ESTIMATION OF CONDITIONAL EXPECTATIONS AND CONDITIONAL VARIANCES WITH STATE-DEPENDENT PARAMETER MODELS

Ratto* M. A. Pagano* and P.C. Young**

*Joint Research Centre,
European Commission

**Lancaster University, UK;
Australian National University, Australia

Summary

European Commission

We aim at the joint estimation of

- conditional expectations and
- conditional variances

using the state-dependent parameter (SDP) metamodelling approach (emulation, non-parametric RS-HDMR).

For applications in global sensitivity analysis (GSA) and in the more general metamodelling framework.

Intro

European Commission

GSA: analyze the mapping $Y=f(X_1,\dots,X_k)$ to quantify the relative contribution of each input factor X_j to the uncertainty of Y.

Meta-modeling: build a statistical approximation $f^*(\mathbf{X})$ to the computational model $f(\mathbf{X})$ that is sufficiently accurate to be used in place of the original one (operationally equivalent) and that, at the same time, can be computed in a much faster way.

Variance-based GSA

Variance based sensitivity indices of single factors or of groups of them are defined as:

$$S_{\mathbf{I}} = \frac{V[E(Y \mid \mathbf{X}_{\mathbf{I}})]}{V(Y)}$$

$$\mathbf{I} = (i_1, \ldots i_g)_{1 \le g \le k}$$

and tell the portion of variance of Y that is explained by $\mathbf{X_I}$

GSA and meta-modelling

European Commission

1. the approximation f*(**X**) can be used to compute sensitivity indices in place of the original computational mapping f(**X**)

2. the variance based sensitivity measures can be interpreted as the non-parametric R^2 or correlation ratio, used in statistics to measure of the explanatory power of covariates in regression

[Memo: the inner argument $E(Y|\mathbf{X}_I)$ in V_I is the function of the subset of input factors $\mathbf{X}_I$ that approximates f(**X**), by minimizing a quadratic loss, i.e. maximizing the R^2]

GSA and meta-modelling

Estimating functions $E(Y|\mathbf{X}_I)$ provides a route for both a model approximation and sensitivity estimation;

Smoothing methods to estimate $E(Y|\mathbf{X}_I)$ are becoming a popular approach to sensitivity analysis:
Storlie et al. review (2006), Li et al. (RS-HDMR, 2002-2006); Ratto et al. (SDP, 2004-2006), Levandowski, (polynomial, 2007).

Kriging and Gaussian emulators (Oakley O'Hagan, 2004; Kleijenen, 2006-2007)

GSA and meta-modelling

SDP modeling is one class of non-parametric smoothing approach first suggested by Young (1993).

The estimation is performed with the help of the `classical' recursive (non-numerical) Kalman filter and associated fixed interval smoothing algorithms and has been applied for sensitivity analysis in Ratto et al. (2004-2006)

SDP meta-modelling

European Commission

In great summary a state dependent model approximating based on a Monte Carlo sample of dimension N can be written as

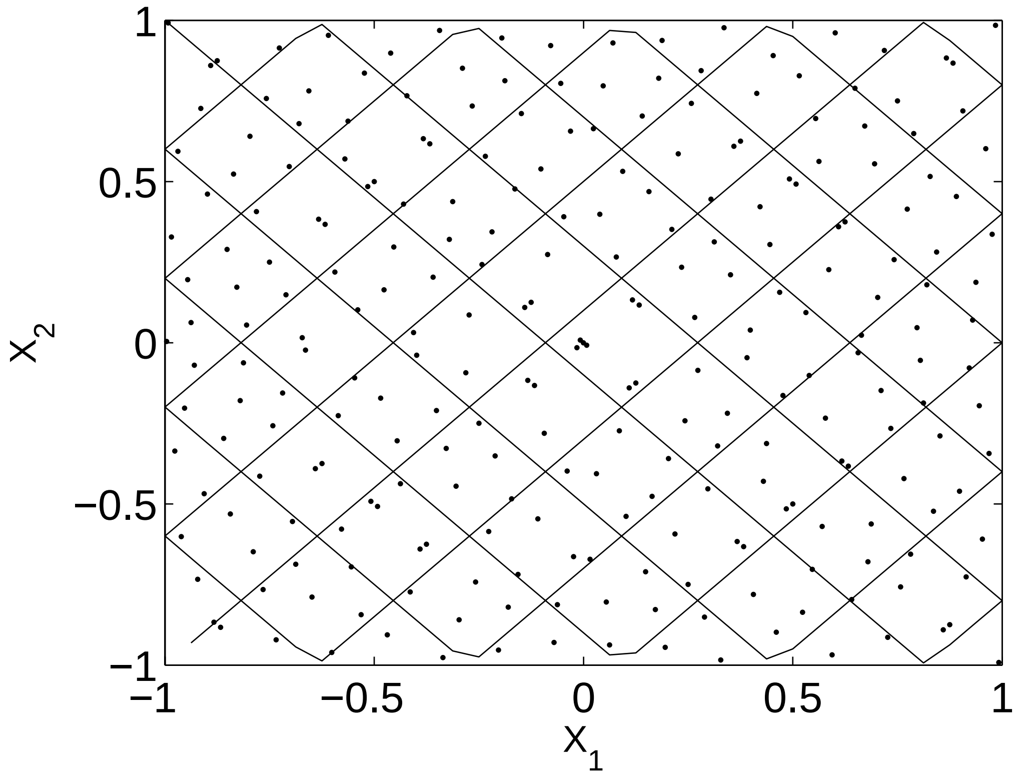
$$Y_t = E(Y \mid \mathbf{X}_{\mathbf{I},t}) + e_t = p_{\mathbf{I},t}(s_{\mathbf{I}}) + e_t$$

e_t: 'observation noise'

$p_{\mathbf{I},t}$: state dependent parameter, depending on a state variable $s_{\mathbf{I}}$ that moves according to a generalised sorting strategy t=1,…,N along the co-ordinates of the single factor or group of factors indexed by ***I***.

