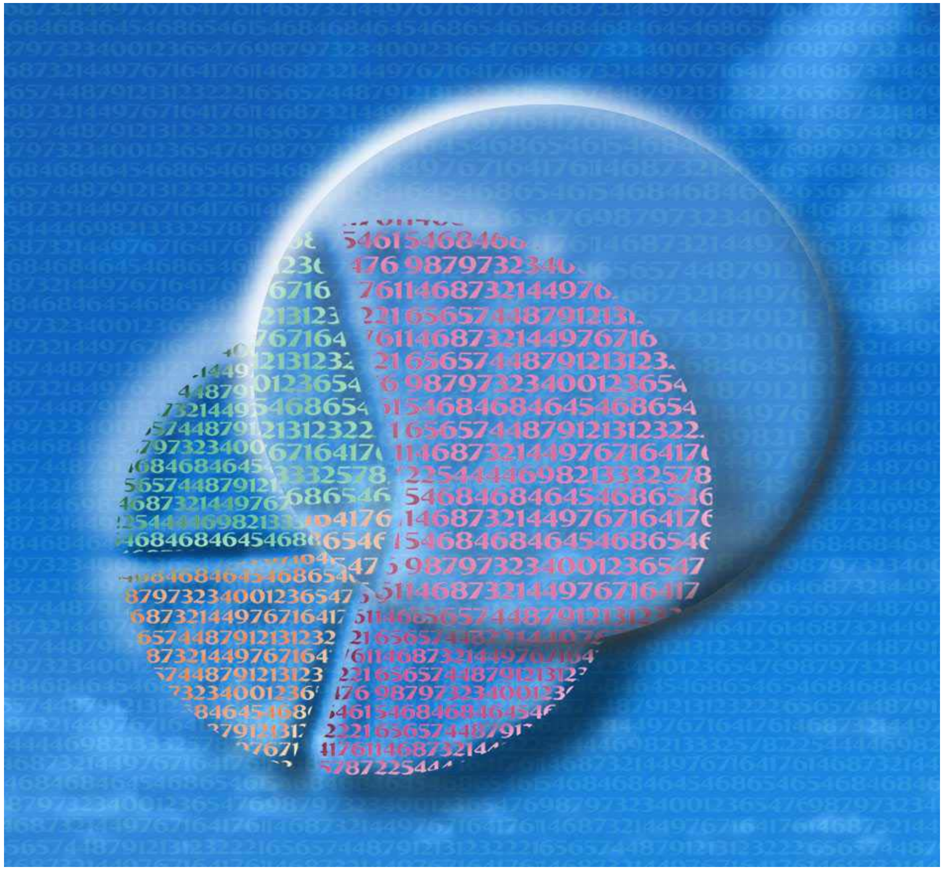
Sensitivity analysis

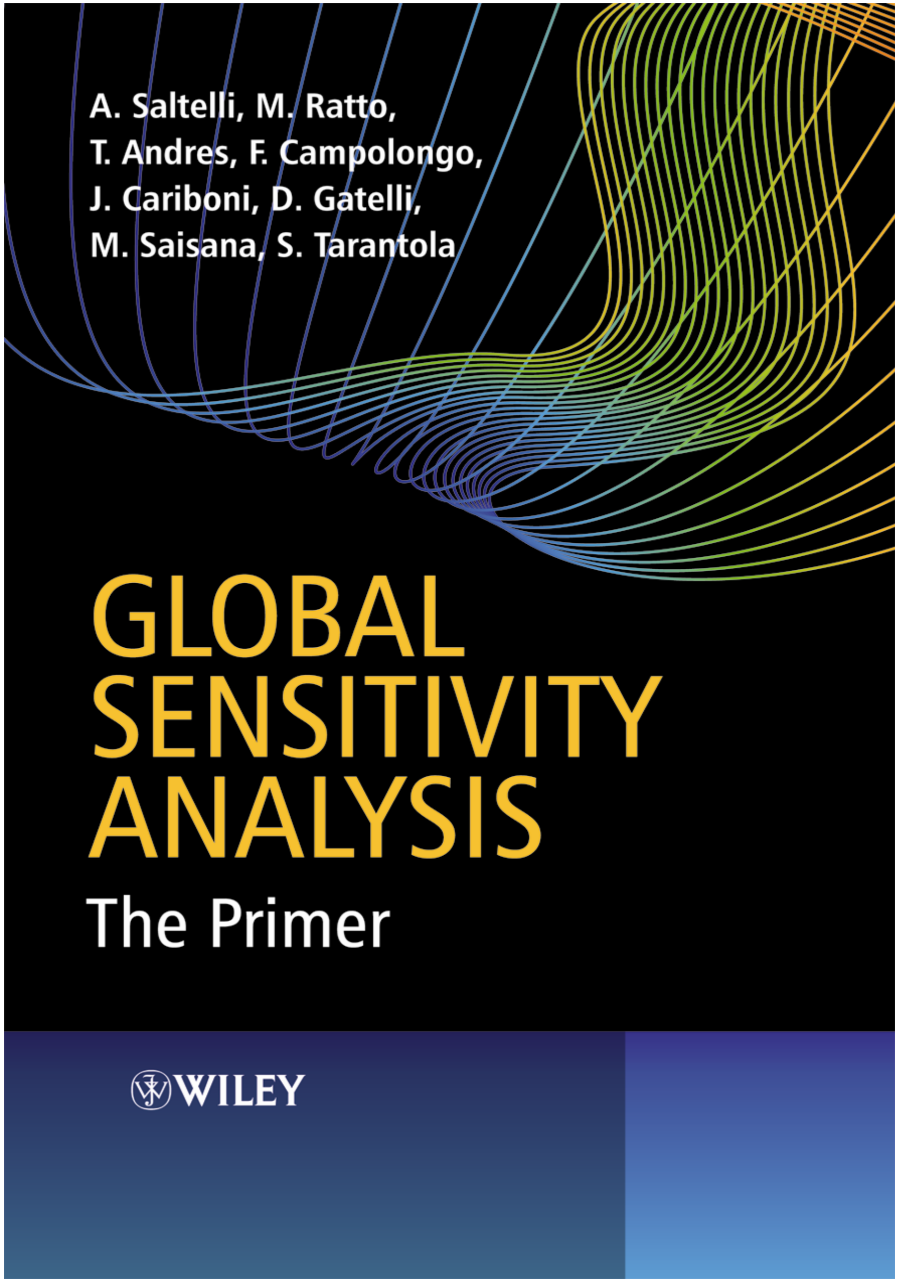


Andrea Saltelli,
European Commission, Joint Research Centre, Ispra (I)

Ecole-chercheurs MEXICO

sur l'analyse de sensibilité et l'exploration de modèles appliquées aux sciences de la nature et de l'environnement

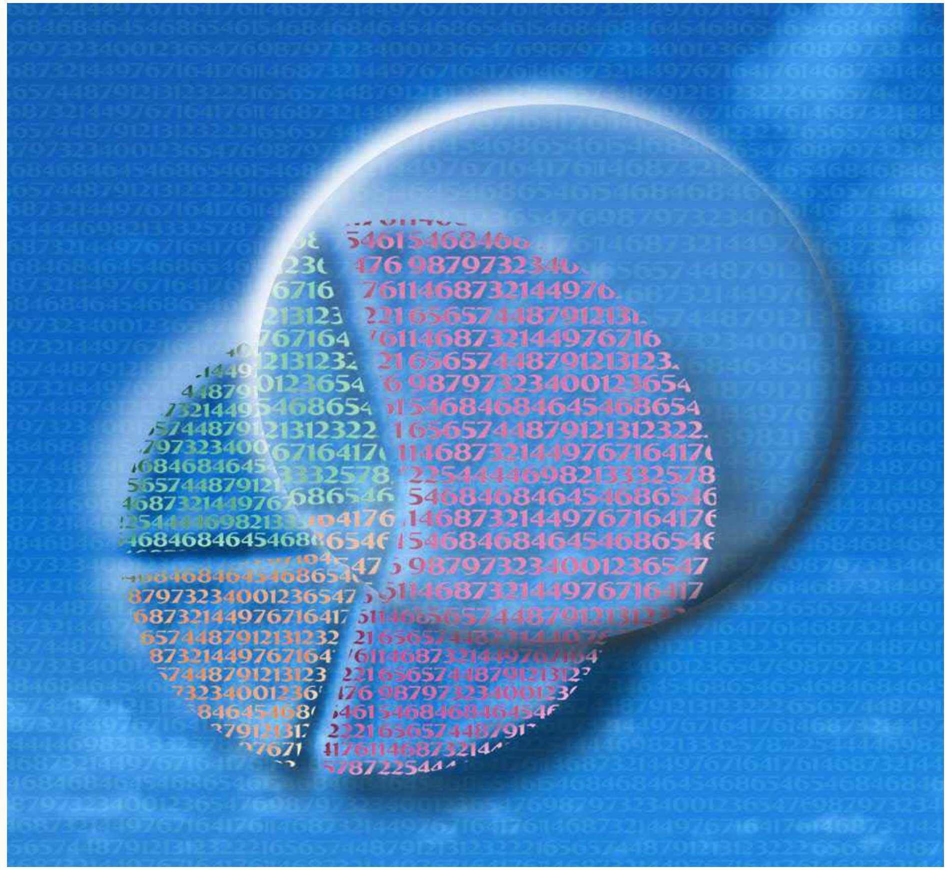
Giens, 11-14 mai 2009,



Partly based on

Global sensitivity analysis.
The Primer

1. Context: the critique of models and what SA has to do with it



NEWSFOCUS

Dueling Visions For a Hungry World

Sparks began to fly when scientists and activists against genetically modified crops came together to assess agricultural knowledge and the role of biotech in development

When economist Carl Pray heard about plans for the first international assessment of agricultural research, a gold standard sprang to mind: the Intergovernmental Panel on Climate Change (IPCC). But things didn't turn out the way he expected.

IPCC has been pivotal in proving that climate change is real and linking it to human activities. As an agricultural economist at Rutgers University who has worked in many poor countries, Pray is convinced that agricultural research—and genetic modification in particular—is key to fighting pervasive

mentally, socially and economically sustainable development through the generation, access to, and use of agricultural knowledge, science and technology?" Critics say this broad mandate made conflict inevitable and stunted the assessment's analytical rigor.

On several key issues, consensus proved elusive. Industry scientists and some academics—mainly agricultural economists and plant biologists—believe the assessment was "hijacked" by participants who oppose genetically modified (GM) crops and other common tools of industrial agriculture. Tensions peaked

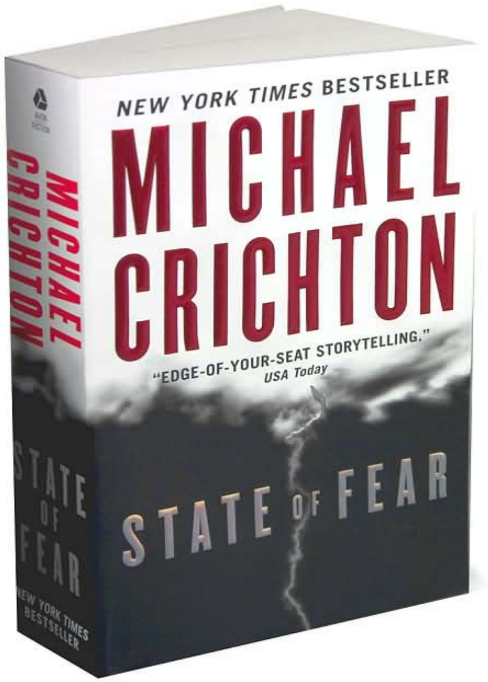
the outcome. They note that th experience of small-scale farmers women, have finally been broug by the assessment. "It really deal ... of power, influence, and benefits," says Marcia Ishii-Eiteman of the Pesticide Action Network North America in San Francisco, California. Toby Kiers, who studies sustainable agriculture at Vrije University in Amsterdam, the Netherlands, agrees. "For technology to be most effective, farmers must be at the center, influencing how it is developed, delivered, and managed," she says.

The IFPRI had raised about $460,000 for the modeling, which would have provided insights to help policymakers […]

[…] But Greenpeace's Haerlin and others objected that the models were not "transparent".

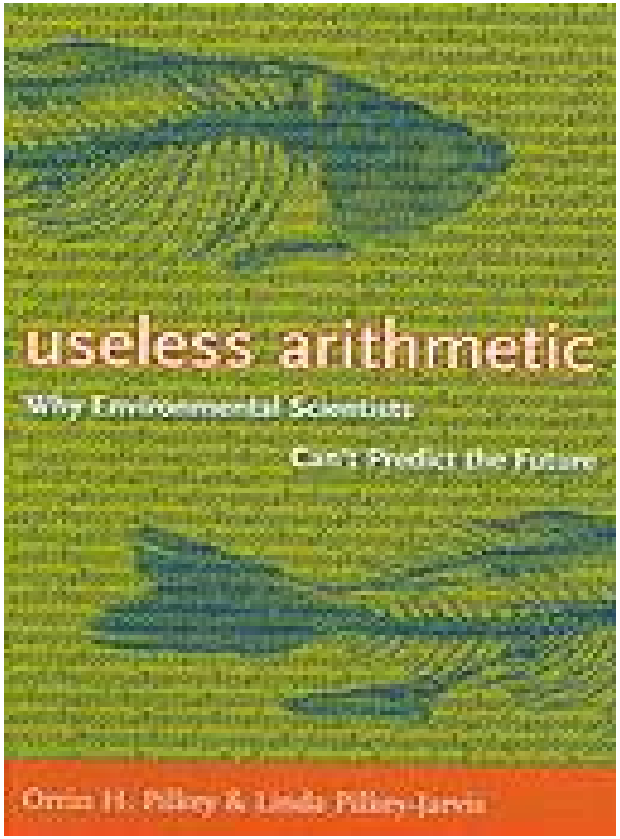
Source: Dueling visions for an hungry world, Erik Stokstad, 14 MARCH 2008, **319** SCIENCE

14



"They talk as if simulation were real-world data. They 're not. That 's a problem that has to be fixed. I favor a stamp: WARNING: COMPUTER SIMULATION – MAY BE ERRONEOUS and UNVERIFIABLE. Like on cigarettes [⋯]"

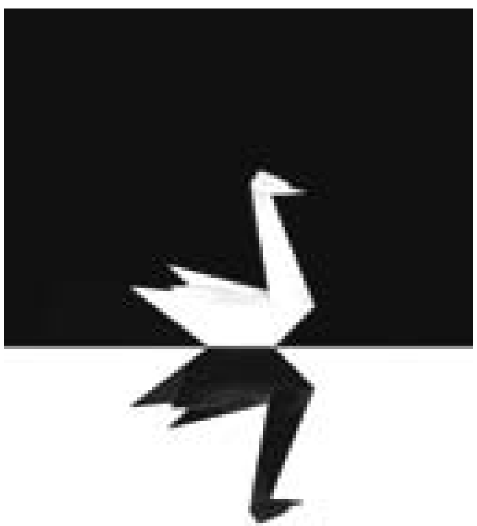
Op. Cit. *p. 556* .



Useless Arithmetic: Why Environmental Scientists Can't Predict the Future
by Orrin H. Pilkey and Linda Pilkey-Jarvis

‘Quantitative mathematical models used by policy makers and government administrators to form environmental policies are seriously flawed’

We just can't predict, and we are victims delusion of uncertainty, of the ludic fallacy, and so on. Modelling is just another attempt to 'Platonify' reality…



Nassim Nichola Taleb, The Black Swan, Penguin, London 2007

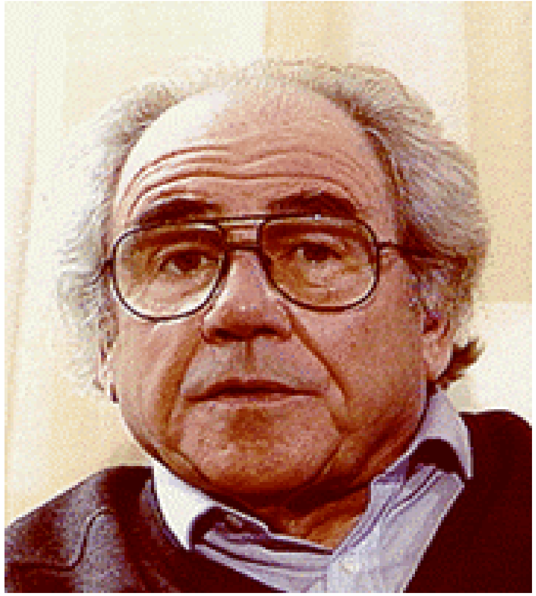
On present day uncertainty assessment practice: "the uncertainties which are more carefully scrutinised are usually those which are the least relevant" (*lampposting*).

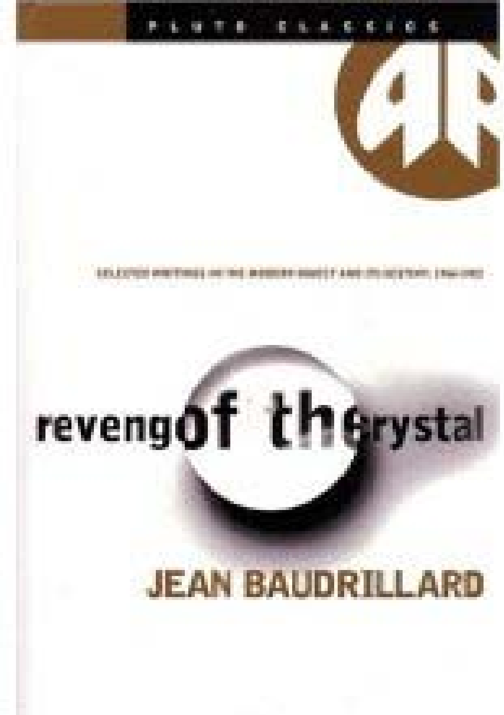
(Taleb's 'delusion of uncertainty') .

Jeroen van der Sluijs, Professor at Utrecht

www.nusap.net

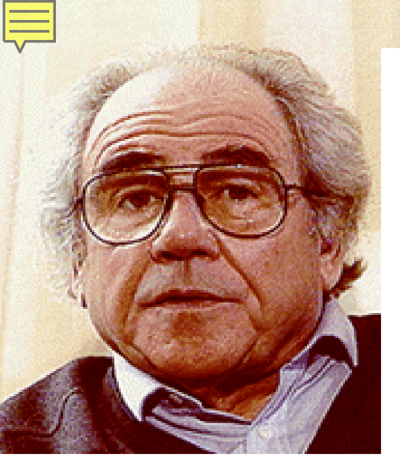
The (post modern) critique of models. The post-modern French thinker Jean Baudrillard (1990) presents 'simulation models' as unverifiable artefact which, used in the context of mass communication, produce a fictitious hyper reality that annihilates truth.





An immense process of simulation has taken place throughout all of everyday life , in the image of those 'simulation models' on which operational and computer sciences are based. One 'fabricates' a model by combining characteristics or elements of the real; and, by making them 'act out' a future event, structure or situation, tactical conclusions can be drawn and applied to reality. It can be used as an analytic tool under controlled scientific conditions. In mass communication, this procedure assumes the force of reality, abolishing and volatilizing the latter in favour of that neo-reality of a model materialized by the medium itself.

Jean Baudrillard, Revenge of the Crystal, PLUTO Press 1999, p. 92



Just philosophy? Maybe not:

A title during the RIVM media scandal (1999):

"RIVM over-exact prognoses based on virtual reality of computer models"

From a paper of JvdS

Other Newspaper headlines:

"Environmental institute lies and deceits"

"Fuss in parliament after criticism on environmental numbers"

"The bankruptcy of the environmental numbers"

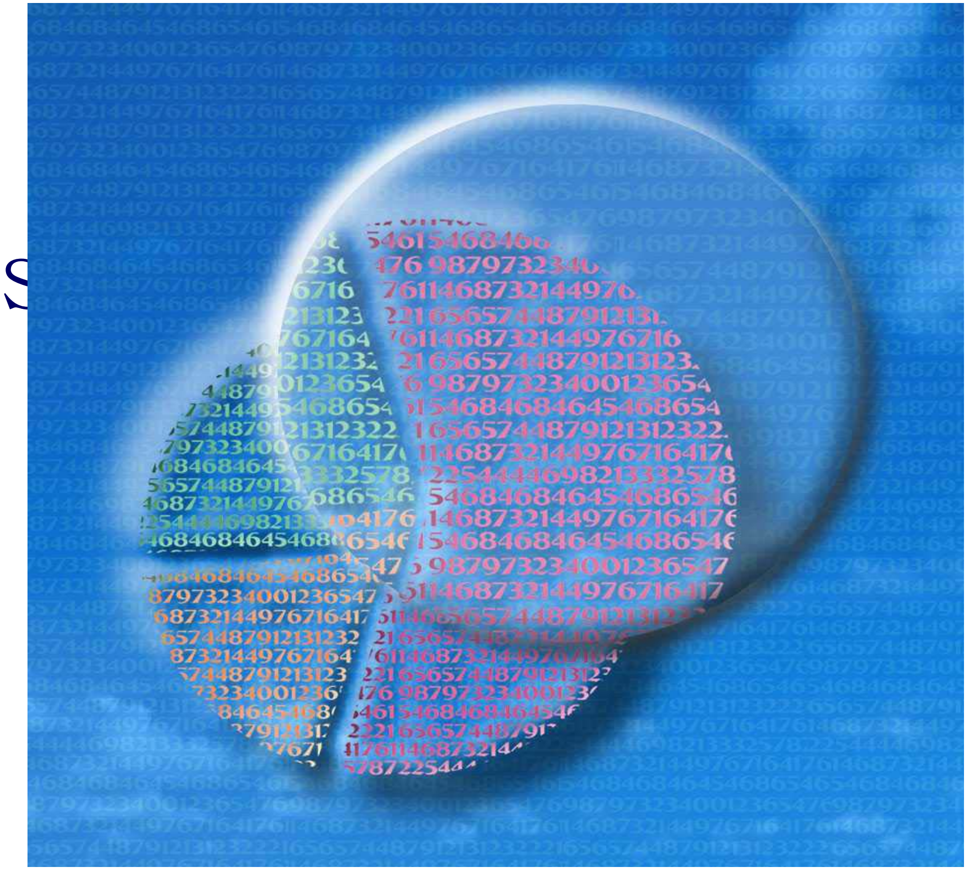
"Society has a right on fair information, RIVM does not provide it"

More on the subject?

Zidek, J. Post-normal Science. J. R. Statist. Soc. A (2006) 169, Part 1, pp. 1-4.

FUNTOWICZ, S. and J. RAVETZ. 1993. "Science for the Post-Normal Age." Futures 25, (7) p. 735-755.

2. Prescriptions



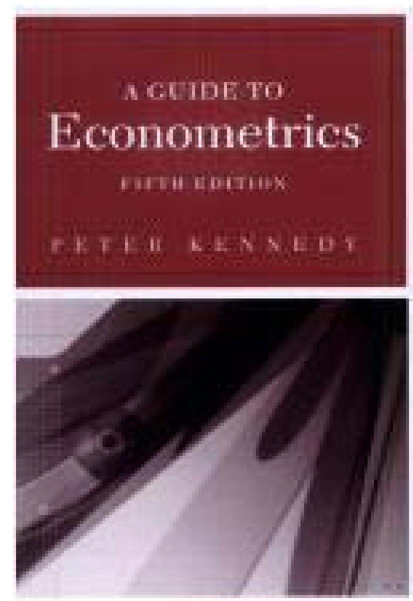
The critique of models <-> Uncertainty

Peter Kennedy, A Guide to Econometrics.

