FRAGILE BELIEFS AND PRICING

Lars Peter Hansen

Tjalling C. Koopmans Lectures, September 2008

1Based in part on Hansen and Sargent (JET, 2006), Hansen (AER,2007), Ely Lecture and Hansen and Sargent, “Fragile beliefs and the price of model uncertainty”
RATIONAL EXPECTATIONS FROM TWO PERSPECTIVES

- Econometricians - impose rational expectations - use cross equation restrictions that assume agent knowledge of parameters

- Economic decision-makers - make investment decisions - forecast the future - arguably not knowing parameters

Should economic agents and econometricians be placed on a more equal footing, or not?
Imposing rational expectations is justified heuristically by appealing to large histories of data. What happens when history is not so informative? This leads me to the questions:

- When is estimation difficult?
- What are the consequences for the econometrician?
- What are the consequences for economic agents and for equilibrium outcomes?
- What are the real time consequences of learning for competitive security markets?
- How is learning altered when decision-makers admit that their models are misspecified or simplified?
Bray-Kreps taxonomy distinguishes between

- Learning about a rational expectations equilibrium
  
  Uses learning rules with at least temporary misspecification, but without agents addressing this misspecification. There is no scope for uncertainty to alter directly equilibrium outcomes.

- Learning within a rational expectations equilibrium
  
I will explore learning within an equilibrium but include a concern for misspecification.
Misspecification and robustness: Probability models are “approximations”. A reference model is associated with a family of model perturbations. Robust statistics and robust control theory.

Ambiguity:

- Multiple priors decision model. Specifying a reference model along with a family of perturbations is one way to construct multiple priors.
- Smooth ambiguity decision model. Risk conditioned on a model is distinct from ambiguity across models. When viewed as a “compound lottery”, this avoids reduction. Do not form (weighted) averages of probability distributions across models prior to ranking random consumption claims.
For models $\nu = 0, 1$.

\[
\begin{align*}
\zeta_{t+1} - \zeta_t &= A(\nu)\zeta_t + C(\nu)w_{t+1} \\
\sigma_{t+1} - \sigma_t &= D(\nu)\zeta_t + G(\nu)w_{t+1}
\end{align*}
\]

and $w_{t+1} \sim \mathcal{N}(0, I)$.

Continuous-time counterpart

\[
\begin{align*}
d\zeta_t &= A(\nu)\zeta_t dt + C(\nu)dW_t \\
d\sigma_t &= D(\nu)\zeta_t dt + G(\nu)dW_t
\end{align*}
\]

where $W$ is a multivariate Brownian motion.
LEARNING CONTINUED

Application of the Kalman filter yields the following innovations representation:

\[ d\bar{\zeta}_t(\iota) = A(\iota)\bar{\zeta}_t(\iota) + K_t(\iota)[ds_t - D(\iota)\bar{\zeta}_t(\iota)] \]

where

\[ K_t(\iota) = [C(\iota)G(\iota)' + \Sigma_t(\iota)D(\iota)'][G(\iota)G(\iota)']^{-1} \]

\[ \frac{d\Sigma_t(\iota)}{dt} = A(\iota)\Sigma_t(\iota) + \Sigma_tA(\iota)' + C(\iota)C(\iota)' - K_t(\iota)[G(\iota)C(\iota)' + D(\iota)\Sigma_t(\iota)] \]

Construct the innovation process:

\[ d\bar{W}_t(\iota) = [\bar{G}(\iota)]^{-1} [ds_t - D(\iota)\bar{\zeta}_t dt] \]

where \( G(\iota)G'(\iota) = \bar{G}(\iota)\bar{G}(\iota)' \) and \( \bar{G}(\iota) \) is nonsingular.
Learning about a model

Two models $\iota = 0, 1$. Assume that $G(\iota)G(\iota)'$ independent of $\iota$. (Otherwise $\iota$ revealed immediately.)

