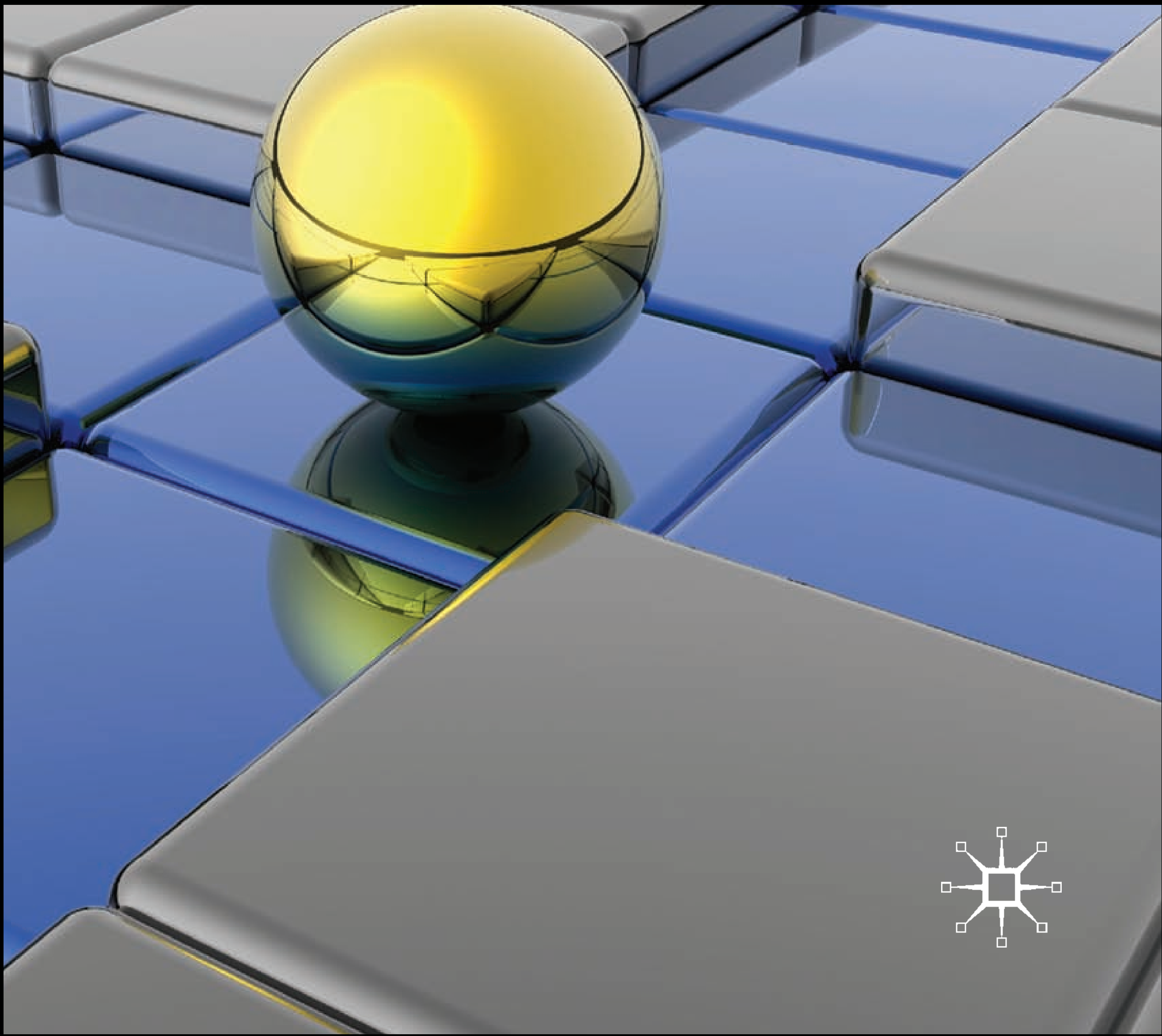
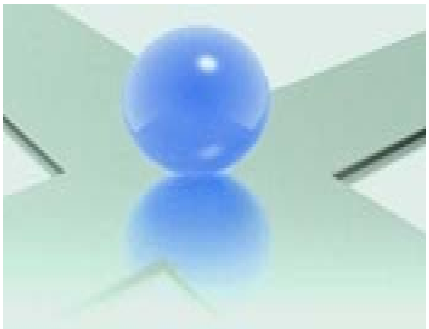


Forecasting
with Poor Data and Models:
An Info-Gap Approach

Yakov Ben-Haim

Technion

Israel Institute of Technology



Contents

1 *Highlights*

1.1 *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.

[0] \lectures\talks\lib\indif5d-shackle-pop.tex 4.6.2010

§ **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.