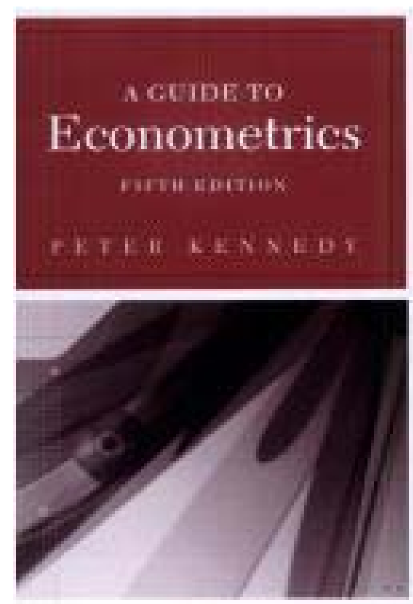
One of the ten commandments of applied econometrics

<<Thou shall confess in the presence of sensitivity.

Corollary: Thou shall anticipate criticism>>



When reporting a sensitivity analysis, researchers should explain fully their specification search so that the readers can judge for themselves how the results may have been affected. This is basically an `honesty is the best policy' approach, advocated by Leamer'.



Sensitivity analysis and the White House

The US the OFFICE OF MANAGEMENT AND BUDGET (OMB) in its controversial 'Proposed Risk Assessment Bulletin' also prescribes how to do a sensitivity analysis.

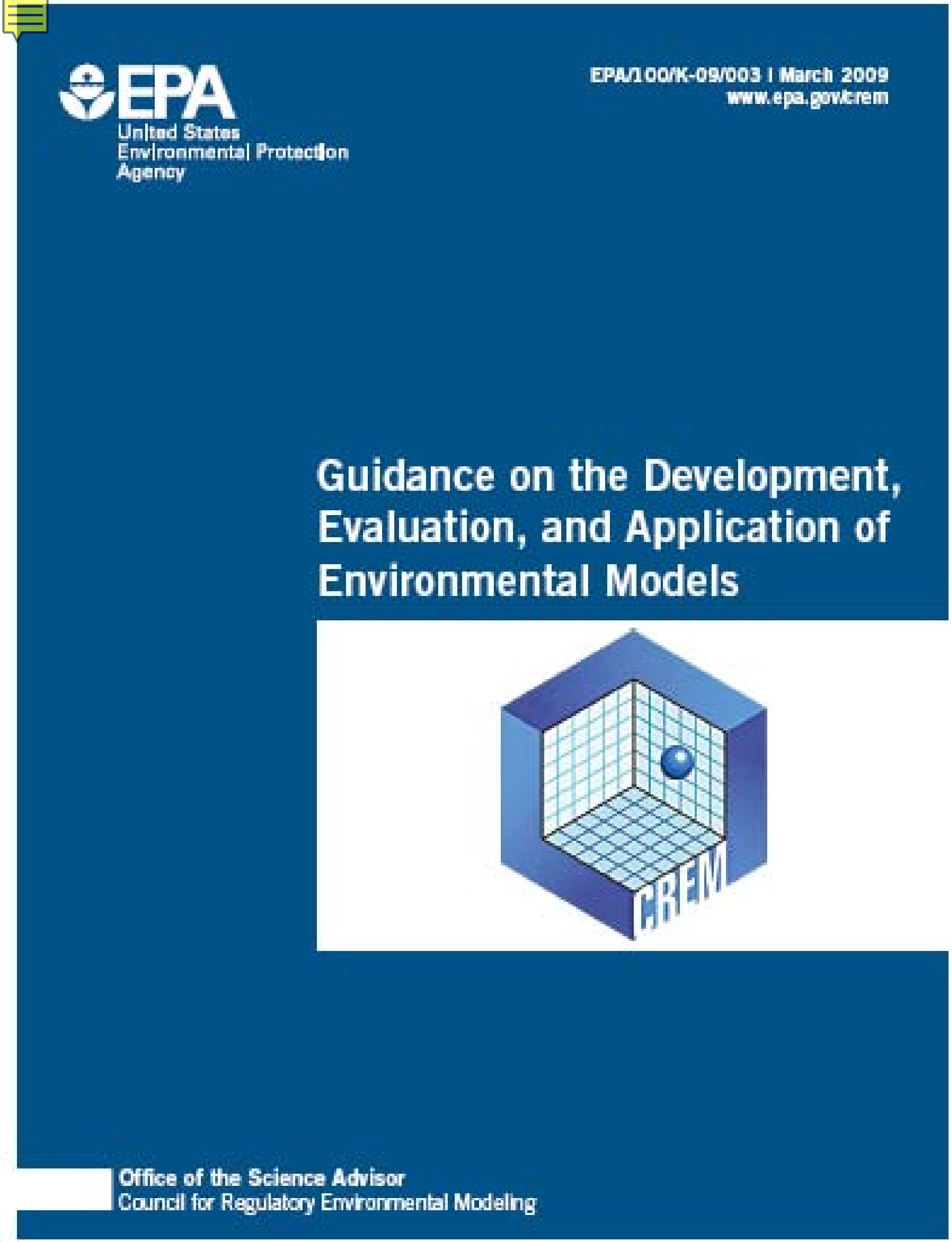
John Graham, one of the authors of the bulletin; for a critical discussion see Colin Macilwain, Safe and sound? *Nature*, 19 July 2006.

Source of the quote: OFFICE OF MANAGEMENT AND BUDGET Proposed Risk Assessment Bulletin (January 9, 2006) http://www.whitehouse.gov/omb/inforeg/

4. Standard for Characterizing Uncertainty

Influential risk assessments should characterize uncertainty with a sensitivity analysis and, where feasible, through use of a numeric distribution.

[…] Sensitivity analysis is particularly useful in pinpointing which assumptions are appropriate candidates for additional data collection to narrow the degree of uncertainty in the results. Sensitivity analysis is generally considered a minimum, necessary component of a quality risk assessment report.



“ [SA] methods should preferably be able to deal with a model regardless of assumptions about a model’s linearity and additivity, consider interaction effects among input uncertainties, […], and evaluate the effect of an input while all other inputs are allowed to ... ll.”

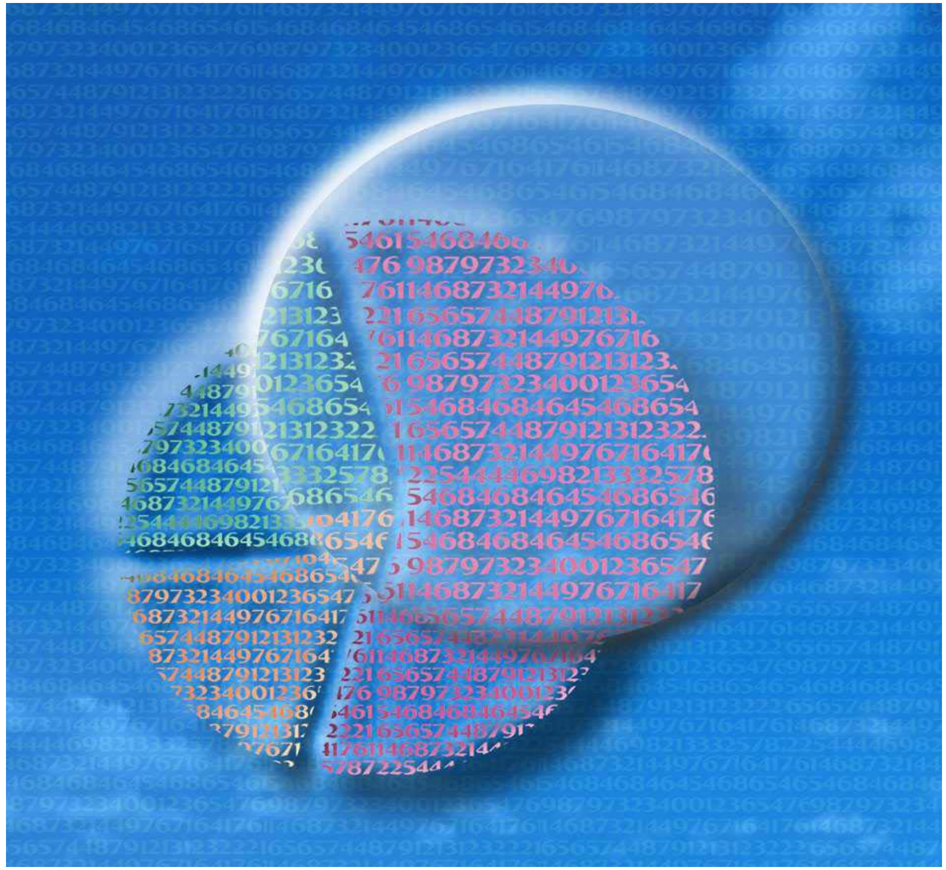
European Commission

IMPACT ASSESSMENT GUIDELINES

15 January 2009
SEC(2009) 92

"… Sensitivity analysis can be used to explore how the impacts of the options you are analysing would change in response to variations in key parameters and how they interact."

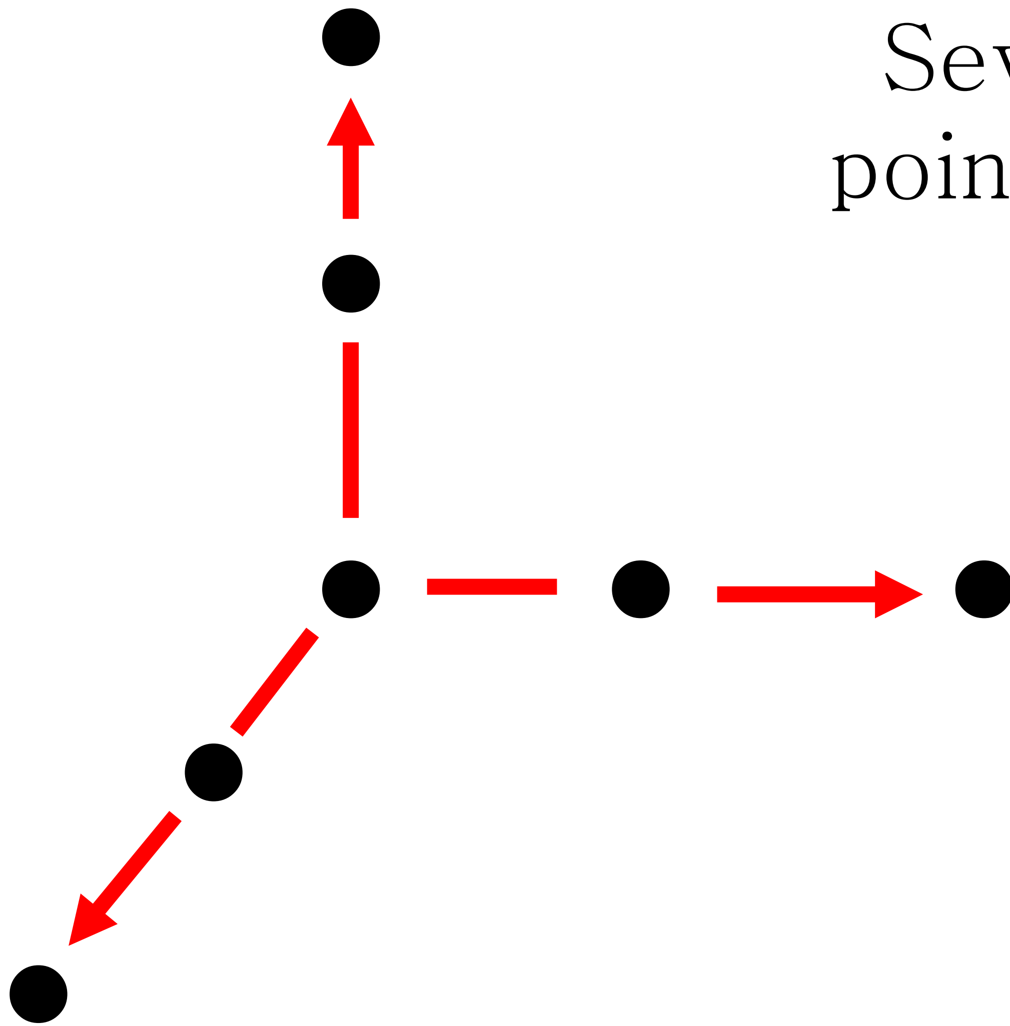
3. Do's and don'ts



Who do these have in common?

J. Campbell, *et al.*, *Science* **322**, 1085 (2008).
R. Bailis, M. Ezzati, D. Kammen, *Science* **308**, 98 (2005).
E. Stites, P. Trampont, Z. Ma, K. Ravichandran, *Science* **318**, 463 (2007).
J. Murphy, et al., *Nature* **430**, 768-772 (2004).
J. Coggan, *et al.*, *Science* **309**, 446 (2005).

OAT



Seven OAT points in a 3D space

Before we go on to discuss OAT a premise:

Before the analysis, we are in a model free setting!

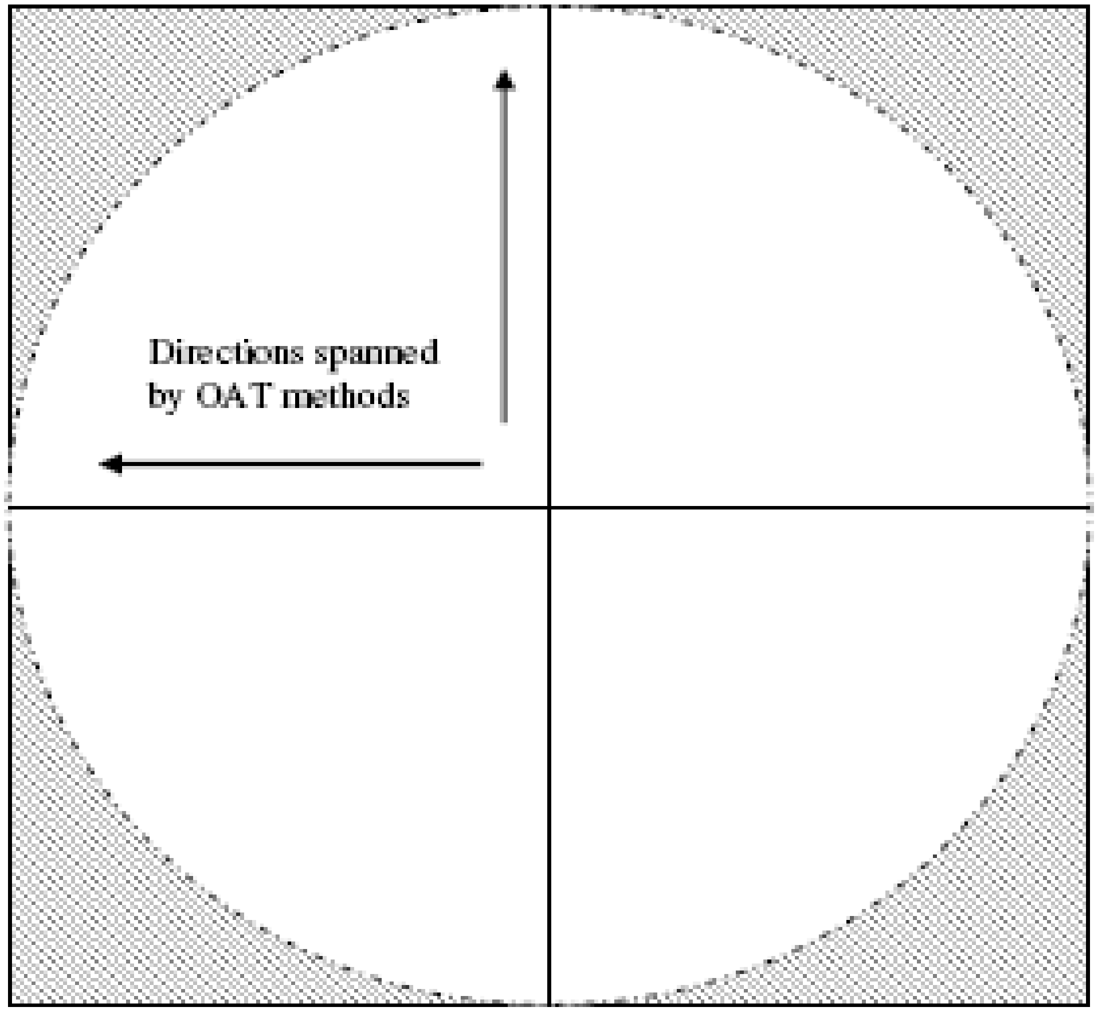
$$\frac{\partial Y_j}{\partial X_i}$$

Otherwise the model could be declared linear or additive (or otherwise well behaved) and one could make it do with derivatives at a single baseline point.

Thus derivates are out, but is OAT OK?

Or how bad is it?

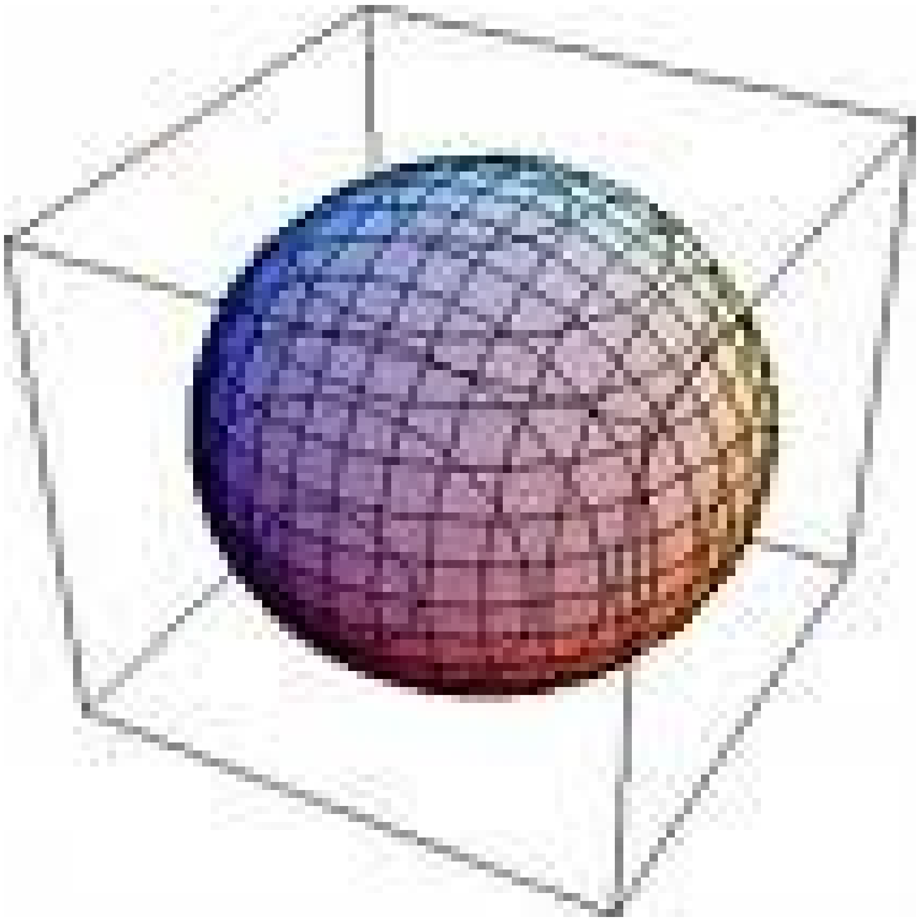
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions



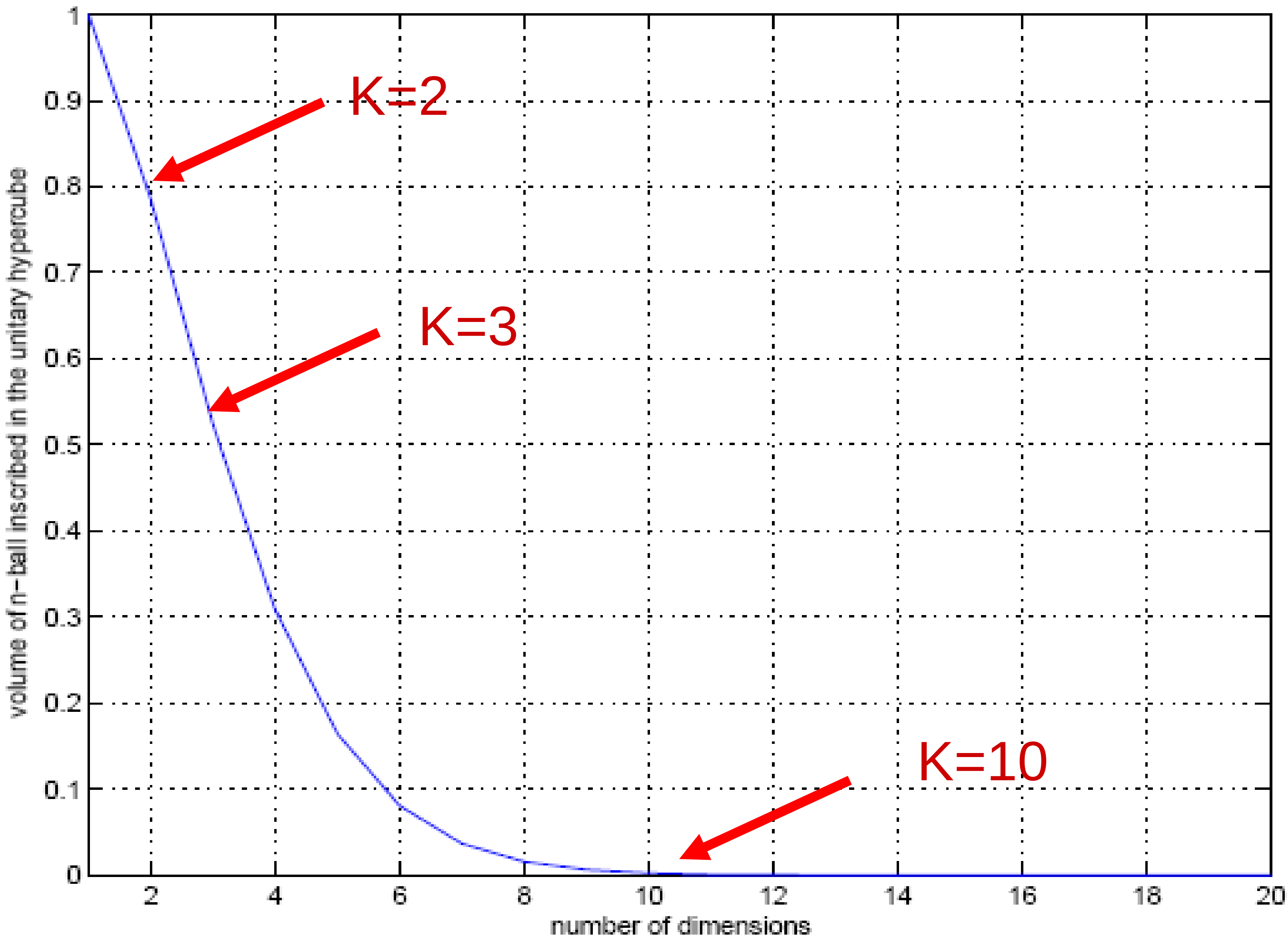
Volume sphere / volume cube =?

~ 1/2

OAT in 10 dimensions

Volume hypersphere / volume ten dimensional hypercube = ~ 0.0025

OAT in k dimensions



Thus OAT is very poor in exploring the space of the factors – it is also non conservative.

Why?

OAT in not roughly right … it is precisely wrong!

Done with the don't**s**.

