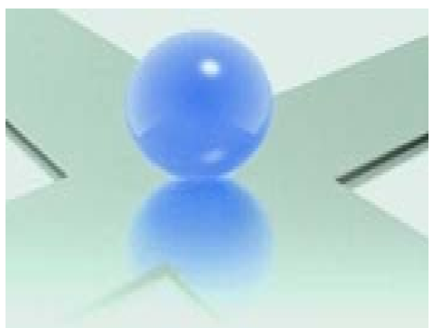
Lecture 5

Info-Gap Analysis of Estimation and Forecasting

Yakov Ben-Haim

Technion

Israel Institute of Technology



[0]lectures\talks\lib\tufts2025lec05-001.tex 25.12.2024

Contents

1 System Identification

[0] lectures talks lib sys_id5intro01.tex 27.1.2016

§ **Optimal system identification:**

Maximize fidelity of model to data.

§

§ **Optimal system identification:**

Maximize fidelity of model to data.

§ **Main thesis:**

Optimal identification has no robustness to structural error in the model.

§

§ Optimal system identification:

Maximize fidelity of model to data.

§ Main thesis:

Optimal identification has no robustness to structural error in the model.

§ Robust-satisficing.

Sub-optimal models can:

- **Satisfy data-fidelity requirements.**
- **Robustify against model uncertainty.**

1.1 *Example: Force-Displacement*

§ Force-displacement data:

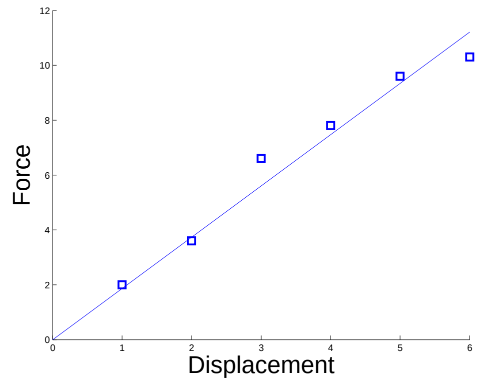


Figure 1: **Force-displacement data with LS fit.**

§ We might use linear model because:

- **Scientific theory.**
- **Computational limit (e.g. large dimension FE).**
- **Insufficient data to verify higher powers.**
- **Lab data is linear.**

---

0 lectures talks lib sys_id5force-disp02.tex 25.12.2024

§ Uncertainty.

The true model might be:

$$f = kx + k_3 x^3$$

- **Suggested by:**
  - **Contextual information.**
  - **Similar systems.**
  - **Aging.**
  - **Competing scientific theory.**
-

§ Uncertainty.

The true model might be:

$$f = kx + k_3 x^3$$

- **Suggested by:**
  - **Contextual information.**
  - **Similar systems.**
  - **Aging.**
  - **Competing scientific theory.**
- **But:**
  - **Cubic term is not modelled.**
  - **Sign and magnitude of k_3 unknown.**
  - **Info-gap model of uncertainty:**

$$\mathcal{U}(h) = \{k_3 : \ |k_3| \le h\}, \quad h \ge 0$$

§ Mean Square Error of linear model:

$$S(k) = \frac{1}{n} \sum_{i=1}^{n} (f_i - kx_i)^2$$

§

§ Mean Square Error of linear model:

$$S(k) = \frac{1}{n} \sum_{i=1}^{n} (f_i - kx_i)^2$$

§ Least squares estimate of linear model:

$$k_{\mathrm{LS}} = \arg \min_{k} S(k)$$

§

§ Mean Square Error of linear model:

$$S(k) = \frac{1}{n} \sum_{i=1}^{n} (f_i - kx_i)^2$$

§ Least squares estimate of linear model:

$$k_{\mathrm{LS}} = \arg \min_{k} S(k)$$

§ Questions:

- **How good is k_{LS} if true model is cubic?**
- **What k has good MSE if true model is cubic?**

§ Mean Square Error of cubic model:

$$S(k, k_3) = \frac{1}{n} \sum_{i=1}^{n} (f_i - kx_i - k_3 x_i^3)^2$$

§ Robustness of linear model, k:

Max uncertainty in k_3 with acceptable MSE.

$$\widehat{h}(k, S_{\rm c}) = \max \left\{ h : \ \left(\max_{k_3 \in \mathcal{U}(h)} S(k, k_3) \right) \le S_{\rm c} \right\}$$

§ Optimize or Robust-Satisfice???

