RATIONAL INATTENTION: BEYOND THE LQ CASE

CHRISTOPHER A. SIMS

1. INTRODUCTION

- References: Two previous papers by me, listed in the references for the one I suggested as background reading. Also the text by Cover and Thomas and a more recent book by David MacKay that combines discussion of information theory with Bayesian inference and learning algorithms.
- The existing literature using information-theoretic ideas in economics has stuck to the case of Gaussian randomness and (usually) the linear-quadratic optimization problem that rationalizes Gaussian ex post uncertainty as optimal.
- Here we discuss a simple two-period non-LQ problem. The example exercises intuition about how the information flow constraint works, and it has some interesting economic implications.
- We will perhaps also have time to discuss how to generalize to fully dynamic and to general equilibrium setups.

2. ENTROPY, MUTUAL INFORMATION

- For these basic qualitative ideas, and how they relate to macroeconomic inertia, see the appendix to "Stickiness".
- Information flow is measured as reduction in uncertainty.
- We begin with a probability distribution over some object Y; we receive information X; we have a new probability distribution $Y \mid X$ that, on average, implies less uncertainty.
- There is just one measure that works naturally here: $H(Y) = -E[\log(p(Y))]$, the **entropy** of Y. This is unique up to choice of base for the logarithm, with $\log_2$ the conventional choice.
- $I(X;Y) = H(Y) - E[H(Y \mid X)]$ is the **mutual information** between X and Y. It is symmetric and invariant to mon0tone transformations of X and Y — so it depends only on their copula.
- Entropy depends on existence of a density, but can be defined for a density against any base measure. So it covers discrete, continuous, and mixed discrete-continuous cases. But being told the exact value of a variable distributed continuously on the real line, or even being given information that

Date: March 15, 2006.

converts continuously distributed uncertainty to discrete uncertainty, represents infinite information flow.

3. CHANNEL, CAPACITY, CODING

- A **channel** has inputs X and outputs Y. It defines the conditional distributions $Y \mid X$. (More generally, Y might be a stochastic process depending on the history of the input X.)
- If there is an input and an output at each date, the **capacity** of the channel is the maximum, over distributions for X, of $I(X;Y)$.
- The **coding** problem is how to convert the class of message we actually want to send into a sequence of X values that has the capacity-maximizing distribution. By proper coding, any sort of message can be sent with arbitrarily low error probability at any rate below capacity.

4. WHY THIS IS INTERESTING FOR ECONOMICS

- If asset prices are continuous time diffusion processes, and if, as implied by standard theory, optimizing agents are continuously adjusting their portfolios and consumption plans in response to these prices, so that their own actions are diffusion processes determined by the asset prices, then people are processing information at an infinite rate, which is not possible.
- For the mapping between the prices and behavior to represent a finite rate of information flow, the behavior must smooth and delay the effects of the price signals, and the behavior must inject signal-processing randomness.
- This is exactly what we see as time series macroeconomic facts. Both prices and quantities are "sticky", with hump-shaped responses to other variables, yet each variable has a less smooth idiosyncratic component.
- If we know channel capacity, we do not need to know the physics (or biology, or psychology) ot the channel if we assume it is being used optimally. Thus we have a route to modelling inertia with optimizing models, without invoking detailed restrictions on behavior based on psychology or sociology.

5. THE GENERAL STATIC CAPACITY-CONSTRAINED DECISION PROBLEM

$$\max_{f(\cdot,\cdot)\in S} E[U(C,W)] \qquad \text{subject to}$$

$$I(C;W) \leq \kappa$$

$$\int f(c,w)\,dc = g(w)\,.$$

- g is a given marginal density of the exogenous random variable W
- f is the joint density, drawn from the set of permissible densities S, of the choice variable C and W

- $I(C;W)$ is the mutual information between C and W.