About the do's:

For the papers just mentioned a better (statistical theory based) alternative is available, be it:

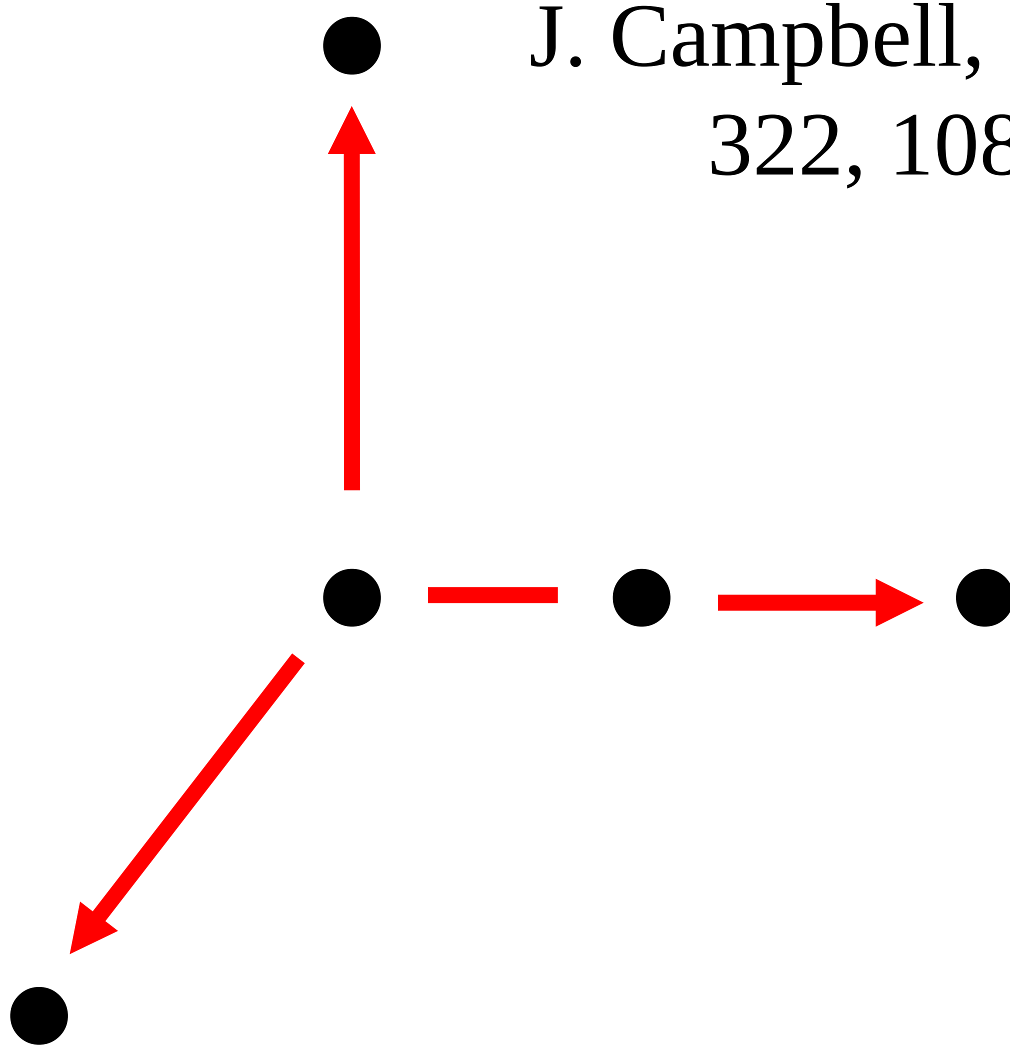
- A two level factorial design,
- A trajectory analysis (a-la-Morris) or
- A linear regression based on a Monte Carlo Sample

<u>Using about the same low number of points.</u>

J. Campbell, *et al.*, *Science* 322, 1085 (2008)

DID:

Three factors are changed OAT-wise for a total of five points as one of the factors is moved twice.

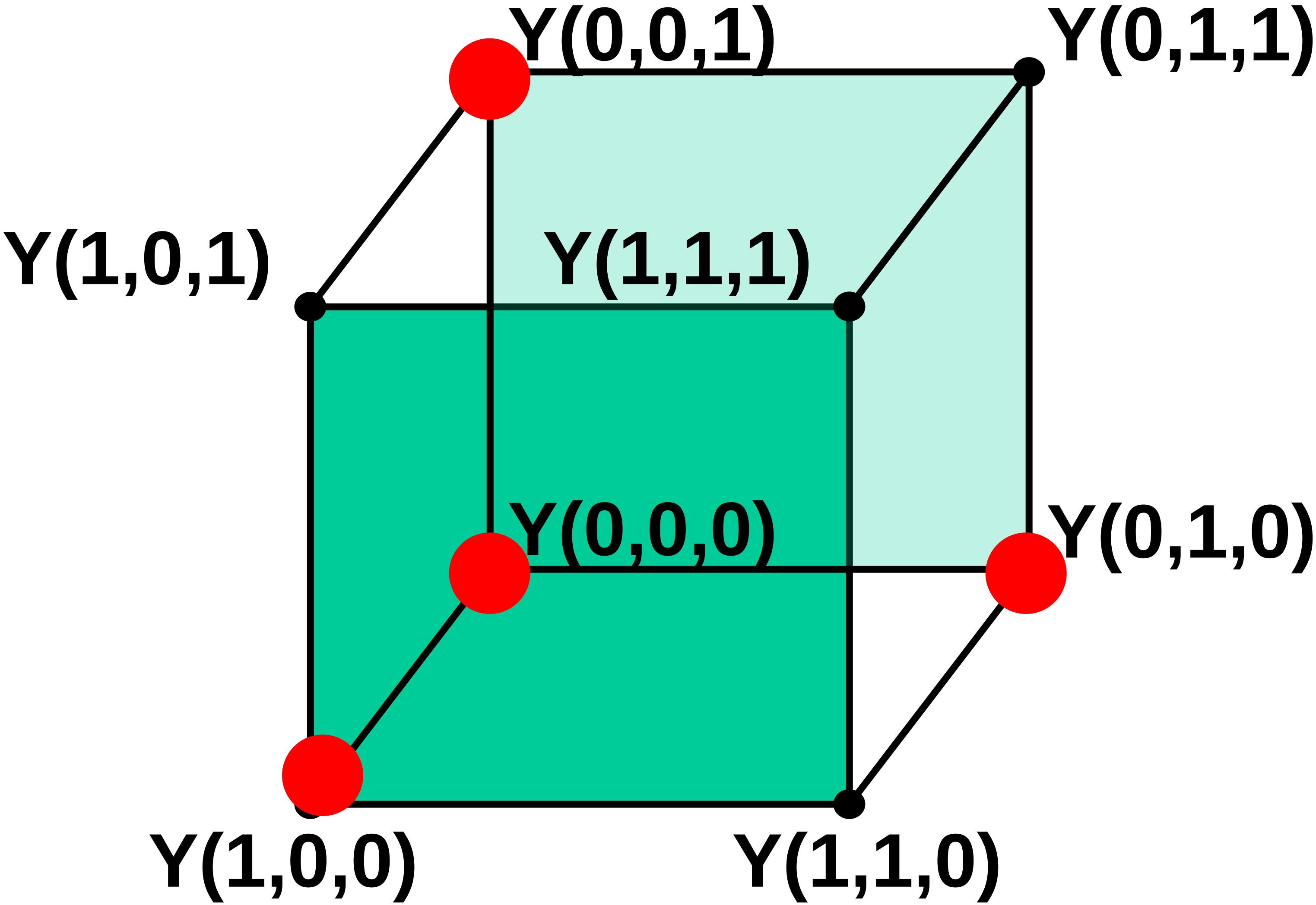


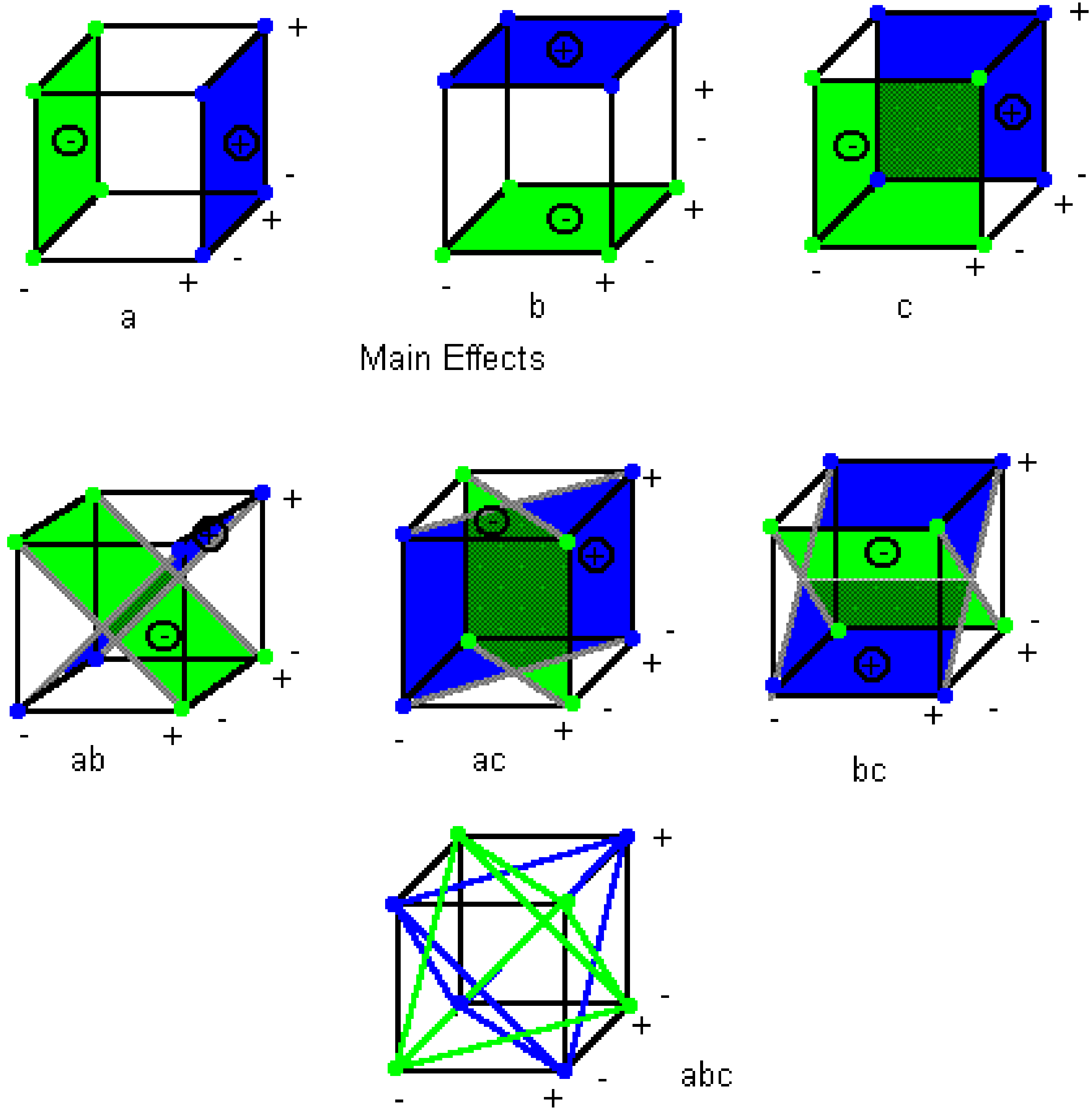
J. Campbell, *et al.*, *Science* 322, 1085 (2008)

J. Campbell, *et al.*, *Science* 322, 1085 (2008)
COULD DO:

With just eight points (instead if five) using a two level design one can obtain an estimate of the main effect plus estimates of the second and third order effects using the same eight points.

Each estimate based on all points!


Y(0,0,1)
Y(0,1,1)
Y(1,0,1)
Y(1,1,1)
Y(0,0,0)
Y(0,1,0)
Y(1,0,0)
Y(1,1,0)


a
b
c
Main Effects
ab
ac
bc
abc

E. Stites, P. Trampont, Z. Ma, K. Ravichandran, *Science* 318, 463 (2007).

DID:

In k=12 dimensions, they use 1 + 2k = 25 points, each factor being moved twice from the baseline.

E. Stites, P. Trampont, Z. Ma, K. Ravichandran, *Science* 318, 463 (2007).

COULD DO:

With two independent trajectories of k + 1 points each (e.g. 26 points instead of 25), one would have two independent(*) estimates of the effect of each factor, and by their difference an idea of the interactions.

(*) OAT ones are on the same axis

J. Coggan, *et al.*, *Science* 309, 446 (2005)

DID:

Roughly forty OAT points are used for four factors.

J. Coggan, *et al.*, *Science* 309, 446 (2005)

COULD DO:

A Monte Carlo design of the same size (using e.g. random points) followed by a regression analysis of the output versus the inputs would give the effect of the factors as well as an idea of the combined effect of model non-linearity and non-additivity.

Another story of SA

Nicholas Stern, London School of Economics

Stern's Review – Technical Annex to postscript

William Nordhaus, University of Yale

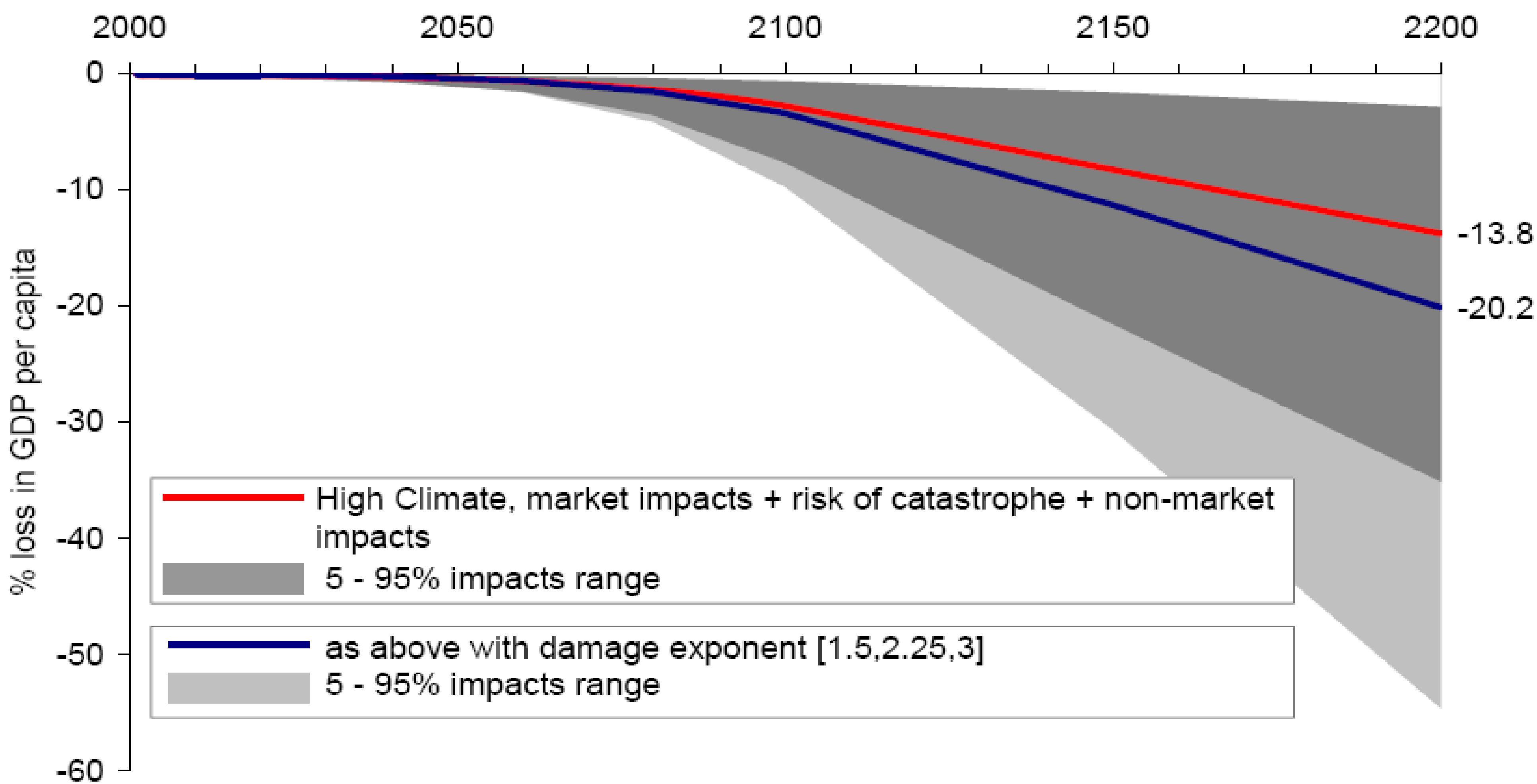
Stern's Review – Technical Annex To postscript
(a sensitivity analysis of a cost benefit analysis)

The Stern – Nordhaus exchange on *SCIENCE*

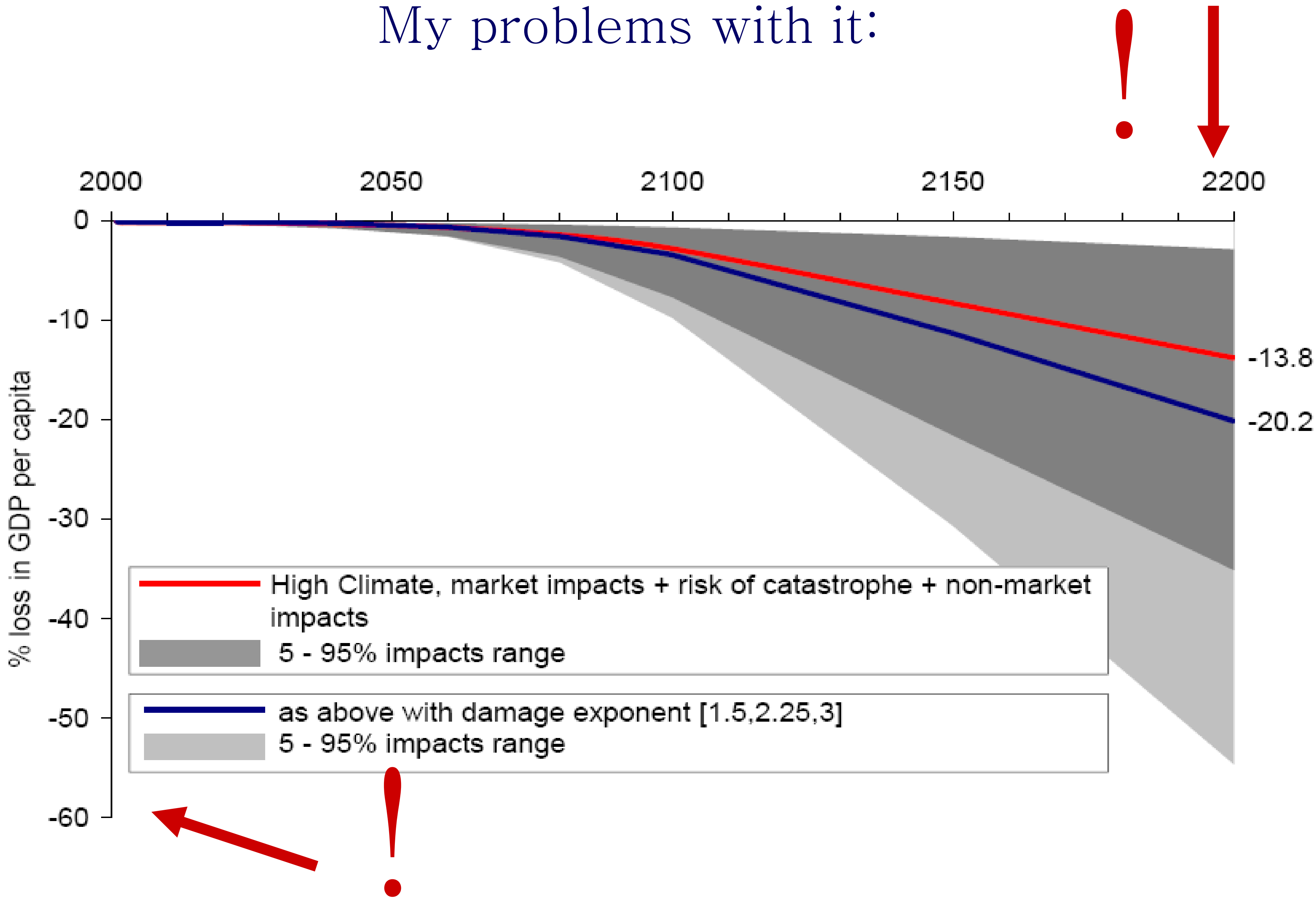
Nordhaus → falsifies Stern based on 'wrong' range of discount rate (~ you GIGOing)

Stern → 'My analysis shows robustness'

From Stern's Review SA



My problems with it:

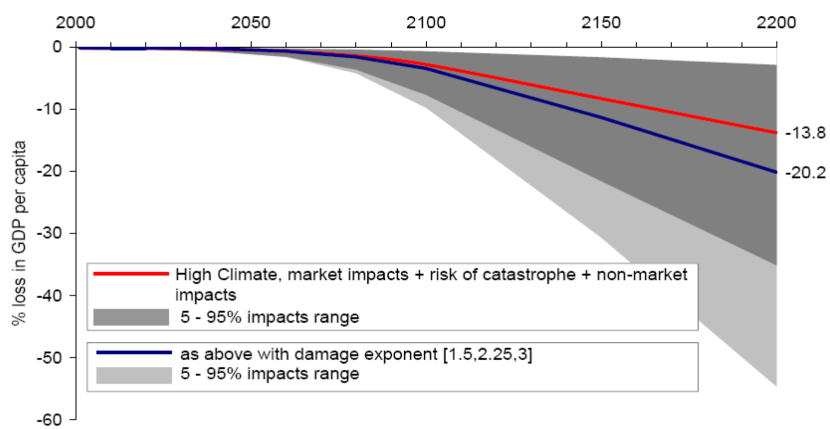


… but foremost he says:

changing assumptions → important effect

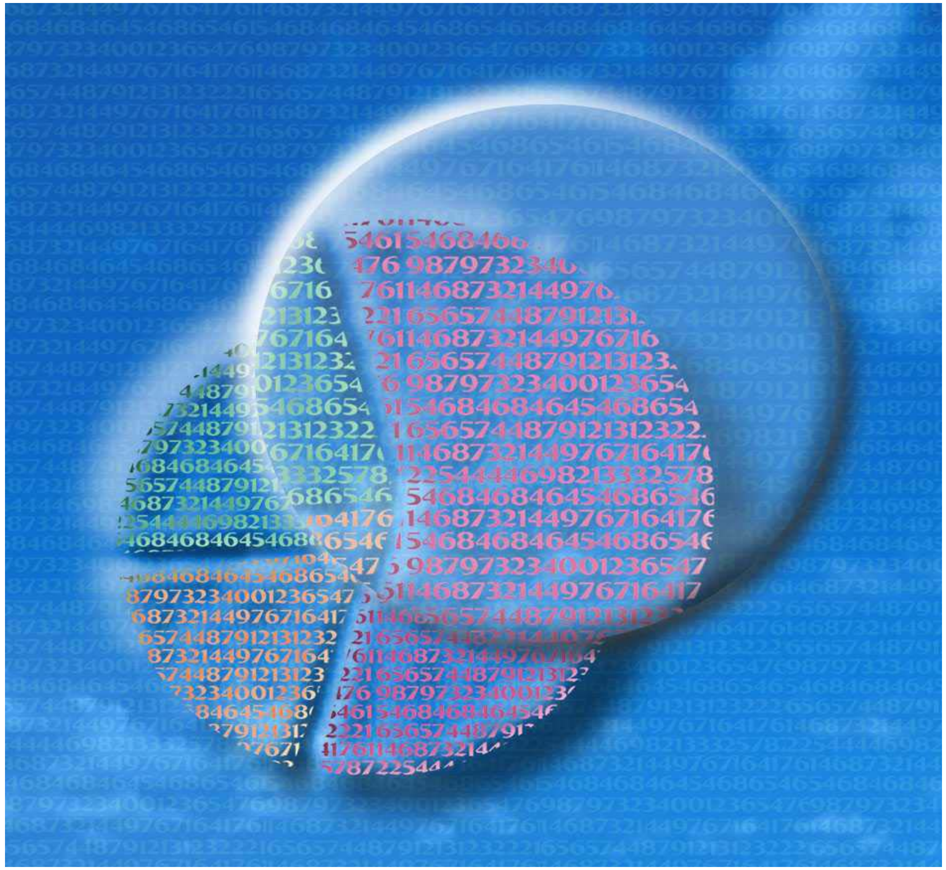
when instead he should admit that:

changing assumptions → all changes a lot

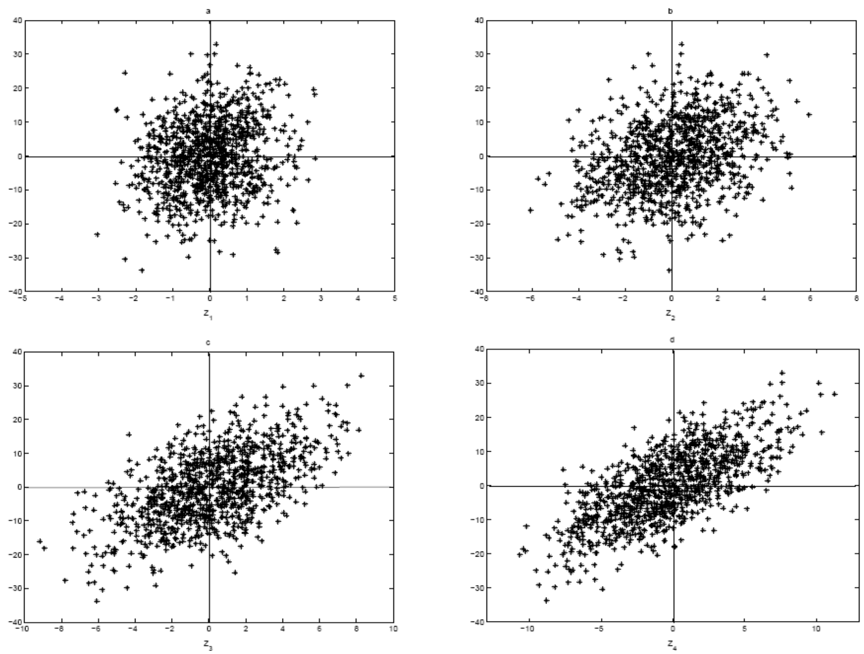


And now …

4. Variance based methods; a best practice?



Mostly based on the work of Ilya M. Sobol' (1990), who extended the work of R.I. Cukier (1973). Further extensions by T. Homma and myself (1996, onward).