Let $\bar{\iota}_t = E(\iota | S_t)$ where $S_t$ is generated by the signal history and

$$d\bar{W}_t = \bar{G}^{-1}(d\zeta_t - \mu_t dt) = \bar{\iota}_t d\bar{W}_t(1) + (1 - \bar{\iota}_t)d\bar{W}_t(2).$$

where

$$\mu_t = [\bar{\iota}_t D(1)\bar{\zeta}_t(1) + (1 - \bar{\iota}_t) D(0)\bar{\zeta}_t(0)].$$

Then

$$d\bar{\iota}_t = \bar{\iota}_t(1 - \bar{\iota}_t)[\bar{\zeta}_t(1)'D(1)' - \bar{\zeta}_t(0)' D(0)'] (\bar{G}')^{-1} d\bar{W}_t.$$
RISK PRICES

- Risk prices are the compensation in terms of expected returns for the exposure to alternative risks.
- In the time separable power utility model these prices are given by the coefficient of relative risk aversion $\gamma$ times the exposure of consumption to macroeconomic shocks. (Breeden)
# Risk Prices and Information

<table>
<thead>
<tr>
<th>Information</th>
<th>Local Risk</th>
<th>Risk Price</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>complete</td>
<td>$dW_t$</td>
<td>$\gamma G(\iota)'H$</td>
<td>$\gamma \sqrt{H'G(\iota)G(\iota)'H}$</td>
</tr>
<tr>
<td>unknown state</td>
<td>$d\tilde{W}_t(\iota)$</td>
<td>$\gamma \tilde{G}(\iota)'H$</td>
<td>$\gamma \sqrt{H'G(\iota)G(\iota)'H}$</td>
</tr>
<tr>
<td>unknown model</td>
<td>$d\tilde{W}_t$</td>
<td>$\gamma \tilde{G}'H$</td>
<td>$\gamma \sqrt{H'GG'H}$</td>
</tr>
</tbody>
</table>

**Table:** When the model is unknown, $G(\iota)G(\iota)'$ is assumed to be independent of $\iota$. The parameter $\gamma$ is the coefficient of relative risk aversion in a power utility model. The entries in the “slope” column are the implied slope of the mean-standard deviation frontier. The consumption growth rate is $dc_t = H'ds_t$. 
The risk prices change across information structures because different risks are being priced. The slope (the risk price magnitude) is the same across information structures.

A concern about model misspecification or ambiguity will alter this table.
A pessimist thinks that good news is temporary but that bad news will endure.

We build a model of endogenous pessimism about components of consumption that are hard to detect.
Motivation

Le doute n’est pas une condition agréable, mais la certitude est absurde. Voltaire. 1767.

Doubt is not a pleasant condition, but certainty is absurd.

- Long-run components of uncertainty make learning challenging.
- Because asset prices are forward looking, they are sensitive to beliefs about long-run components.
- Agents don’t trust their models. This makes their beliefs appear especially fragile with respect to the arrival of news.
Investors have two models – one imputes a predictable component to consumption the other does not. The models are difficult to distinguish. Investors treat the models as approximations.

Distinguish:

- misspecification conditional on the current hidden state
- misspecification of probabilities assigned to the hidden states

Investors’ corrections for fears of these two types of misspecification are the sources of fragile beliefs.
“Risk Prices” reconsidered

- Risk prices are the compensation in terms of expected returns for the exposure to alternative risks.
- Concern about model misspecification alters equilibrium “risk prices” through the introduction of (constrained) worst-case models. It adds a model uncertainty component.
- Worst-case models are imputed by solving a robust social planner’s problem.
- Equilibrium risk and uncertainty prices support decentralized Arrow security markets.

I characterize and compute the uncertainty components. Pure risk components are very small.
Robust Social Planner under Full Information

Temporarily abstract from model choice $i$.

I interpret Tallarini (2000) as replacing discounted log utility with

$$V(\zeta, c) = \lambda' \zeta + \kappa + c$$

where

$$V(\zeta, c) = (1 - \beta)c + T^1 [\beta V(\zeta^*, c^*)]$$

$T^1$ is a risk-sensitivity operator:

$$T^1 [\beta V(\zeta^*, c^*)] = -\theta_1 \log E \left[ \exp \left( \frac{-\beta V(\zeta^*, c^*)}{\theta_1} \right) \right]$$

$$= \min_{m(w^*) \geq 0, E(m(w^*)) = 1} E \left( m(w^*) \left[ \beta V(\zeta^*, c^*) + \theta_1 \log m(w^*) \right] \right)$$

where $\zeta^* = (I - A)\zeta + Cw^*$ and $c^* - c = H'D\zeta + Gw^*$. 
Value function is \( V(\zeta, c) = \lambda'\zeta + \kappa + c \)

**Proposition**

The value function shares the same \( \lambda \) with the expected utility model, and

\[
\kappa = \frac{-\beta^2}{2(1 - \beta)\theta_1} |\lambda' C + G|^2.
\]

The associated worst distribution for \( w^* \) is normal with mean

\[
\mu^* = \frac{-\beta}{\theta_1} (C' \lambda + G')
\]

and covariance matrix \( I \).

