Forecasting
with Poor Data and Models:
An Info-Gap Approach

Yakov Ben-Haim
Technion
Israel Institute of Technology
## Contents

1 Highlights (royal-holloway2010-01.tex)  
   1.1 Severe Uncertainties: Info-Gaps  
   1.2 Shackle-Popper Indeterminism (indif5d-shackle-pop.tex)  

2 Info-Gap Forecasting (forecast02.tex)  
   2.1 1-D Example  
   2.2 European Central Bank: Overnight Rate  
   2.3 Crossing Robustness Curves: General Case  
   2.4 Robustness and Probability of Forecast Success  
   2.5 Regression Prediction  

3 Summary (royal-holloway2010-01.tex)
1 Highlights
1.1  \textit{Severe Uncertainties: Info-Gaps}

- **Models:**
  - Conflicting.
  - Simplistic.
  - Incomplete.

- **Data:**
  - Random.
  - Biased, unknown correlations.
  - Subject to revision.

- **Time:**
  - Past may not reflect future.
  - Laws may change.
§ The art of designing, deciding, planning:

Use the **wrong model and data**

to make the **right decision**

(when the right model is unknown).
§ Info-gap decision strategies:

- **Robust-satisficing:**
  
  protect against uncertainty.

- **Opportune-windfalling:**
  
  exploit uncertainty.
1.2 *Shackle-Popper Indeterminism*

§ **Intelligence:**

What people know,

influences how they behave.

§ **Discovery:**

What will be discovered tomorrow

cannot be known today.

§ **Indeterminism:**

Tomorrow’s behavior cannot be

modelled completely today.
§ **Information-gaps**, indeterminisms, sometimes cannot be modelled probabilistically.

§ **Ignorance** is not probabilistic.
2 INFO-GAP FORECASTING

Yakov Ben-Haim, 2009,
Info-gap forecasting,
*European Journal of Operational Research.*

Yakov Ben-Haim, 2010,
*Info-Gap Economics: An Operational Introduction,*
Palgrave-Macmillan.
2.1 1-D Example

§ “True” scalar system:

\[ y_t = \lambda_t y_{t-1} \]

§ Historical data: \( \lambda_t = \bar{\lambda} \) for \( t \leq T \).

§ Contextual understanding:

\( \lambda \) could drift upwards.

§ Fractional-error info-gap model:

\[ \mathcal{U}(h, \bar{\lambda}) = \left\{ \lambda_t, t > T : 0 \leq \frac{\lambda_t - \bar{\lambda}}{\lambda} \leq h \right\}, \quad h \geq 0 \]

- Unbounded family of sets.
- No worst case.
§ **Slope-adjusted (erroneous) forecaster:**

\[ y_t^s = \ell y_{t-1}^s \]

§ **Contrast with historically estimated model:**

\[ y_t = \tilde{\lambda}_t y_{t-1} \]

How to choose \( \ell \geq \tilde{\lambda} \)?

§ **Robust satisficing:**

Satisfice the forecast error:

\[ |y_{T+k}^s - y_{T+k}| \leq \varepsilon_c \]

Maximize robustness to future surprise.
§ Robustness of forecast $\ell$:

Max $h$ up to which all $\lambda_{T+i}$ in $U(h, \bar{\lambda})$

satisfice forecast error at $\varepsilon_c$:

$$\bar{h}_s(\ell, \varepsilon_c) = \max \left\{ h : \left( \max_{\lambda_{T+i} \in U(h, \bar{\lambda})} \left| y^s_{T+k} - y_{T+k} \right| \right) \leq \varepsilon_c \right\}$$

§ Preference: $\ell \succ \ell'$ if $\bar{h}_s(\ell, \varepsilon_c) > \bar{h}_s(\ell', \varepsilon_c)$
\( \ell_k e^\lambda_k \)  

\[ \ell = \lambda \quad \ell > \lambda \]  

\( (\ell^k - \lambda^k)_{g_T} \)  

\( \epsilon_c \)  

\( h_n(\ell, \epsilon_c) \)

\section*{Trade off: robustness vs. forecast error.}

\section*{Zeroing:} Estim outcome has 0 robustness.

\section*{Crossing robustness curves:}  
\( \ell \succ \bar{\lambda} \).

\begin{itemize}
  \item Preference reversal.
  \item Robustness-advantage of sub-optimal (erroneous) model.
\end{itemize}

\section*{Robustness is proxy for success-probability.
§ Numerical example.