•

§ Optimize or Robust-Satisfice???

- **Optimize fidelity to data:**

$$k_{\mathrm{LS}} = \arg\min_{k} S(k)$$

  - **Maximal fidelity to data.**
  - **Zero robustness to model error.**

-

§ Optimize or Robust-Satisfice???

- **Optimize fidelity to data:**

$$k_{\text{LS}} = \arg\min_k S(k)$$

  - **Maximal fidelity to data.**
  - **Zero robustness to model error.**

- **Robust-satisfice the fidelity:**

$$k_{\text{RS}} = \arg\max_k \widehat{h}(k, S_{\text{c}})$$

  - **Adequate fidelity to data.**
  - **Maximal robustness to model error.**

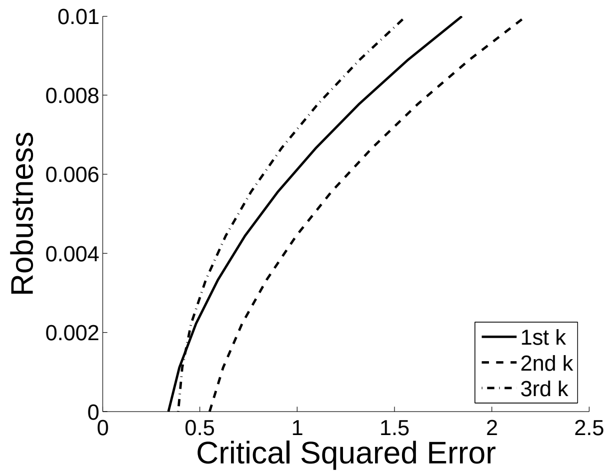


Figure 2: **Robustness curves.** $k_{\rm LS} = 1.8681$ **(solid),** $k = 1.75$ **(- -),** $k = 1.81$ **(-.).**

§ Trade-off: low MSE (good) $\Longleftrightarrow$ low robustness (bad).

$$S_{\rm c} < S'_{\rm c} \quad \Longleftrightarrow \quad \widehat{h}(k, S_{\rm c}) < \widehat{h}(k, S'_{\rm c})$$

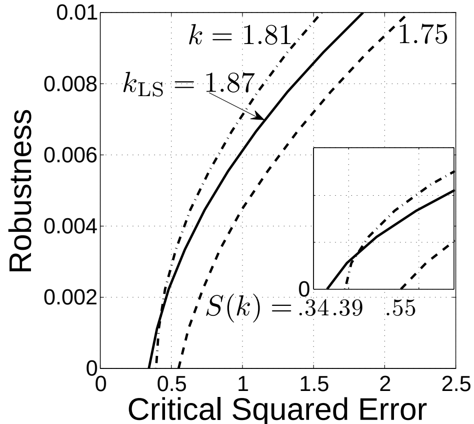


Figure 3: **Robustness curves.** $k_{\rm LS} = 1.8681$ **(solid)**, $k = 1.75$ **(- -)**, $k = 1.81$ **(-.)**.

§ **Trade-off: low MSE (good) $\Longleftrightarrow$ low robustness (bad).**

$$S_{\rm c} < S_{\rm c}' \quad \Longleftrightarrow \quad \widehat{h}(k, S_{\rm c}) < \widehat{h}(k, S_{\rm c}')$$

§ **Zero robustness: Nominal model.**

$$S_{\rm c} = S(k) \quad \Longrightarrow \quad \widehat{h}(k, S_{\rm c}) = 0$$

No model performs "as advertised".

