exact, but would be exact if there were no indirect effects, which are nearly zero in the RANK model anyway.)

The pure "sticky information" friction can be hard to interpret or deal with, because it is not clear what it means for an agent not to be aware of a change in, say, real interest rates or output that has already happened. Presumably, these changes affected a household's budget constraint—does the household not know how much assets it has?[6] It's simpler to suppose that the household is aware of everything that has happened so far, and that the friction only affects its expectations about the future. This is the so-called *sticky expectations* variant of the sticky information model, introduced by Carroll, Crawley, Slacalek, Tokuoka and White (2020). Under sticky expectations, at date $t$, a household in the fraction $\theta^t$ that has not yet updated since date 0 is still aware of any changes in, say, real interest rates or income that have happened at dates $0,\dots,t$, and how those changes have affected its current level of assets.[7]

Sticky information and sticky expectations tend to generate *lags* and *humps* in impulse responses. For instance, if we assume a very persistent (but not permanent) real interest rate shock, its effect will start closer to zero because few households have yet updated their information to realize that rates are changing. Then the effect will increase with $t$ as a larger share $1-\theta^t$ of households become aware that rates have changed. Finally, the effect will dwindle down as the shock itself goes away. This kind of "hump-shaped" response is what we tend to see in empirical studies, e.g. Christiano et al. (2005), and it also seems realistic that the effect of a shock to rates doesn't happen immediately. If we want, we can interpret the sticky information/expectations friction as a lag not just in forming expectations but also in *reoptimizing a plan*—e.g. maybe you become aware that mortgage rates have changed, but it takes a while for you to rethink your housing decisions to reflect that.

**Summary.** So far, we have two puzzles (too much forward-lookingness, and lack of "humps"), and two deviations from FIRE that address them (the cognitive discounting and sticky information/expectations frictions, respectively).

It's simple enough to solve for how these matter in a RANK economy, but once we have HANK, things become more difficult, because there can be many feedbacks: for instance, the frictions affect equilbrium output, and then that in turn affects consumption (but through non-FIRE expectations), and so on. This is sometimes thought to be a prohibitive complication, but it will turn out that there is a simple way to modify our sequence-space Jacobian-based analysis (i.e. the $\mathbf{M}$ matrices) that incorporates this.[8]

## A general framework for deviations from FIRE

**Summarizing frictions with the "E" matrix.** We'll now introduce a more general framework that encompasses both cognitive discounting and sticky information/expectations, in addition to some other deviations from FIRE.

We define the matrix $\mathbf{E} = [E_{ts}]$ to give the expectation of agents at date $t$ of deviations of variables from steady state at date $s$, as a fraction of the FIRE belief. (Date 0 is, as usual, defined as the date at which a FIRE agent would have perfect foresight.)

[6] Maybe this seems possible, but it introduces other complications we'd rather not deal with: for instance, a household that doesn't know its level of assets might want to spend an amount that violates its borrowing constraint.

[7] Date $t$ itself is a boundary case, and sometimes we will assume that the household doesn't know about date $t$.

[8] This general approach was first introduced in Auclert, Rognlie and Straub (2020), but the notation here is a generalization that is not yet in the public version of the paper.

The **E** matrix for cognitive discounting is

$$\mathbf{E}^{cd} \equiv \begin{pmatrix} 1 & \alpha & \alpha^2 & \cdots \\ 1 & 1 & \alpha & \cdots \\ 1 & 1 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{6}$$

Here, we see the distinctive pattern of cognitive discounting: agents "discount" any variation around the steady state that will happen in the future (the upper triangle of the matrix, where $t < s$), by an amount that depends on how far in the future the change is (i.e. how far to the right of the main diagonal we are).