6. More explicit mathematical structure

$$\max_{f(\cdot,\cdot)} \int U(c,w) f(c,w)\, dc\, dw \qquad \text{subject to}$$

$$-\int \log(g(w))g(w)\,dw - \int \log\left(\int f(c,w)\,dw\right) f(c,w)\,dc\,dw + \int \log(f(c,w))f(c,w)\,dc\,dw \leq \kappa$$

$$\int f(c,w)\,dc = g(w)$$

$$f \in S$$

Note: linear objective function, concave constraint set

7. FOC

$$\log(U(c,w)) = \lambda\left(1 + \log f(c,w) - 1 - \log\left(\int_c^\infty f(c,w)\,dw\right)\right) + \mu(w) + \psi(c,w)$$

or

$$q(w \mid c) = \nu(w)\exp(U(c,w)/\lambda)$$

8. Equation-solving solution

Find $\nu(w)$ function such that

$$\int \nu(w) \exp\left(\frac{U(c,w)}{\lambda}\right) dw = 1 \qquad \text{all } c$$

and $h(c)$ (its marginal pdf) such that

$$\int q(w \mid c)h(c)\,dc = g(w) \qquad \text{all } w\,.$$

Discretized, this is two linear equation systems that should be recursively solvable. But: Fredholm integral equations of the first kind. Notoriously ill behaved. We find numerically that h is not a density, but has discrete elements.

9. More stable approach

- Discretize c, w space.
- Treat the values of $q(c \mid w)$ at these points as free parameters, subject to the constraint that they sum to one for each w.
- Grind ahead with optimization, not using the FOC.

10. A ONE-SHOT SAVING PROBLEM

Specialize to $U(c, w)$ a standard utility function in c, first-period consumption, and $w - c$, second-period consumption. The domain of f includes only the part of the positive orthant with $w > c$. For example, with log utility:

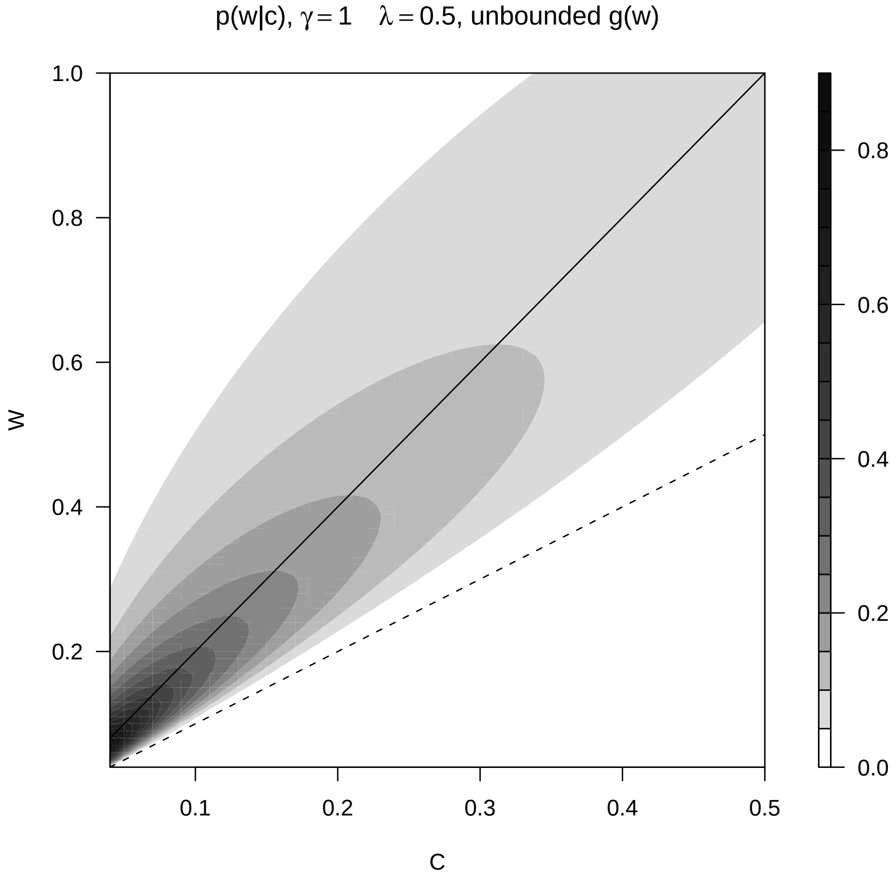
$$\max_f \int_{0<c<w} \log\big(c \cdot (w - c)\big) f(c, w) \, dw \, dc \tag{1}$$

- Wealth is uncertain, both before and after information-collection.
- The information constraint controls how precisely consumption can respond to wealth.