Scatterplots' notation:

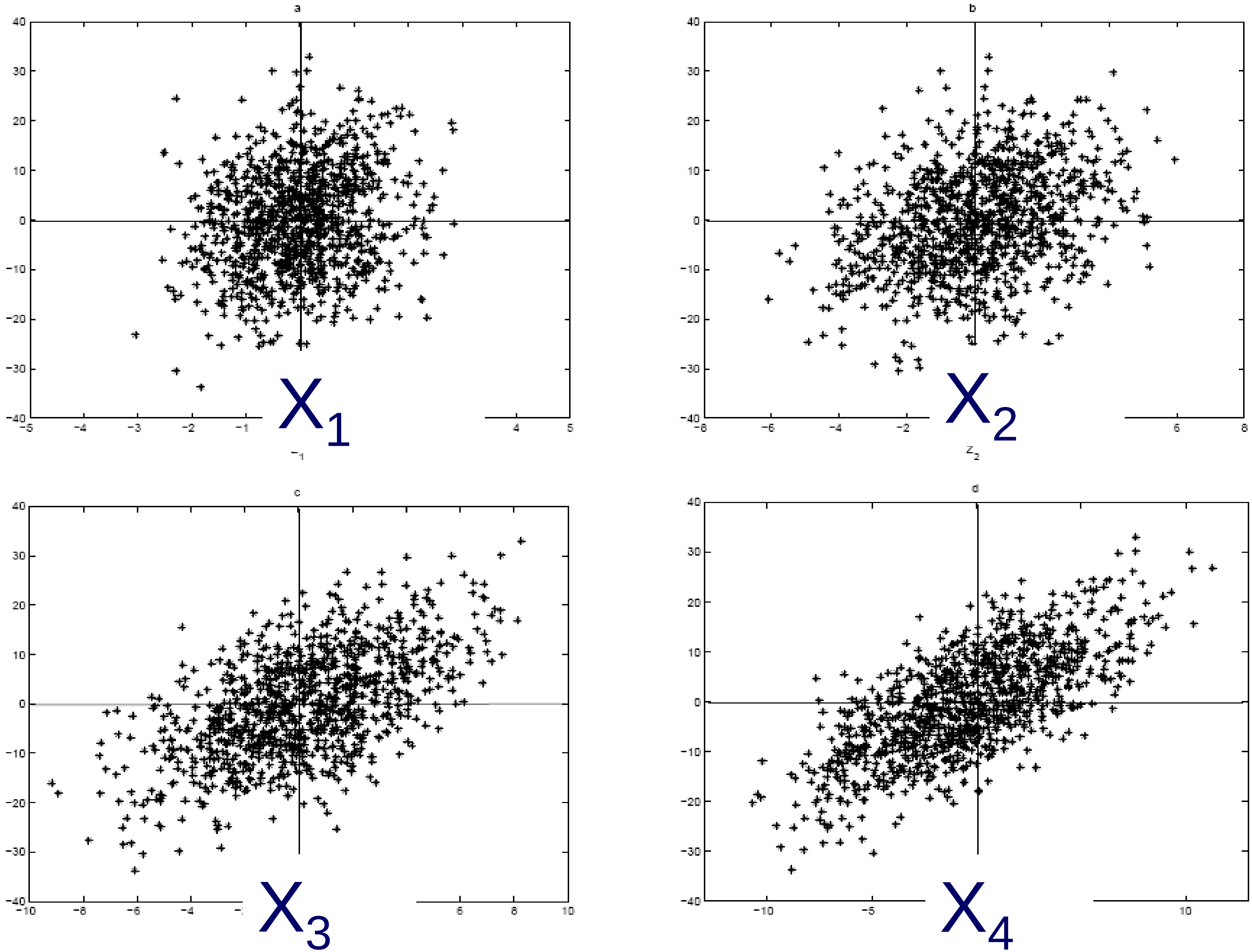
$$Y = f\left(X_1, X_2, \ldots X_k\right)$$

$$f_0 = E\left(Y\right)$$

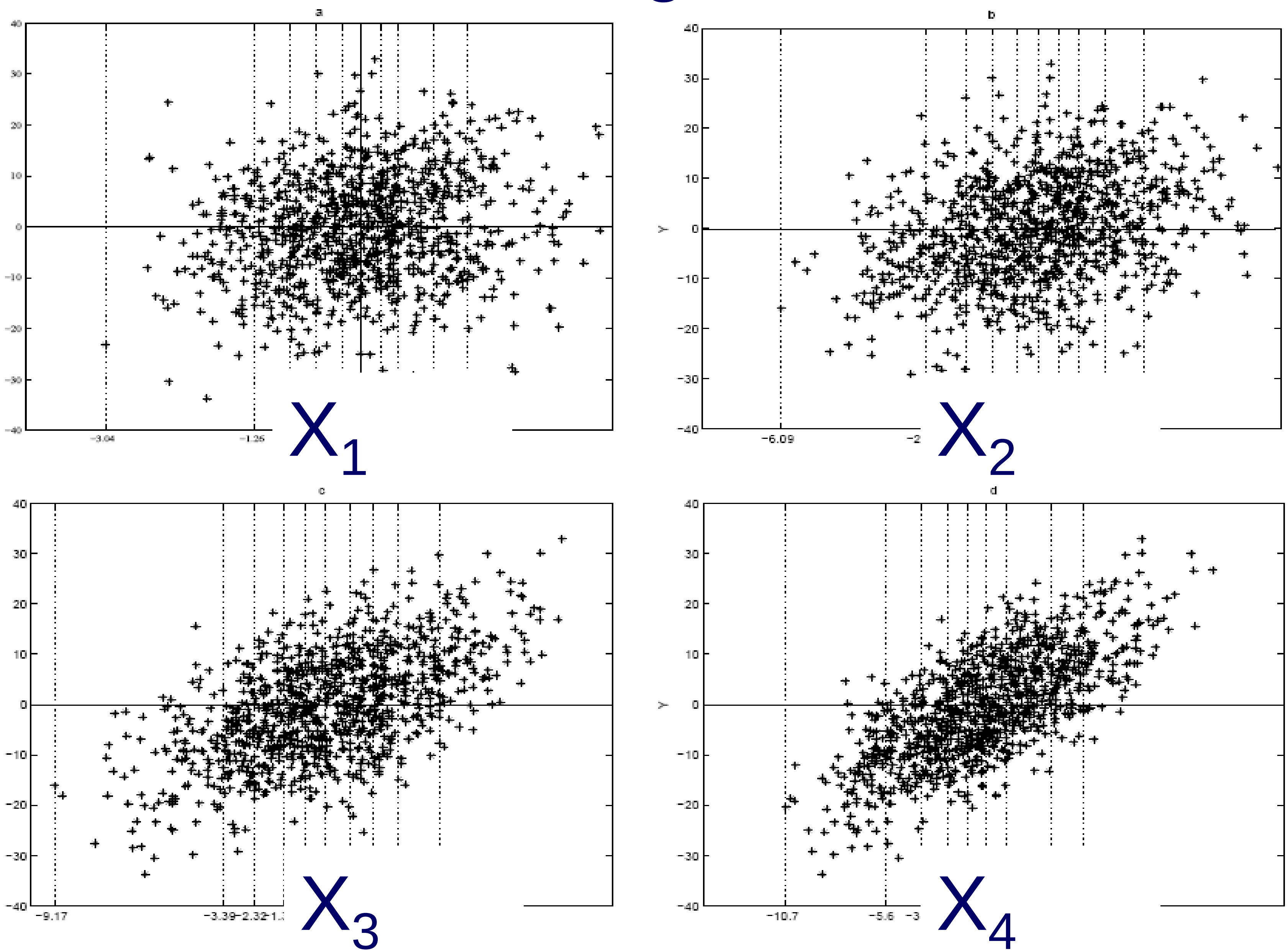
The ordinate axis is always Y

The abscissa are the various factors X_i in turn.

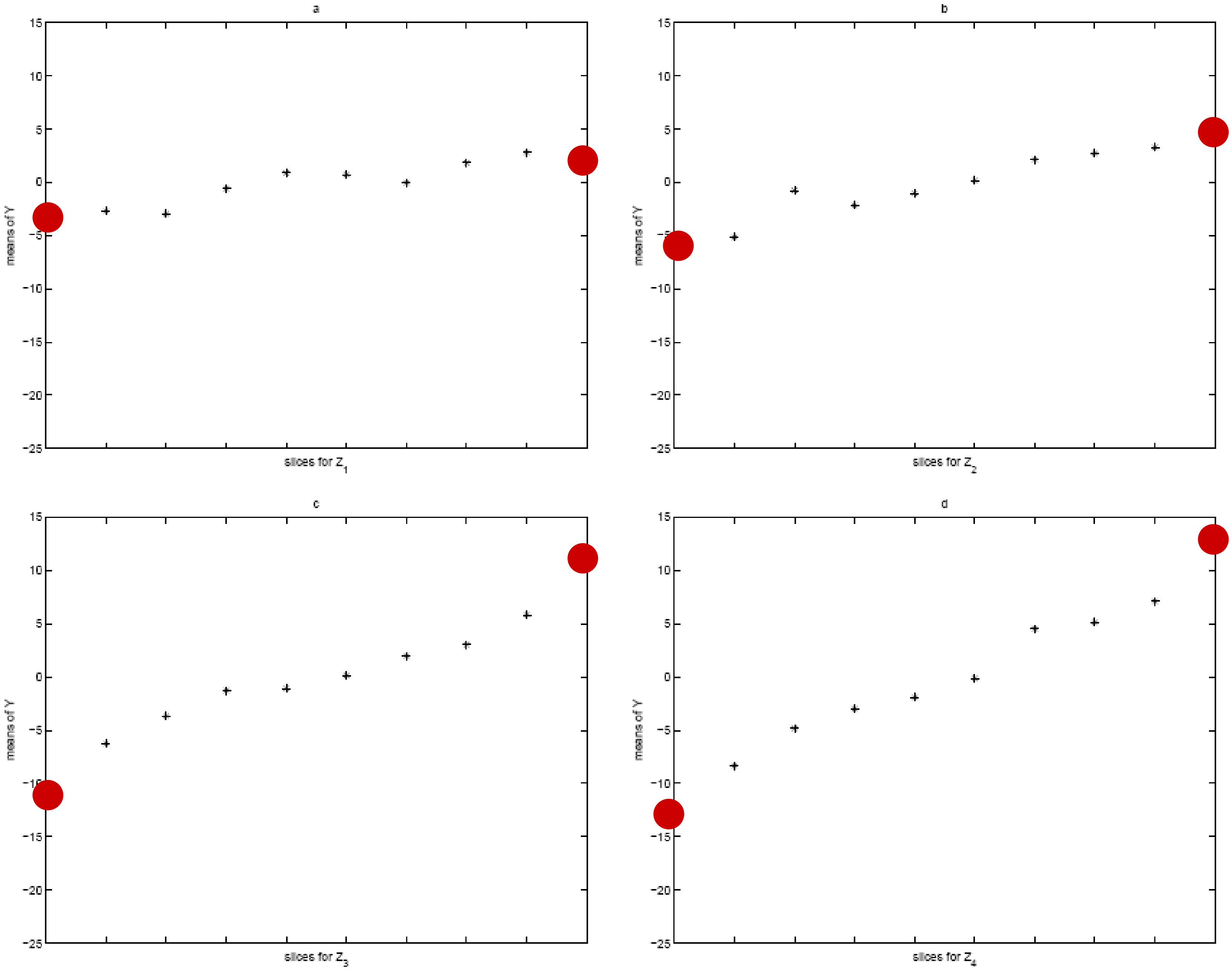
The points are always the same!

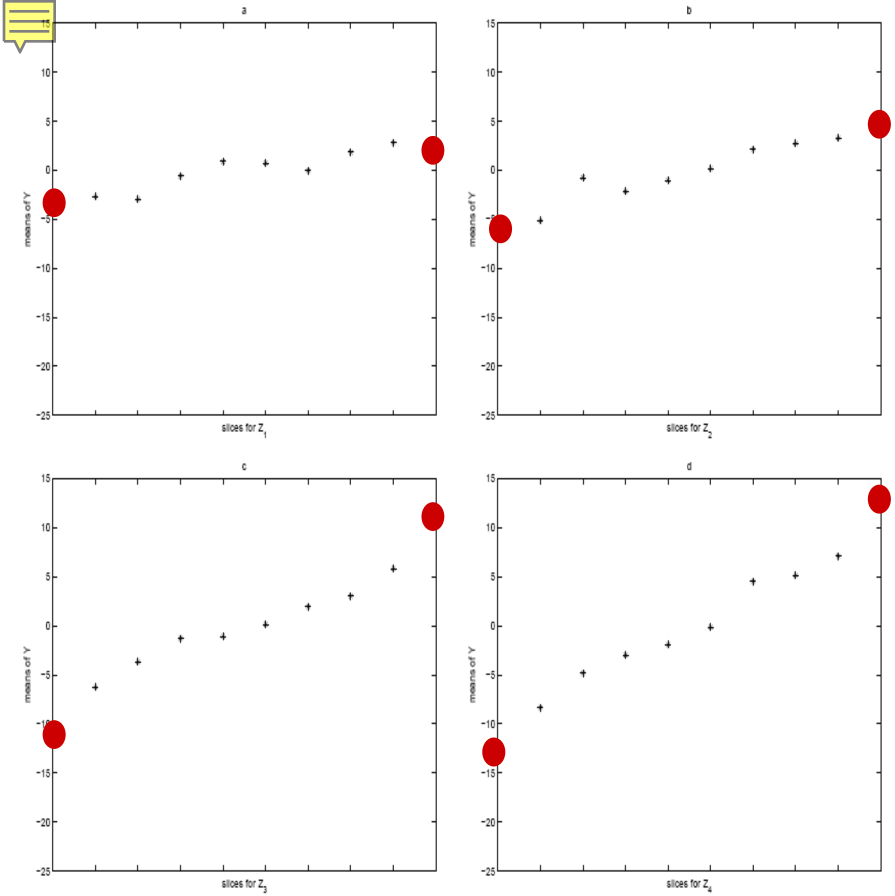

a
X$_1$
b
X$_2$
c
X$_3$
d
X$_4$

Cutting into slices…



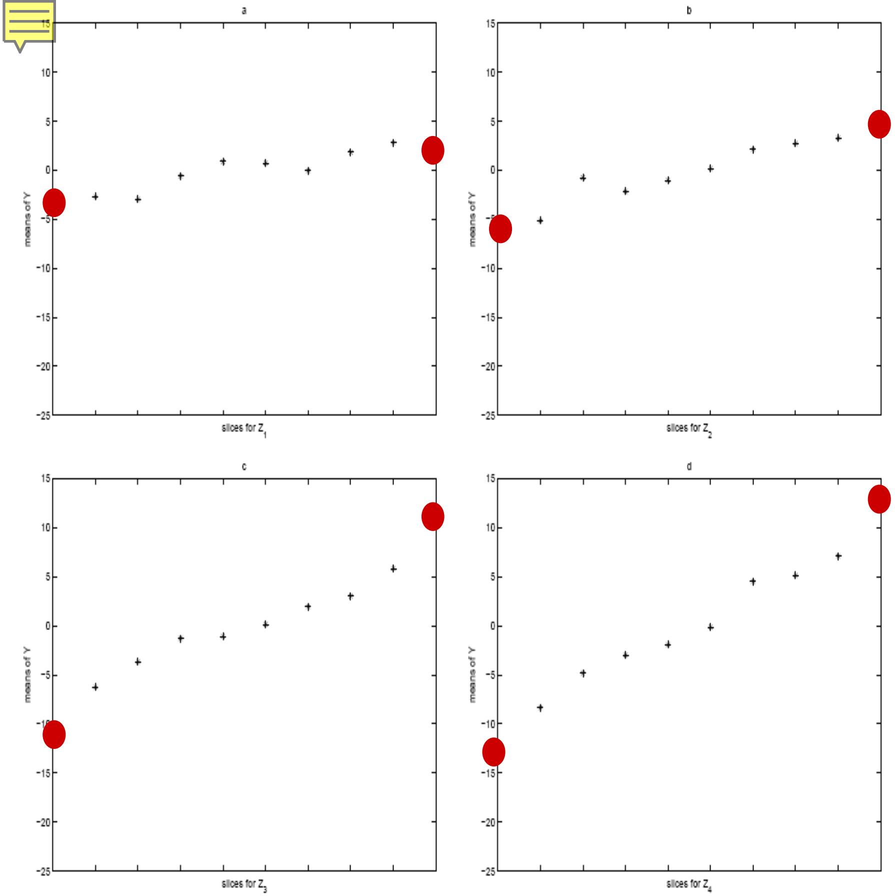
Average of Y versus X_i – same scale for Y





This shows the variance of Y across the slices: greater for X_4 than for X_1

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$



If the model is linear:

$$\frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)}{V} \approx \beta_{X_i}^2$$

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$

First order effect, or top marginal variance=

= the expected reduction in variance than would be achieved if factor Xi could be fixed.

For additive systems one can decompose the total variance as a sum of first order effects

$$\sum_i V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right) = V(Y)$$

… and a powerful variance based measure is also available for non-additive models …

From this ⋯ ⋯ to this

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)$$

This is a first order effect, or top marginal variance.

The expected reduction in variance than would be achieved if factor Xi could be fixed.

This is a total order effect, or bottom marginal variance.

The expected variance than would be left if all factors but Xi could be fixed.

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)$$

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$

This has an intuitive interpretation (the scatterplots)

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

How About this?

Variance decomposition (ANOVA)

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right) = V_i$$

$$V_{X_i X_j}\left(E_{\mathbf{X}_{\sim ij}}\left(Y \middle| X_i X_j\right)\right) =$$

$$= V_i + V_j + V_{ij}$$

$$\ldots$$

Variance decomposition (ANOVA)

$$V(Y) =$$

$$\sum_{i} V_i + \sum_{i,\, j>i} V_{ij} + \ldots + V_{123\ldots k}$$

Variance decomposition (ANOVA)

When the factors are independent the total variance can be decomposed into main effects and interaction effects up to the order *k*, the dimensionality of the problem.

When the factors are not independent the decomposition loses its unicity (and hence its appeal!)

This variance decomposition can be related to (and built from) a functional decomposition

$$Y = f\left(X_1, X_2, \ldots X_k\right) =$$

$$\sum_i f_i + \sum_{i,j>i} f_{ij} + \ldots + f_{123\ldots k}$$

where: $$f_{s_1 s_2 \ldots s_j} =$$

$$= f_{s_1 s_2 \ldots s_j}\left(X_{s_1}, X_{s_2}, \ldots X_{s_j}\right)$$

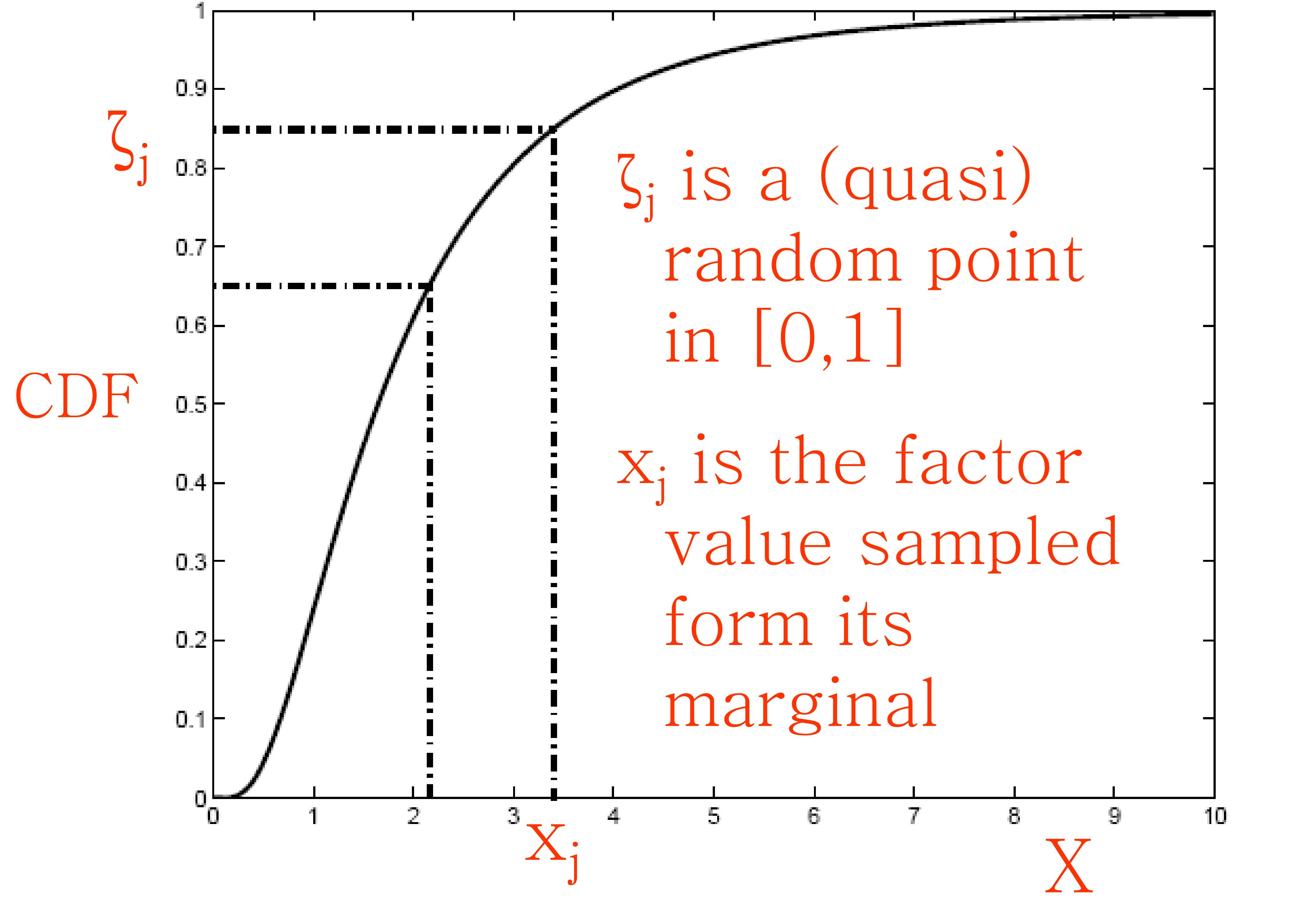
Here and in the following we assume without loss of generality *$X_i \sim U[0,1]$*

Thus we can write

$$\iint \cdots \int f_{s_1 s_2 \ldots s_j} \, dX_{s_1}, dX_{s_2}, \ldots dX_{s_j}$$

instead of

$$\iint \cdots \int f_{s_1 s_2 \ldots s_j} \, p_{s_1}\left(X_{s_1}\right) p_{s_2}\left(X_{s_2}\right) \ldots p_{s_j}\left(X_{s_j}\right) dX_{s_1}, dX_{s_2}, \ldots dX_{s_j}$$


ζ_j
CDF
x_j
X
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8
9
10
ζ_j is a (quasi) random point in [0,1]
x_j is the factor value sampled form its marginal

Functions:

$$f_{s_1 s_2 \dots s_j} =$$

$$= f_{s_1 s_2 \dots s_j}\left(X_{s_1}, X_{s_2}, \dots X_{s_j}\right)$$

Must must satisfy (for the decomposition to be unique)

$$\iint \dots \int f_{s_1 s_2 \dots s_j} dX_{s_1}, dX_{s_2}, \dots dX_{s_j} = 0$$

How can these be computed?