Figure 1: Robustness vs. normalized forecast error for $\ell = 1.05, 1.1, 1.15, 1.2$ from bottom to top curve. $\lambda = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- Preference reversal at all time horizons, $k$.
- Robustness premium decreases with $k$.
- Reversal-$\varepsilon_c$ increases with $k$.
Robustness & Probability of Forecast Success

§ Future growth coefficients: $\lambda_{T,k} = (\lambda_{T+1}, \ldots, \lambda_{T+k})$.

$\lambda_{T,k}$ is random vector on domain $D$.

$F(\lambda_{T,k}) =$ cumulative probability distrib.

§ Forecast success set:

$\mathcal{Y}(\ell) = \{ \lambda_{T,k} \in D : |y_{T+k}^s(\ell) - y_{T+k}| \leq \varepsilon_c \}$

§ Probability of success:

$P_s(\ell) = F[\mathcal{Y}(\ell)]$
§ Goal:

Choose $\ell$ to maximize success prob.

§ Problem:

$F(\lambda_{T,k}) =$ is unknown.

§ Solution:

• $\tilde{h}_s(\ell, \varepsilon_c)$ is known.
• $\tilde{h}_s(\ell, \varepsilon_c)$ proxies for success prob.
Theorem.

Probability of successful forecast, $P_s(\ell)$, increases with increasing info-gap robustness, $\tilde{h}_s(\ell, \varepsilon_c)$.

Given: (a) The domain of $F(\cdot)$ is contained in the info-gap model of eq.(2.1). (b) $y_T > 0$, $\tilde{\lambda} > 0$. (c) $\ell$ and $\ell'$ are two slope parameters for which:

$$\tilde{h}_s(\ell, \varepsilon_c) > \tilde{h}_s(\ell', \varepsilon_c) > 0$$

Then:

$$P_s(\ell) \geq P_s(\ell')$$

Robustness is proxy for success-probability.
Summary so far:

§ **Forecasters** do better if they robust-satisfice.

§ **Satisficing** is **not a last resort**.

   It is **strategically advantageous**.
2.2 **European Central Bank: Overnight Rate**

<table>
<thead>
<tr>
<th>Date</th>
<th>Interest rate</th>
<th>Implied λ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Jan 1999</td>
<td>4.50</td>
<td></td>
</tr>
<tr>
<td>9 Apr 1999</td>
<td>3.50</td>
<td>0.778</td>
</tr>
<tr>
<td>5 Nov 1999</td>
<td>4.00</td>
<td>1.143</td>
</tr>
<tr>
<td>4 Feb 2000</td>
<td>4.25</td>
<td>1.063</td>
</tr>
<tr>
<td>17 Mar 2000</td>
<td>4.50</td>
<td>1.059</td>
</tr>
<tr>
<td>28 Apr 2000</td>
<td>4.75</td>
<td>1.056</td>
</tr>
<tr>
<td>9 Jun 2000</td>
<td>5.25</td>
<td>1.105</td>
</tr>
<tr>
<td>28 Jun 2000</td>
<td>5.25</td>
<td>1.000</td>
</tr>
<tr>
<td>1 Sep 2000</td>
<td>5.50</td>
<td>1.048</td>
</tr>
<tr>
<td>6 Oct 2000</td>
<td>5.75</td>
<td>1.045</td>
</tr>
<tr>
<td>11 May 2001</td>
<td>5.50</td>
<td>0.957</td>
</tr>
<tr>
<td>31 Aug 2001</td>
<td>5.25</td>
<td>0.955</td>
</tr>
</tbody>
</table>