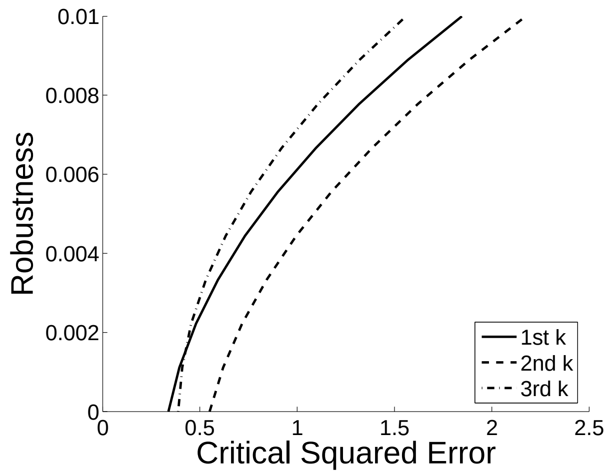


Figure 4: **Robustness curves.** $k_{\rm LS} = 1.8681$ **(solid)**, $k = 1.75$ **(- -)**, $k = 1.81$ **(-.)**.

§ **Trade-off: low MSE $\Longleftrightarrow$ low robustness.**

$$S_{\rm c} < S'_{\rm c} \quad \Longleftrightarrow \quad \widehat{h}(k, S_{\rm c}) < \widehat{h}(k, S'_{\rm c})$$

§ **Zero robustness: Nominal model.**

$$S_{\rm c} = S(k) \quad \Longrightarrow \quad \widehat{h}(k, S_{\rm c}) = 0$$

No model performs "as advertised".

§ **Preference reversal: Crossing robustness curves.**

Sub-optimal more robust than "optimal" model.

1.2 *An Interpretation: Foci of Uncertainty*

[0] lectures talks lib sys_id5foci-unc01.tex 27.1.2016

An interpretation: Foci of uncertainty

§ **Least-squares estimation focusses on managing statistical error in data:**

$$\textbf{Minimize:} \quad \frac{1}{n}\sum_{i=1}^{n}(f_i - kx_i)^2$$

An interpretation: Foci of uncertainty

§ **Least-squares estimation focusses on managing**

statistical error in data:

$$\textbf{Minimize:} \quad \frac{1}{n}\sum_{i=1}^{n}(f_i - kx_i)^2$$

§ **Info-gap estimation focusses on managing**

- **statistical error in data:**

$$\textbf{Satisfice:} \quad \frac{1}{n}\sum_{i=1}^{n}(f_i - kx_i)^2$$

- **epistemic error in model:**

$$\textbf{Maximize:} \quad \widehat{h}(k, S_{\rm c}).$$

§ Info-gap robust satisficing:

- **Satisfice the fidelity to data:**
  - **Ameliorate statistical error.**
  - **Anchor in historical data.**
-

§ Info-gap robust satisficing:

- **Satisfice the fidelity to data:**
  - Ameliorate statistical error.
  - Anchor in historical data.
- **Robustify against model error:**
  - Ameliorate changing model structure.
  - Reduce overfitting to data.
  - Generalize to new realizations.
-

§ Info-gap robust satisficing:

- **Satisfice the fidelity to data:**
  - **Ameliorate statistical error.**
  - **Anchor in historical data.**
- **Robustify against model error:**
  - **Ameliorate changing model structure.**
  - **Reduce overfitting to data.**
  - **Generalize to new realizations.**
- **Prioritize model updates.**

1.3 *Robustness and Opportuneness*

[0] lectures talks lib sys_id5rbs-ops01.tex 27.1.2016

§ Robustness of model kx:

how wrong can kx be

without exceeding acceptable fidelity?

$$\widehat{h}(k, S_{\rm c}) = \max \left\{ h : \ \left(\max_{k_3 \in \mathcal{U}(h)} S(k, k_3) \right) \le S_{\rm c} \right\}$$

§

§ Robustness of model kx:

how wrong can kx be

without exceeding acceptable fidelity?