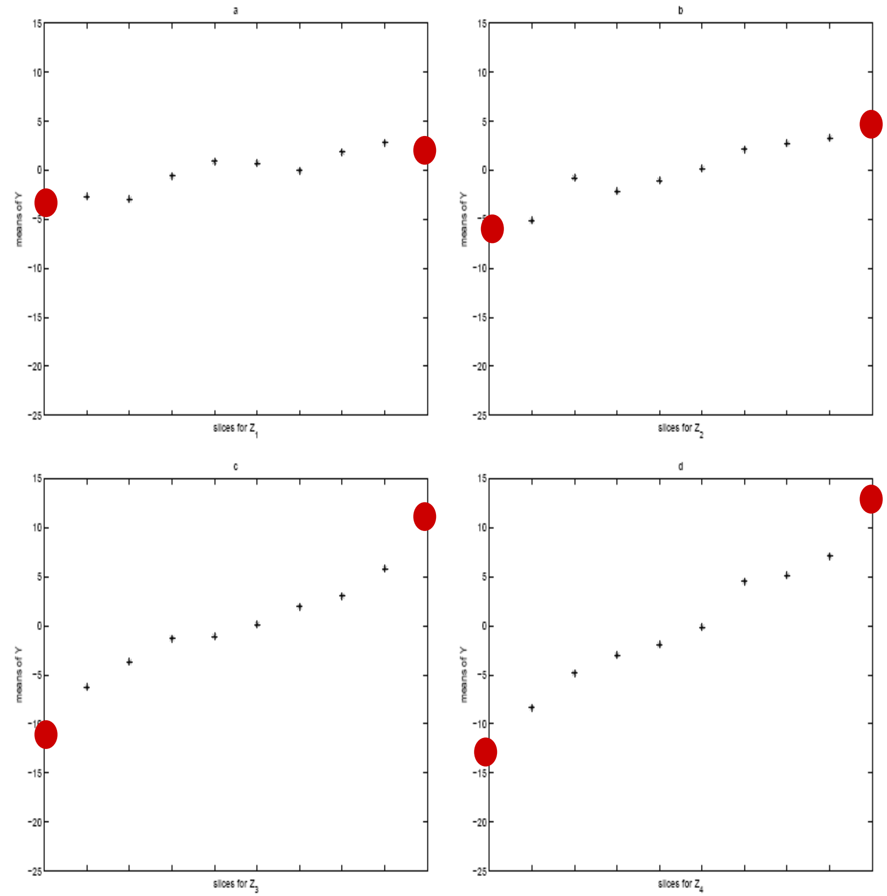
$$f_0 = E(Y)$$

$$f_i(X_i) = E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right) - f_0$$

$$f_{ij}(X_i X_j) =$$

$$= E_{\mathbf{X}_{\sim ij}}\left(Y \middle| X_i X_j\right) - f_i - f_j - f_0$$

Functions such as $f_{s_1 s_2 \ldots s_j}$



can be used for sensitivity analysis …

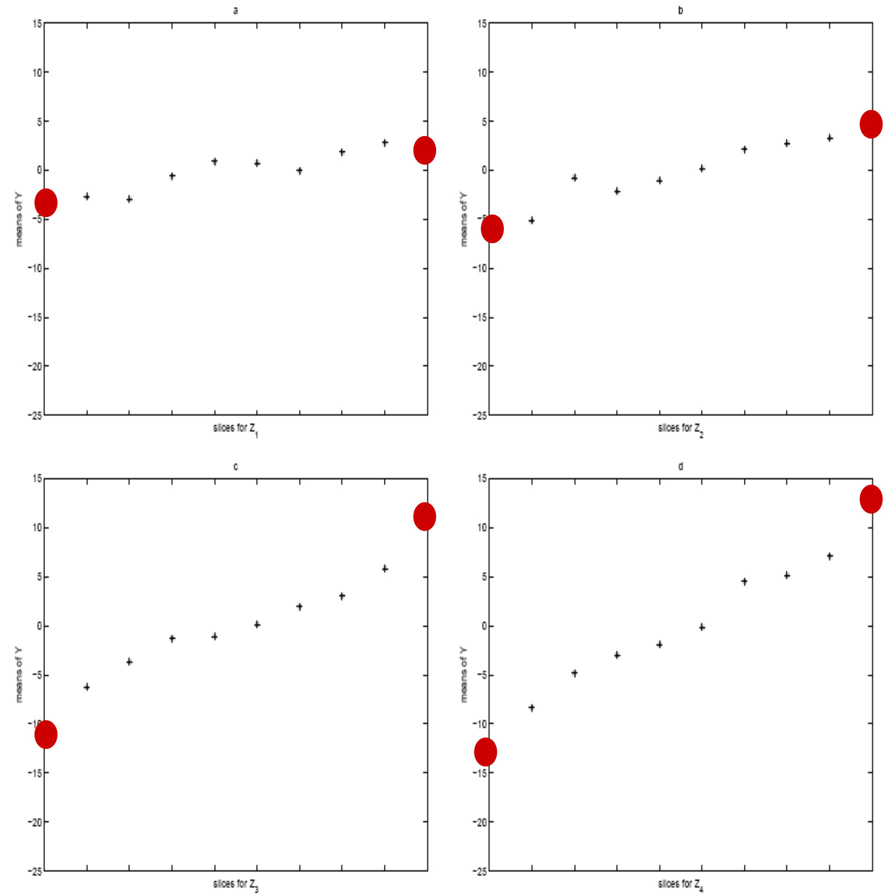
Sacks, Welch, Mitchell, and Wynn, 1989

… although these plots are not precisely $f_1(X_1), f_2(X_2), f_3(X_3), f_4(X_4)$

$$f_i(X_i) = E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right) - f_0$$

Functions $f_{s_1 s_2 \ldots s_j}$

have a drawback



These decompositions (of V and well as of f) have their own dimensionality curse: the terms are 2^k !

… thus we go back to our preferred measures

From main effect to total effect

From

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$

Main effect of factor X_i

replacing X_i with $\mathbf{X}_{\sim i}$

To main effect of non- X_i

$$V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

BUT:

$$V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y|\mathbf{X}_{\sim i}\right)\right)+$$
$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y|\mathbf{X}_{\sim i}\right)\right)=V(Y)$$

Easy to prove using $V(\bullet)=E(\bullet)^2-E^2(\bullet)$

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

Main effect on non-X_i

$$V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

… all remaining variance must be due to X_i and its interactions

Main effects

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$

$$V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

Residuals

$$E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

Main (or first order) effect of X_i

Main effects

Residuals

$$\boxed{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)} + E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right) = \mathbf{V(Y)}$$

$$V_{\mathbf{X}_{\sim i}}\left(E_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right) + \boxed{E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)} = \mathbf{V(Y)}$$

Total (or total order) effect of X_i

Rows add up to V(Y); diagonal terms equal for additive models.

$$\frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)}{V(Y)} = S_i$$

$$\frac{E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)}{V(Y)} = S_{Ti}$$

Rescaled to [0,1], under the name of first order and total order sensitivity coefficient

How to compute the indices

Let us apply the relation

$V(\bullet)=E(\bullet)^2-E^2(\bullet)$ to $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)$

It becomes

$$\int E^2_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)dX_i \; - \left(\int E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)dX_i\right)^2$$

The second term is easy:

$$\left(\int E_{\mathbf{X}_{\sim i}} \left(Y \middle| X_i \right) dX_i \right)^2 = f_0^{\,2}$$

The first term is a bit more laborious:

$$E^2_{\mathbf{X}_{\sim i}}\left(Y\middle|X_i\right)=$$

$$=E_{\mathbf{X}_{\sim i}}\left(Y\middle|X_i\right)E_{\mathbf{X}'_{\sim i}}\left(Y'\middle|X_i\right)=$$

$$=\iint\ldots\int f\left(X_1,\ldots,X_k\right)$$

$$f\left(X'_1,\ldots X'_{i-1},X_i,X'_{i+1},\ldots X'_k\right)$$

$$d\mathbf{X}_{\sim i}d\mathbf{X}'_{\sim i}$$

Once one integrates over X_i the integral

$$\int\int \dots \int f\left(X_1, \dots, X_k\right)$$

$$f\left(X_1', \dots X_{i-1}', X_i, X_{i+1}', \dots X_k'\right)$$

$$d\mathbf{X}_{\sim i} d\mathbf{X}_{\sim i}'$$

One obtains

$$
\iint\cdots\int f\left(X_1,\ldots,X_k\right)
$$

$$
f\left(X'_1,\ldots X'_{i-1},X_i,X'_{i+1},\ldots X'_k\right)
$$

$$
\boxed{d\mathbf{X}_{\sim i}}\,d\mathbf{X}'_{\sim i}\,\boxed{dX_i} =
$$

$$
=\iint\cdots\int f\left(X_1,\ldots,X_k\right)
$$

$$
f\left(X'_1,\ldots X'_{i-1},X_i,X'_{i+1},\ldots X'_k\right)
$$

$$
\boxed{d\mathbf{X}}\,d\mathbf{X}'_{\sim i}
$$

$$\int\int\dots\int f\left(X_1,\dots,X_k\right)$$

$$f\left(X'_1,\dots X'_{i-1},X_i,X'_{i+1},\dots X'_k\right)$$

$$d\mathbf{X}d\mathbf{X}'_{\sim i}$$

This is the expectation in k+k-1 dimensions of:

$$f\left(X_1,\dots,X_k\right)f\left(X'_1,\dots X'_{i-1},X_i,X'_{i+1},\dots X'_k\right)$$

Two function values where in the second all is re-sampled but factor X_i

Wrapping all together

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right) =$$

$$= E_{\mathbf{X}\mathbf{X}'_{\sim i}}\left(f f'\right) - f_0^2$$

And this can be computed as follows – generate a (quasi) random numbers matrix of row dimension *2k* and column length *N*

$$
\begin{array}{cccc}
x_{11} & x_{12} & \cdots & x_{1(2k)} \\
x_{21} & x_{22} & \cdots & x_{2(2k)} \\
\cdots & \cdots & \cdots & \cdots \\
x_{N1} & x_{N2} & \cdots & x_{N(2k)}
\end{array}
$$

Split into two:

$$\mathbf{A} = \begin{matrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{Nk} \end{matrix} \quad \mathbf{B} = \begin{matrix} x_{1(k+1)} & x_{1(k+2)} & \cdots & x_{1(2k)} \\ x_{2(k+1)} & x_{2(k+2)} & \cdots & x_{2(2k)} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N(k+1)} & x_{N(k+2)} & \cdots & x_{N(2k)} \end{matrix}$$

And generate a third matrix which is all-A but one column (column i) which is from B

$$\mathbf{A}_i^B = \begin{array}{cccccc} x_{11} & x_{12} & \cdots & \boxed{x_{1(k+i)}} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2(k+i)} & \cdots & x_{2k} \\ \cdots & \cdots & & \cdots & & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{N(k+i)} & \cdots & x_{Nk} \end{array}$$

(call it a quasi-A matrix)

Finally we compute $V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right) =$

$$= E_{\mathbf{XX}'_{\sim i}}\left(ff'\right) - f_0^{\,2} =$$

as:

$$= \frac{1}{N}\sum_{j=1}^{N} f_j^{\,B}\left(f_j^{\,A_i^B} - f_j^{\,A}\right)$$

Where:

f_j^B is computed from row j of

$$\mathbf{B} = \begin{matrix} x_{1(k+1)} & x_{1(k+2)} & \cdots & x_{1(2k)} \\ x_{2(k+1)} & x_{2(k+2)} & \cdots & x_{2(2k)} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N(k+1)} & x_{N(k+2)} & \cdots & x_{N(2k)} \end{matrix}$$

and $f_j^{A_i^B}$ from the quasi-A matrix:

$$\mathbf{A}_i^B = \begin{matrix} x_{11} & x_{12} & \cdots & x_{1(k+i)} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2(k+i)} & \cdots & x_{2k} \\ \cdots & \cdots & & \cdots & & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{N(k+i)} & \cdots & x_{Nk} \end{matrix}$$

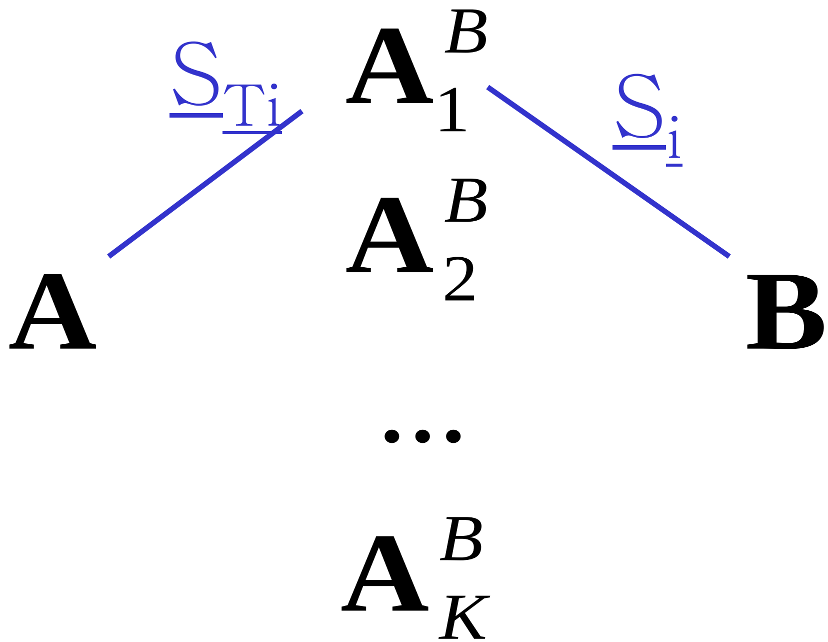
In summary one can compute the first order terms from one matrix A and B each and k matrices A_i^B i.e. using function values

$$f_j^A \qquad f_j^B \qquad f_j^{A_i^B}$$

The entire story can be repeated for the total effect index, which can be computed from

$$f_j^A \qquad f_j^{A_i^B}$$

Thus with k quasi-A matrices and the two matrices A and B one can compute for a total of k+2 matrices all total and first order effects



In three dimensions (k=3), three points (N=3)

$$\mathbf{A} = \begin{matrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{matrix} \quad B = \begin{matrix} x_{1(3+1)} & x_{1(3+2)} & x_{1(3+3)} \\ x_{2(3+1)} & x_{2(3+2)} & x_{2(3+3)} \\ x_{3(3+1)} & x_{3(3+2)} & x_{3(3+3)} \end{matrix}$$

Rewriting B: $$B = \begin{matrix} x_{14} & x_{15} & x_{16} \\ x_{24} & x_{25} & x_{26} \\ x_{34} & x_{35} & x_{36} \end{matrix}$$

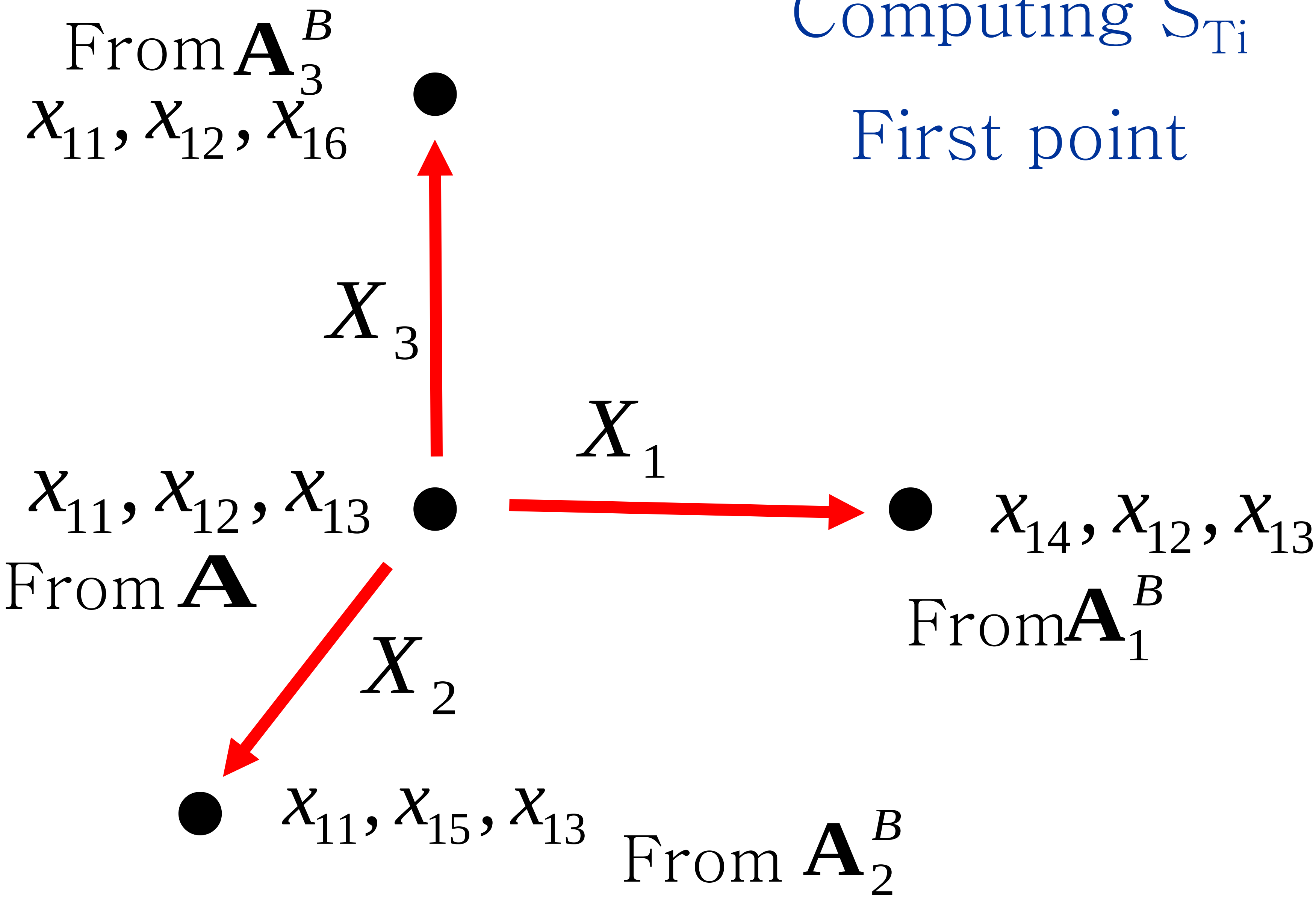
Generate the 3 quasi-A matrices

$$\mathbf{A}_1^B = \begin{matrix} \boxed{\begin{matrix} x_{14} \\ x_{24} \\ x_{34} \end{matrix}} & \begin{matrix} x_{12} \\ x_{22} \\ x_{32} \end{matrix} & \begin{matrix} x_{13} \\ x_{23} \\ x_{33} \end{matrix} \end{matrix}$$

$$\mathbf{A}_2^B = \begin{matrix} \begin{matrix} x_{11} \\ x_{21} \\ x_{31} \end{matrix} & \boxed{\begin{matrix} x_{15} \\ x_{25} \\ x_{35} \end{matrix}} & \begin{matrix} x_{13} \\ x_{23} \\ x_{33} \end{matrix} \end{matrix} \qquad \mathbf{A}_3^B = \begin{matrix} \begin{matrix} x_{11} \\ x_{21} \\ x_{31} \end{matrix} & \begin{matrix} x_{12} \\ x_{22} \\ x_{32} \end{matrix} & \boxed{\begin{matrix} x_{16} \\ x_{26} \\ x_{36} \end{matrix}} \end{matrix}$$