Worst-case model implied by the minimizing \( m(w^*) \) changes the mean but not the variance of the normal shock. This implies constant uncertainty premia.
RECURSIVE UTILITY OR ROBUSTNESS

\[ V(\zeta, c) = (1 - \beta)c + T^1 [\beta V(\zeta^*, c^*)] \]

\[ T^1 [\beta V(\zeta^*, c^*)] = -\theta_1 \log E \left[ \exp \left( \frac{-\beta V(\zeta^*, c^*)}{\theta_1} \right) \right]_{\zeta, c} \]

\[ = \min_{m(w^*) \geq 0, Em(w^*) = 1} E \left( m(w^*) [\beta V(\zeta^*, c^*) + \theta_1 \log m(w^*)] \right)_{\zeta, c}. \]

- Special case of Koopmans/Kreps-Porteus recursive utility motivated by the literature on risk sensitive control theory. Hansen and Sargent (IEEE, 1995).
- Jacobson, Whittle and others relate risk-sensitive control theory to robust control theory. Petersen, James and Dupuis (IEEE, 2000).
- Axiomatic justifications in Maccheroni, Marinacci and Rustinchini (Econometrica, 2006) and Strzalecki (2008).
# Where are we?

<table>
<thead>
<tr>
<th>Information</th>
<th>Local Risk</th>
<th>Risk Price</th>
<th>Uncertainty Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>$dW_t$</td>
<td>$G(\nu)'H$</td>
<td>$\frac{1}{\theta_1}[C(\nu)'\lambda(\nu) + G(\nu)'H]$</td>
</tr>
<tr>
<td>Unknown state</td>
<td>$d\bar{W}_t(\nu)$</td>
<td>$\bar{G}(\nu)'H$</td>
<td>?</td>
</tr>
<tr>
<td>Unknown model</td>
<td>$d\bar{W}_t$</td>
<td>$\bar{G}'H$</td>
<td>?</td>
</tr>
</tbody>
</table>

**Table:** When the model is unknown, $G(\nu)G(\nu)'$ is assumed to be independent of $\nu$. The consumption growth rate is $dc_t = H'ds_t$.

Model misspecification price is time invariant.
AMBIGUOUS OR ROBUST STATE ESTIMATION

Yesterday’s posterior is today’s prior.

Explore robustness to prior.
This risk-sensitive adjustment is

\[ T^2 \left[ V(\zeta, c) | \bar{\zeta}, c, \Sigma \right] = -\theta_2 \log \int \exp \left( \frac{-V(\zeta, c)}{\theta_2} \right) \phi(\zeta | \bar{\zeta}, \Sigma) d\zeta \]

\[ = \min_{h(\zeta) \geq 0, \int h(\zeta) \phi(\zeta | \bar{\zeta}, \Sigma) = 1} \int [V(\zeta, c) + \theta_2 \log h(\zeta)] h(\zeta) \phi(\zeta | \bar{\zeta}, \Sigma) d\zeta \]

where \( \theta_2 \) is another risk-sensitivity parameter and \( \phi(\zeta | \bar{\zeta}, \Sigma) \) is a Gaussian density with mean \( \bar{\zeta} \) and covariance matrix \( \Sigma \).
**Solution**