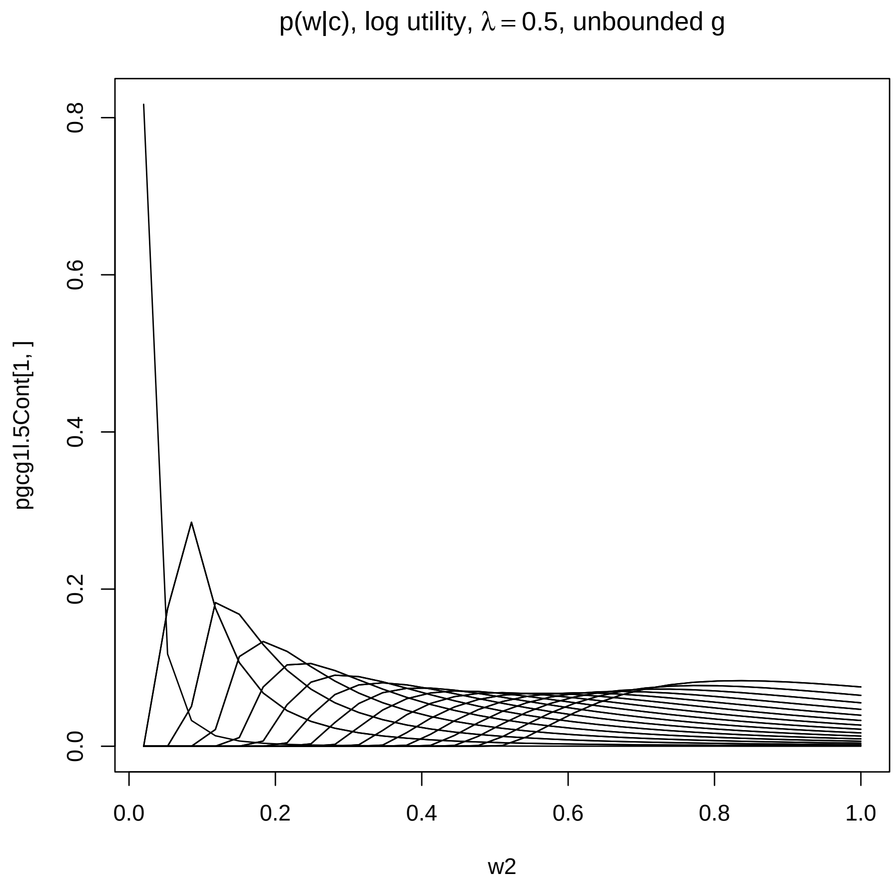
11. ANALYTIC SOLUTION FOR LOG UTILITY

- Setting $q(w - c \mid c)$ to an $F(2\alpha + 2, 2\alpha)$ density, where $\alpha = 1/\lambda$, solves the FOC's.
- Normalized to have constant spread, this does approach normality as $\alpha = 1/\lambda$ approaches infinity
- The peak is at $(2 - 1/(\alpha - 1))c$, the mean at $(2 + 1/(\alpha - 1))c$, for $\alpha > 1$.
- The distribution is more tightly concentrated around $w = 2c$ the larger is $\alpha = 1/\lambda$.
- I.e., as the shadow price on the information constraint declines, we come closer and closer to the certainty solution.

12.


p(w|c), γ = 1 λ = 0.5, unbounded g(w)
1.0
0.8
0.6
0.4
0.2
w
0.1
0.2
0.3
0.4
0.5
c
0.8
0.6
0.4
0.2
0.0

13.