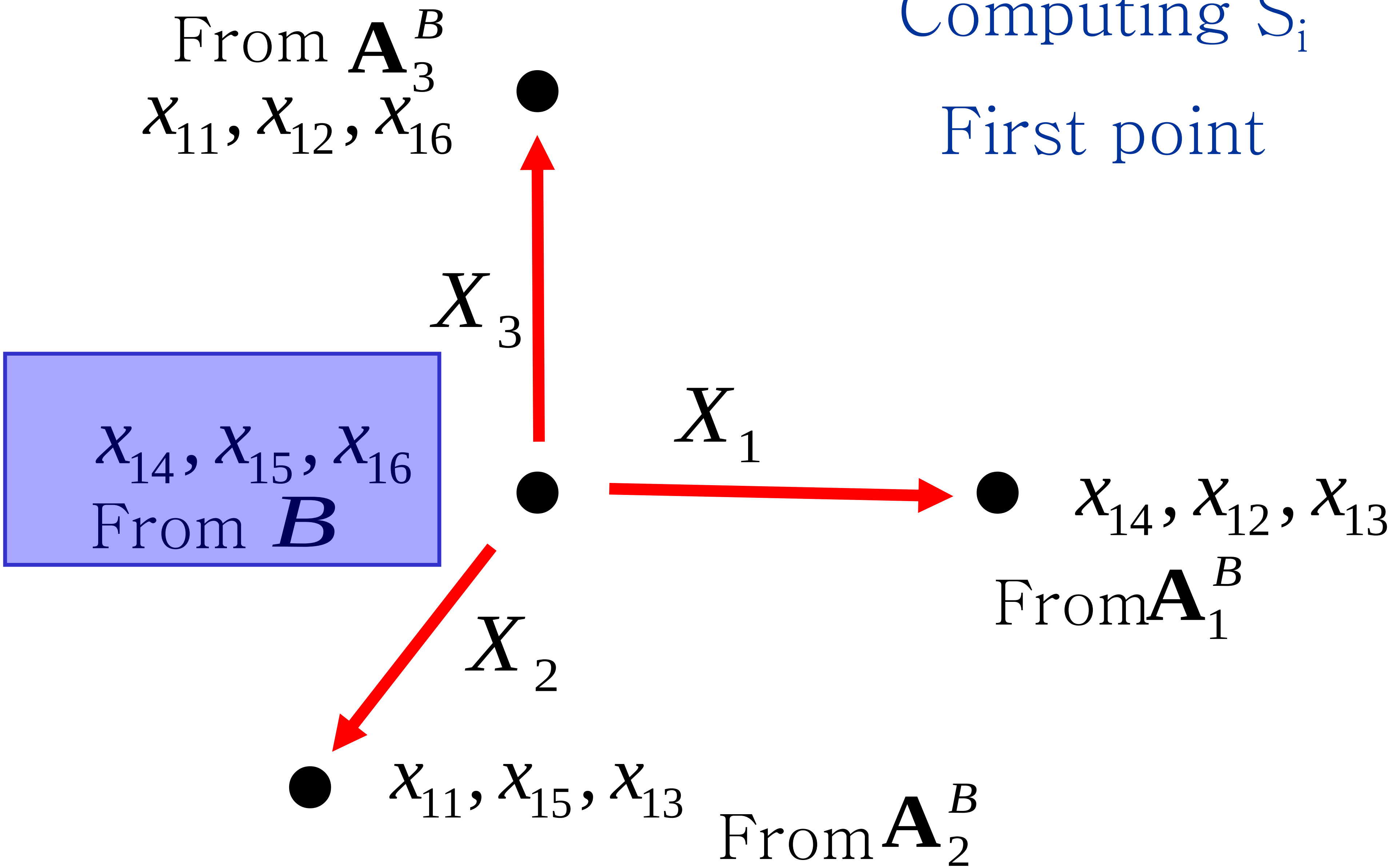
Computing S_{Ti}

First point



Computing S_i

First point



One last words on estimators: first order

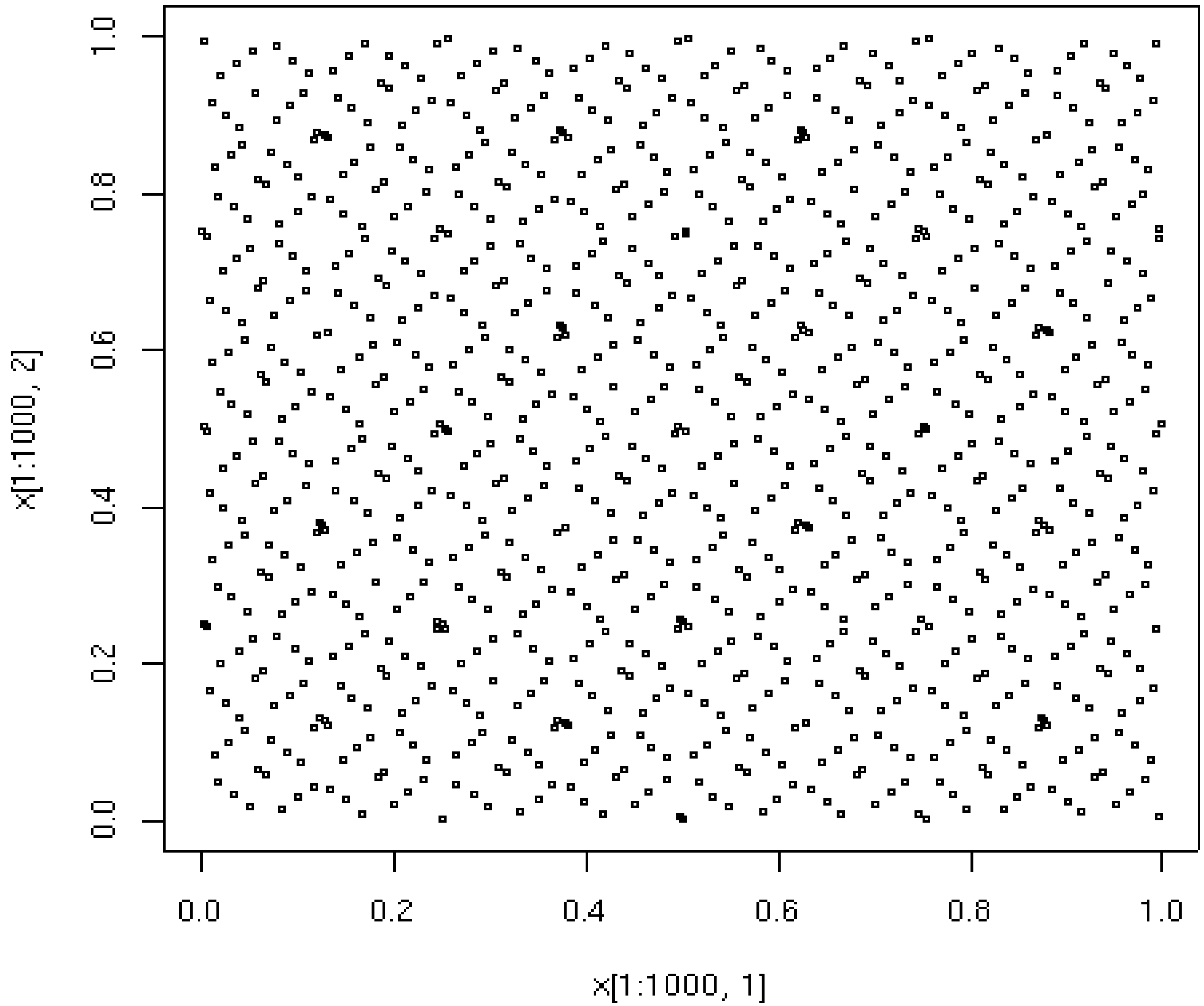
$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right) \approx \frac{1}{N}\sum_{j=1}^{N} f_j^{B}\left(f_j^{A_i^B} - f_j^{B}\right)$$

Extensive testing shows that for the total effect the best measure is (Jansen 1999):

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right) \approx \frac{1}{2N}\sum_{j=1}^{N}\left(f_j^{A} - f_j^{A_i^B}\right)^2$$

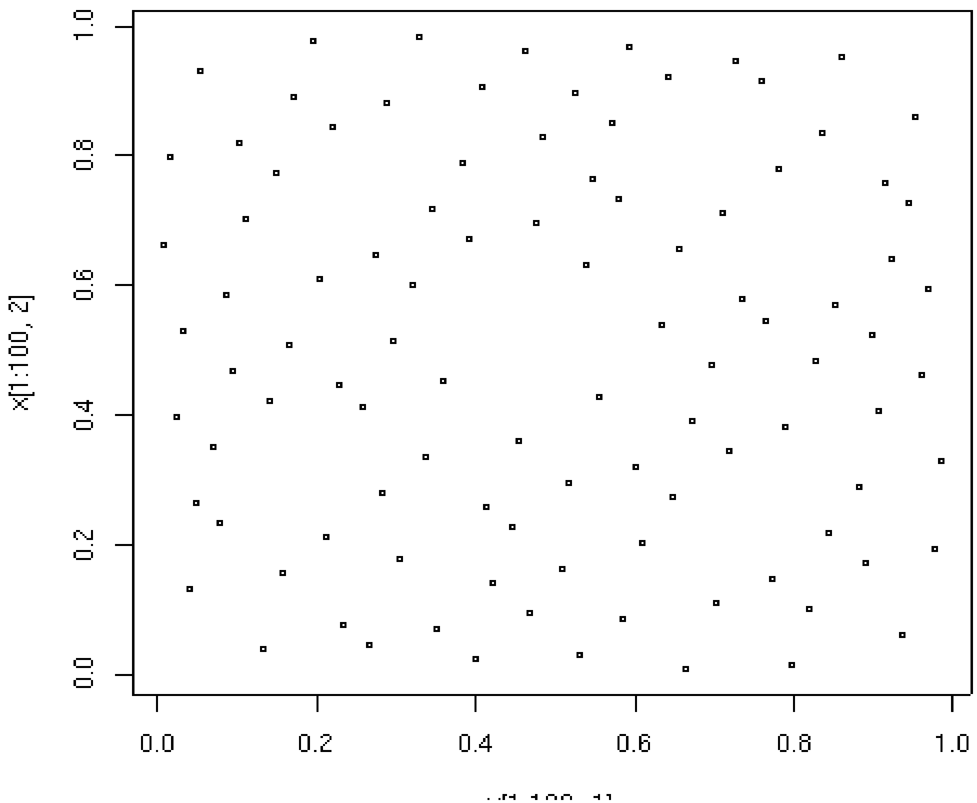
What you have seen so far has been optimized as to have a maximum of coordinates from A and a minimum of coordinates from B.

Why?

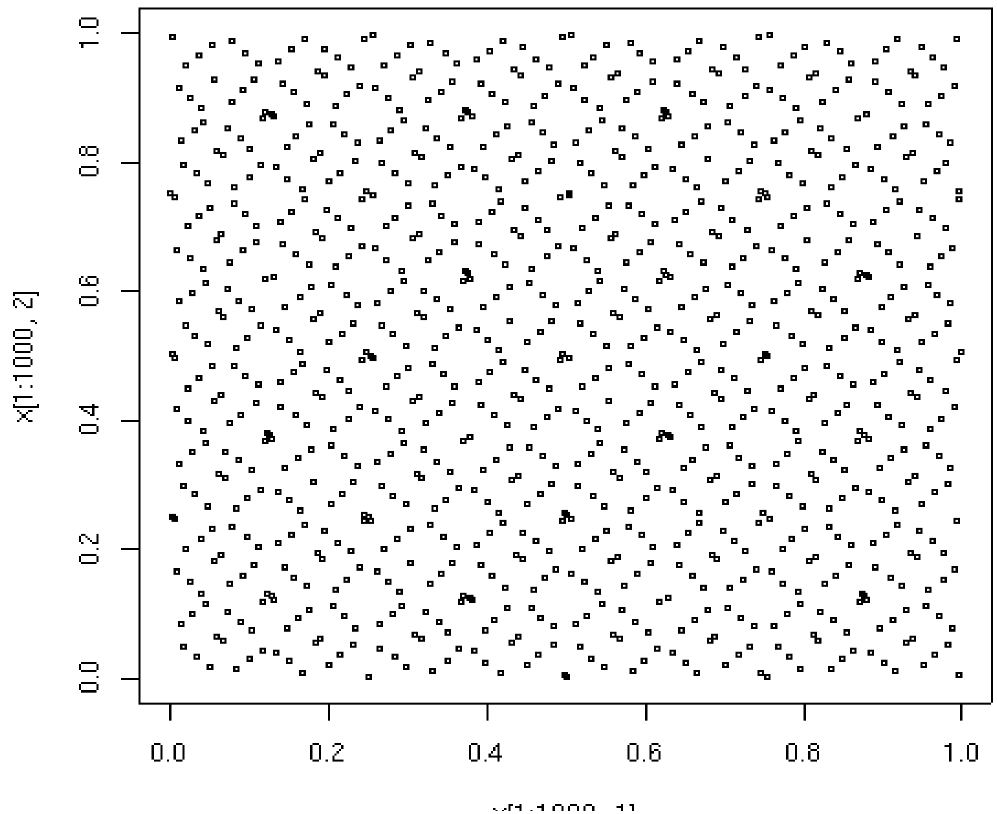


An LP_τ sequence

We normally use low discrepancies sequences developed by I.M Sobol' – these are known as LP-TAU sequences

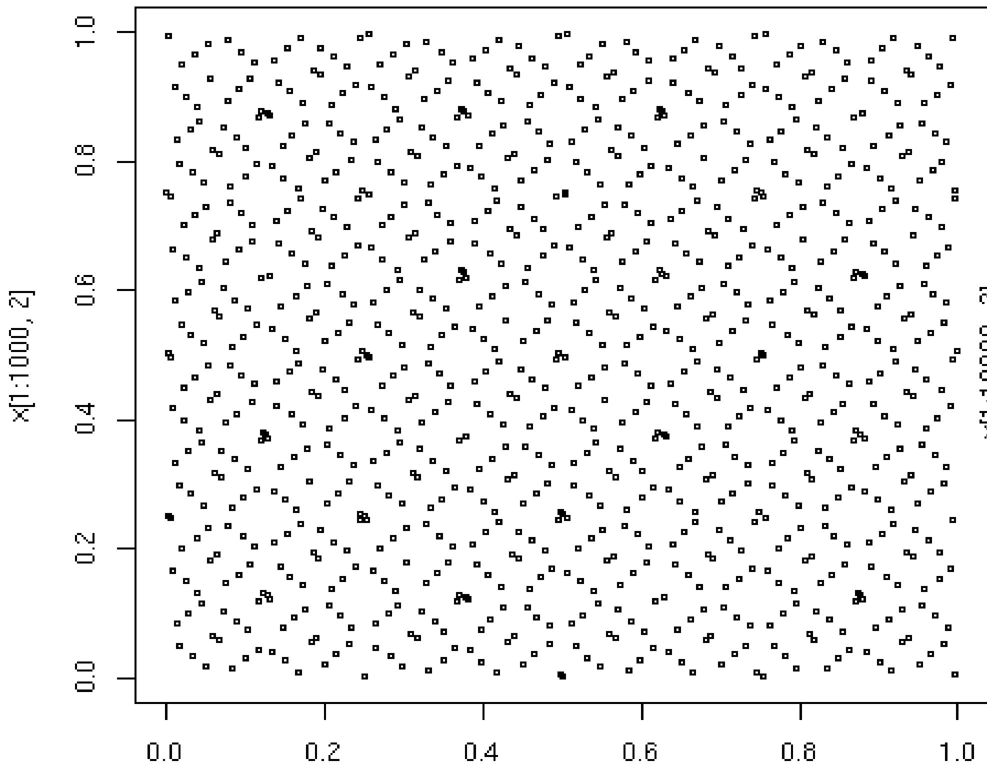


X1,X2 plane, 100 Sobol' points

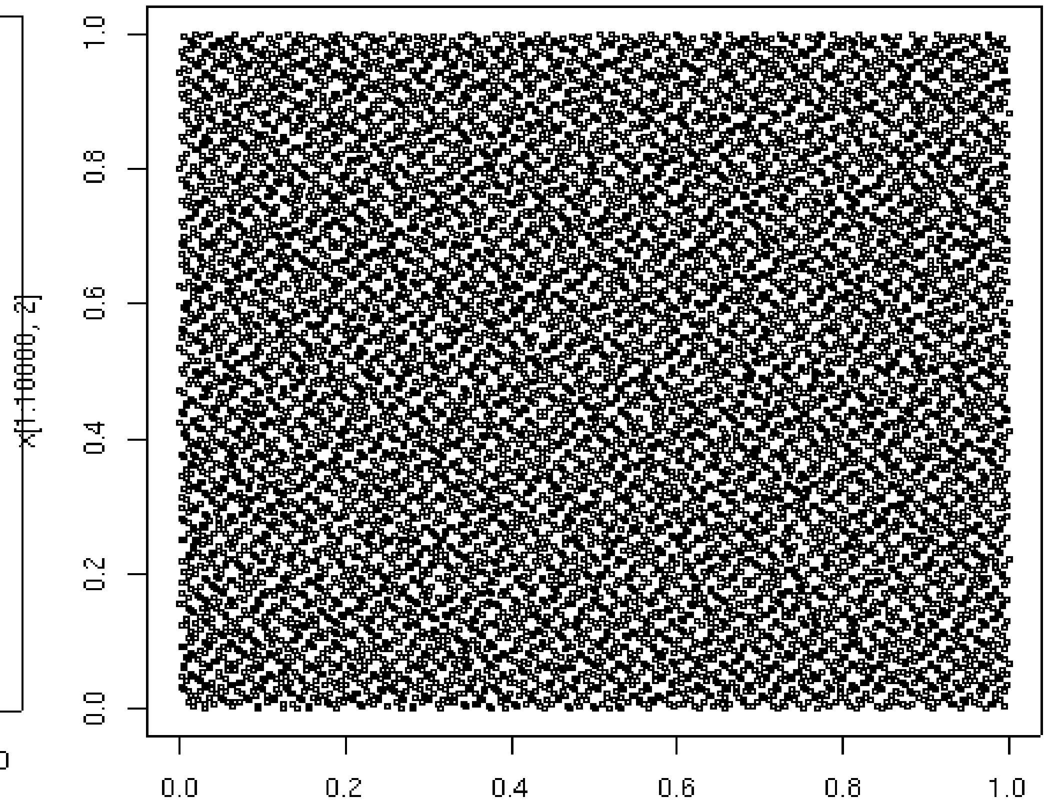


X1,X2 plane, 1000 Sobol' points

Sobol' sequences of quasi-random points

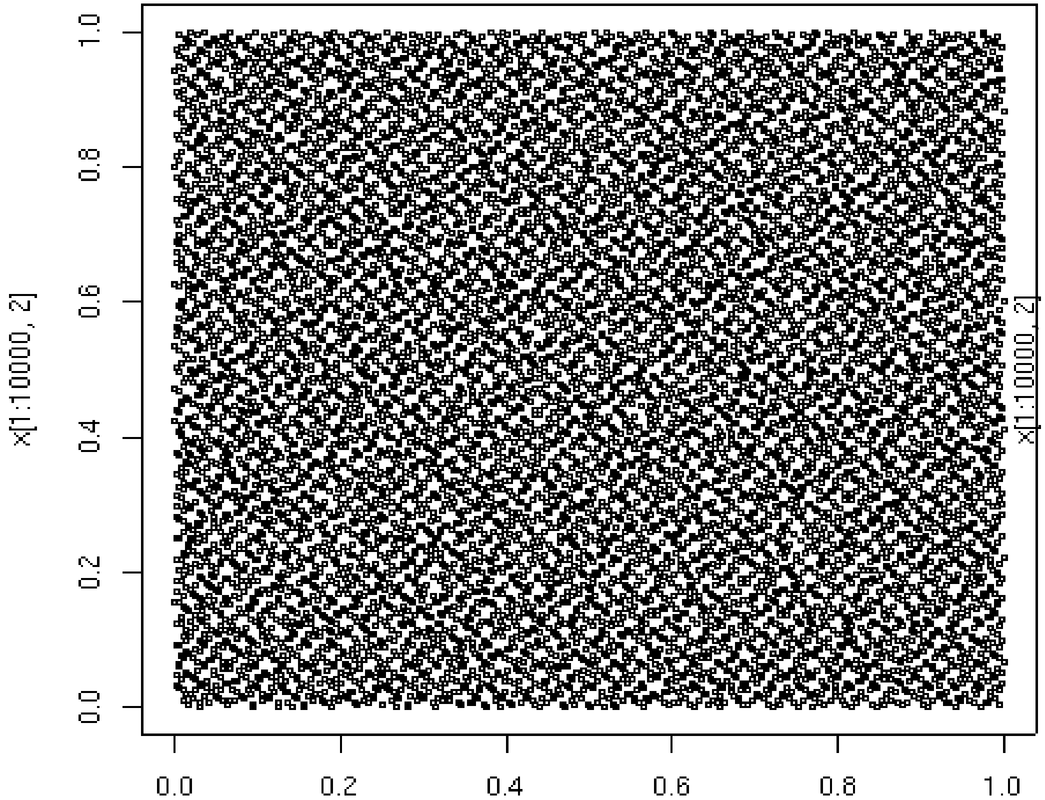


X1,X2 plane, 1000 Sobol’ points

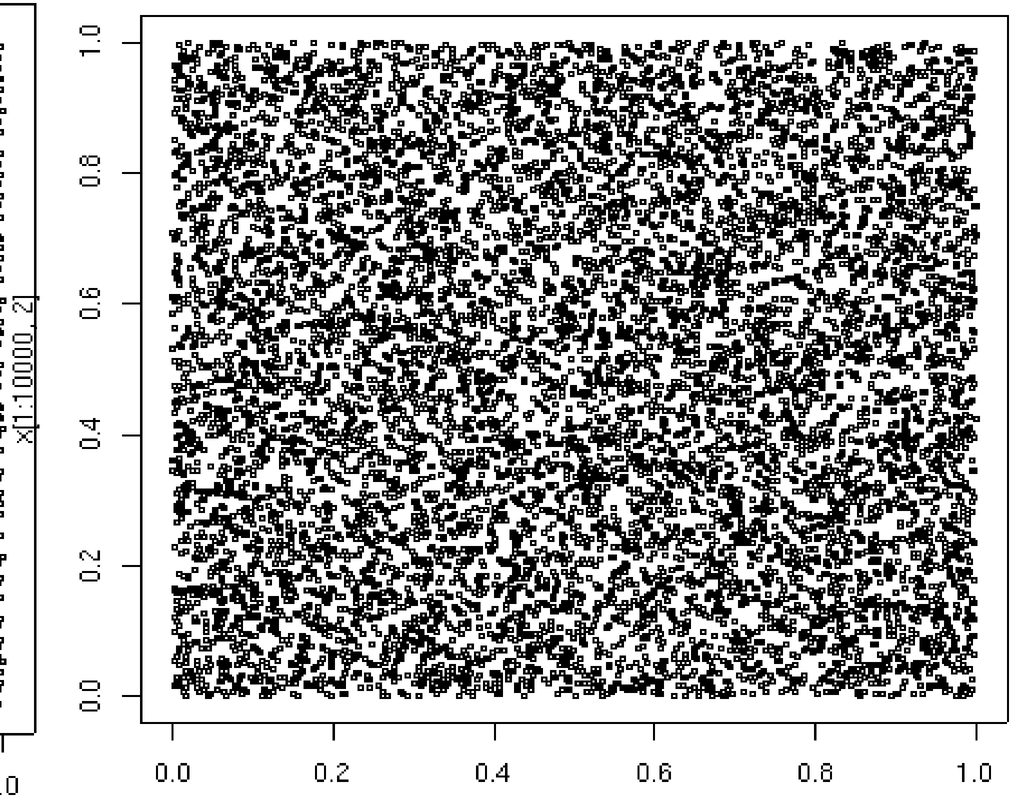


X1,X2 plane, 10000 Sobol’ points

Sobol’ sequences of quasi-random points



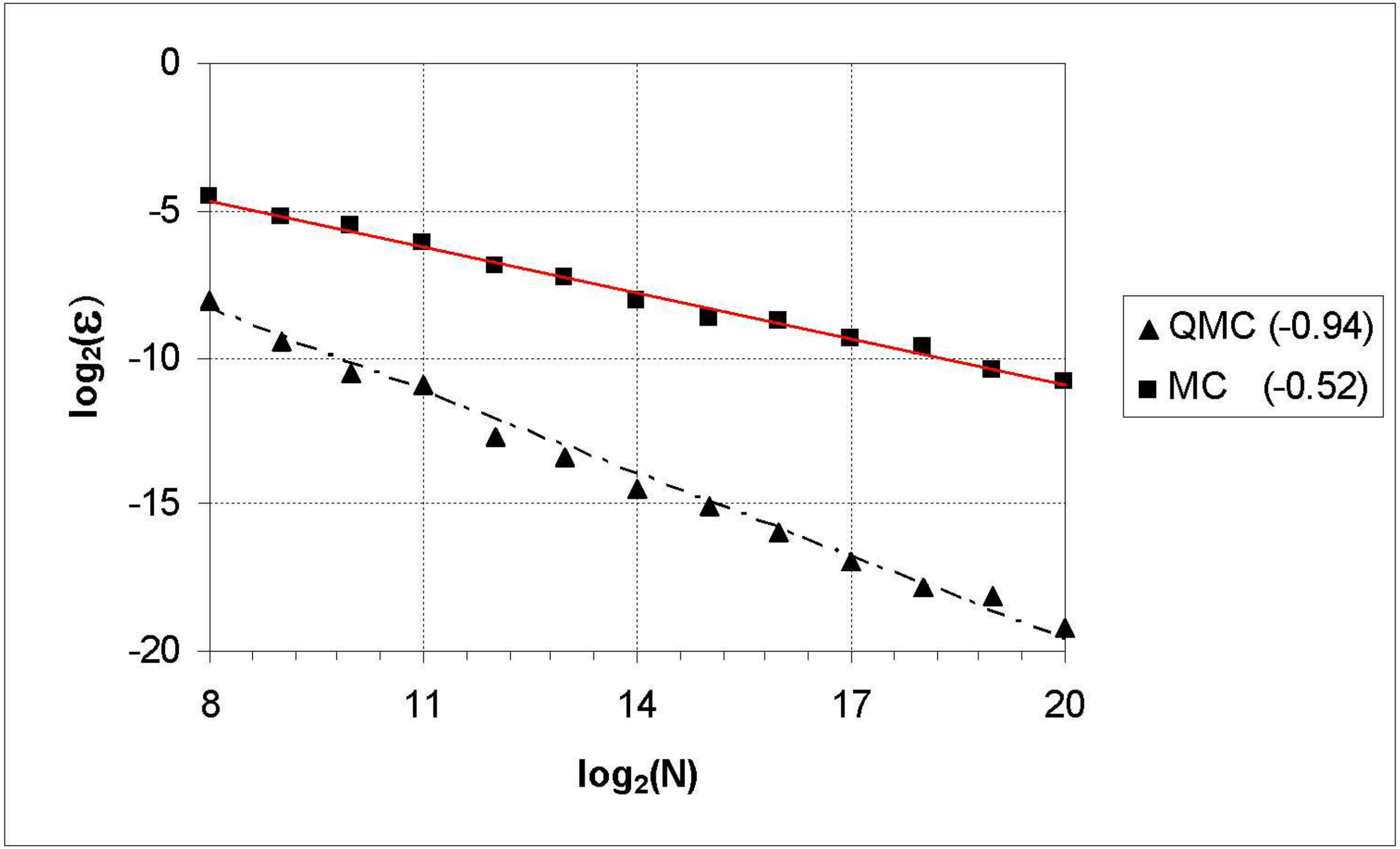
X1,X2 plane, 10000 Sobol' points



X1,X2 plane, 10000 random points

Sobol' sequences of quasi-random points against random points

Why quasi-random



Source: Mauntz and Kucherenko, 2005

Why estimate using as much as possible from A and quasi-A matrices?

