Noisy Memory and Over-Reaction to News

By Rava Azeredo da Silveira and Michael Woodford*

The hypothesis of rational expectations (RE) proposes that decisions are based on expectations that make use of all available information in an optimal way: that is, those that would be derived by correct Bayesian inference from an objectively correct prior and the data that has been observed to that date. Yet both surveys of individual forecasts of macroeconomic and financial variables and forecasts elicited in experimental settings are both more heterogeneous than this hypothesis should allow, and errors that are predictable on the basis of variables observable by the forecasters, contrary to this hypothesis. And it is arguable that the dynamics of macroeconomic and financial variables are more easily explained on the hypothesis of expectations that respond to news in systematically biased ways, as proposed for example by Andreas Fuster, Ben Hébert and David Laibson (2011).

One reason for expectations to be both heterogeneous and biased is inattention to current conditions on the part of decision makers (DMs). In the "rational inattention" model of Christopher A. Sims (2003), decisions are assumed to be optimal, conditional on their having to be based on an imprecise internal representation of the situation, rather than on the DM's true situation; and the nature of the imprecision in the DM's perceptions is also optimal (given the decision problem), subject to a constraint on the complexity of possible internal representations.

Such a theory can easily account for insensitivity or delayed reaction to changing conditions, as when prices appear to be "sticky" in response to a monetary disturbance. But often people appear instead to *over-react* to news, relative to the decisions that would be made in the RE case. As discussed below, Pedro Bordalo *et al.* (2018) present evidence of over-reaction in professionals' forecasts of a variety of macroeconomic and financial series, while Augustin Landier, Yueran Ma, and David Thesmar (2017) show this even more directly in the case of forecasts by laboratory subjects. Similarly, Fuster, Hébert and Laibson (2011) discuss aspects of macroeconomic and financial dynamics that they attribute to over-extrapolation of short-run trends in economic series. Such patterns of over-reaction might seem to require sources of bias other than an optimal response to a fuzzy perception of one's situation.

We show instead that over-reaction of the kind documented by these authors is consistent with a model of rational inattention. Our model differs, however, from the kind proposed by Sims (2003). In the Sims model, the information constraint limits the precision of new observations of one's situation, but a DM is assumed to have perfect memory of all past observations; and there are perfect records of all past data in the external environment as well, so that events are equally easily observed anytime after they happen. In our model, instead, the crucial cognitive constraint is on the precision of memory, which is furthermore the only source of access to past events.

I. A Model of Noisy Memory

As an illustration of our theory, consider a problem in which a DM observes realizations of an i.i.d. random variable, y_t, drawn from a Gaussian distribution with an unknown mean μ but (for simplicity) a known variance σ_y^2. The DM's prior is that μ is drawn from a distribution $N(0, \Omega)$. Each period, after observing y_t the DM must produce an estimate z_t of a forward-looking

* Azeredo da Silveira: Ecole Normale Supérieure, Paris, rava@ens.fr. Woodford: Columbia University, New York, mw2230@columbia.edu. We thank Yeji Sung for research assistance, and Ben Hébert, David Laibson, Yueran Ma and Andrei Shleifer for helpful discussions.

moving average of the state, the true (ex post) value of which will be

$$z_t^* \;\equiv\; (1-\beta)\sum_{j=0}^{\infty}\beta^j y_{t+j},$$

where $0 < \beta < 1$ is a discount factor.

Since the optimal information structure depends on the penalty for inaccurate estimates, we assume that the DM's goal is to minimize the discounted mean squared error (MSE) of her estimate. This can be justified as corresponding to discounted expected utility maximization in a consumption-smoothing problem. In this application, y_t is the DM's income (other than from assets) in period t, and her real asset holdings a_t evolve according to

$$a_{t+1} \;=\; \beta^{-1}\,[a_t \,+\, y_t \,-\, c_t],$$

where c_t is consumption, and the real return on assets is assumed to equal the DM's rate of time preference. The DM's problem is to choose a state-contingent path of consumption so as to maximize the expected discounted value of utility, where in each period $u(c_t) \;=\; -(c_t \,-\, c^*)^2$, subject to a transversality condition on asset holdings. If we let $z_t \equiv c_t - (1-\beta)a_t$ be consumption in excess of interest income, then maximization of expected discounted utility is equivalent to minimizing the expected discounted value of $(z_t - z_t^*)^2$. In this application, z_t can be understood as an estimate of the DM's "permanent income" apart from financial assets.