**Proposition**

The adjusted value function is:

\[
T^2 \left[ V(\zeta, c)|\bar{\zeta}, c, \Sigma \right] = \lambda' \bar{\zeta} + \kappa - \frac{1}{2\theta_2} \lambda' \Sigma \lambda + c,
\]

and the mean of the worst-case normal distribution for \( \zeta_t - \bar{\zeta}_t \) is:

\[
u^* = -\frac{1}{\theta_2} \Sigma \lambda
\]
**Smooth ambiguity or robustness**

\[ T^2 \left[ V(\zeta, c) | \bar{\zeta}, c, \Sigma \right] = -\theta_2 \log \int \exp \left( \frac{-V(\zeta, c)}{\theta_2} \right) \phi(\zeta | \bar{\zeta}, \Sigma) d\zeta \]

\[ = \min_{h(\zeta) \geq 0, \int h(\zeta) \phi(\zeta | \bar{\zeta}, \Sigma) = 1} \int \left[ V(\zeta, c) + \theta_2 \log h(\zeta) \right] h(\zeta) \phi(\zeta | \bar{\zeta}, \Sigma) d\zeta \]

**Smooth ambiguity:** Klibanoff, Marinacci and Mukerji (Econometrica, 2005) and previously Segal (Econometrica, 1990).

**Robustness:** Hansen and Sargent (JET, 2006).
**WHERE ARE WE?**

<table>
<thead>
<tr>
<th>information</th>
<th>local risk</th>
<th>risk price</th>
<th>uncertainty price</th>
</tr>
</thead>
<tbody>
<tr>
<td>complete</td>
<td>$dW_t$</td>
<td>$G(\iota)'H$</td>
<td>$\frac{1}{\theta_1}[C(\iota)'\lambda(\iota) + G(\iota)'H]$</td>
</tr>
<tr>
<td>unknown state</td>
<td>$d\tilde{W}_t(\iota)$</td>
<td>$\tilde{G}(\iota)'H$</td>
<td>$\frac{1}{\theta_2} \tilde{G}(\iota)^{-1} D(\iota) \Sigma_t(\iota) \lambda(\iota)$</td>
</tr>
<tr>
<td>unknown model</td>
<td>$d\tilde{W}_t$</td>
<td>$\tilde{G}'H$</td>
<td>?</td>
</tr>
</tbody>
</table>

**Table:** When the model is unknown, $G(\iota)G(\iota)'$ is assumed to be independent of $\iota$. The consumption growth rate is $dc_t = H'ds_t$.  

- Learning makes worst-case mean time-dependent but **not** state-dependent.
- Add the two distortions.

$$
\tilde{G}(\iota)^{-1} \left( \frac{1}{\theta_1} G(\iota)[C(\iota)'\lambda(\iota) + G(\iota)'H] + \frac{1}{\theta_2} D(\iota) \Sigma_t(\iota) \lambda(\iota) \right)
$$
Robust model-averaging

Apply a second $T^2$ operator that makes a robust adjustment of the model probabilities $\bar{\nu}_t$ and $1 - \bar{\nu}_t$ based on the continuation values of the two models.

\[
(1 - \tilde{\nu}_t) \propto (1 - \bar{\nu}_t) \exp \left( - \frac{U[0, \bar{\zeta}(0), \Sigma_t(0)]}{\theta_2} \right)
\]

\[
\tilde{\nu}_t \propto \bar{\nu}_t \exp \left( - \frac{U[1, \bar{\zeta}(1), \Sigma_t(1)]}{\theta_2} \right)
\]

The investor slants probabilities towards the model with the worse utility consequences.
EXAMPLES IN THE LITERATURE

- Gollier - Does Ambiguity Reinforce Risk Aversion?
- Ju and Miao - Ambiguity, Learning and Asset Returns
- Collard, Mukerji, Sheppard and Tallard - Ambiguity and the Historical Equity Premium
- Hansen and Sargent - Fragile Beliefs and the Price of Model Uncertainty
CONSUMPTION GROWTH IID, $\nu = 0$

\[
\begin{align*}
\zeta_{t+1}(0) &= \zeta_t(0) \\
\kappa_{t+1} &= \zeta_t(0) + [0 \quad \sigma_2(0)] \, w_{t+1}.
\end{align*}
\]

Investors estimate constant mean growth rate $\zeta_t(0)$. 
Consumption growth predictable, $t = 1$

\[
\begin{bmatrix}
\zeta_{1,t+1}(1) \\
\zeta_{2,t+1}(1)
\end{bmatrix}
= \begin{bmatrix}
\rho & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
\zeta_{1,t}(1) \\
\zeta_{2,t}(1)
\end{bmatrix}
+ \begin{bmatrix}
\sigma_1 & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
w_{1,t+1} \\
w_{2,t+1}
\end{bmatrix}
\]

\[
s_{t+1} - s_t = \begin{bmatrix} 1 & 1 \end{bmatrix}
\begin{bmatrix}
\zeta_{1,t}(1) \\
\zeta_{2,t}(1)
\end{bmatrix}
+ \begin{bmatrix} 0 & \sigma_2(1) \end{bmatrix}
\begin{bmatrix}
w_{1,t+1} \\
w_{2,t+1}
\end{bmatrix}
\]