The lower the column number the better its discrepancy property

➔ quasi-MC trick: if possible put important variables on the left

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \middle| X_i\right)\right)$$

Equal to one another when the model is additive

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \middle| \mathbf{X}_{\sim i}\right)\right)$$

Why these two measures?

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right)$$

Factors prioritization

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y \mid \mathbf{X}_{\sim i}\right)\right)$$

Fixing (dropping) non important factors

Even when factors are not independent!

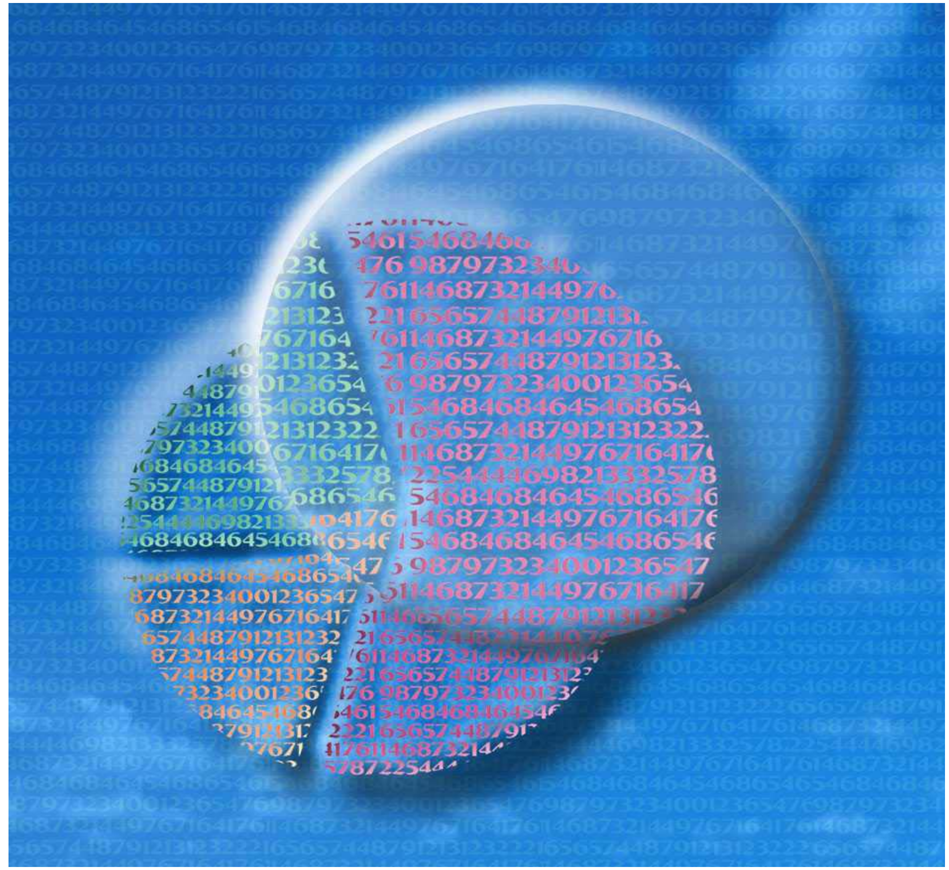
Computational details:

1. Easy-to-code, Monte Carlo – better on quasi-random points. Estimate of the error available.

2. The main effect can be made cheap; its computational cost does not depend upon k.

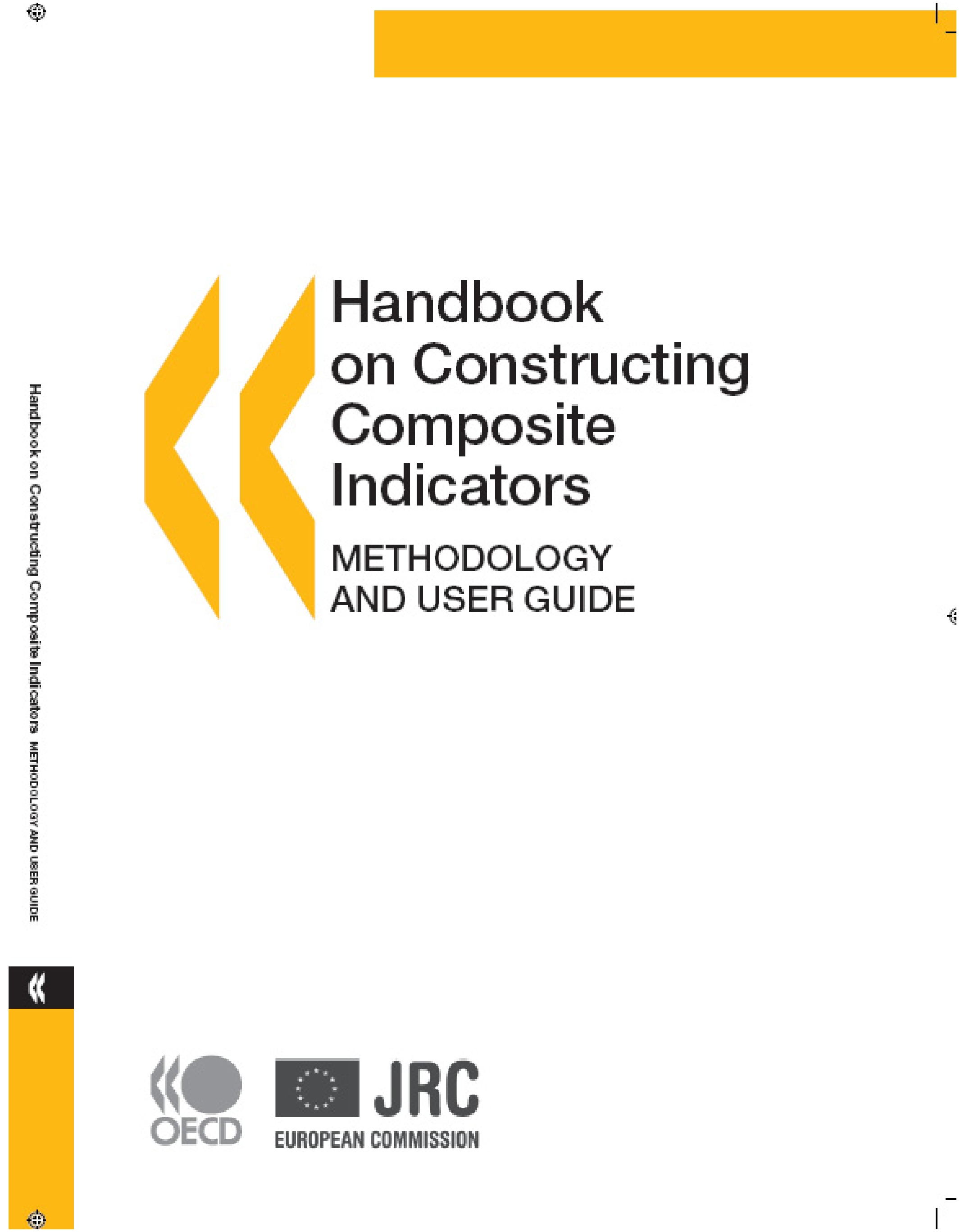
3. The total effect is expensive; its computational cost is $(k+1)N$ where N is one of the order of one thousand.

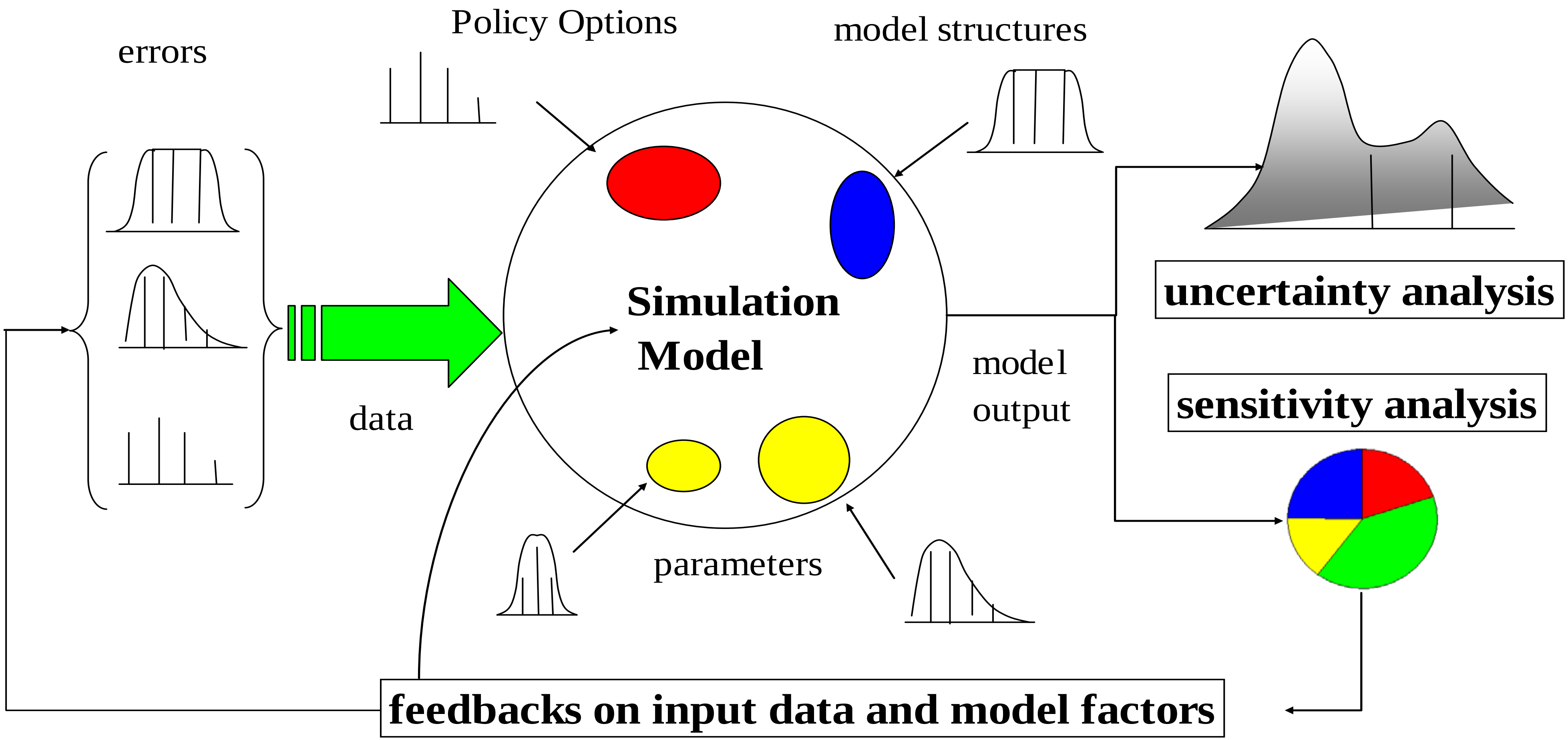
5. Applications



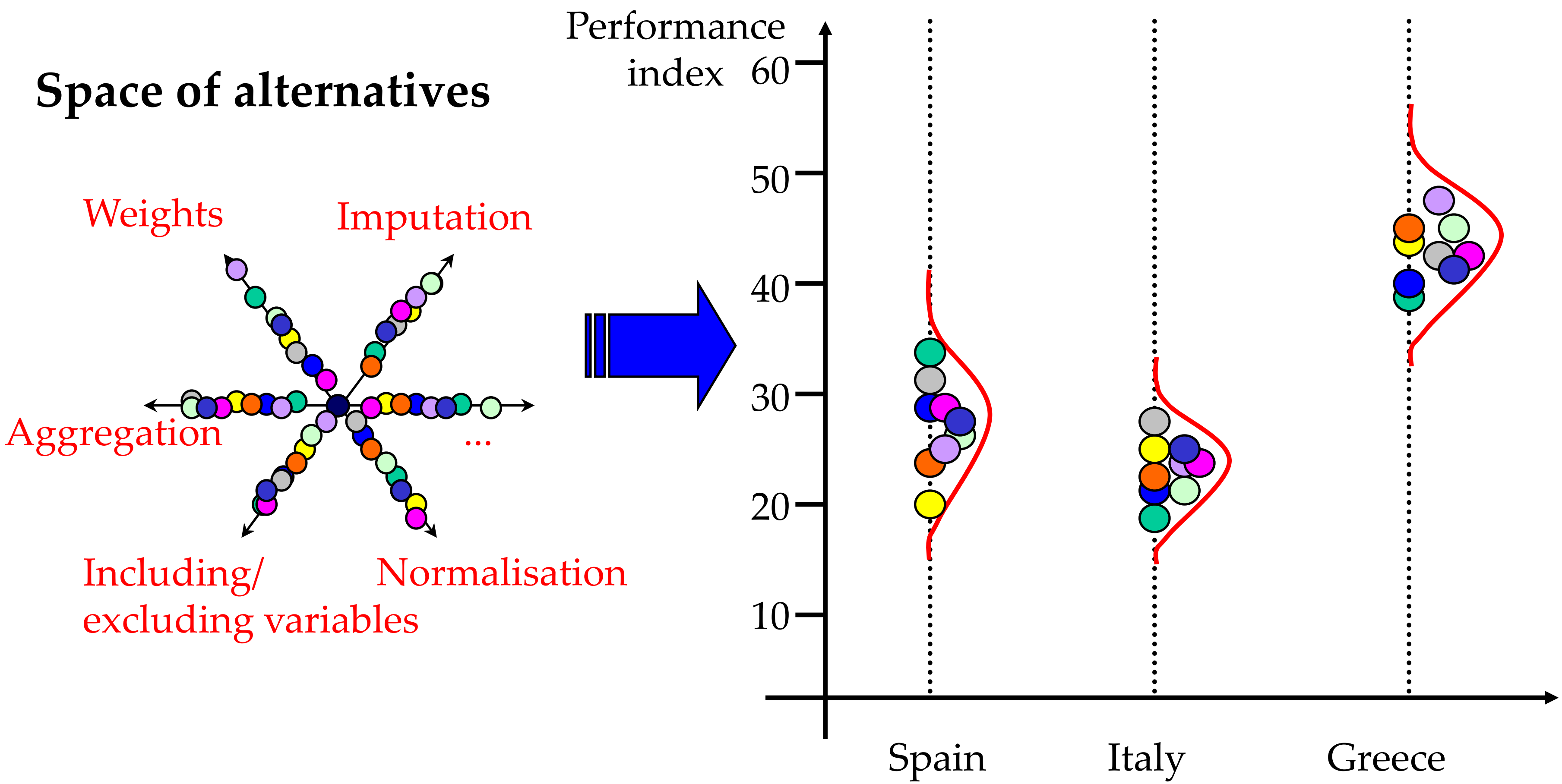
Methodology from:
Joint OECD-JRC **handbook**.

- 5 years of preparation,
- 2 rounds of consultation with OECD high level statistical committee,
- finally endorsed March 2008 with one abstention

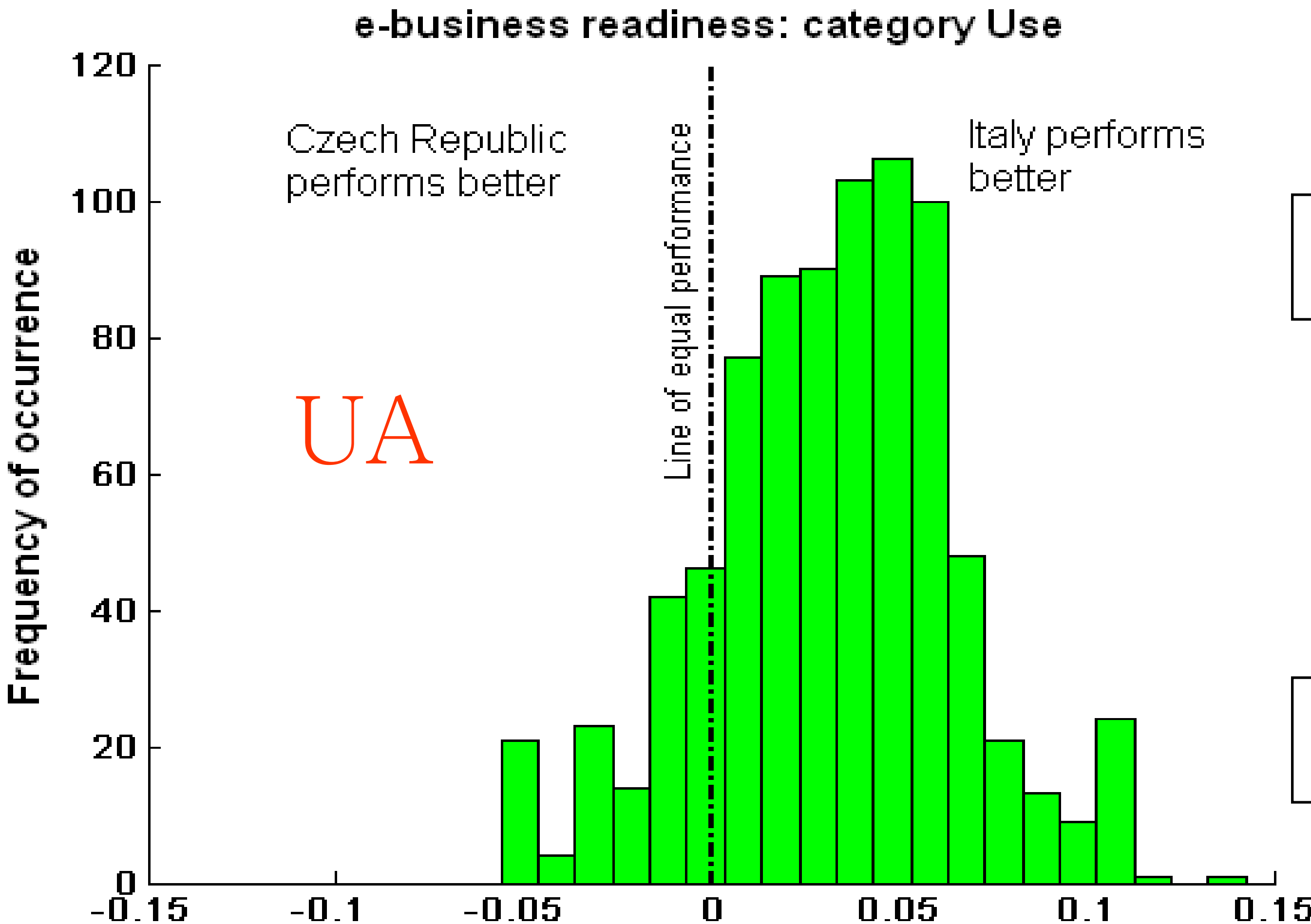


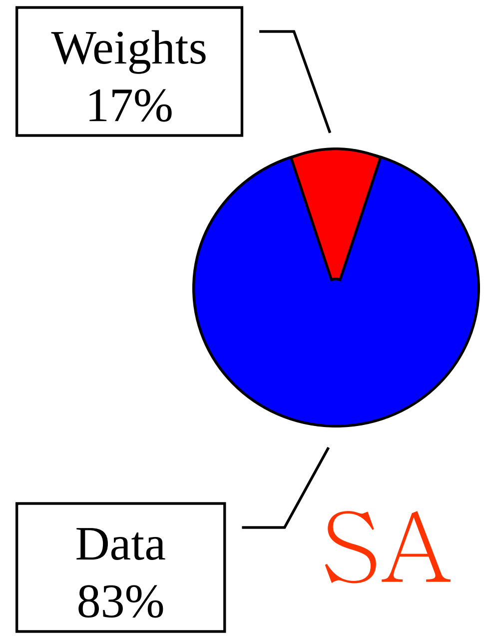

errors
Policy Options
model structures
Simulation Model
data
model output
uncertainty analysis
sensitivity analysis
parameters
feedbacks on input data and model factors

Uncertainty analysis can be used to assess the robustness of composite indicators …