$$\widehat{h}(k, S_\mathrm{c}) = \max \left\{ h : \ \left(\max_{k_3 \in \mathcal{U}(h)} S(k, k_3) \right) \le S_\mathrm{c} \right\}$$

§ Opportuneness of model kx:

how wrong must kx be to enable windfall fidelity?

$$S_\mathrm{w} \ll S_\mathrm{c}$$

$$\widehat{\beta}(k, S_\mathrm{w}) = \min \left\{ h : \ \left(\min_{k_3 \in \mathcal{U}(h)} S(k, k_3) \right) \le S_\mathrm{w} \right\} \quad \textbf{(Ops.)}$$

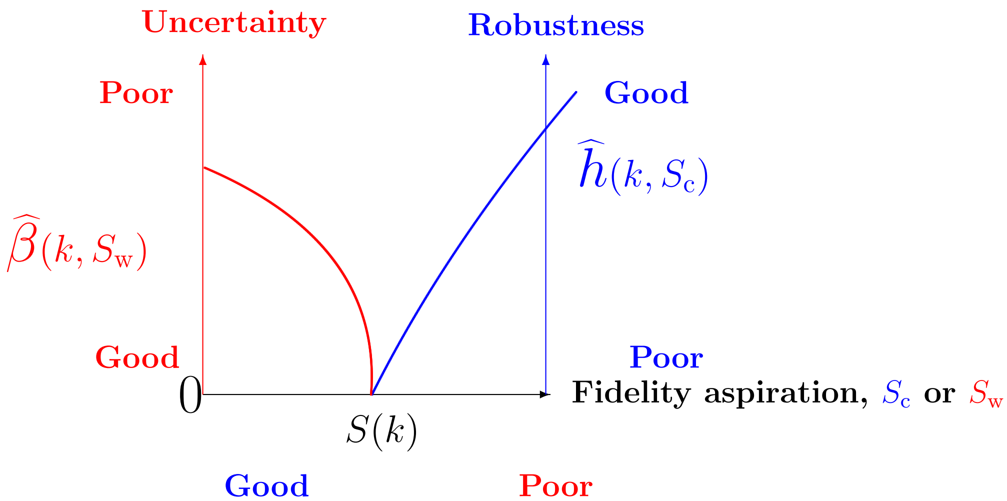
$$\widehat{h}(k, S_\mathrm{c}) = \max \left\{ h : \ \left(\max_{k_3 \in \mathcal{U}(h)} S(k, k_3) \right) \le S_\mathrm{c} \right\} \quad \textbf{(Rbs.)}$$

§ Preferences:

- **Robustness function:**
  - **Immunity to failure.**
  - **Satisficing at critical fidelity.**
  - **Bigger is better**
-

§ Preferences:

- **Robustness function:**
  - **Immunity to failure.**
  - **Satisficing at critical fidelity.**
  - **Bigger is better**
- **Opportuneness function:**
  - **Immunity to windfall.**
  - **Windfalling at wildest-dream fidelity.**
  - **Big is bad.**



§ Trade-offs:

- **Robustness vs. critical fidelity.**
- **Opportuneness vs. windfall fidelity.**

§ **Sympathetic immunities:**

change in model, k, which improves $\widehat{h}$

also improves $\widehat{\beta}$.

$$\frac{\partial \widehat{h}}{\partial k} \frac{\partial \widehat{\beta}}{\partial k} < 0$$

§

§ Sympathetic immunities:

change in model, k, which improves $\widehat{h}$

also improves $\widehat{\beta}$.

$$\frac{\partial\widehat{h}}{\partial k}\frac{\partial\widehat{\beta}}{\partial k} < 0$$

§ Antagonistic immunities:

change in model, k, which improves $\widehat{h}$

also degrades $\widehat{\beta}$.

$$\frac{\partial\widehat{h}}{\partial k}\frac{\partial\widehat{\beta}}{\partial k} > 0$$