The **E** matrix for pure sticky information is

$$\mathbf{E}^{si} \equiv \begin{pmatrix} 1-\theta & 1-\theta & 1-\theta & \cdots \\ 1-\theta^2 & 1-\theta^2 & 1-\theta^2 & \cdots \\ 1-\theta^3 & 1-\theta^3 & 1-\theta^3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{7}$$

and the more practical sticky expectations version, which is what we'll use, imposes full information (1s) on and below the main diagonal,[9], giving

$$\mathbf{E}^{si} \equiv \begin{pmatrix} 1 & 1-\theta & 1-\theta & \cdots \\ 1 & 1 & 1-\theta^2 & \cdots \\ 1 & 1 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{8}$$

Here, expectations of future changes vs. steady state are dampened by an amount that depends on how many periods it has been since date 0.

Note that here we are not talking here about *which* households have updated their information sets. That's because we are going to be looking at linear solutions, and in that case, all that we need to know is the average belief, assuming that beliefs are uncorrelated with all other aspects of a household (i.e. its state). Assuming that beliefs are uncorrelated with household characteristics, of course, is a limitation. We can move beyond it by assuming multiple permanent types of households and giving them different **E** matrices, but going beyond this and further endogenizing information becomes much more difficult.

We can, in principle, imagine different **E** matrices for different variables: perhaps households are more aware of future changes in interest rates (because those are nicely summarized in market long-term interest rates) than future changes in output.

**Detour: what kinds of frictions can't be captured using this framework?** This is a useful framework to capture simple frictions like cognitive discounting and sticky expectations. But it requires that our expectations about a variable (say, income) changing at date $s$ in response to an MIT shock at date 0 is, for any $t$, proportional to the actual change in the variable at date $s$. We deviate from FIRE only by assuming that

[9] As mentioned earlier, sometimes the status of the main diagonal is ambiguous: depending on the shock, we can sometimes suppose people don't know what is happening today.

this constant of proportionality doesn't have to be one.[10]

There are many conceivable deviations from FIRE that are not captured by this framework. Perhaps the most prominent case would be "adaptive" expectations, where agents form expectations over income (or other variables) based on the income they've actually observed. In this case, expectations about income at date $s$ are based on realized incomes at many dates in the past, which is not consistent with the $\mathbf{E}$ matrix framework. Similar techniques can potentially deal with this case, but they're more complicated.

**Jacobian modification using the E matrix: the cognitive discounting case.** Let's go back to the setting where the $\mathbf{E}$ matrix applies. Now we get to the tricky, key innovation.

Before we consider the general case, let's recall that the $\mathbf{M} = [M_{ts}]$ matrix summarizes the aggregate consumption response at date $t$ to an income shock at date $s$:

$$\mathbf{M} \equiv \begin{pmatrix} M_{00} & M_{01} & M_{02} & \cdots \\ M_{10} & M_{11} & M_{12} & \cdots \\ M_{20} & M_{21} & M_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{9}$$

Let's assume that (9) is the matrix for FIRE households. What happens if those households cognitively discount income?

First, there should be no change in column 0, because there is never any discounting of income at date 0: it's received and known right away.

Column 1 should be different, because at date 0, households cognitively discount the date-1 income, so that their consumption response will be only $\alpha M_{01}$. Then, at date 1, they will receive the full income, a fraction $1-\alpha$ of which will come as a surprise to them. The consumption response to a *surprise* income shock is just $M_{00}$, which is what the reaction here will look like. The response at date 1 will therefore be $\alpha M_{11} + (1-\alpha)M_{00}$. This will carry through to date 2, where the response will be $\alpha M_{21} + (1-\alpha)M_{10}$: a fraction $\alpha$ of the income was known 1 period in advance, but a fraction $1-\alpha$ of the income was a surprise.