Under the RE assumption, this would be a consumption-smoothing problem of the kind treated by Thomas J. Sargent (1987, chap. XII). Instead, we assume that in period t the DM's choices can depend only on her current *cognitive state,* a vector consisting of a *memory state* m_t and the current observation y_t (assumed for simplicity to be observable with perfect precision). Thus the choice of z_t must be some function of (m_t, y_t). In addition, the conditional probability of different possible memory states m_{t+1} in the following period must also be some function $p_t(m_{t+1}|m_t, y_t)$ of the current cognitive state. The actual state m_{t+1} will be a random draw from this conditional distribution; the randomness reflects the limited precision with which m_{t+1} can reflect the DM's cognitive state at time t.

We do not arbitrarily specify a form for the memory state or its dynamics, but we assume a cost of greater precision. Specifically, we assume a cost each period of storing a memory for retrieval in the following period (or alternatively, a cost of retrieval) that is proportional to I_t, the *mutual information* between the cognitive state (m_t, y_t) and the memory state m_{t+1}; this is a measure of the informativeness of the subsequent memory state about the prior cognitive state (Thomas M. Cover and Joy A. Thomas, 2006).[1] Our hypothesis is that the decision rule $z_t(m_t, y_t)$ each period and the stochastic transition law $p_t(m_{t+1}|m_t, y_t)$ for the memory state minimize the objective

$$\sum_{t=0}^{\infty}\beta^t\,\left\{\mathrm{E}[(z_t - z_t^*)^2] \,+\, \theta\, I_t\right\},$$

where $\theta \,>\, 0$ indexes the cost of greater memory precision. Here the expectation is over the value of μ, the sequence of realizations $\{y_t\}$, and the stochastic evolution of the memory state.

Azeredo da Silveira and Woodford (2018) show that in the solution to this problem, the optimal structure for memory implies that the posterior distribution for μ implied by any memory state m_t will be a Gaussian distribution with a variance that depends only on elapsed time; we can thus index memory states by the implied posterior mean value for μ (so that m_t is a real number). At least for small enough values of the information cost θ, the posterior uncertainty about μ decreases over time (i.e., with additional observations of y_t). In the perfect-memory case ($\theta = 0$), the posterior uncertainty evolves according to the usual Kalman filter formulas, and converges to zero as t becomes large; asymptotically,

[1]For Sims (2003), instead, the cognitive state is a history of subjective observations $(s_t, s_{t-1}, \ldots)$, and the information cost each period is proportional to the mutual information between the new observation s_t and the entire history of objective states $(y_t, y_{t-1}, \ldots)$.

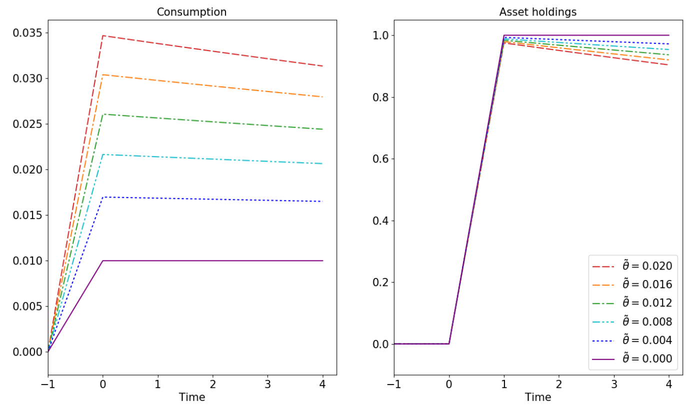

FIGURE 1. RESPONSES OF CONSUMPTION AND ASSET HOLDINGS TO AN INCOME SHOCK

Note: Responses to a unit income innovation at $t = 0$, for several possible values of $\tilde{\theta}$.

the value of μ is known, and observations of y_t have no effect on expectations, as in the RE solution. But when memory is imperfect ($\theta > 0$), the posterior uncertainty asymptotes to a positive level, no matter how many observations have been made. In this case, fluctuations in y_t continue to cause beliefs about μ to fluctuate, forever.

