Implications of Rational Inattention

C.A. Sims

Cambridge (underground) macro-reading group, 17th January 2012

Some introductory stuff.

- Sims is looking for another way to move from models of seamless markets and rational actors.
- He proposes the idea of information constraints, which is nice because:
  - Chimes with reality.
  - Can account for a wide-range of phenomena via a simple mechanism.
  - One does not need to know how information is processed only its stochastic form.
- These models operate in a fashion whereby the capacity to absorb information is exogenous.
- Given this constraint optimising agents assign endogenous signal noise ratios to relevant state variables and then choose optimal behaviour.
- This means it has a lot of similarities with the signal extraction literature: e.g the Lucas Island model.

Concepts of information theory

- Information flow can be considered as a rate of change in uncertainty.
- In this setting in entropy, $-E(log_2(p(x))$, captures uncertainty, where $p(.)$ represents a density function.
  - Entropy measured in bits: $-0.5log_2(0.5) - 0.5log_2(0.5) = 1\,bit$
  - e.g. the outcome of a coin toss.
- Therefore the information from receiving signal Z can be expressed as the change in entropy:

$$-E(log_2(p(X \mid Z)) \mid Z) + E(log_2(p(x))$$

- The models work by choosing $p(X \mid Z)$ such that expected change in entropy is less than the information constraint.
- Various other considerations: coding, continuous variables, non-negative information etc. but they are .

Information Constrained Optimisation

A stylistic example: Let X be a random variable that is observed through an information constrained channel, such that the maximum information from the signal over X is C. The agent wants to make action Y in order to minimise $E(Y-X)^2$. What is the optimal conditional distribution $Y \mid X$? The problem can be formulated as:

$$min_{q(.)}\{E(Y-X)^2 = \int (y-x)^2 q(y|x)p(x)dydx\}$$

s.t.

$$\int q(y|x)dy = 1, \ \forall x$$

$$-E[E(log_2(q(y \mid x)) \mid x)] + E(log_2(\int q(y|x)p(x)dx) < C$$

The first constraint is to ensure that the density function is well behaved. The second is the information constraint such that the expected change in entropy from the signal on X must be less than an exogenously given value C.

Information Constrained Optimisation (some intuition)

- The mutual information principle means that the information flow is this same whether we condition on observations of X or Y.
- Hence this model reduces to something like a signal extraction problem: the optimal behaviour of Y is consistent with X being observed with a noisy iid signal.
- In the Gaussian case: the choice of $q(.)$, the conditional density function, essentially determines the signal noise ratio.

Information Constrained Optimisation: Numerical Example

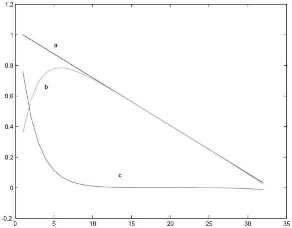

Fig. 1. $C = 0.641$ bpt, $R^2 = 0.856$, linear a.

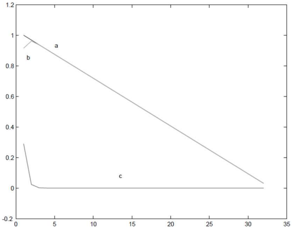

Fig. 2. $C = 3.56$ bpt, $R^2 = 0.992$, linear a.

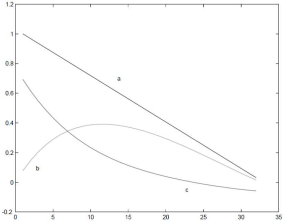

Fig. 3. $C = 0.111$ bpt, $R^2 = 0.577$, linear a.

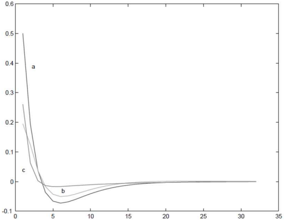

Fig. 4. $C = 0.711$ bpt, $R^2 = 0.443$, rational whipsaw a.

Summary of ideas & comparisons

- Main distinction from signal extraction: the nature of noise is determined endogenously and optimally from the information capacity constraints. Given multiple variables of interest information processing capacity can be reassigned between variables implying the noise of series can potentially fluctuate with economic conditions.
- Time series: impulse responses a exhibit smooth path in response to other variables shocks, consistent with information constraints. Erratic moves in own responses can also be justified by correlated information shocks.

PIH: An analytical example

Consider the following modification to the permanent income model:

$$max[\sum_{t=0}^{\infty}\beta^t(C_t-0.5C_t^2)$$

s.t.

$$W_t=R(W_{t-1}-C_{t-1})+Y_t$$

with $Y_t \sim NID(\bar{Y},\omega^2)$ an exogenously given income process, W is wealth, C is consumption and R is an exogenously given gross interest rate (let $\beta R=1$). This implies W_t can be considered a state variable which the household monitors inattentively, $E(W_t)=\hat{W}_t$, with an information constraint κ. Noting the certainty equivalence of the utility function, solving forward the budget constraint yields:

$$C_t=(1-\beta)\hat{W}_t+\beta\bar{Y}$$

PIH: An analytical example (contd)

Sims says that, given the form of the information constraints and the process for Y_t, $W_t|\mathscr{I}_t \sim NID(\hat{W}_t, \sigma_t)$. Using the budget constraint:

$$E(W_{t+1}) = \hat{W}_t$$

$$Var(W_{t+1}) = R^2\sigma_t^2 + \omega^2$$

The agent chooses sigma such that the wealth signal does not violate the information constraint:

$$\kappa = 0.5(log_2(R^2\sigma_t^2 + \omega) - log_2(\sigma_{t+1}^2)) \tag{1}$$

Which has a steady state: $\sigma^2 = \frac{\omega^2}{e^{2\kappa} - R^2}$ .This is equivalent to the household receiving a noisy signal over its true wealth:

$$\hat{W}_{t+1} = W_{t+1} + \xi_{t+1}$$

With $var(\xi_{t+1}) = \frac{\sigma^2(\omega^2 + \sigma^2 R^2)}{(R^2-1)\sigma^2 + \omega^2}$.

Some thoughts

- The elephant in the room in is the size of the information constraint:
  - How big should it actually be? How do you go about measuring it? This can have very important implications for the results, as in the example.
  - 1 bit is essentially the result of a coin toss, as Sims points out the information capacity of a person over any reasonable time interval is many multiples of this. For these models to work agents must optimally assign little of their attention to monitoring the economy. In which case, can the costs of rational inattention really be so large if the agents behaves this way?
- There is also the point of justifying what the agent is inattentive to. In the PIH model, agents didn't know their own wealth - is this realistic, given that for most people it is just a matter of looking at a bank statement. Macro-variables are obviously more difficult to observe.