Finally, let's consider column 2. Now the date 0 consumption response is only $\alpha^2 M_{02}$, because of the cognitive discounting of the income two periods later by $\alpha^2$. Then, at date 1, the cognitive discounting shrinks to $\alpha$. It follows that a fraction $\alpha - \alpha^2$ will be treated as an income surprise, still expected to come in the next period, at that date. The consumption response is then $\alpha^2 M_{12} + (\alpha - \alpha^2)M_{01}$. Finally, at date 2, the cognitive discounting disappears, and there is a positive income shock of $1-\alpha$. Here, the consumption response is $\alpha^2 M_{22} + (\alpha-\alpha^2)M_{11} + (1-\alpha)M_{00}$.

The matrix is therefore modified to become

$$\widetilde{\mathbf{M}}^{cd} \equiv \begin{pmatrix} M_{00} & \alpha M_{01} & \alpha^2 M_{02} & \cdots \\ M_{10} & \alpha M_{11} + (1-\alpha)M_{00} & \alpha^2 M_{12} + (\alpha-\alpha^2)M_{01} & \cdots \\ M_{20} & \alpha M_{21} + (1-\alpha)M_{10} & \alpha^2 M_{22} + (\alpha-\alpha^2)M_{11} + (1-\alpha)M_{00} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{10}$$

The idea is that revisions of expectations about the future cause changes in the consumption plan that are given by the original $\mathbf{M}$ matrix, where date 0 was the date at which information arrived.

[10] Note, however, that we could always have *over*reaction: we could choose some elements of $E$ to be greater than 1.

**The general case with the E matrix.** In the general case, if we have some **E** matrix, then in response to an income change that will happen at date $s$, at each date $\tau$ prior to $s$, there will be an update in expectations vs. the previous period of

$$E_{\tau,s} - E_{\tau-1,s}$$

This response to this (including in future periods) will be the same as the response of FIRE agents to a shock at $s - \tau$ that they learn about at date 0, which is in the column $s - \tau$ of the original $\mathbf{M}$.

In response to income at $s$, we can think of there as being multiple "shocks" (here we're assuming that there is complete information at date $s$, i.e. all 1s below the main diagonal):

- First, at date 0, agents' original knowledge $E_{0,s}$ of the income change
- Next, at date 1, their update $E_{1,s} - E_{0,s}$, which acts as an informational shock
- ...
- Finally, at date $s$, the actual income change occurs, of which the fraction $E_{s,s} - E_{s-1,s}$ comes as an informational shock

We can add up the contribution of each one of these "shocks" to outcomes at date $t$, using the reasoning above:

- First, we have $E_{0,s}M_{t,s}$, which is just the FIRE response to the income at date $s$, scaled by the actual knowledge of that income at date 0.
- Next, for $t \geq 1$, we have $(E_{1,s} - E_{0,s})M_{t-1,s-1}$, which is the response to the informational update $E_{1,s} - E_{0,s}$ at date 1 to the shock that is now $s - 1$ periods in the future.
- Next, for $t \geq 2$, we have $(E_{2,s} - E_{1,s})M_{t-2,s-2}$.
- ...
- Finally, for $t \geq s$, we have $(E_{s,s} - E_{s-1,s})M_{t-s,0}$.

If we sum up all these contributions, we'll get the following formula for the response of consumption at date $t$ to income at date $s$, which we denote by the "modified" matrix $\widetilde{M}_{t,s}$.

$$\widetilde{M}_{t,s} = \sum_{\tau=0}^{\min(t,s)} (E_{\tau,s} - E_{\tau-1,s})M_{t-\tau,s-\tau} \tag{11}$$

where we adopt the convention $E_{-1,s} = 0$ for all $s$ for the $\tau = 0$ term.

To give some examples of (11), for $\widetilde{M}_{2,3}$ it is

$$\widetilde{M}_{2,3} = E_{0,3}M_{2,3} + (E_{1,3} - E_{0,3})M_{1,2} + (E_{2,3} - E_{1,3})M_{0,1} \tag{12}$$

and for $\widetilde{M}_{4,2}$ it is

$$\widetilde{M}_{4,2} = E_{0,2}M_{4,2} + (E_{1,2} - E_{0,2})M_{3,1} + (E_{2,2} - E_{1,2})M_{2,0} \tag{13}$$

where we require that $E_{2,2} = 1$.