- § Typical change: 25 basis points.
- § Largest change: 100 basis points.
<table>
<thead>
<tr>
<th>Date</th>
<th>Interest rate</th>
<th>Implied ( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Jan 1999</td>
<td>4.50</td>
<td></td>
</tr>
<tr>
<td>9 Apr 1999</td>
<td>3.50</td>
<td>0.778</td>
</tr>
<tr>
<td>5 Nov 1999</td>
<td>4.00</td>
<td>1.143</td>
</tr>
<tr>
<td>4 Feb 2000</td>
<td>4.25</td>
<td>1.063</td>
</tr>
<tr>
<td>17 Mar 2000</td>
<td>4.50</td>
<td>1.059</td>
</tr>
<tr>
<td>28 Apr 2000</td>
<td>4.75</td>
<td>1.056</td>
</tr>
<tr>
<td>9 Jun 2000</td>
<td>5.25</td>
<td>1.105</td>
</tr>
<tr>
<td>28 Jun 2000</td>
<td>5.25</td>
<td>1.000</td>
</tr>
<tr>
<td>1 Sep 2000</td>
<td>5.50</td>
<td>1.048</td>
</tr>
<tr>
<td>6 Oct 2000</td>
<td>5.75</td>
<td>1.045</td>
</tr>
<tr>
<td>11 May 2001</td>
<td>5.50</td>
<td>0.957</td>
</tr>
<tr>
<td>31 Aug 2001</td>
<td>5.25</td>
<td>0.955</td>
</tr>
</tbody>
</table>

\textbf{§ 9Jun’00–31Aug’01:} \( \mu = 5.4\%, \ 
\sigma = 0.19\% \).

\textbf{§ On 9/12/2001, (1 day after 9/11)}

\textit{predict next interest rate.}

Rate down, but by how much?
\section*{Historical model:}

\[ y_t = \lambda_t y_{t-1}, \quad \lambda_t = \bar{\lambda} = 1 \]

\section*{Info-gap model:}

\[ \mathcal{U}(h, \bar{\lambda}) = \{ \lambda_T : (1 - h)\bar{\lambda} \leq \lambda_T \leq \bar{\lambda} \}, \quad h \geq 0 \]

\section*{Robustness:}

\[ \hat{h}(\ell, \varepsilon_c) = \begin{cases} 
0 & \varepsilon_c < \varepsilon_x \\
\frac{\varepsilon_c}{\lambda y_T} + \frac{\bar{\lambda} - \ell}{\lambda} & \varepsilon_x \leq \varepsilon_c
\end{cases} \]

\begin{itemize}
\item $\ell < \bar{\lambda}$
\item $\bar{\lambda}$
\item $\varepsilon_c$
\end{itemize}
Figure 2: Robustness vs normalized forecast error. \( \tilde{\lambda} = 1, y_T = 5.25 \).

\[ \ell = 1.0 \implies 0\% \text{ robustness at 0\% error.} \]

\[ \ell = 0.95 \implies 10\% \text{ robustness at 5\% error.} \]

\[ \ell = 0.9 \implies 20\% \text{ robustness at 10\% error.} \]
Figure 3: Robustness vs normalized forecast error. $\tilde{\lambda} = 1$, $y_T = 5.25$.

§ $\ell = 0.9 \implies 20\%$ robustness at 10\% error.

§ **Forecast:** $y_{T+1}^S = 0.9y_T = 4.725.$
Figure 4: Robustness vs normalized forecast error. $\tilde{\lambda} = 1, y_T = 5.25$.

$\ell = 0.9 \implies 20\%$ robustness at $10\%$ error.

**Forecast:** $y_{T+1}^s = 0.9y_T = 4.725$.

**Outcome:**
- $y_{T+1} = 4.75$ on 18.9.2001.
- $-0.5\%$ forecast error.
2.3 Crossing Robustness Curves: General Case

Recall numerical example:

\[ \hat{h}_a(\ell, \varepsilon_c) \]

Crossing robustness curves:

Advantage of sub-optimal model.

How general?
§ **System:**

\[ y_t = A_t y_{t-1}, \quad y_t \in \mathbb{R}^N \]

§ **Incorporate** inputs into state vector.

§ **Ignore** zero-mean, additive, random disturbances.

§ **Goal:**

Given \( y_T \) and historical \( \bar{A} \),

predict \( y_{T+k} \) within \( \pm \varepsilon_c \).

§ **Problem:** Uncertain future \( A_t \).
§ **Info-gap model**, $\mathcal{U}(h, \bar{A})$, $h \geq 0$.