[0] lectures talks lib forecast02.tex 22.6.2010

2.1 *1-D Example*

§ **"True" scalar system:**

$$y_t = \lambda_t y_{t-1}$$

§ **Historical data:** $\lambda_t = \widetilde{\lambda}$ **for** $t \le T$**.**

§ **Contextual understanding:**

λ **could drift upwards.**

§ **Fractional-error info-gap model:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t, t > T : \ 0 \le \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \le h \right\}, \quad h \ge 0$$

- **Unbounded family of sets.**
- **No worst case.**

§ Slope-adjusted (erroneous) forecaster:

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}}$$

§ Contrast with historically estimated model:

$$y_t = \widetilde{\lambda}_t y_{t-1}$$

How to choose $\ell \ge \widetilde{\lambda}$?

§ Robust satisficing:

Satisfice the forecast error:

$$\left| y_{T+k}^{\mathrm{s}} - y_{T+k} \right| \le \varepsilon_{\mathrm{c}}$$

Maximize robustness to future surprise.

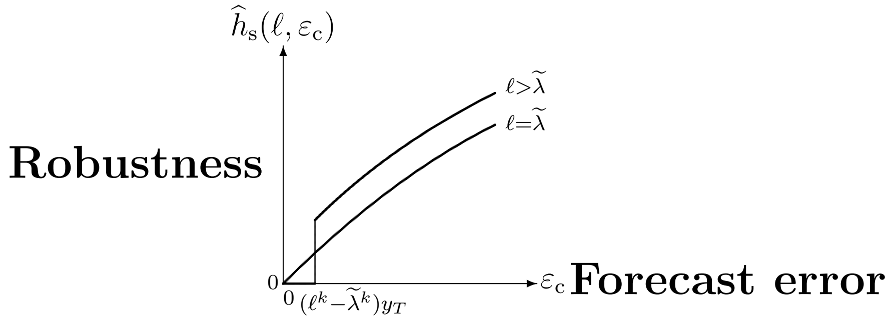
§ Robustness of forecast ℓ:

Max h up to which all λ_{T+i} in $\mathcal{U}(h, \widetilde{\lambda})$

satisfice forecast error at $\varepsilon_{\rm c}$:

$$\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) = \max\left\{ h : \left(\max_{\substack{\lambda_{T+i} \in \mathcal{U}(h, \widetilde{\lambda}) \\ i=1, \ldots, k}} \left| y^{\rm s}_{T+k} - y_{T+k} \right| \right) \le \varepsilon_{\rm c} \right\}$$

§ Preference: $\ell \succ \ell'$ **if** $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell', \varepsilon_{\rm c})$



§ **Trade off:** **robustness vs. forecast error.**

§ **Zeroing:** **Estim outcome has 0 robustness.**

§ **Crossing robustness curves:** $\ell \succ \widetilde{\lambda}$.

- **Preference reversal.**
- **Robustness-advantage of sub-optimal (erroneous) model.**

§ **Robustness is proxy for success-probability.**

§ Numerical example.

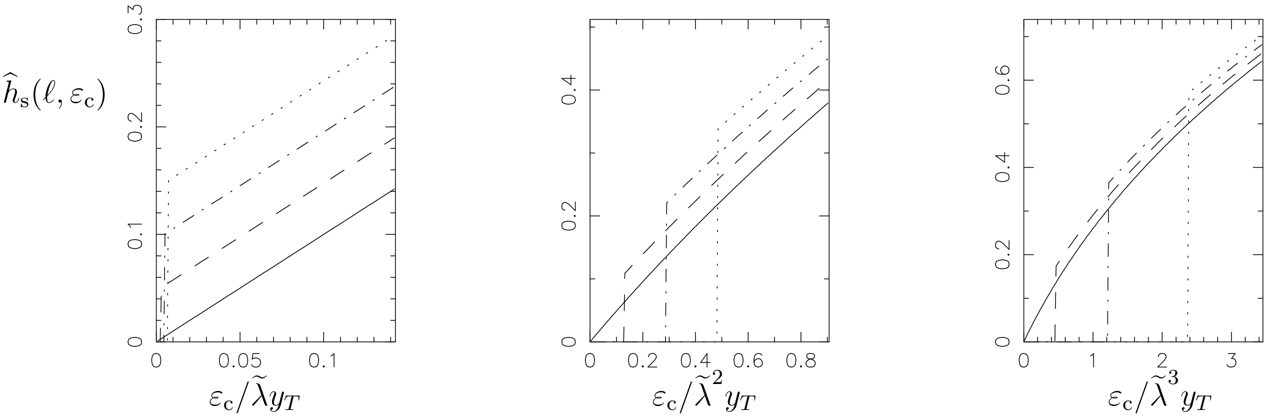


Figure 1: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- **Preference reversal at all time horizons, k.**
- **Robustness premium decreases with k.**
- **Reversal-ε_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:**

$$\mathcal{Y}(\ell) = \{\lambda_{T,k} \in D : \ |y^{\mathrm{s}}_{T+k}(\ell) - y_{T+k}| \le \varepsilon_{\mathrm{c}}\}$$

§ **Probability of success:**

$$P_{\mathrm{s}}(\ell) = F[\mathcal{Y}(\ell)]$$

§ Goal:

Choose ℓ to maximize success prob.

§ Problem:

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

§ Solution:

- $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$ **is known.**
- $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$ **proxies for success prob.**

§ Theorem.

Probability of successful forecast, $P_{\rm s}(\ell)$, increases with increasing info-gap robustness, $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$.

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

$$\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell', \varepsilon_{\rm c}) > 0$$

Then:

$$P_{\rm s}(\ell) \geq P_{\rm s}(\ell')$$

§ 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*

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

§ Typical change: 25 basis points.

§ Largest change: 100 basis points.

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

§ 9Jun'00–31Aug'01: $\mu = 5.4\%$, $\sigma = 0.19\%$.

§ On 9/12/2001, (1 day after 9/11)

predict next interest rate.

Rate down, but by how much?

§ Historical model:

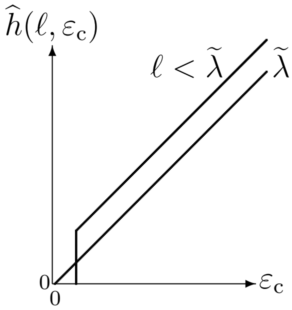
$$y_t = \lambda_t y_{t-1}, \qquad \lambda_t = \widetilde{\lambda} = 1$$

§ Info-gap model:

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_T : \ (1-h)\widetilde{\lambda} \le \lambda_T \le \widetilde{\lambda} \right\}, \qquad h \ge 0$$

§ Robustness:

$$\widehat{h}(\ell, \varepsilon_{\rm c}) = \begin{cases} 0 & \varepsilon_{\rm c} < \varepsilon_\times \\ \dfrac{\varepsilon_{\rm c}}{\widetilde{\lambda} y_T} + \dfrac{\widetilde{\lambda} - \ell}{\widetilde{\lambda}} & \varepsilon_\times \le \varepsilon_{\rm c} \end{cases}$$



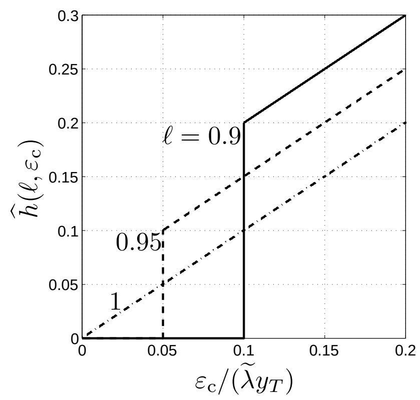


Figure 2: **Robustness vs normalized forecast error.** $\widetilde{\lambda} = 1$, $y_T = 5.25$.

§ $\ell = 1.0 \Longrightarrow$ **0% robustness at 0% error.**

§ $\ell = 0.95 \Longrightarrow$ **10% robustness at 5% error.**

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

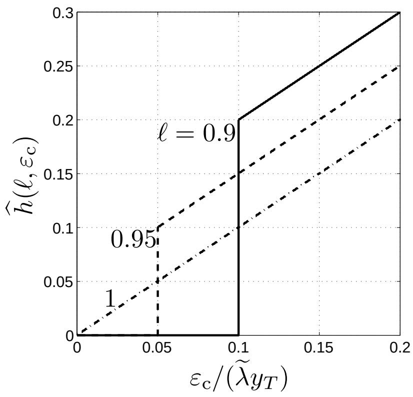


Figure 3: **Robustness vs normalized forecast error.** $\widetilde{\lambda} = 1$, $y_T = 5.25$.

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

§ **Forecast:** $y^{\rm s}_{T+1} = 0.9 y_T =$ **4.725.**

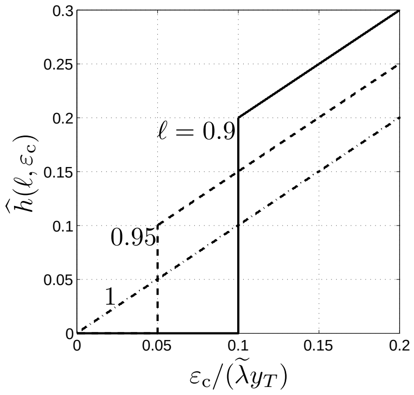


Figure 4: **Robustness vs normalized forecast error.** $\widetilde{\lambda} = 1$, $y_T = 5.25$.

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

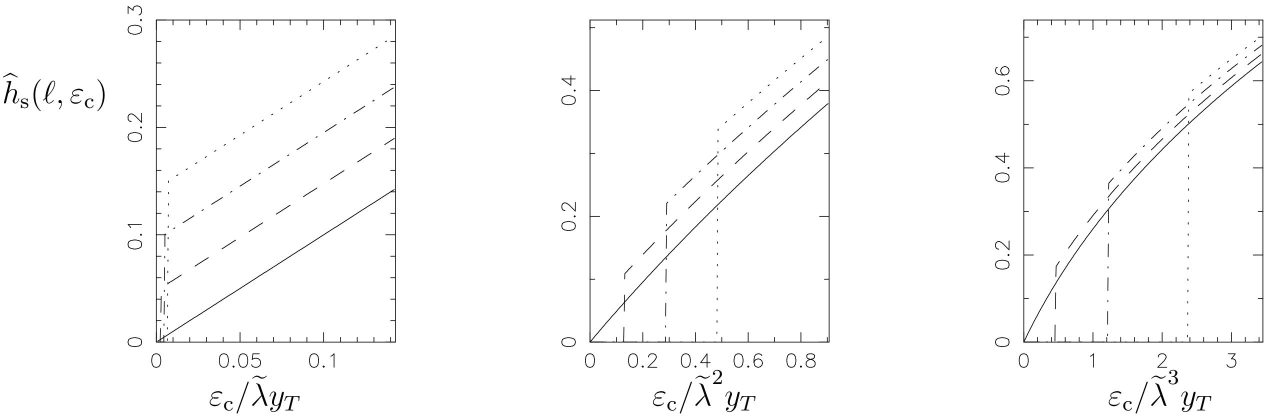
§ **Forecast:** $y^{s}_{T+1} = 0.9 y_T =$ **4.725.**

§ **Outcome:**

- $y_{T+1} =$ **4.75 on 18.9.2001.**
- -0.5% **forecast error.**

2.3 *Crossing Robustness Curves: General Case*

§ Recall numerical example:



§ Crossing robustness curves:

Advantage of sub-optimal model.

§ How general?

§ **System:**

$$y_t = A_t y_{t-1}, \quad y_t \in \Re^N$$

§ **Incorporate inputs into state vector.**

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

§ **Goal:**

Given y_T **and historical** $\widetilde{A}$**,**

predict y_{T+k} **within** $\pm\varepsilon_{\mathrm{c}}$**.**

§ **Problem: Uncertain future** A_t**.**

§ **Info-gap model,** $\mathcal{U}(h, \widetilde{A})$**,** $h \geq 0$**.**

Axioms:

Nesting: $h < h'$ **implies** $\mathcal{U}(h, \widetilde{A}) \subset \mathcal{U}(h', \widetilde{A})$

Contraction: $\mathcal{U}(0, \widetilde{A}) = \{\widetilde{A}\}$

§ **Two levels of uncertainty:**

- **Horizon of uncertainty,** h**, unknown.**
- **Realization unknown.**

§ Example: Unbounded-interval info-gap model:

$$\mathcal{U}(h, \widetilde{A}) = \Big\{ A_t, \ t > T : \quad \widetilde{A}_{ij} - h v_{ij} \le [A_t]_{ij} \le \widetilde{A}_{ij} + h w_{ij}, \quad i, j = 1, \ldots, N \Big\}, \quad h \ge 0$$

§ Historically estimated model:

$$y_t = \widetilde{A} y_{t-1}$$

§ Slope-adjusted predictor:

$$y_t^{\mathrm{s}} = B y_{t-1}^{\mathrm{s}}$$

B **chosen by forecaster.**

§ Forecast error:

$$\eta_k(B, A_t) = y_{T+k}^{\mathrm{s}} - y_{T+k} = \left(B^k - \prod_{i=1}^{k} A_{T+i} \right) y_T$$

§ Satisficing forecast requirement:

$$|\eta_{k,m}(B, A_t)| \le \varepsilon_{\rm c}$$

§ Robustness: Max tolerable uncertainty.

$$\widehat{h}(B, \varepsilon_{\rm c}) = \max \left\{ h : \left(\max_{\substack{A_{T+i} \in \mathcal{U}(h, \widetilde{A}) \\ i=1, \ldots, k}} |\eta_{k,m}(B, A_t)| \right) \le \varepsilon_{\rm c} \right\}$$

§ k-Step transition matrices:

$$\mathcal{U}_k(h, \widetilde{A}^k) = \left\{ A = \prod_{i=1}^{k} A_i : \; A_i \in \mathcal{U}(h, \widetilde{A}) \right\}, \qquad h \geq 0$$

Lemma 1 ***If*** $\mathcal{U}(h, \widetilde{A})$ ***is an info-gap model. Then*** $\mathcal{U}_k(h, \widetilde{A}^k)$ ***is an info-gap model.***

§ 1-step thm based on

nesting and contraction axioms

is k-step thm.

§ Robustness-premium theorem. Define:

$$\theta_{\rm c}(h) = \max_{A_{T+1} \in \mathcal{U}(h, \widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

$$\theta_{\rm a}(h) = - \min_{A_{T+1} \in \mathcal{U}(h, \widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

- **Contraction:** $\theta_{\rm a}(0) = \theta_{\rm c}(0) = 0$.
- **Nesting:** $\theta_{\rm a}(h)$ **and** $\theta_{\rm c}(h)$ **increase with** h.
- $\theta_{\rm c}(h)$ **large: IGM coherent with** y_T.
- $\theta_{\rm a}(h)$ **large: IGM anti-coherent with** y_T.

Theorem 1 ***Sub-optimal models are more robust than optimal models.***

Given:

- $y_T \neq 0$.
- $\mathcal{U}(h, \widetilde{A})$ ***is an info-gap model.***
- $\theta_{\rm c}(h)$ ***and*** $\theta_{\rm a}(h)$ ***are continuous, at least one is unbounded, and either*** $\theta_{\rm c}(h) \geq \theta_{\rm a}(h)$ ***or*** $\theta_{\rm a}(h) \geq \theta_{\rm c}(h)$ ***for all*** $h > 0$.

Then: ***for any*** $\varepsilon_{\rm c} > 0$ ***for which*** $\delta_{\times}(\varepsilon_{\rm c}) \neq 0$, ***there is a*** B ***such that:***

$$\widehat{h}(B, \varepsilon_{\rm c}) > \widehat{h}(\widetilde{A}, \varepsilon_{\rm c})$$

§ Evaluating the robustness. Define:

$$\theta_{\rm c}(h) = \max_{A_{T+1}\in\mathcal{U}(h,\widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

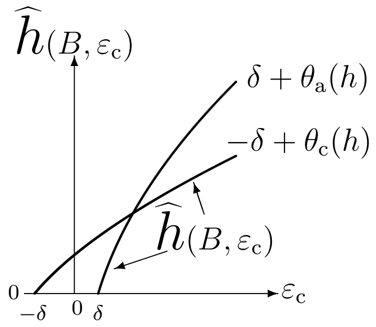
$$\theta_{\rm a}(h) = -\min_{A_{T+1}\in\mathcal{U}(h,\widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

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

- **Contraction:** $\theta_{\rm a}(0) = \theta_{\rm c}(0) = 0$.
- **Nesting:** $\theta_{\rm a}(h)$ **and** $\theta_{\rm c}(h)$ **increase with** h.
- $\theta_{\rm c}(h)$ **large: IGM coherent with** y_T.
- $\theta_{\rm a}(h)$ **large: IGM anti-coherent with** y_T.
- δ **controlled by forecaster.**

§ Robustness function:

$$\widehat{h}(B,\varepsilon_{\rm c}) = \max\left\{h:\ \ \delta+\theta_{\rm a}(h)\le\varepsilon_{\rm c}\ \textbf{and}\ \ -\delta+\theta_{\rm c}(h)\le\varepsilon_{\rm c}\right\}$$



§ When is the robustness zero?

§ Anticipated 1-step prediction error:

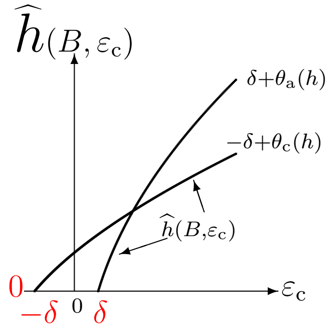
$$\eta_{1,m}(B,\widetilde{A}) = \underbrace{\sum_{n=1}^{N}\left(B-\widetilde{A}\right)_{mn} y_{T,n}}_{\delta}$$

§ When is the robustness zero?

§ Anticipated 1-step prediction error:

$$\eta_{1,m}(B,\widetilde{A}) = \underbrace{\sum_{n=1}^{N} \left(B-\widetilde{A}\right)_{mn} y_{T,n}}_{\delta}$$

$\varepsilon_c = \delta$ **has zero robustness:**



§ Positive robustness only at

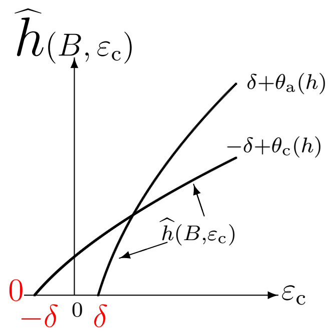
greater-than-predicted forecast error.

§ Choose $B = \widetilde{A}$ to minimize anticipated prediction error, $\eta_{1,m}(B, \widetilde{A})$???

§ **Choose $B = \widetilde{A}$ to minimize**

anticipated prediction error, $\eta_{1,m}(B, \widetilde{A})$**???**

- **Predicted error is** $\delta = 0$**.**
- $\varepsilon_{\rm c} = \delta$ **still has zero robustness:**



§ **Positive robustness only at**

greater-than-predicted forecast error.

2.4 *Robustness and Probability of Forecast Success*

§ Probability of Forecast Success.

- **Forecast success:** $|\eta_{1,m}| \le \varepsilon_{\rm c}$**, or:**

$$\delta - \varepsilon_{\rm c} \le \underbrace{\sum_{n=1}^{N} \left[A_{T+1} - \widetilde{A} \right]_{mn} y_{T,n}}_{u} \le \delta + \varepsilon_{\rm c}$$

$u =$ **random variable.**

$F(u) =$ **cdf (unknown).**

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

- **Probability of forecast success:**

$$P_{\rm s}(B) = F(\delta + \varepsilon_{\rm c}) - F(\delta - \varepsilon_{\rm c})$$

Thus:

$$\frac{{\rm d}P_{\rm s}(B)}{{\rm d}\delta} > 0 \quad \textbf{if and only if} \quad f(\delta + \varepsilon_{\rm c}) > f(\delta - \varepsilon_{\rm c})$$

§ $\theta_{\rm c}(h) > \theta_{\rm a}(h)$ **implies:**

- $[A_{T+1} - \widetilde{A}]_{mn}$ **coherent with sgn**(y_{T_n})**.**
- u **tends to be positive.** $u = \Sigma_{n=1}^{N} \left[A_{T+1} - \widetilde{A}\right]_{mn} y_{T,n}$
- $f(u)$ **tends to increase around** $u = 0$**.**

Definition 1 $\mathcal{U}(h, \widetilde{A})$ ***and*** $F(u)$ ***coherent at*** $(\delta, \varepsilon_{\rm c})$ ***if:***

$$[\theta_{\rm c}(h) - \theta_{\rm a}(h)]\,[f(\delta + \varepsilon_{\rm c}) - f(\delta - \varepsilon_{\rm c})] \geq 0 \;\; \textit{for all}\; h > 0$$

- **Coherence:**

  The info-gap model

  weakly reveals the pdf.

Theorem 2 ***Robustness is a proxy for probability of forecast success. Given:***

- $y_T \neq 0$, $\varepsilon_\mathrm{c} \geq 0$, $\delta_\times(\varepsilon_\mathrm{c}) \neq 0$, $|\delta| < |\delta_\times(\varepsilon_\mathrm{c})|$.
- $\widehat{h}(B, \varepsilon_\mathrm{c}) > 0$.
- $\theta_\mathrm{a}(h)$ ***and*** $\theta_\mathrm{c}(h)$ ***are continuous and at least one is unbounded.***
- $\mathcal{U}(h, \widetilde{A})$ ***and*** $F(u)$ ***are coherent.***

Then:

$$\frac{\mathrm{d}\widehat{h}(B, \varepsilon_\mathrm{c})}{\mathrm{d}\delta} > 0 \quad \textit{if and only if} \quad \frac{\mathrm{d}P_\mathrm{s}(B)}{\mathrm{d}\delta} > 0$$

§ Importance of proxy thm:

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

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

P_s **is not known but can be optimized.**

§ Many proxy theorems.

- A = **model, data: uncertain.**
- $\mathcal{U}(h, \widetilde{A})$ = **info-gap model for uncertainty.**
- B = **design, decision, strategy.**
- $\eta(B, A)$ = **outcome:** $\eta(B, A) \le \varepsilon_{\mathrm{c}}$**.**
- $\widehat{h}(B, \varepsilon_{\mathrm{c}})$ = **robustness function.**
- $P_{\mathrm{s}}(B)$ = **probability of success.**
- **Proxy "theorem":**

$$\left(\frac{\partial \widehat{h}(B, \varepsilon_{\mathrm{c}})}{\partial B}\right)\left(\frac{\partial P_{\mathrm{s}}(B)}{\partial B}\right) \ge 0$$

- **Fine print: e.g.** $\mathcal{U}(h, \widetilde{A})$ **&** $F(A)$ **coherent.**

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.

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

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.

  E.g. $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \lambda_\star / u^2, & u_t \geq \lambda_\star \end{cases}$.
- **Forecaster:** $x_t^{\mathrm{s}} = \ell x_{t-1}^{\mathrm{s}}$.
- δ = **anticipated forecast error** = $(\ell - \widetilde{\lambda}) y_T$.

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.
- **Forecaster:** $x_t^{\rm s} = \ell x_{t-1}^{\rm s}$.
- δ = **anticipated forecast error** = $(\ell - \widetilde{\lambda}) y_T$.
- $\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda : \ \widetilde{\lambda} \leq \lambda \leq (1+h)\widetilde{\lambda} \right\}, \ h \geq 0$.
- $\mathcal{U}(h, \widetilde{\lambda})$ **and** $F(u)$ **coherent if:**

$$\widetilde{\lambda} + \lambda_\star - \frac{\varepsilon_{\rm c}}{y_T} \ < \ \ell \ < \ \widetilde{\lambda} + \lambda_\star + \frac{\varepsilon_{\rm c}}{y_T}.$$

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.
- **Forecaster:** $x_t^{\mathrm{s}} = \ell x_{t-1}^{\mathrm{s}}$.
- $\delta =$ **anticipated forecast error** $= (\ell - \widetilde{\lambda}) y_T$.
- $\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda : \ \widetilde{\lambda} \leq \lambda \leq (1+h)\widetilde{\lambda} \right\}, \ h \geq 0$.
- $\mathcal{U}(h, \widetilde{\lambda})$ **and** $F(u)$ **coherent if:**

  $$\widetilde{\lambda} + \lambda_\star - \frac{\varepsilon_{\mathrm{c}}}{y_T} \ < \ \ell \ < \ \widetilde{\lambda} + \lambda_\star + \frac{\varepsilon_{\mathrm{c}}}{y_T}.$$
- **Then:** $\dfrac{\partial P_{\mathrm{s}}}{\partial \ell} > 0$.
- $P_{\mathrm{s}}(\ell)$ **unknown but improvable.**

2.5 *Regression Prediction*

Yakov Ben-Haim,

Info-Gap Economics:

An Operational Introduction,

2010, Palgrave-Macmillan, section 6.1.

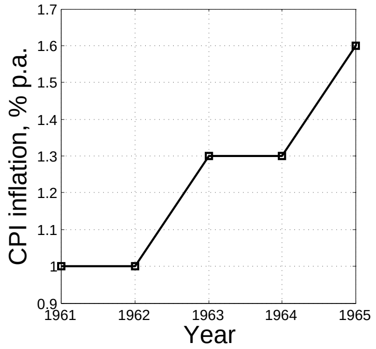


Figure 5: **US inflation vs. year, 1961–1965.**

§ US inflation '61–'65: Linear?

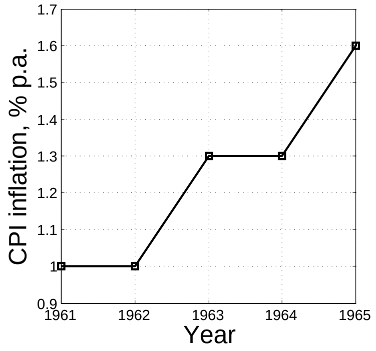


Figure 6: **US inflation vs. year, 1961–1965.**

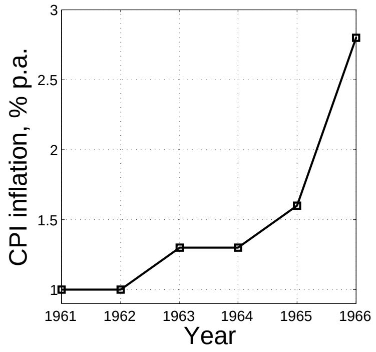


Figure 7: **US inflation vs. year, 1961–1966.**

§ US inflation ’61–’65: Linear?

§ US inflation ’61–’66: Quadratic?

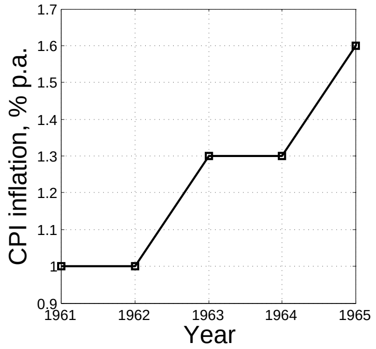


Figure 8: **US inflation vs. year, 1961–1965.**

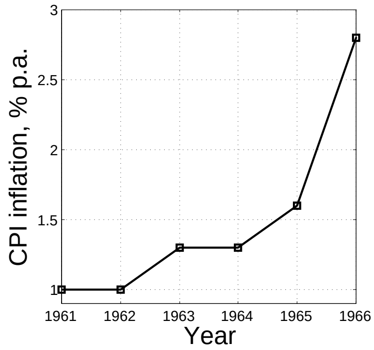


Figure 9: **US inflation vs. year, 1961–1966.**

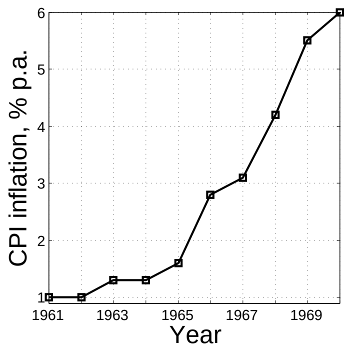


Figure 10: **US inflation vs. year, 1961–1970.**

§ US inflation '61–'65: Linear?

§ US inflation '61–'66: Quadratic?

§ US inflation '61–'70: Piece-wise linear?

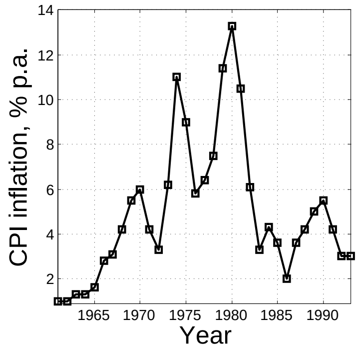


Figure 11: **US inflation vs. year, 1961–1993.**

§ US inflation '61–'93: A mess?

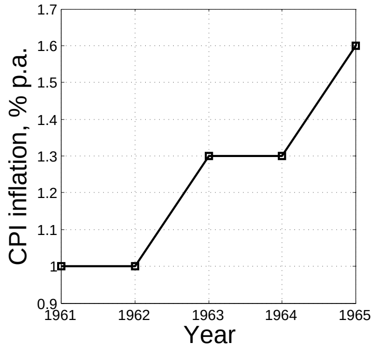


Figure 12: **US inflation vs. year, 1961–1965.**

§ US inflation '61–'65:

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

$$y_i^{\mathrm{r}} = c_0 + c_1 t_i + c_2 t_i^2$$

§ System model: MSE.

$$S_N^2(c) = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - y_i^{\mathrm{r}}\right)^2$$

$N = 5$ **for ’61–’65.**

§ If we knew y_{N+1} ('66):

$$S^2_{N+1}(c) = \frac{1}{N+1}\sum_{i=1}^{N+1}(y_i - y_i^{\rm r})^2$$

$$= \frac{N}{N+1}S^2_N(c) + \frac{(y_{N+1} - y^{\rm r}_{N+1})^2}{N+1}$$

§ **If we knew** y_{N+1} **('66):**

$$S^2_{N+1}(c) = \frac{1}{N+1}\sum_{i=1}^{N+1}(y_i - y_i^{\mathrm{r}})^2$$

$$= \frac{N}{N+1}S^2_N(c) + \frac{(y_{N+1} - y^{\mathrm{r}}_{N+1})^2}{N+1}$$

§ **All we know is contextual info:**

y_{N+1} **may well exceed prediction,** $\widetilde{y}^{\mathrm{r}}_{N+1}$**.**

§ **If we knew** y_{N+1} **('66):**

$$S^2_{N+1}(c) = \frac{1}{N+1}\sum_{i=1}^{N+1}(y_i - y_i^{\mathrm{r}})^2$$
$$= \frac{N}{N+1}S^2_N(c) + \frac{(y_{N+1} - y^{\mathrm{r}}_{N+1})^2}{N+1}$$

§ **All we know is soft info:**

y_{N+1} **may well exceed prediction,** $\widetilde{y}^{\mathrm{r}}_{N+1}$**.**

§ **Info-gap model of uncertain** y_{N+1}**:**

$$\mathcal{U}(h) = \left\{y_{N+1} : \ 0 \le y_{N+1} - \widetilde{y}^{\mathrm{r}}_{N+1} \le h\right\}, \quad h \ge 0$$

- **Unbounded family of nested sets.**
- **No worst case.**

§ Performance requirement:

$$S_{N+1}(c) \le S_{\mathrm{c}}$$

§ Performance requirement:

$$S_{N+1}(c) \le S_{\mathrm{c}}$$

§ Robustness of regression c:

Greatest tolerable uncertainty.

$$\widehat{h}(c, S_{\mathrm{c}}) = \max\left\{h : \ \left(\max_{y_{N+1}\in\mathcal{U}(h)} S_{N+1}(c)\right) \le S_{\mathrm{c}}\right\}$$

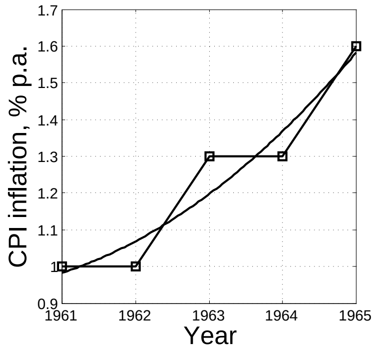


Figure 13: **US inflation vs. year, 1961–1965, and least squares fit.**

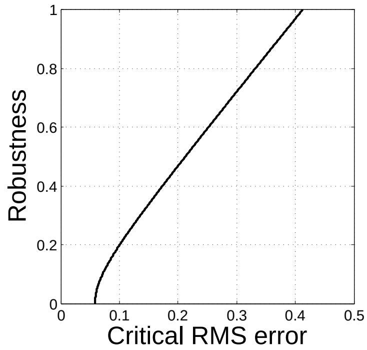


Figure 14: **Robustness vs. critical root mean squared error for inflation 1961–1965.**

§ **Least squares fit: fig. 13.**

§ **Robust of LS fit: fig. 14.**

§ **Trade off: Greater rbs. ≡ greater MSE.**

§ **Zeroing: No robustness of est. MSE.**

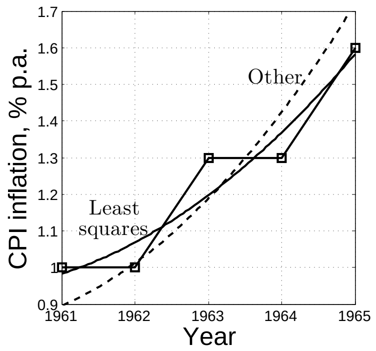


Figure 15: **US inflation vs. year, 1961–1965, and least squares fit (solid) and other fit (dash).**

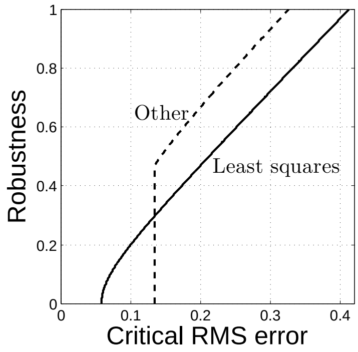


Figure 16: **Robustness vs. critical root mean squared error for inflation 1961–1965 for least squares fit (solid) and other fit (dash).**

§ Least squares and other fit: fig. 15.

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

Curve-crossing: preference reversal.

3 *SUMMARY*

§ **Models:**

Attributes of model correspond to attributes of reality.

§ **Model-based decision:**

Adapt decision to attributes of model.

§ **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 sur$^{\text{p}}$rises.**

§ 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.