Cognitive discounting is just a special case of (11). Let's check to make sure the formulas are consistent for $\widetilde{M}^{cd}_{2,2}$. There, we have

$$\begin{aligned}\widetilde{M}^{cd}_{2,2} &= E^{cd}_{0,2}M_{2,2} + (E^{cd}_{1,2} - E^{cd}_{0,2})M_{1,1} + (E^{cd}_{2,2} - E^{cd}_{1,2})M_{0,0} \\ &= \alpha^2 M_{22} + (\alpha - \alpha^2)M_{11} + (1-\alpha)M_{00}\end{aligned}$$

which is indeed exactly the 2,2 entry of (10).

**How can we use this?** The beauty of the general formula (11) is that once we use it to modify **M** matrices (and other sequence-space Jacobians), we can use them exactly as before to calculate the effects of fiscal and monetary policy: just substitute the modified matrix where before we would have used the FIRE one. See the accompanying **lecture 10 computations.ipynb** notebook for applications.

**Addendum: alternative formulation in terms of the fake news matrix (optional).** If we let **F** be the "fake news matrix" associated with **M** (and similarly $\widetilde{\mathbf{F}}$ with $\widetilde{\mathbf{M}}$), defined as before by $F_{t,s} = M_{t,s} - M_{t-1,s-1}$ for $t, s > 0$ and $F_{t,s} = M_{t,s}$ otherwise, then it is possible to rewrite (11) as

$$\widetilde{M}_{t,s} = \sum_{\tau=0}^{\min(t,s)} E_{\tau,s} F_{t-\tau,s-\tau} \tag{14}$$

Recall that $F_{t,s}$ can be interpreted as the effect in period $t$ of *knowing* in period 0 that there will be a shock at period $s$. (If we want the effect in period $t$ of knowing in period 1 that there will be a shock at $s$, then we would use $F_{t-1,s-1}$; knowing in period 2 would be $F_{t-2,s-2}$; and so on. The matrix is defined with the convention that information is learned at date 0.)

Equation (14) writes the spending at period $t$ out of income $s$ under non-rational expectations as the sum of these terms, times the fraction of people who actually knew at each date. For instance, we can write $\widetilde{M}_{2,3} = E_{0,3}F_{2,3} + E_{1,3}F_{1,2} + E_{2,3}F_{0,1}$: the sum of the fraction $E_{0,3}$ who knew about the date-3 shock at date 0 times the effect $F_{2,3}$ of their knowledge at date 2, plus the fraction $E_{1,3}$ who knew at date 1 times the effect $F_{1,2}$ of their knowledge, plus the fraction $E_{2,3}$ who know in date 2 times the effect $F_{0,1}$ of their knowledge.

We note that in the rational expectations case with $E_{t,s} \equiv 1$, (14) reduces to our original formula for deriving the Jacobian from the fake news matrix.

## References


Auclert, A., Rognlie, M. and Straub, L. (2020). Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model. Working paper, National Bureau of Economic Research.

Carroll, C.D., Crawley, E., Slacalek, J., Tokuoka, K. and White, M.N. (2020). Sticky Expectations and Consumption Dynamics. *American Economic Journal: Macroeconomics* 12(3):40–76.

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

Gabaix, X. (2020). A Behavioral New Keynesian Model. *American Economic Review* 110(8):2271–2327.

Mankiw, N.G. and Reis, R. (2002). Sticky Information Versus Sticky Prices: a Proposal to Replace the New Keynesian Phillips Curve. *The Quarterly Journal of Economics* 117(4):1295–1328.

McKay, A., Nakamura, E. and Steinsson, J. (2016). The Power of Forward Guidance Revisited. *American Economic Review* 106(10):3133–58.