**Axioms:**

- **Nesting:** $h < h'$ implies $\mathcal{U}(h, \bar{A}) \subset \mathcal{U}(h', \bar{A})$
- **Contraction:** $\mathcal{U}(0, \bar{A}) = \{\bar{A}\}$

§ **Two levels of uncertainty:**

- Horizon of uncertainty, $h$, unknown.
- Realization unknown.
Example: Unbounded-interval info-gap model:

\[ \mathcal{U}(h, \tilde{A}) = \{ A_t, \ t > T : \]

\[ \tilde{A}_{ij} - hv_{ij} \leq [A_t]_{ij} \leq \tilde{A}_{ij} + hw_{ij}, \]

\[ i, j = 1, \ldots, N \}, \quad h \geq 0 \]
§ Historically estimated model:

\[ y_t = \tilde{A} y_{t-1} \]

§ Slope-adjusted predictor:

\[ y^s_t = B y^s_{t-1} \]

\[ B \text{ chosen by forecaster.} \]

§ Forecast error:

\[ \eta_k(B, A_t) = y^s_{T+k} - y_{T+k} = \left( B^k - \prod_{i=1}^{k} A_{T+i} \right) y_T \]
§ Satisficing forecast requirement:

\[ |\eta_{k,m}(B, A_t)| \leq \varepsilon_c \]

§ Robustness: Max tolerable uncertainty.

\[ \bar{h}(B, \varepsilon_c) = \max \left\{ h : \left( \max_{A_{T+i} \in \mathcal{U}(h,A)} \left| \eta_{k,m}(B, A_t) \right| \right) \leq \varepsilon_c \right\} \]
§ k-Step transition matrices:
\[ U_k(h, \overline{A}^k) = \left\{ A = \prod_{i=1}^{k} A_i : A_i \in U(h, \overline{A}) \right\}, \quad h \geq 0 \]

Lemma 1 If \( U(h, \overline{A}) \) is an info-gap model. Then \( U_k(h, \overline{A}^k) \) is an info-gap model.

§ 1-step thm based on

- nesting and contraction axioms

is k-step thm.
§ Robustness-premium theorem. Define:

\[ \theta_c(h) = \max_{A_{T+1} \in \mathcal{U}(h, \bar{A})} \sum_{n=1}^{N} [A_{T+1} - \bar{A}]_{mn} y_{T,n} \]

\[ \theta_a(h) = -\min_{A_{T+1} \in \mathcal{U}(h, \bar{A})} \sum_{n=1}^{N} [A_{T+1} - \bar{A}]_{mn} y_{T,n} \]

- **Contraction:** \( \theta_a(0) = \theta_c(0) = 0 \).
- **Nesting:** \( \theta_a(h) \) and \( \theta_c(h) \) increase with \( h \).
- \( \theta_c(h) \) large: IGM coherent with \( y_T \).
- \( \theta_a(h) \) large: IGM anti-coherent with \( y_T \).
Theorem 1 *Sub-optimal models are more robust than optimal models.*

**Given:**
- $y_T \neq 0$.
- $U(h, A)$ is an info-gap model.
- $\theta_c(h)$ and $\theta_a(h)$ are continuous, at least one is unbounded, and either $\theta_c(h) \geq \theta_a(h)$ or $\theta_a(h) \geq \theta_c(h)$ for all $h > 0$.

**Then:** for any $\varepsilon_c > 0$ for which $\delta_x(\varepsilon_c) \neq 0$, there is a $B$ such that:

$$\tilde{h}(B, \varepsilon_c) > \tilde{h}(A, \varepsilon_c)$$
§ Evaluating the robustness. Define:

\[ \theta_c(h) = \max_{A_{T+1} \in U(h, \bar{A})} \sum_{n=1}^{N} [A_{T+1} - \bar{A}]_{mn}y_{T,n} \]

\[ \theta_a(h) = -\min_{A_{T+1} \in U(h, \bar{A})} \sum_{n=1}^{N} [A_{T+1} - \bar{A}]_{mn}y_{T,n} \]

\[ \delta = \sum_{n=1}^{N} [B - \bar{A}]_{mn}y_{T,n} \]

- **Contraction:** \( \theta_a(0) = \theta_c(0) = 0 \).
- **Nesting:** \( \theta_a(h) \) and \( \theta_c(h) \) increase with \( h \).
- **\( \theta_c(h) \) large:** IGM coherent with \( y_T \).
- **\( \theta_a(h) \) large:** IGM anti-coherent with \( y_T \).
- **\( \delta \) controlled by forecaster.**
§ Robustness function:

\[ \tilde{h}(B, \varepsilon_c) = \max \{ h : \delta + \theta_a(h) \leq \varepsilon_c \text{ and } -\delta + \theta_c(h) \leq \varepsilon_c \} \]
§ When is the robustness zero?