14. PRECAUTIONARY SAVING?

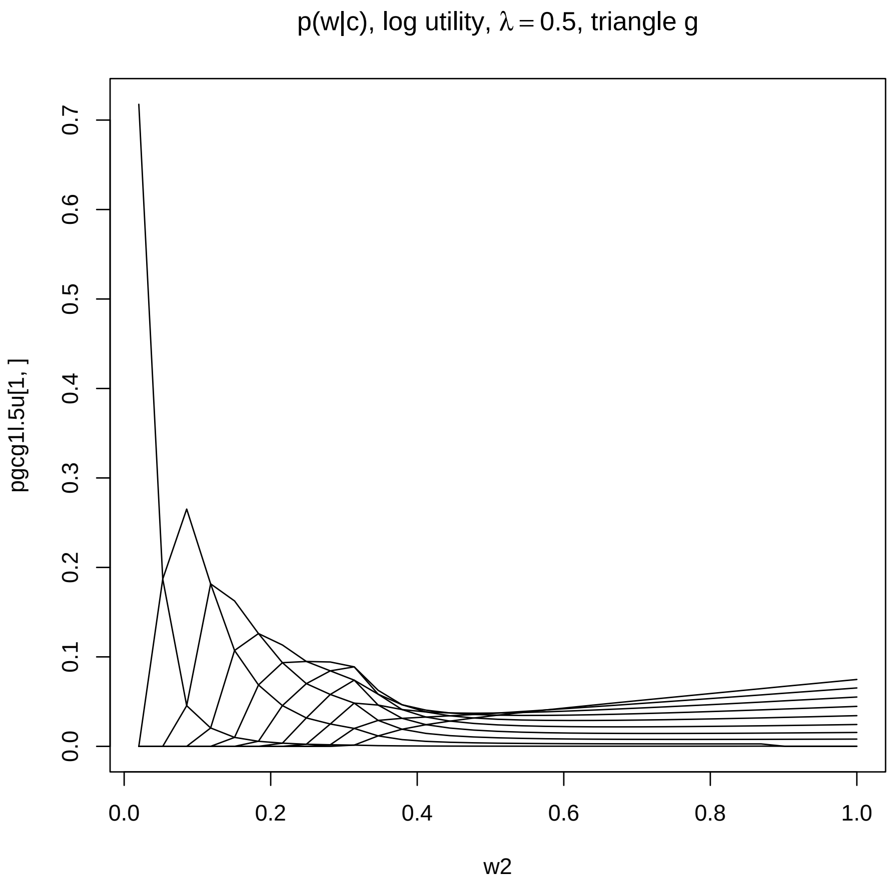
- Yes, with log utility or with discounted linear utility.

15. NUMERICAL SOLUTIONS FOR CRRA UTILITY

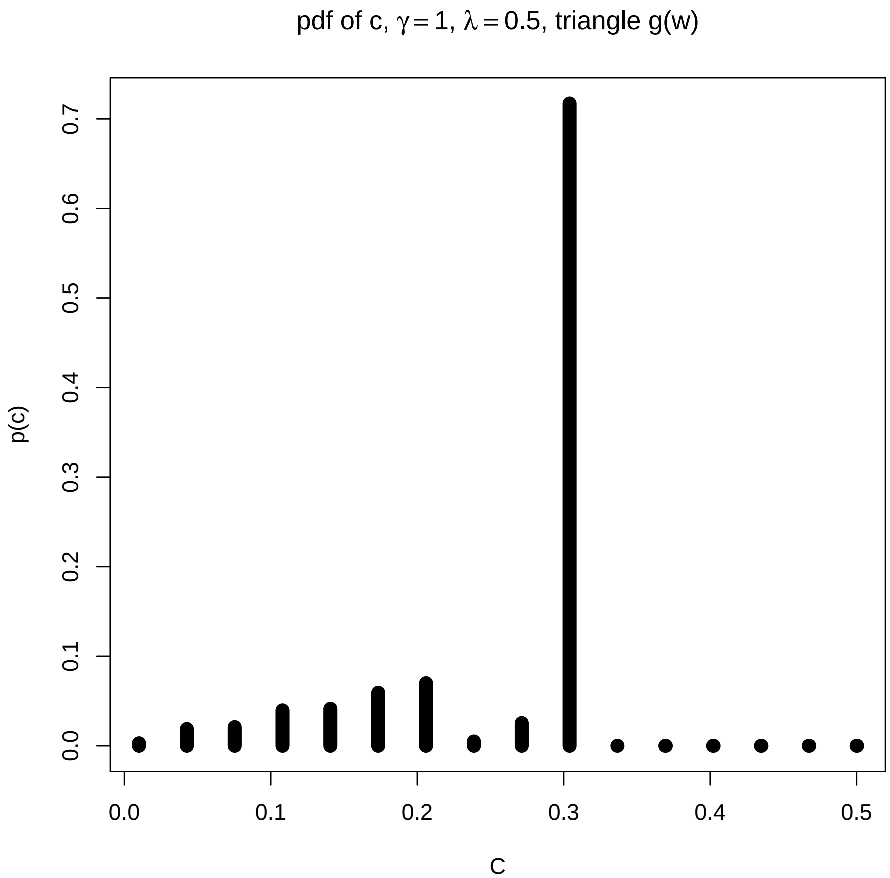
$$U(c,w) = \frac{c^{1-\gamma} + (w-c)^{1-\gamma}}{1-\gamma} .$$

- equi-spaced grid with 16 c values ranging from .01 to .5 and 31 w values ranging from .02 to 1.
- The marginal pdf g for w is given in each case as $g(w) = 2w$.
- After adding-up and $c < w$ constraints imposed, 345 free parameters.
- Convergence in 450 iterations, with each iteration taking about 1.5 seconds (making the whole computation take about 11 minutes) on a 3GHz Pentium 4 running Linux, coding in **R** (an interpreted language, like Matlab).

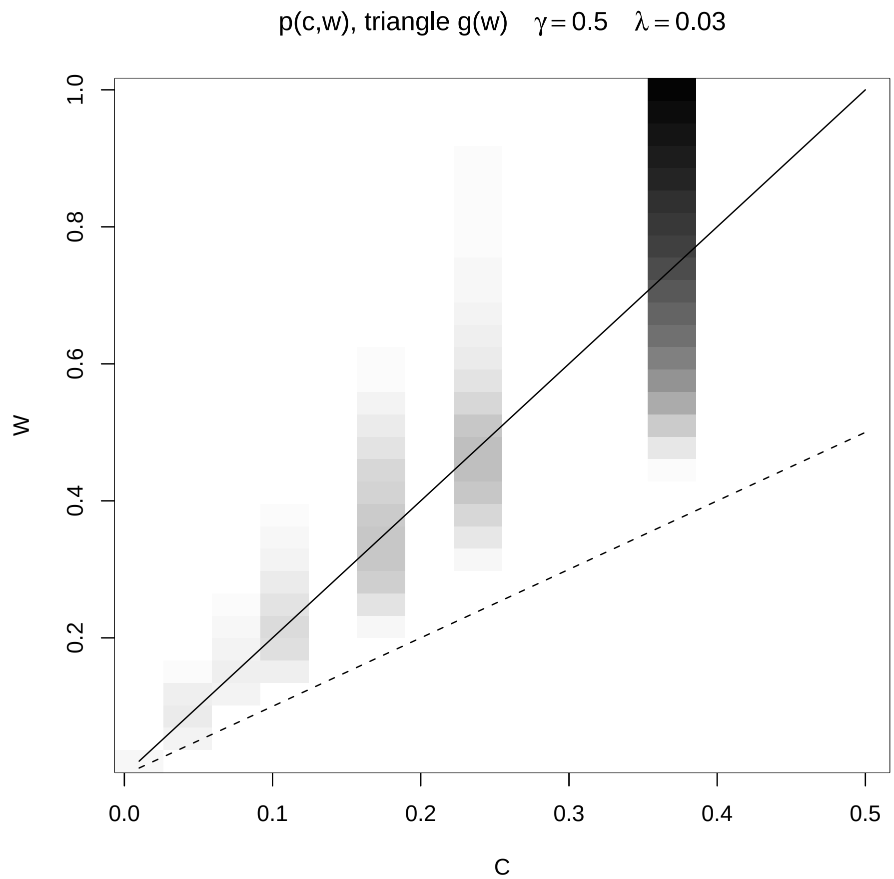
16.


p(w|c), log utility, λ = 0.5, triangle g
pgcg1l.5u[1,]
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.0
0.2
0.4
0.6
0.8
1.0
w2

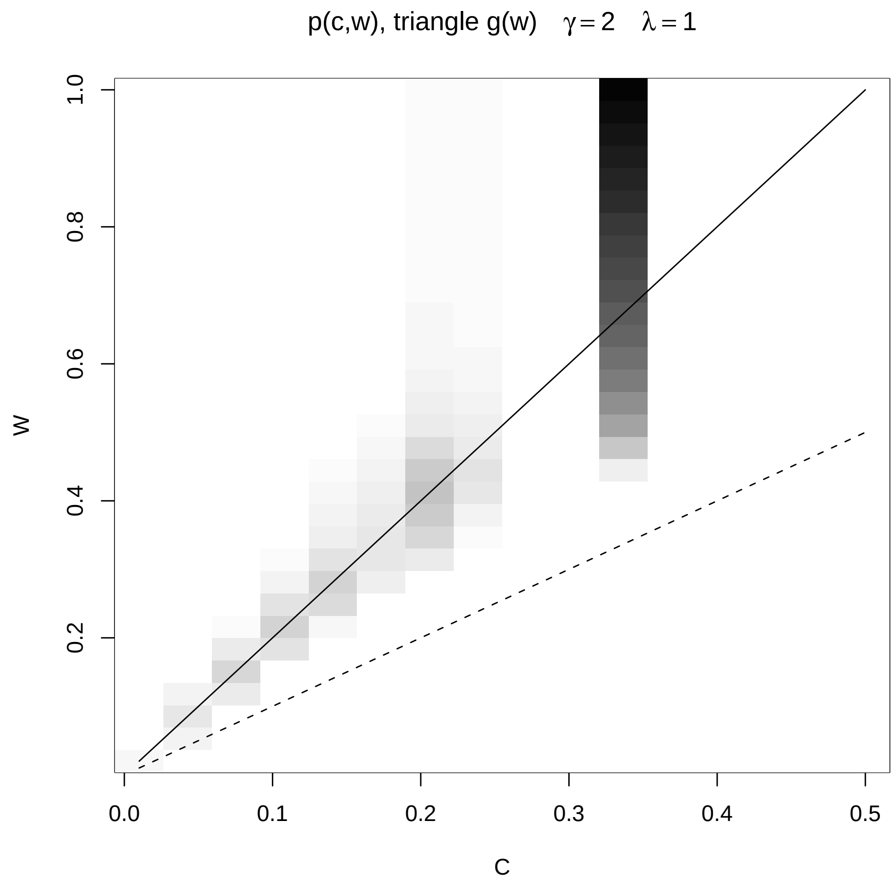
17.


pdf of c, γ = 1, λ = 0.5, triangle g(w)
p(c)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.0
0.1
0.2
0.3
0.4
0.5
C

18.



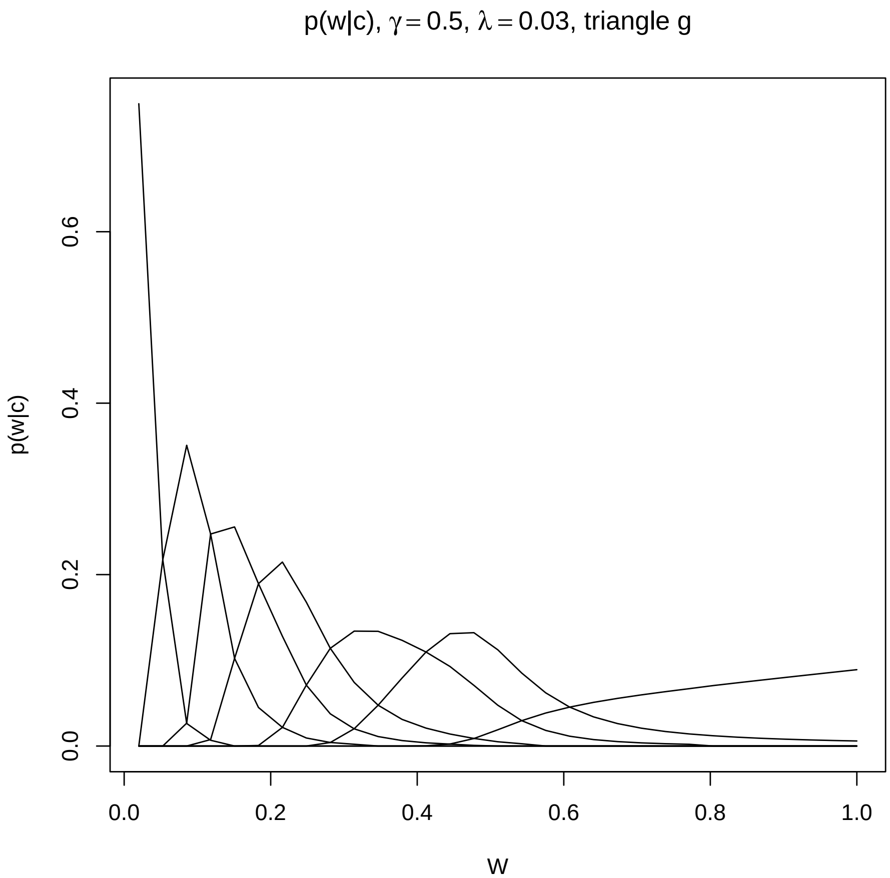
19.



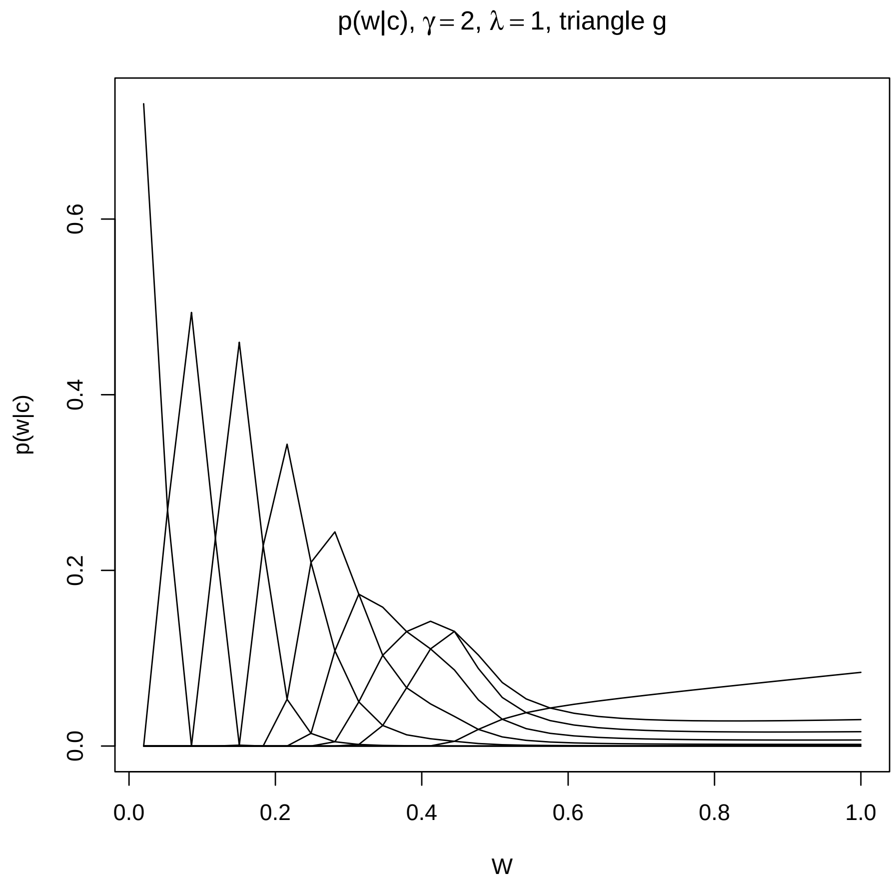
20. Discreteness

- Discreteness in the distribution of c and finite support, varying with c, in $q(w \mid c)$ show up in both these standard cases, despite a continuous marginal distribution for w
- Any numerical approach to solution of rational inattention models must allow for these possibilities.
- They are not an artifact of the triangular $g(w)$.
- A challenge for model-solvers.
- But promising as an explanation for discrete responses of individuals to continuously varying signals, even where there are no apparent fixed costs of acting.

21.



22.



23. ON TO A FULLY DYNAMIC NON-LQG MODEL

Here is the Bellman equation for a dynamic programming problem with Shannon capacity as a constraint, the current pdf g of w as the state variable, and $f(c,w)$ as the control:

$$V(g) = \max_{f(\cdot,\cdot)} \int U(c) f(c,w)\, dc\, dw + \beta \int V\left(\int h(\cdot; c, w) f(w \mid c)\, dw \right) f(c,w)\, dc\, dw$$

subject to

$$\int f(c,w)\, dc = g(w), \text{ all } w$$
$$f(c,w) \geq 0, \text{ all } c, w$$
$$I(C;W) \leq \kappa\,.$$

This is in the form of a standard dynamic programming problem, except that the state and control variables are both in principle infinite-dimensional. But computing solutions to such problems should be feasible.

24. RATIONAL INATTENTION MODELS OF GENERAL EQUILIBRIUM

- van Niewerburgh and Veldkamp and Mondria calculate market equilibrium with capacity constraints.
- But these models assume costless and perfect observation of market prices, which is both contrary to the notion of finite-capacity agents and a source of anomalous results.
- This is a modeling choice, not an easily corrected oversight.
- Modeling a market equilibrium with agents who do not know exactly what prices are requires being explicit about aspects of market microstructure that are not standard parts of the economic theory toolbox and about which we have few stylized facts or modeling conventions to guide us.

25. DIFFICULTIES WITH RIGE

- Prices will not clear markets, at least not via individual supply and demand behavior.
- Agents will not have perfect knowledge of prices as they take decisions that affect economic exchange and production.
- Inventories, retailers, wholesalers, demand deposits, cash, and credit cards, are all devices that allow agents to make transactions in which quantities and price are known and chosen only approximately.
- Few of our models have explicit roles for retailers and wholesalers, our models of inventory behavior are only modestly successful, and microfounded models of money that connect to data are non-existent.
- The idea of Shannon capacity may be of some help in modeling these phenomena, but they are inherently difficult, long-standing problems, so realistic general equilibrium models with capacity-constrained agents may not emerge for some time.

26. MACRO MODELING

- applying rational attention to the representative agent equilibrium models that now constitute mainstream macro will take some time.
- In the meantime, though, we can see even from simple linear-quadratic examples that there are implications for current modeling practice. In the linear-quadratic framework, rational inattention behavior is a constrained special case of the behavior of an agent who observes state variables with error. While there are some examples of such models in the macroeconomic literature (Lucas, 1973; Woodford, 2001), the rational inattention idea should encourage us to pay more attention to such models.

- The objection that there is no physical interpretation for the observation error such models postulate is answered by the rational inattention framework
- The RI framework gives us some guidance as to reasonable properties for the observation error, even when we cannot derive it analytically.

27. Public and private information models

- Recently, following the paper by Morris and Shin (2002), there have been a number of papers considering models with public and private information (Hellwig, 2004; Angeletos and Pavan, 2004).
- Raises interesting questions and arrives at conflicting conclusions about the value of "transparency", depending on assumptions about externalities that are difficult to calibrate against an actual economy.
- This literature assumes that there are private information sources with exogenously given attributes and a public information source whose stochastic character we can imagine controlling.
- From a rational inattention perspective, this is a strange setup. Capacity-limited agents will act as if observing the state of the economy with error even if some public authority announces it exactly.
- The amount of the error will depend both on the stochastic properties of the state itself and of any noise in public signals about it.
- If private signals carry a lot of information about privately important variables, information they contain about an aggregate state may be ignored or reacted to very slowly and erraticly.

References

Angeletos, G.-M., and A. Pavan (2004): "Transparency of Information and Coordination in Economies with Investment Complementarities," *American Economic Review*, 94.

Hellwig, C. (2004): "Heterogeneous Information and the Benefits of Transparency," Discussion paper, UCLA.

Lucas, Robert E., J. (1973): "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63(3), 326–334.

Morris, S., and H. S. Shin (2002): "The Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

Woodford, M. (2001): "Imperfect Common Knowledge and the Effects of Monetary Policy," Discussion paper, Princeton University, http://www.princeton.edu/~woodford/.