Investors estimate time varying conditional mean of the consumption growth rate $\zeta_{1,t}(1) + \zeta_{2,t}(1)$. 
INTERPRETATION OF STATE VARIABLES IN TWO SUBMODELS

<table>
<thead>
<tr>
<th>$\iota$</th>
<th>$\zeta(\iota)$</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\zeta(0)$</td>
<td>$E[(c_{t+1} - c_t)</td>
</tr>
<tr>
<td>1</td>
<td>$\zeta_{1,t}(1)$</td>
<td>persistent component</td>
</tr>
<tr>
<td>1</td>
<td>$\zeta_2(1)$</td>
<td>$E[(c_{t+1} - c_t)</td>
</tr>
</tbody>
</table>

Submodel $\iota = 1$, is an example of the “long-run risk model” used in finance.

Goal: Reduce investor confidence in this submodel while maintaining its importance in pricing.
CONDITIONAL MEANS FOR TWO MODELS

Figure 2: Top panel: Consumption growth and $\zeta_1(t) + \zeta_2(t)$; bottom panel: consumption growth and $\zeta_1(t)$.

Figure 3: Bayesian probability $\pi_t$ attached to model 1 for U.S. quarterly consumption (non-durables plus services) per capita for $\pi_0 = 0.5$ (solid blue line) and worst case probability $\tilde{\pi}_t$ associated with $\theta_1 = 20$, $\theta_2 = 0.2$ (dashed red line).
Figure 2: Top panel: Consumption growth and \( \hat{\zeta}_1(t) + \hat{\zeta}_2(t) \); bottom panel: consumption growth and \( \hat{\zeta}_2(t) \).

Figure 3: Bayesian probability \( \hat{p}_t \) attached to model 1 for U.S. quarterly consumption (non-durables plus services) per capita for \( \hat{p}_0 = 0.5 \) (solid blue line) and worst case probability \( \tilde{p}_t \) associated with \( \theta_1 = 20 \), \( \theta_2 = 0.2 \) (dashed red line).
There are more channels for misspecification of submodel \( \iota = 1 \), the model with consumption predictability.

Submodel \( \iota = 1 \) confronts an investor/statistician with a more challenging estimation/inference problem.
**Prior Robustness, \( \theta_1 = \infty \)**

- **Solid line**: \( \theta_2 = 1 \)
- **Dashed line**: \( \theta_2 = 1/2 \)
- **Dotted-dashed line**: \( \theta_2 = 1/4 \)
Prior Robustness, $\theta_1 = .02$

top plot: solid line: $\theta_2 = 2$ dashed line: $\theta_2 = 4$;
bottom plot: solid line: $\theta_2 = 1/4$ dashed line: $\theta_2 = 1/2$
SOURCE OF “COUNTERCYCLICAL” UNCERTAINTY PREMIA

- Consumer’s concern about misspecification causes him to calculate worst case probabilities that depend on value functions.
- The value functions for the two submodels respond to shocks in ways that bring them closer together after positive consumption growth shocks and farther apart after negative shocks.

Therefore, our cautious consumer slants probability more towards the model with predictable consumption growth rates when recent observations of consumption growth have been lower than average than when these observed growth rates are higher than average.
FINAL OBSERVATIONS

Model serves as a useful illustration but

- “Calibration” of robustness parameters using statistical detection. Smooth ambiguity formulations consider other calculations.
- Other macro variables may be useful in forecasting consumption. This complicates the estimation problem so that there are more channels for misspecification.
- IES = 1 implies constant consumption-wealth ratio. Smooth “consumption-wealth” ratios are a “challenge” more generally for recursive utility models.
- There are other sources of time variation in risk prices, for example, stochastic volatility or volatility regime shifts. Learning provides an endogenous mechanism for such variation.
- What are the price implications over longer horizons? Remains to be computed and/or characterized.