§ Anticipated 1-step prediction error:

\[ \eta_{1,m}(B, \bar{A}) = \sum_{n=1}^{N} \left( B - \bar{A} \right)_{mn} \frac{y_{T,n}}{\delta} \]
§ When is the robustness zero?

§ Anticipated 1-step prediction error:

\[ \eta_{1,m}(B, \bar{A}) = \sum_{n=1}^{N} (B - \bar{A})_{mn} y_{T,n} \delta \]

\[ \varepsilon_c = \delta \text{ has zero robustness:} \]

\[ \tilde{h}(B, \varepsilon_c) \]

\[ 0 \quad 0 \quad \delta \quad \delta \]

§ Positive robustness only at greater-than-predicted forecast error.
§ Choose $B = \bar{A}$ to minimize anticipated prediction error, $\eta_{1,m}(B, \bar{A})$. 

---

Info-Gap Theory
§ Choose $B = \bar{A}$ to minimize anticipated prediction error, $\eta_{1,m}(B, \bar{A})$.

- Predicted error is $\delta = 0$.
- $\varepsilon_c = \delta$ still has zero robustness:

\[ \tilde{h}_{\delta}(B, \varepsilon_c) = \begin{cases} \delta + \theta_a(h) & \text{if } \varepsilon_c > \delta \\ \delta - \theta_c(h) & \text{if } \varepsilon_c < -\delta \\ 0 & \text{if } \varepsilon_c = \delta \end{cases} \]

§ Positive robustness only at greater-than-predicted forecast error.
2.4 Robustness and Probability of Forecast Success
§ Probability of Forecast Success.

- Forecast success: $|\eta_{1,m}| \leq \varepsilon_c$, or:

$$\delta - \varepsilon_c \leq \sum_{n=1}^{N} \left[ A_{T+1} - \bar{A} \right]_{mn} y_{T,n} \leq \delta + \varepsilon_c$$

$u =$ random variable.

$$F(u) = \text{cdf (unknown)}.$$

$\delta(B) =$ anticip. forecast error (chosen).

- Probability of forecast success:

$$P_s(B) = F(\delta + \varepsilon_c) - F(\delta - \varepsilon_c)$$

Thus:

$$\frac{dP_s(B)}{d\delta} > 0 \quad \text{if and only if} \quad f(\delta + \varepsilon_c) > f(\delta - \varepsilon_c)$$
\( \theta_c(h) > \theta_a(h) \) implies:

- \([A_{T+1} - \bar{A}]_{mn}\) coherent with \(\text{sgn}(y_{Tn})\).
- \(u\) tends to be positive. \(u = \sum_{n=1}^{N} [A_{T+1} - \bar{A}]_{mn} y_{Tn}\)
- \(f(u)\) tends to increase around \(u = 0\).

Definition 1 \(U(h, \bar{A})\) and \(F(u)\) coherent at \((\delta, \varepsilon_c)\) if:

\[
[\theta_c(h) - \theta_a(h)] [f(\delta + \varepsilon_c) - f(\delta - \varepsilon_c)] \geq 0 \text{ for all } h > 0
\]

- Coherence:

  The info-gap model weakly reveals the pdf.
Theorem 2 *Robustness is a proxy for probability of forecast success.* Given:

- $y_T \neq 0$, $\varepsilon_c \geq 0$, $\delta \times (\varepsilon_c) \neq 0$, $|\delta| < |\delta \times (\varepsilon_c)|$.
- $\bar{h}(B, \varepsilon_c) > 0$.
- $\theta_a(h)$ and $\theta_c(h)$ are continuous and at least one is unbounded.
- $\mathcal{U}(h, \bar{A})$ and $F(u)$ are coherent.

Then:

$$\frac{d\bar{h}(B, \varepsilon_c)}{d\delta} > 0 \quad \text{if and only if} \quad \frac{dP_s(B)}{d\delta} > 0$$
§ Importance of proxy thm:

- $P_s(B)$ unknown.
- $\hat{h}(B, \varepsilon_c)$ known.
- $B$ chosen by forecaster.

§ Coherence of $\mathcal{U}(h, \bar{A})$ and $F(u)$ implies:

$P_s$ is not known but can be optimized.
§ Many proxy theorems.

- $A =$ model, data: uncertain.
- $\mathcal{U}(h, \overline{A}) =$ info-gap model for uncertainty.
- $B =$ design, decision, strategy.
- $\eta(B, A) =$ outcome: $\eta(B, A) \leq \varepsilon_c$.
- $\tilde{h}(B, \varepsilon_c) =$ robustness function.
- $P_s(B) =$ probability of success.
- Proxy “theorem”:
  $$\left( \frac{\partial \tilde{h}(B, \varepsilon_c)}{\partial B} \right) \left( \frac{\partial P_s(B)}{\partial B} \right) \geq 0$$
- Fine print: e.g. $\mathcal{U}(h, \overline{A}) \& F(A)$ coherent.
§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. Historically: $\lambda_t = \bar{\lambda}$.

- **Future:** $\lambda_t = \bar{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_* \\ \text{unknown}, & u_t \geq \lambda_* \end{cases}$.

  E.g. $f(u_t) = \begin{cases} 0, & u_t < \lambda_* \\ \lambda_* / u^2, & u_t \geq \lambda_* \end{cases}$. 
§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \bar{\lambda}$.

- **Future:** $\lambda_t = \bar{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda^* \\ \text{unknown}, & u_t \geq \lambda^* \end{cases}$.
  
  E.g. $f(u_t) = \begin{cases} 0, & u_t < \lambda^* \\ \lambda^*/u^2, & u_t \geq \lambda^* \end{cases}$.

- **Forecaster:** $x^s_t = \ell x^s_{t-1}$.

- $\delta = \text{anticipated forecast error} = (\ell - \bar{\lambda})y_T$. 
§ Coherence: Example.

- **System:** \( x_t = \lambda_t x_{t-1} \). Historically: \( \lambda_t = \bar{\lambda} \).

- **Future:** \( \lambda_t = \bar{\lambda} + u_t \), \( f(u_t) = \begin{cases} 0, & u_t < \lambda_* \\ \text{unknown}, & u_t \geq \lambda_* \end{cases} \).

- **Forecaster:** \( x^s_t = \ell x^s_{t-1} \).

- \( \delta = \) anticipated forecast error = \((\ell - \bar{\lambda})y_T\).

- \( \mathcal{U}(h, \bar{\lambda}) = \{ \lambda : \bar{\lambda} \leq \lambda \leq (1 + h)\bar{\lambda} \}, \ h \geq 0. \)

- \( \mathcal{U}(h, \bar{\lambda}) \) and \( F(u) \) coherent if:
  \[ \bar{\lambda} + \lambda_* - \frac{\varepsilon_c}{y_T} < \ell < \bar{\lambda} + \lambda_* + \frac{\varepsilon_c}{y_T}. \]
§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \bar{\lambda}$.

- **Future:** $\lambda_t = \bar{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_* \\ \text{unknown}, & u_t \geq \lambda_* \end{cases}$.

- **Forecaster:** $x^s_t = \ell x^s_{t-1}$.

- $\delta = \text{anticipated forecast error} = (\ell - \bar{\lambda})y_T$.

- $\mathcal{U}(h, \bar{\lambda}) = \{ \lambda : \bar{\lambda} \leq \lambda \leq (1 + h)\bar{\lambda} \}, \ h \geq 0$.

- $\mathcal{U}(h, \bar{\lambda})$ and $F(u)$ **coherent** if:
  $\bar{\lambda} + \lambda_* - \frac{\varepsilon_c}{y_T} < \ell < \bar{\lambda} + \lambda_* + \frac{\varepsilon_c}{y_T}$.

- Then: $\frac{\partial P^s}{\partial \ell} > 0$.

- $P^s(\ell)$ **unknown** but **improvable**.
2.5 Regression Prediction

Yakov Ben-Haim,

*Info-Gap Economics: An Operational Introduction*,
Figure 5: US inflation vs. year, 1961–1965.

§ US inflation ’61–’65: Linear?
Figure 6: US inflation vs. year, 1961–1965.
Figure 7: US inflation vs. year, 1961–1966.

§ **US inflation ’61–’65:** Linear?

§ **US inflation ’61–’66:** Quadratic?
§ **US inflation ’61–’65**: Linear?

§ **US inflation ’61–’66**: Quadratic?

§ **US inflation ’61–’70**: Piece-wise linear?
Figure 11: US inflation vs. year, 1961–1993.

§ US inflation ’61–’93: A mess?
Figure 12: US inflation vs. year, 1961–1965.

§ **US inflation ’61–’65:**

- Linear? Quadratic?
- Model ’61–’65 for predicting ’66:

\[ y_i^r = c_0 + c_1 t_i + c_2 t_i^2 \]
§ System model: MSE.

\[
S_N^2(c) = \frac{1}{N} \sum_{i=1}^{N} (y_i - y_i^r)^2
\]

\(N = 5 \text{ for '61–'65.}\)
§ If we knew $y_{N+1}$ ('66):

\[
S_{N+1}^2(c) = \frac{1}{N+1} \sum_{i=1}^{N+1} (y_i - y_i^r)^2
\]

\[
= \frac{N}{N+1} S_N^2(c) + \frac{(y_{N+1} - y_{N+1}^r)^2}{N+1}
\]
§ If we knew $y_{N+1}$ ('66):

$$S_{N+1}^2(c) = \frac{1}{N+1} \sum_{i=1}^{N+1} (y_i - y_{i}^{r})^2$$

$$= \frac{N}{N+1} S_N^2(c) + \frac{(y_{N+1} - y_{N+1}^{r})^2}{N+1}$$

§ All we know is contextual info:

$y_{N+1}$ may well exceed prediction, $\bar{y}_{N+1}^{r}$. 
§ If we knew \( y_{N+1} \) (’66):
\[
S^2_{N+1}(c) = \frac{1}{N + 1} \sum_{i=1}^{N+1} (y_i - y_i^r)^2
\]
\[
= \frac{N}{N + 1} S^2_N(c) + \frac{(y_{N+1} - y_{N+1}^r)^2}{N + 1}
\]

§ All we know is soft info:
\( y_{N+1} \) may well exceed prediction, \( y_{N+1}^r \).

§ Info-gap model of uncertain \( y_{N+1} \):
\[
\mathcal{U}(h) = \{ y_{N+1} : 0 \leq y_{N+1} - y_{N+1}^r \leq h \}, \quad h \geq 0
\]
- Unbounded family of nested sets.
- No worst case.
\section*{Performance requirement:}

\[ S_{N+1}(c) \leq S_c \]
§ **Performance requirement:**

\[ S_{N+1}(c) \leq S_c \]

§ **Robustness of regression \( c \):**

Greatest tolerable uncertainty.

\[ \tilde{h}(c, S_c) = \max \left\{ h : \left( \max_{y_{N+1} \in \mathcal{U}(h)} S_{N+1}(c) \right) \leq S_c \right\} \]
Figure 13: US inflation vs. year, 1961–1965, and least squares fit.

Figure 14: Robustness vs. critical root mean squared error for inflation 1961–1965.

§ Least squares fit: fig. 13.

§ Robust of LS fit: fig. 14.

§ Trade off: Greater rbs. $\equiv$ greater MSE.

§ Zeroing: No robustness of est. MSE.
Figure 15: US inflation vs. year, 1961–1965, and least squares fit (solid) and other fit (dash).

Figure 16: Robustness vs. critical root mean squared error for inflation 1961–1965 for least squares fit (solid) and other fit (dash).

§ Least squares and other fit: fig. 15.

§ Robust of LS and other fit: fig. 16.

Curve-crossing: preference reversal.
3 SUMMARY

\section{Models:}
Attributes of model correspond to attributes of reality.

\section{Model-based decision:}
Adapt decision to attributes of model.

\section{Optimization:}
Use best model to choose decision with best outcome.
§ **Uncertainty:**

- **Randomness:** structured uncertainty.
- **Info-gaps:** surprises, ignorance.
§ Fallacy of optimal model-based decision:

- Severe uncertainty:
  - Best model errs seriously.
  - Some model attributes are correct.
  - Some model attributes err greatly.

- Best-model optimization
  - Exploits all model attributes to extreme.
  - Vulnerable to model error.
§ Resolution: Info-gap decision theory.
  • Satisfice performance.
  • Optimize robustness to uncertainty.
  • Model and manage surPrises.

§ Robust-satisficing syllogism:
  • Adequate performance is necessary.
  • More reliable adequate performance is better than less reliable adeq. perf.
  • Thus maximum reliability is best.

§ Proxy theorems:
  max robustness ≡ max survival prob.
Sources: http://info-gap.com

Applications of info-gap theory:

- Biological conservation.
- Public policy and regulation.
- Climate change.
- Sampling, assay design.
- Medical decision making.
- Engineering design.
- Fault detection and diagnosis.
- Project management.
- Homeland security.
- Statistical hypothesis testing.
- Monetary economics.
- Financial stability.