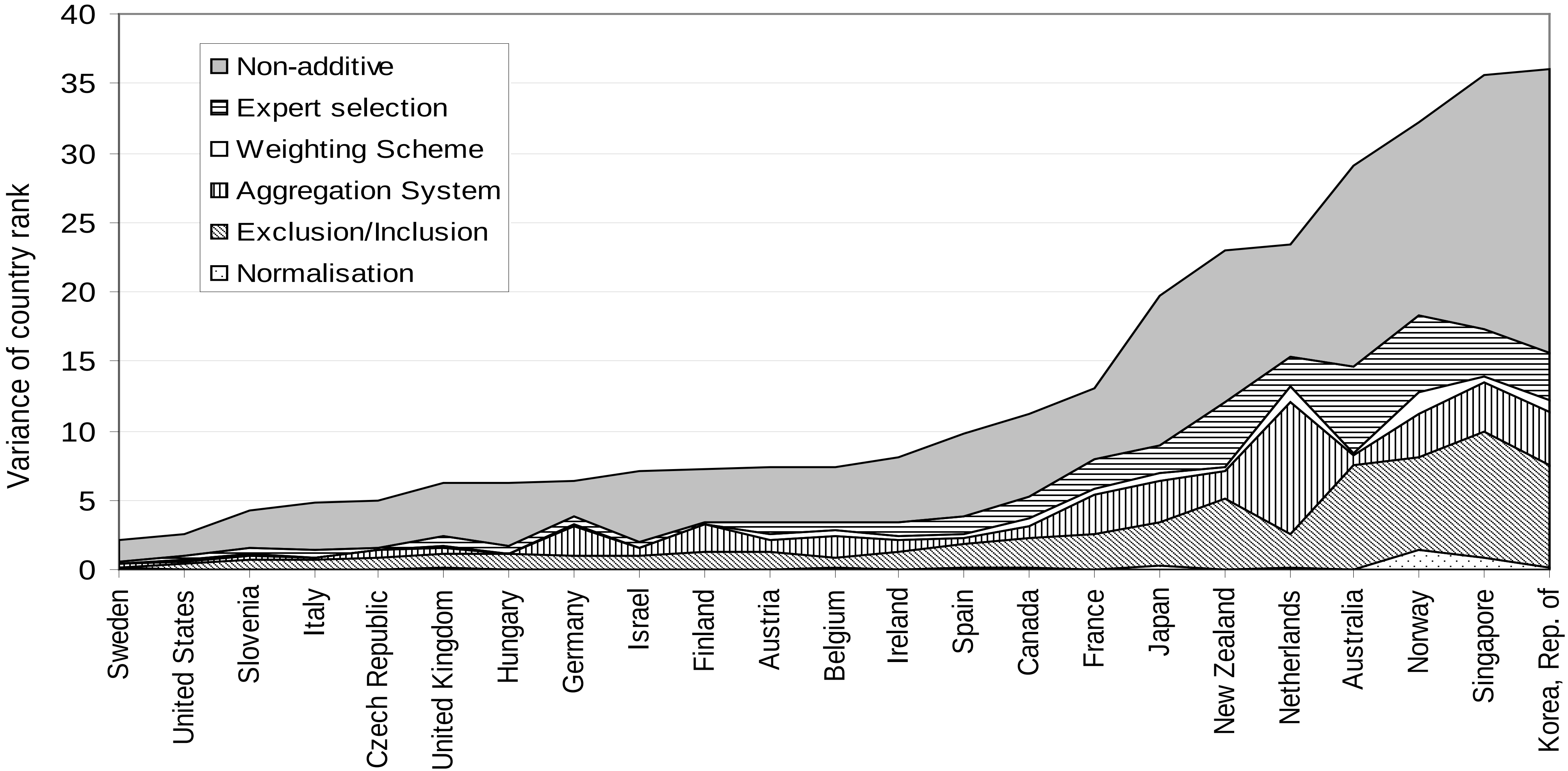
Uncertainty and sensitivity (UA, SA)



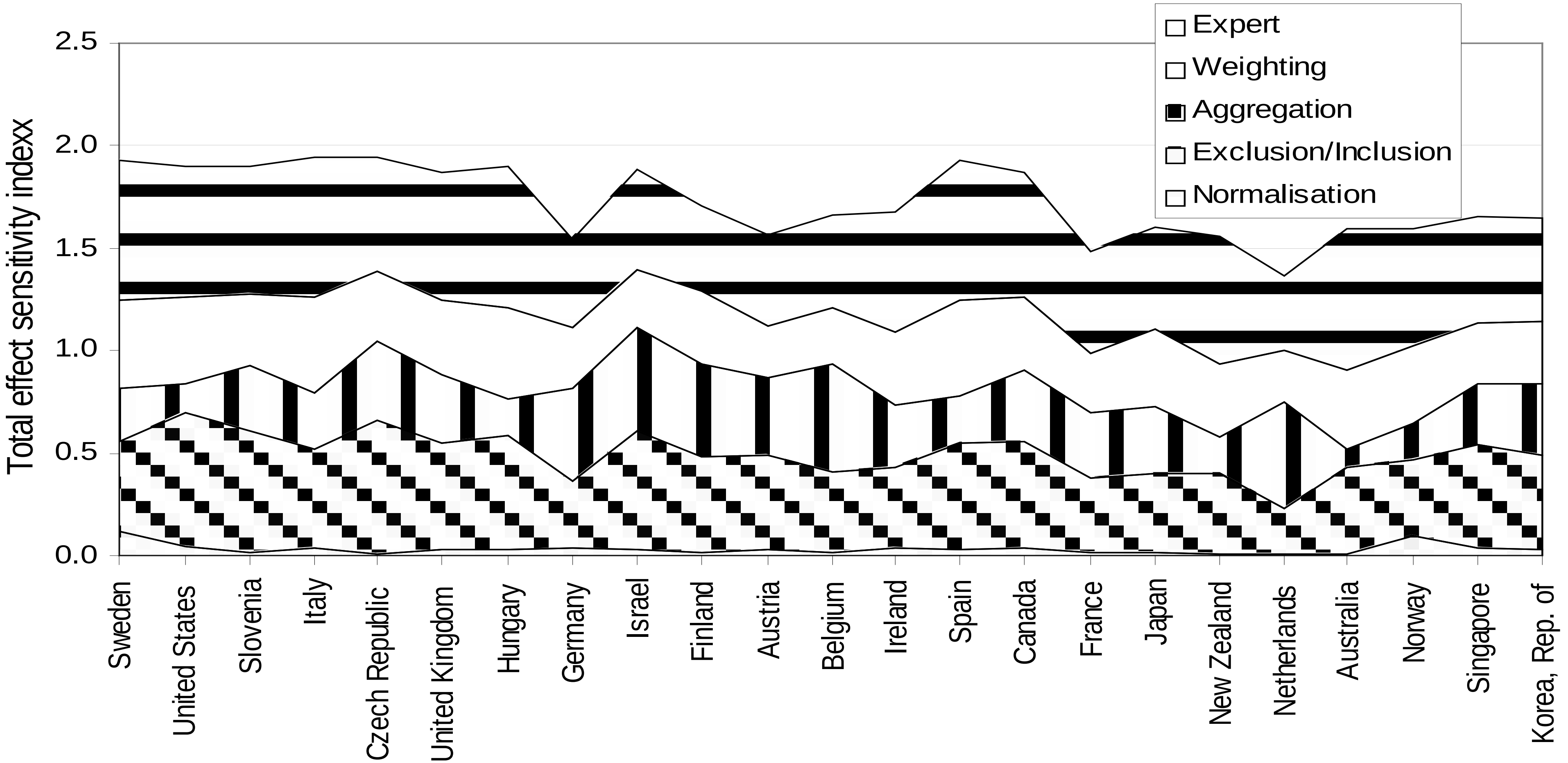


Index (IT) – Index (CZ)

Sensitivity of Technology Achievement Index (JRC-OECD Handbook) – Decomposition of variance into first order and interaction



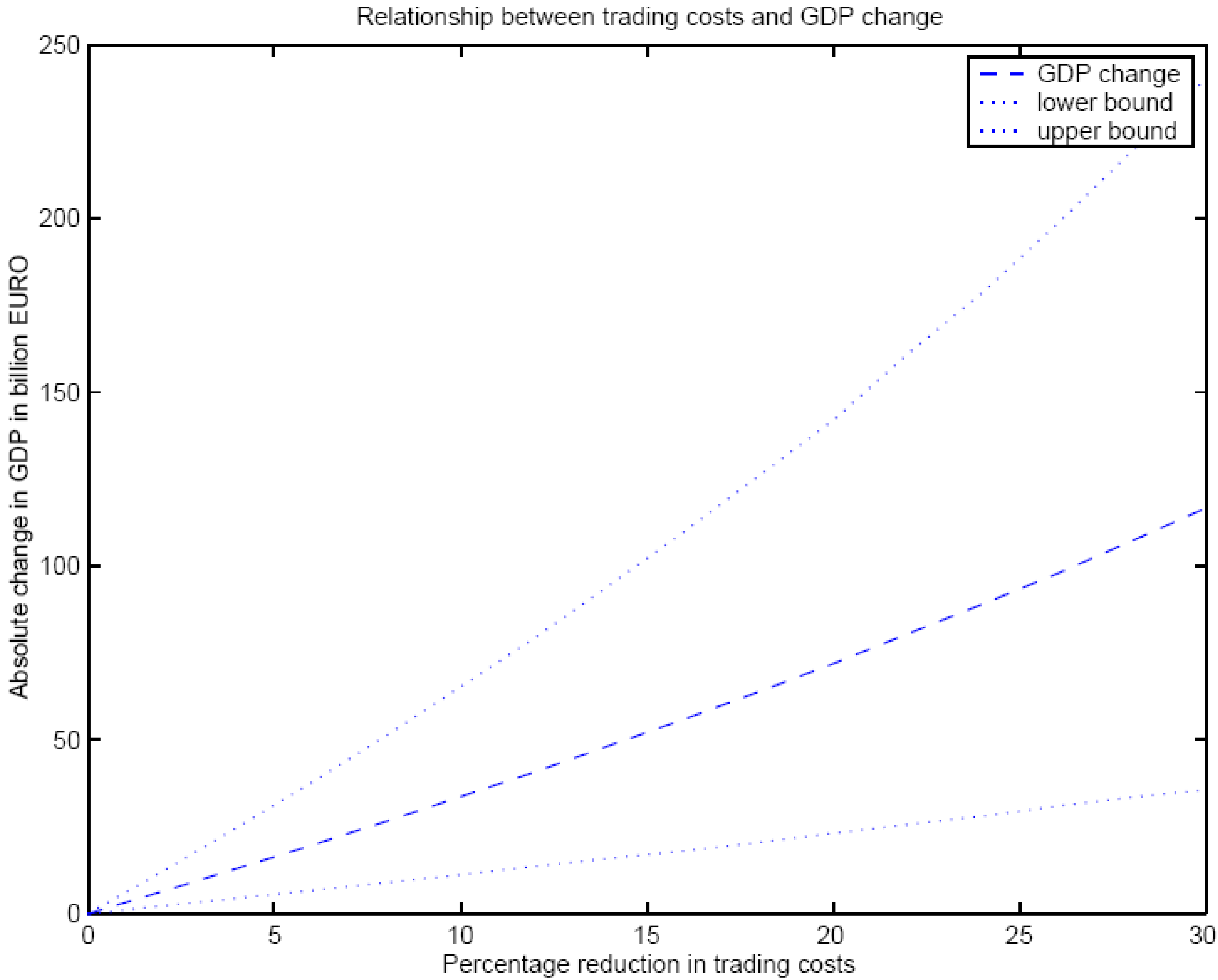
Sensitivity of Technology Achievement Index (JRC-OECD Handbook) - Total indices



Not always what works with the ‘customer’ is the method you like the best!

An example of
Impact assessment by JRC

Uncertaintv analvsis



International ranking of universities

THES World University Rankings

Jiao Tong ranking of World Universities

► Classement des meilleures universités dans le monde, par discipline

Rang	Institution	Pays	Note sur 100
Sciences naturelles et mathématiques			
1	Harvard	Etats-Unis	100
2	Berkeley	Etats-Unis	95,6
3	Princeton	Etats-Unis	92,9
24	Paris XI	France	62,1
Ingénierie et informatique			
1	MIT	Etats-Unis	100
2	Stanford	Etats-Unis	89,6
3	Urbana-Illinois	Etats-Unis	84,9
61	Bordeaux I	France	–
Sciences de la vie			
1	Harvard	Etats-Unis	100
2	MIT	Etats-Unis	75,6
3	California Un.	Etats-Unis	75,4
77	Paris VI	France	–
Médecine et pharmacie*			
1	Harvard	Etats-Unis	100
2	California Un.	Etats-Unis	82,8
3	Washington-Seattle	Etats-Unis	77,3
Sciences sociales			
1	Harvard	Etats-Unis	100
2	Chicago	Etats-Unis	95,5
3	Stanford	Etats-Unis	82,9
42	Insead	France	44,4

Source : Université Jiao Tong *France hors classement

What can we do to improve our position on the international scene?

Is it relevant to today’s discourse on HE reform?

In the press:

•France: moaning or questioning the validity of the proposed rankings.

•Italy: self-flagellation.

•Spain: a success (one Spanish university in the top 200)….

1. Number of alumni of an institution winning Nobel Prizes and Fields Medals, (10% weight);

2. Number of staff of an institution winning Nobel Prizes and Fields Medals, (20%);

3. Number of highly cited researchers in 21 broad subject categories (1981-1999), (20%);

4. Number of articles published in Nature and Science in 2007, (20%);

5. Number of articles in Science Citation Index-expanded, Social Science Citation Index in 2007 (20%);

6. Academic Performance (weighted scores of the above 5 indicators divided by the number of full-time equivalent academic staff, (10%).

1. Assumption on the weighting scheme:

Four weighting schemes:

- SJTU
- equal weighting
- factor analysis derived weights, and
- university-specific weighting (DEA).

Our analysis of the 2008 data

2. Assumption on the aggregation rule:

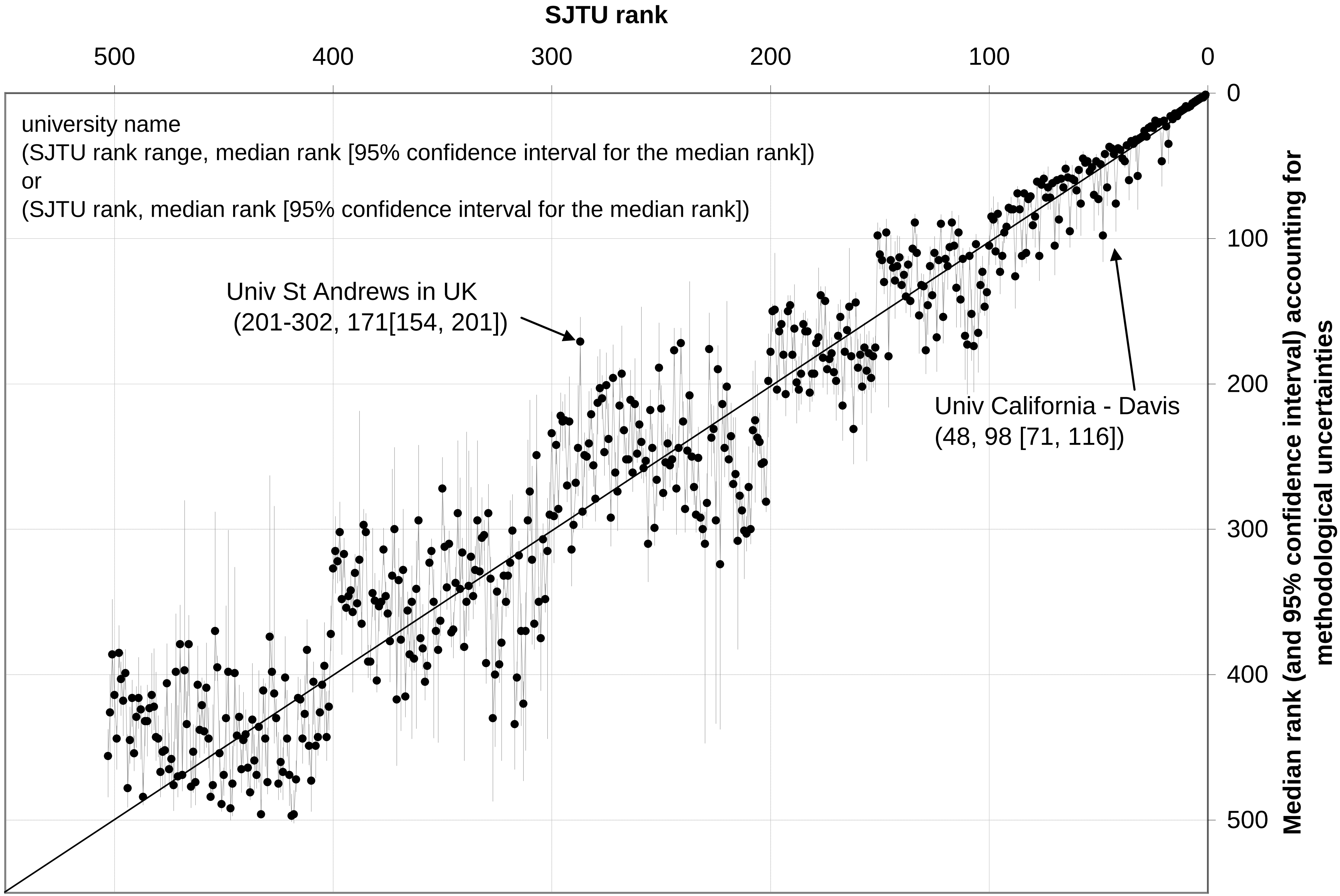
- original SJTU
- a geometric weighted average
- Borda (a multi-criteria analysis method)

Our analysis of the 2008 data

3. Assumption on the indicators:

- either kept all six indicators or exclude one-at-a-time (jackknife)

Before looking at the results … is this experiment too extreme? Only if we think that giving the 6 original SJTU variables to different teams would not give on return radically different rankings …


SJTU rank
500
400
300
200
100
0
university name
(SJTU rank range, median rank [95% confidence interval for the median rank])
or
(SJTU rank, median rank [95% confidence interval for the median rank])
Univ St Andrews in UK
(201-302, 171[154, 201])
Univ California - Davis
(48, 98 [71, 116])
0
100
200
300
400
500
Median rank (and 95% confidence interval) accounting for methodological uncertainties

Simulated rank range

	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50	51-55	56-60	61-65	66-70	71-75	76-80	81-85	86-90	91-95	96-100	101-105	106-110	111-115	116-120	121-125	126-130	SJTU rank	
Harvard Univ	100																										1	USA
Stanford Univ	89	11																									2	USA
Univ California - Berkeley	97	3																									3	USA
Univ Cambridge	90	10																									4	UK
Massachusetts Inst Tech (MIT)	74	26																									5	USA
California Inst Tech	27	53	19	1																							6	USA
Columbia Univ	23	77																									7	USA
Princeton Univ		71	9	11	7	1																					8	USA
Univ Chicago		51	34	13	1																						9	USA
Univ Oxford		99	1																								10	UK
Yale Univ		47	53																								11	USA
Cornell Univ		27	73																								12	USA
...	...																										...	
Univ California - San Francisco			14	9	14	3	11	3	7	10	4	3	3	3		6			1	6		1					18	USA
...	...																										...	
Duke Univ					10	6	13	11		6	3	7	6	3	1	3	1	9	9	7	1	3			1		32	USA
Rockefeller Univ				4	10	23	26	1		3	3	3	3	3	4	4	6	3	1	1				1			32	USA
Univ Colorado - Boulder						19	39	30	11	1																	34	USA
Univ British Columbia						20	60	20																			35	Canada
Univ California - Santa Barbara						9	9	10	3	10	6	7	6		11	4	6	3	4	7		1	1				36	USA
Univ Maryland - Coll Park						6	37	44	9	4																	37	USA
...	...																										...	
Ecole Normale Super Paris						7	9	4	6	7	6	4	9	6	7	4	3	3	4	3	3			1	6	4	73	France
Univ Melbourne												1	20	17	31	23	1	6									73	Australia
Univ Rochester								1	10	7	16	24	14	10	10	6	1										73	USA
Univ Leiden									3	6	9	23	24	13	14	9											76	Netherlands
...	...																										...	
Univ Sheffield										1	21	26	21	9	13	7	1										77	UK
Tohoku Univ										4	1	7	1		4	17	19	3	3	3		19	7	3	4	4	79	Japan
Univ Utah											4	4	6	1	4	9	6	16	7	13	4	9	6	6	1		79	USA
King's Coll London											4	6	9	29	17	14	10	1	6	3	1						81	UK
Univ Nottingham											1	6	10	21	21	10	17	7	4	1							82	UK
Boston Univ													3	1	6	3	6	11	1	4	3	13	14	10	10	10	83	USA
...	...																										...	

Legend:

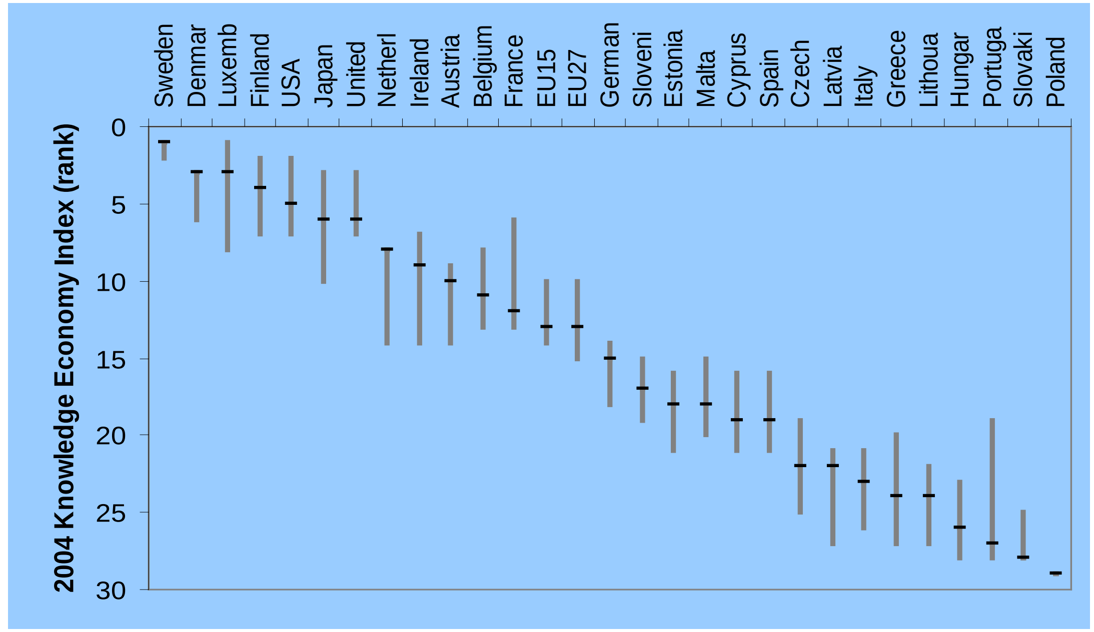
Frequency lower 15%

Frequency between 15 and 30%

Frequency between 30 and 50%

Frequency greater than 50%

Is this fragile? Compare with another index



KEI

Knowledge Economy Index

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6	Rank 7	Rank 8	Rank 9	Rank 10	Rank 11	Rank 12	Rank 13	Rank 14	Rank 15	Rank 16	Rank 17	Rank 18	Rank 19	Rank 20	Rank 21	Rank 22	Rank 23	Rank 24	Rank 25	Rank 26	Rank 27	Rank 28	Rank 29
Sweden	54	46																											
Denmark			55	30	14																								
Luxembourg	36	4	14	25	4	7	7	4																					
Finland		18	23	29	9	11	11																						
USA	11	32	2	4	39	9	4																						
Japan			4	7	18	32	36			4																			
UK			2	5	16	38	39																						
Netherlands								86	4		4			7															
Ireland							4		61	14	4	9	9																
Austria									18	50	18	7		7															
Belgium								11	4	11	57	16	2																
France						4			14	18	11	54																	
EU15											4		57	39															
EU25										4	4	14	32	39	7														
Germany														7	79	4	7	4											
Slovenia															7	41	38	14											
Estonia																4	36	25	21	11	4								
Malta															7	13	9	21	23	27									
Cyprus																36	7	4	23	23	7								
Spain																4	4	32	25	29	7								
Czech. Rep.																			4	7	30	39	5	7	7				
Latvia																					20	36	11	21	7		5		
Italy																					29	18	9	29	9	7			
Greece																				4	4	4	29	18	21	7	14		
Lithuania																						4	41	13	32	11			
Hungary																							2	13	13	57	2	14	
Portugal																			4				4		7	11	61	14	
Slovakia																									4	7	18	71	
Poland																													100

Legend:

Frequency lower 15%

Frequency between 15 and 30%

Frequency between 30 and 50%

Frequency greater than 50%

Back to SJTU: It is beyond doubt that Harvard, Stanford, Berkley, Cambridge, and MIT are top 5

(both in the original SJTU and in more than 80% of our simulations) ...

… Still for 96% of the universities, the range of ranks is greater than 10 positions.

Examples of rank variation

- •92 positions (Univ Autonoma Madrid) and 277 positions (Univ Zaragoza) in Spain,
- •71 positions (Univ Milan) and 321 positions (Polytechnic Inst Milan) in Italy,
- •22 positions (Univ Paris 06) and 386 positions (Univ Nancy 1) in France.

questions: