Shannon capacity-constrained behavior

Christopher A. Sims

June 20, 2014

Outline

Intro

Examples

Macro stylized facts

- For macro time series that are not auction-market prices and are not linked by accounting identities, typical impulse responses show a substantial weight at zero lag for own shocks, but a smooth hump shape for cross-variable shocks.
- To explain the smooth hump shapes, economic modelers introduce "adjustment costs", which have no clear connection to anything measurable in microeconomic data.
- The adjustment costs themselves imply that own-responses should also be smooth. Thus additional shocks, that change choice variables without incurring adjustment costs, are also required.

Micro stylized facts

- Most people, most of the time, do not respond immediately to price signals, and indeed may not even be fully aware of price changes.

Micro stylized facts

- Most people, most of the time, do not respond immediately to price signals, and indeed may not even be fully aware of price changes.
- For example: A person driving down Route 1 in New Jersey has dozens of opportunities to buy gas, at prices that are posted on large signs so she can see them as she drives by.
- A fully rational agent would keep track of the price distribution and the amount of gas in the gas tank, continually re-optimizing to decide where to buy gas.
- An actual person probably is listening to the radio and talking with her companions instead, stopping to buy gas at the first station to show up when the low-gas signal lights up, so long as the price doesn't look too outrageous.

Information processing costs: Shannon's measure

- It seems realistic to recognize that even freely available information is sometimes ignored, because processing it is in some sense costly.

Information processing costs: Shannon's measure

- It seems realistic to recognize that even freely available information is sometimes ignored, because processing it is in some sense costly.
- An economist wants to include this cost in the rational agent's decision problem.

Information processing costs: Shannon's measure

- It seems realistic to recognize that even freely available information is sometimes ignored, because processing it is in some sense costly.
- An economist wants to include this cost in the rational agent's decision problem.
- But how to measure such a cost? Shannon had an answer.
- If you believe information flow should be modeled as reducing uncertainty about some random quantities, and if you believe that information flows from observing two random quantities in succession should "add up" to the amount in the two random quantities observed jointly:

Information processing costs: Shannon's measure

- It seems realistic to recognize that even freely available information is sometimes ignored, because processing it is in some sense costly.
- An economist wants to include this cost in the rational agent's decision problem.
- But how to measure such a cost? Shannon had an answer.
- If you believe information flow should be modeled as reducing uncertainty about some random quantities, and if you believe that information flows from observing two random quantities in succession should "add up" to the amount in the two random quantities observed jointly: You end up with Shannon's measure.

Shannon's measure is all around us

- We're used to measuring he speed of an internet connection in megabits per second: This is a measure of the maximal rate of information flow over the connection, in Shannon's units.
- It is usable across arbitrary variations in hardware — one doesn't need to know whether the connection is fiber, coax cable, wireless, etc.
- This is one reason it is promising for economic modeling — it abstracts from the detail of people's "hardware".

Shannon capacity constraint's generic implications for dynamics

- With a capacity constraint preventing full adjustment of a decision variable x to a target y, we find that the relation between y and x differs from the unconstrained case exactly along the lines of the macro stylized facts.
- The own-responses must be less smooth than the y to x responses if the rate of mutual information flow between y and x is finite.
- These ideas were laid out in my 2003 JME paper.

This paper

- Authors: Junehyuk Jung, Jeong-Ho Kim, Filip Matejka, and Chris Sims
- Subject: Understanding economic behavior with an information-processing constraint in non-Gaussian, non-LQ cases.
- Results: In broad classes of cases, the information constraint converts continuously distributed optimizing behavior into behavior that is either entirely discretely distributed or is distributed over a lower-dimensional set than the unconstrained behavior.

The static information-constrained decision problem

$$
\begin{aligned}
\max_{f,\mu_x} \quad & \int U(x,y) f(x,y)\,\mu_x(dx)\,\mu_y(dy) \\
- \quad & \alpha^{-1}\Bigg(\int \log(f(x,y)) f(x,y)\,\mu_x(dx)\,\mu_y(dy) \\
& \quad + \int \log\Big(\int f(x,y')\,\mu_y(dy')\Big) f(x,y)\,\mu_x(dx)\,\mu_y(dy)\Bigg)
\end{aligned} \tag{1}
$$

$$
\text{subject to } \int f(x,y)\,\mu_x(dx) = g(y)\,,\ \text{a.s. } \mu_y \tag{2}
$$

$$
f(x,y) \geq 0\,,\ \text{all } x, y\,, \tag{3}
$$

Implications of the FOC's

At all values of x where $p(x) > 0$, The FOC's of the problem with respect to f imply that at all values of x, y with $f(x, y) > 0$ and $g(y) > 0$

$$U(x,y) = \theta(y) + \alpha^{-1} \log\left(\frac{f(x,y)}{\int f(x,y)\,dy}\right) \tag{4}$$

$$\therefore f(x,y) = p(x)e^{\alpha U}h(y) \tag{5}$$

$$\therefore \quad \int p(x)e^{\alpha U(x,y)}\,dx \cdot h(y) = g(y)\,, \tag{6}$$

The function $C(x) = \int e^{\alpha U(x,y)} h(y)\, dy$

- It has to be one whenever $p(x) > 0$, because for these values of x it is the integrand is the conditional pdf of $y \mid x$.
- So $B = \{x \mid C(x) = 1\}$ contains the support of x.
- If the objective function U is analytic in x on an open set S, it is often easy to show that $C(x)$ is also analytic in x.
- Then B is either the whole of S or it contains no open sets. When S is one-dimensional, B is either the whole of S or a countable collection of points with no limit points in S.

Outline

Intro

Examples

Monopolist with random costs

- x is price; y is unit cost.
- $U(x, y) = q(x) \cdot (x - c)$
- Say $q(x) = a - bx$ and x restricted to $(0, a/b)$.
- Then price has a distribution with support a finite set of points within $(0, a/b)$
- This model was solved numerically in earlier work by Matejka, where the discreteness was apparent.

What the pricing model might explain

- Micro data on prices show that for a given product they tend to remain constant for a while, but jump around between a finite set of values.
- If Shannon information-processing costs were the explanation, this would suggest not trying to connect the frequency of price adjustments to any measurable physical "menu cost".

Portfolio choice

- Fixed wealth of 1 to be allocated over a risk free asset and a collection of risky assets with yields $z = y + \varepsilon$
- x: portfolio weights (sum to one)
- ε: "hard" uncertainty; information-processing can't reduce it. Only y is reducible.
- $U(x, y) = E_\varepsilon[V(x'y)]$

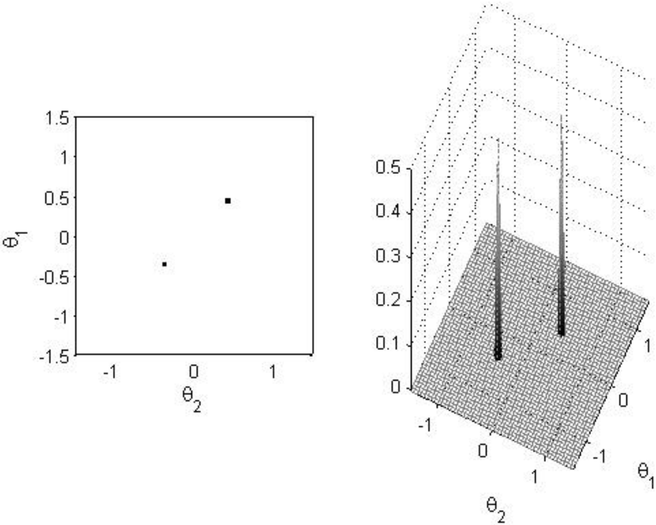
Results for quadratic utility, Gaussian randomness

- This is not an "LQ" problem, because objective is a quadratic function of a quadratic.
- We can show analytically that the solution concentrates on a set of less than two dimensions.
- We assume two risky assets, with yields independent of each other and identically normally distributed, plus a riskless asset.
- The plots show weights on the two risky assets.

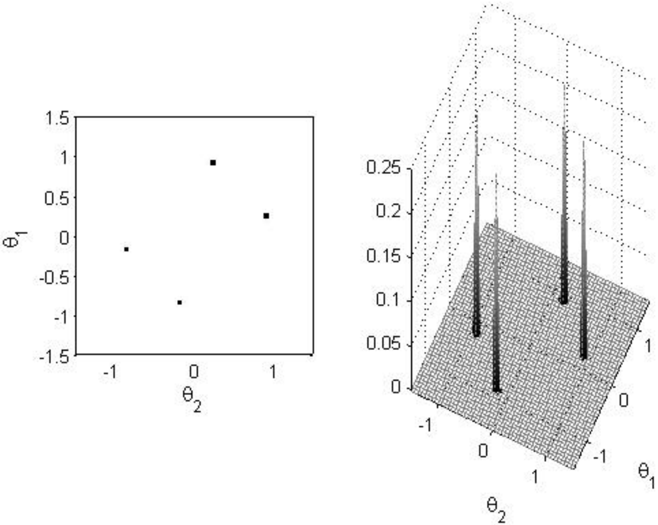
Interpretation of results

- At high information costs, the agent just chooses to go long or short risky assets, keeping the weights on the two fixed.
- At moderate information costs, the weights are distributed over four points, corresponding to relatively long or short risky asset 1 and long or short risky assets generally.
- At low information costs the risky asset weights distribute on a circle. Note that this implies that the riskless asset has a u-shaped pdf, so that there is still a tendency for portfolios to be at the extremes of being long or short on risk.