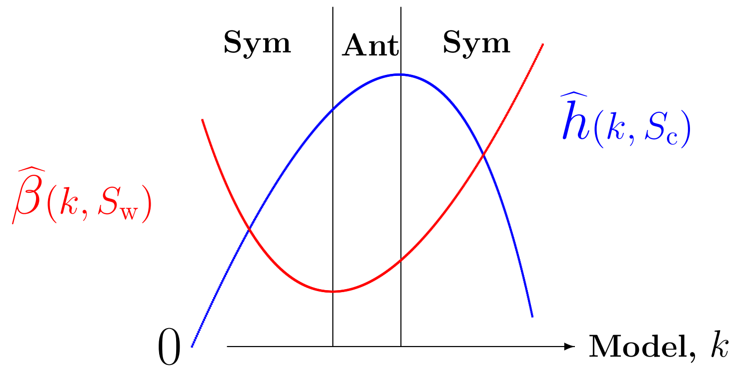


Figure 5: **Schematic immunity curves**

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 forecast02intro.tex 26.7.2015

2.1 *1-D Example*

0 lectures talks lib forecast02-1D.tex 27.1.2016

§ Dynamic state variable, y_t:

Demand, supply, displacement, force, etc.

§

§ **Dynamic state variable, y_t:**

Demand, supply, displacement, force, etc.

§ **Linear 1-dimensional system:**

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

§

§ **Dynamic state variable, y_t:**

Demand, supply, displacement, force, etc.

§ **Linear 1-dimensional system:**

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

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

§

§ **Dynamic state variable, y_t:**

Demand, supply, displacement, force, etc.

§ **Linear 1-dimensional system:**

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

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

§ **Contextual understanding: λ could drift upwards.**

§

§ **Dynamic state variable, y_t:**

Demand, supply, displacement, force, etc.

§ **Linear 1-dimensional system:**

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

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

§ **Contextual understanding:** λ **could drift upwards.**

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

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

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

§

§ **Dynamic state variable, y_t:**

Demand, supply, displacement, force, etc.

§ **Linear 1-dimensional 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.**

§ **Goal: Predict future y_t.**

§

§ **Dynamic state variable, y_t:**

Demand, supply, displacement, force, etc.

§ **Linear 1-dimensional 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.**

§ **Goal:** **Predict future** y_t.

§ **Problem:** **What value of** λ **to use?**

§ Slope-adjusted (erroneous) forecaster:

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

§

§ **Slope-adjusted (erroneous) forecaster:**

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

§ **Contrast with historically estimated model:**

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

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

§

§ **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 $\ell \geq \widetilde{\lambda}$**?**

§ **Robust satisficing:**

Satisfice the forecast error:

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

. . .

§ **Slope-adjusted (erroneous) forecaster:**

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

§ **Contrast with historically estimated model:**

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

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

§ **Robust satisficing:**

Satisfice the forecast error:

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

Maximize robustness to future surprise.

§ Robustness of forecast ℓ:

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

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

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

§

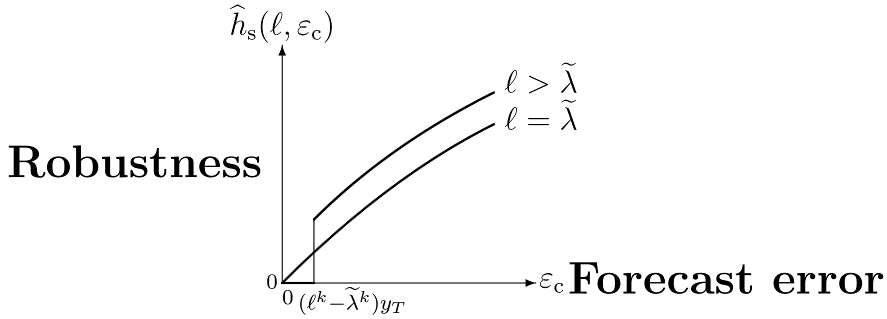
§ Robustness of forecast ℓ:

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

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

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

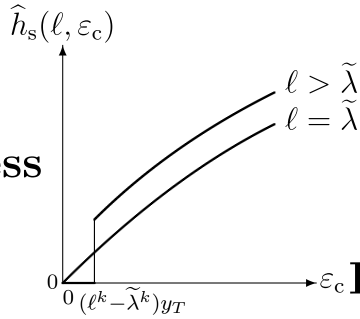
§ Preference: $\ell \succ \ell'$ **if** $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell', \varepsilon_{\rm c})$



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

§

Robustness



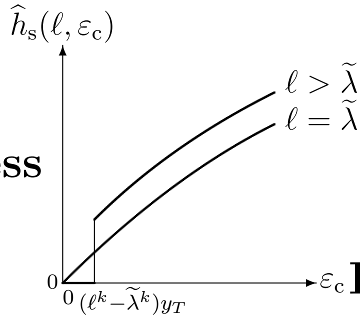
Forecast error

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

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

§

Robustness



Forecast error

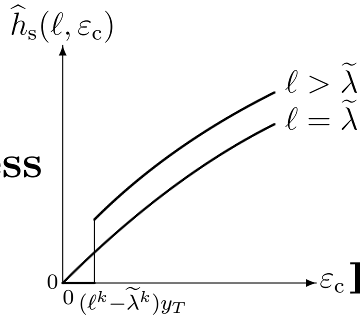
§ **Trade off:** **robustness vs. forecast error.**

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

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

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

§

Robustness 

Forecast error

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

§ **Zeroing:** Estimated 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.

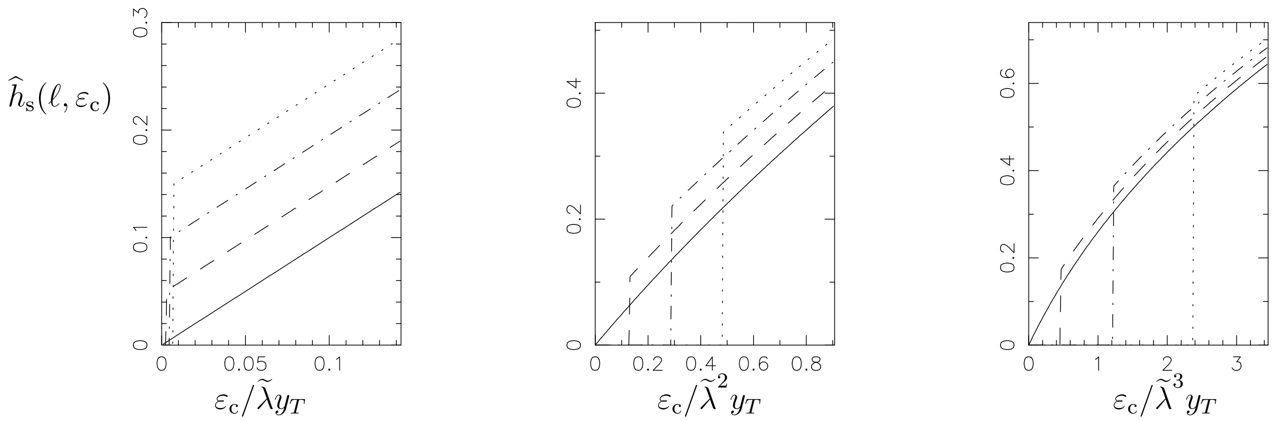


Figure 6: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

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

2.2 *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 distribution.**

§

§ **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 distribution.**

§ **Forecast success set:**

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

§

§ **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 distribution.**

§ **Forecast success set:**

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

§ **Probability of success:**

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

§ Goal:

Choose ℓ to maximize success probability.

§

§ Goal:

Choose ℓ to maximize success distribution.

§ Problem:

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

§

§ Goal:

Choose ℓ to maximize success distribution.

§ 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 distribution.**

§ Theorem.

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

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

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

Then:

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

§ Robustness is proxy for success-probability.

Summary:

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

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

It is strategically advantageous.

. . .

Any Questions?