In the asymptotic limit, the joint dynamics of the external state y_t and the DM's memory state m_t are described by a linear system with constant coefficients and Gaussian innovations. Each period, the DM's estimate of μ after observing y_t will be given by

$$\hat{\mu}_t \;\equiv\; \mathrm{E}[\mu\,|m_t, y_t] \;=\; (1-\gamma)m_t \,+\, \gamma y_t,$$

where the *gain coefficient* $0 < \gamma < 1$ depends on the degree of imprecision of memory, and the optimal action will then be

$$z_t \;=\; \mathrm{E}[z_t^*\,|m_t, y_t] \;=\; (1-\beta)y_t \,+\, \beta\hat{\mu}_t.$$

Since the accuracy of future actions will depend only on the accuracy of one's future estimates of μ, the optimal use of finite memory capacity is to store as precise as possible a record of one's current posterior for μ, which is completely summarized by the value of $\hat{\mu}_t$. Thus

$$m_{t+1} \;=\lambda\hat{\mu}_t \,+\, \omega_{t+1},$$

where ω_{t+1} is a mean-zero Gaussian disturbance term representing idiosyncratic memory noise, and both the coefficient $0 < \lambda < 1$ and the variance of the noise term depend on the degree of imprecision of memory.

II. Over-Reaction of Consumption to Income News

We first illustrate the implications of this solution for the consumption problem discussed above. (In these numerical solutions, we assume that $\Omega/\sigma_y^2 = 100$, and consider a variety of possible values for the ratio $\tilde{\theta} \equiv \theta/\sigma_y^2$.) Figure 1 shows the responses of consumption c_t and asset holdings a_t to a unit positive income surprise y_0 at $t = 0$. In the case of perfect memory (the case $\tilde{\theta} = 0$, shown by the solid lines in the figure), consumption permanently increases by an amount equal to fraction $1-\beta$ of the income innovation; there is also a permanent increase in asset balances, by an amount equal to the income innovation, starting at $t = 1$. Thus both consumption and asset holdings are martingales in this RE solution.

Instead, when $\tilde{\theta} > 0$, consumption over-reacts, in the sense that it increases by more than the model-consistent increase in permanent income. However, the increase is no longer permanent; because assets are not increased to the extent that would be necessary to finance a permanent increase in consumption of that size, both consumption and assets fall over time, eventually returning (beyond the borders of the figures) to their initial levels. Thus both consumption and asset holdings are stationary processes in this case (without any need to invoke an endogenous real rate of return or non-time-separable preferences). Moreover, consumption changes are more volatile than they would be under RE.

The model with noisy memory implies that changes in consumption should be forecastable, contrary to the prediction of the RE permanent-income hypothesis stressed by Robert E. Hall (1978). In particular, as shown in Azeredo da Silveira and Woodford (2018), it predicts that consumption changes should be negatively autocorrelated over the medium run, as Fuster, Hébert and Laibson (2011) find to be true of aggregate US nondurable consumer expenditure. As in the model proposed by these authors, the negative autocorrelation reflects over-reaction of consumption to transitory income variation. Our explanation differs from their hypothesis of "natural expectations," however, in that over-reaction is predicted to occur in our model even when the income process is a very low-order autoregressive process (in the numerical illustration here, it is not persistent at all). Under the hypothesis of natural expectations, expectations should be perfectly model-consistent when the true dynamics are given by a low-order autoregressive process; in our model, instead, the prediction of negative autocorrelation is more robust.

III. Over-Reaction of Forecasts

The forecastability of consumption changes reflects the existence of systematic bias in the DM's estimates of the hidden state μ. Figure 2 shows the impulse response of the minimum-MSE estimate $\hat{\mu}_t$ to the same income shock as in Figure 1, for each of the several possible values of $\tilde{\theta}$. In the case of perfect memory (the RE case), there would be no response at all of beliefs (or of forecasts of future income) to current income observations. Instead, when $\tilde{\theta} > 0$, the current income realization is extrapolated into the future to a greater extent than would occur under rational expectations (so that a higher y_t results in a higher forecast of future income). Moreover, the forecast bias (that is, departures of $\hat{\mu}_t$ from the correct value μ) is predicted to be persistent; as shown in the figure, a transitory shock to income results in a shock to the forecast bias that decays only slowly.

Both of these features of the dynamics of forecast bias are observed in the experimental data of Landier, Ma and Thesmar (2017). Letting Δ_t^i be the discrepancy between subject i's forecast of the future value in question and the RE forecast at time t ($\hat{\mu}_t^i - \mu$, in our model), and s_t the difference between the variable observed at time t and the RE forecast of that variable before its realization ($y_t - \mu$, in our model), they regress Δ_t^i on Δ_{t-1}^i and s_t, and find significantly positive values (less than 1) for both coefficients. Our model predicts that these regression coefficients should equal $\lambda(1-\gamma)$ and γ respectively, and thus both lie between 0 and 1.

Bordalo *et al.* (2018) argue that professional forecasters' individual forecasts of many macroeconomic and financial time series also exhibit over-reaction to news. They draw this conclusion from a regression of the demonstration that revisions of a forecaster's forecast of a given variable can predict the difference between the revised forecast and the (eventually revealed) correct value: an upward revision, say, of the forecast increases the extent to which the later forecast is likely to be too high. Our model predicts that this should be observed. In our model, $\hat{\mu}_t^i$ should exhibit stationary fluctuations around the true value μ; hence the covariance between the forecast error $\mu - \hat{\mu}_t^i$ and the forecast revision $\hat{\mu}_t^i - \hat{\mu}_t^i$ should equal $-(1-\rho)$ times the variance of the forecast error process, where $\rho \equiv \lambda(1-\gamma)$ is the coefficient of serial correlation of the forecast errors, a quantity between 0 and 1.

Unlike many models of extrapolative forecast bias in the literature (including the "natural expectations" of Fuster, Hébert and Laibson or the "diagnostic expectations" of Bordalo *et al.*), our model also

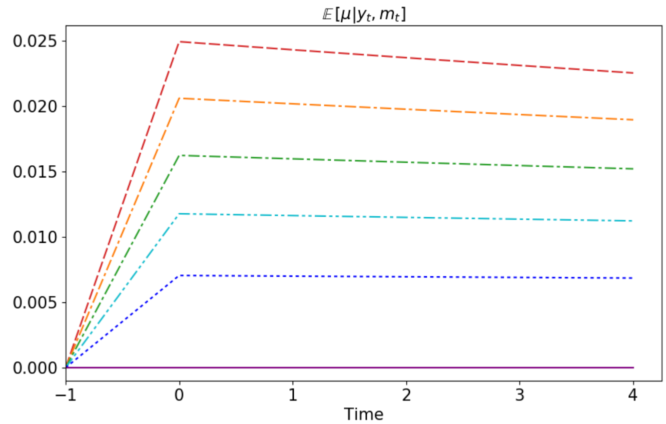

FIGURE 2. RESPONSES OF BELIEFS ABOUT μ TO AN INCOME SHOCK

Note: Responses to a unit income innovation at $t = 0$, for several possible values of $\tilde{\theta}$. The values of $\tilde{\theta}$ are the same as in Figure 1.

provides an explanation for the heterogeneity of individual forecasts, even when different forecasters observe the same information (as is clearly the case in the experiment of Landier, Ma and Thesmar); this follows from the idiosyncratic noise in the evolution of the memory state.

Our model also predicts that the discrepancy between any individual DM's forecast $\hat{\mu}_t^i$ and the average forecast $\hat{\mu}_t$ (averaging over all possible realizations of the idiosyncratic memory noise) should predict subsequent revisions of DM i's forecast, as is shown to be true in surveys of professional forecasters by Fuhrer (2018). In our model, the existence of transitory idiosyncratic variation in beliefs of the kind indicated by Fuhrer's results is not unrelated to the common bias resulting from overreaction to news; instead, in our model, it is the noise in individual forecasters' memories that results in the common bias. Thus the hypothesis of limited memory precision provides a parsimonious explanation for both the common and idiosyncratic bias in individual forecasts.

REFERENCES

Azeredo da Silveira, Rava, and Michael Woodford. 2018. "A Model of Noisy Memory." Unpublished.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2018. "Over-Reaction in Macreconomic Expectations." NBER Working Paper no. 24932, August.

Cover, Thomas M., and Joy A. Thomas. 2006. *Elements of Information Theory.* New York: Wiley, 2d ed.

Fuhrer, Jeffrey. 2018. "Intrinsic Expectations Persistence." Unpublished.

Fuster, Andreas, Ben Hébert, and David Laibson. 2011. "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing." *NBER Macroeconomics Annual*, 26, 1-48.

Hall, Robert E. 1978. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy*, 86, 971-987.

Landier, Augustin, Yueran Ma, and David Thesmar. 2017. "New Experimental Evidence on Expectations Formation." Working Paper.

Sargent, Thomas J. 1987. *Macroeconomic Theory.* Boston: Academic Press, 2d ed.

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics.* 50: 665-690.