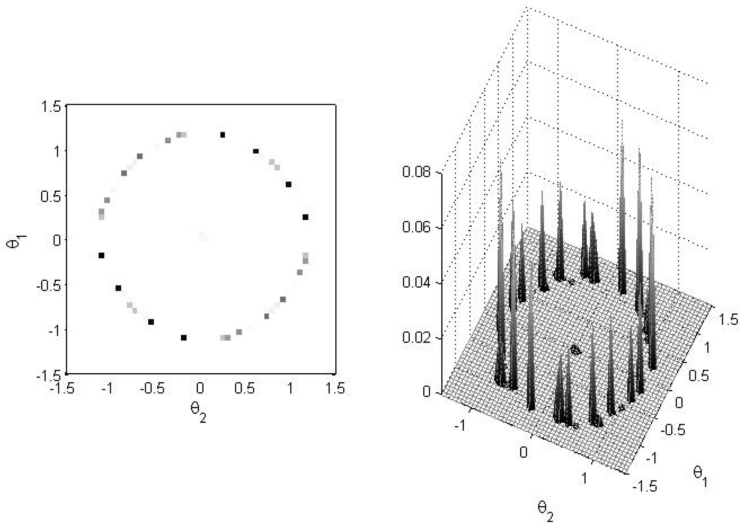
Portfolio choice, high information cost



Portfolio choice, medium information cost



Portfolio choice, low information cost



Linear-quadratic tracking

$$\max E[-(x-y)'A(x-y) - \theta I(x,y) \quad \text{subject to}$$
$$y \sim N(0,\Sigma)$$

$$I(x,y) = \tfrac{1}{2}\log|\Sigma| - \tfrac{1}{2}\log|\mathrm{Var}(y \mid x)|$$
$$= \tfrac{1}{2}\log|\Sigma| - \tfrac{1}{2}\log|\Sigma - \mathrm{Var}(x)|$$

Certainty-equivalence: optimally, $E[y \mid x] = x$. It will be optimal to have x jointly normally distributed with y. Of course without the information constraint, $y \equiv x$, so $x \sim N(0,\Sigma)$ and its distribution has the whole space as support.

Low-dimensional behavior

- When y and x are one-dimensional, if θ is high enough, it will be optimal to collect no information, so $\Sigma = \text{Var}(y \mid x)$ and the distribution of x is degenerate, concentrated on the point $x = E[y]$.
- When θ becomes small enough, the distribution of x in this LQ case immediately has full support.
- When x is multi-dimensional, so we are tracking several y's, it is again true that for large enough θ $x \equiv E[y]$.
- But now, when θ falls just below the threshold, the distribution of x goes from being 0-dimensional to being 1-dimensional. Then as θ falls further there is a switch to two-dimensional x, etc.

Water-filling

- If $A = I$ and Σ is diagonal, with $\sigma_{ii} > \sigma_{jj}$ when $i > j$, the solution has this form: For high enough θ (information cost), it is optimal to collect no information, i.e. concentrate the distribution of x entirely on the point $E[y]$.
- As θ drops, we start to get $\mathrm{Var}(x_n) > 0$, while $\mathrm{Var}(x_i) = 0$ for $i < n$.
- As θ drops further, we start to get $\mathrm{Var}(x_{n-1}) > 0$, with $\mathrm{Var}(y_1 \mid x) = \mathrm{Var}(y_2 \mid x)$.
- When the distribution of x has full support, $\mathrm{Var}(y \mid x) = \theta I$.
- This is a classic result in engineering literature on "rate distortion theory", of which our information-constrained decision problem is a generalization.

Conclusion

- "Stickiness" is pervasive in economic behavior, and attempts to model it as due to high physical costs of rapid change are mostly misguided.
- These "adjustment costs" are not directly measurable with micro data and hard to calibrate (maybe because they do not exist) even by anecdote or introspection.
- Whether stickiness is due to adjustment costs or instead to information processing costs may have policy-relevant implications for calculating the costs of business cycles and for predicting the effects on behavior of policy changes or exogenous changes in the stochastic environment. For example the standard rational expectations approach to modeling the effects of a change in the policy rule will not be correct if based on physical adjustment costs.

The road forward

- Models in which agents optimize subject to Shannon information processing costs are at this point (and maybe forever) hard to solve. This does not mean we should ignore the insights they provide.
- For example, models with ad hoc inertia, qualitatively motivated by information costs, may be more reliable than models that try to "micro-found" inertia using physical adjustment costs.
- Research on cheap approximate methods of solving models with information costs would be valuable.