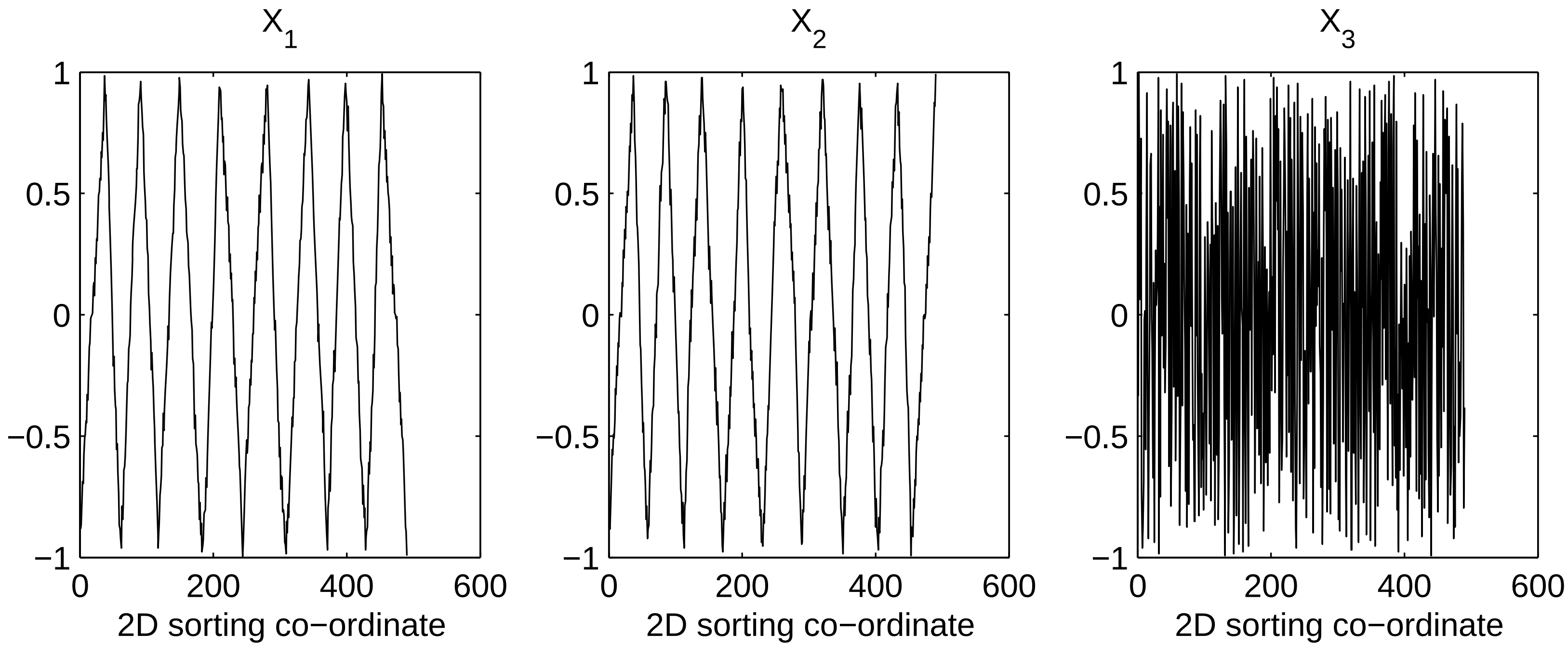
SDP 'generalised sorting'

European Commission



SDP 'generalised sorting' (cnt'd)

European Commission



SDP meta-modelling

European Commission

Alternative/generalised specification

$$Y_t = E(Y \mid \mathbf{X}_{\mathbf{I},t}) + e_t = p_{\mathbf{I},t}(s_{\mathbf{I}})\phi(\mathbf{X}_{\mathbf{I},t}) + e_t$$

e_t: 'observation noise'

$p_{\mathbf{I},t}$: state dependent parameter,

$\phi(X_I)$: may be a function of X_I (e.g. polynomial) used to approximate Y

SDP meta-modelling (cnt'd)

European Commission

The estimation of $E(Y|\mathbf{X}_I)$ reduces to the extraction of the low frequency component of the sorted output Y. The SDP's are modelled by one member of generalised random walk (GRW) class of non-stationary processes.

For instance, the integrated random walk (IRW) process turns out to provide good results, since it ensures that the estimated SDP relationship has the smooth properties of a cubic spline.

SDP model

European Commission

Given the IRW characterisation, the SDP model is put into state space form as:

$$Y_t = p_{\mathbf{I},t}\phi(\mathbf{X}_{\mathbf{I},t}) + e_t$$

$$p_{\mathbf{I},t} = p_{\mathbf{I},t-1} + d_{\mathbf{I},t-1}$$

$$d_{\mathbf{I},t} = d_{\mathbf{I},t-1} + \eta_{\mathbf{I},t}$$

e_t (observation noise) ~ $N(0,\sigma^2)$

η_t (system disturbances) ~ $N(0, \sigma(\eta_I)^2)$

SDP model estimation

European Commission

Noise-variance-ratio: NVR = $\sigma(\eta_I)^2/\sigma^2$

NVR's are optimised by maximum likelihood (ML), using prediction error decomposition.

The SDP approach is couched with optimal ML estimation. For example, Random Walk (RW) or Smoothed Random Walk (SRW) might be preferable in certain circumstances because they yield less smooth estimates. Indeed, if any sharp changes or jumps seem possible, then these can be handled using `variance intervention' (see Young and Ng, 1989).

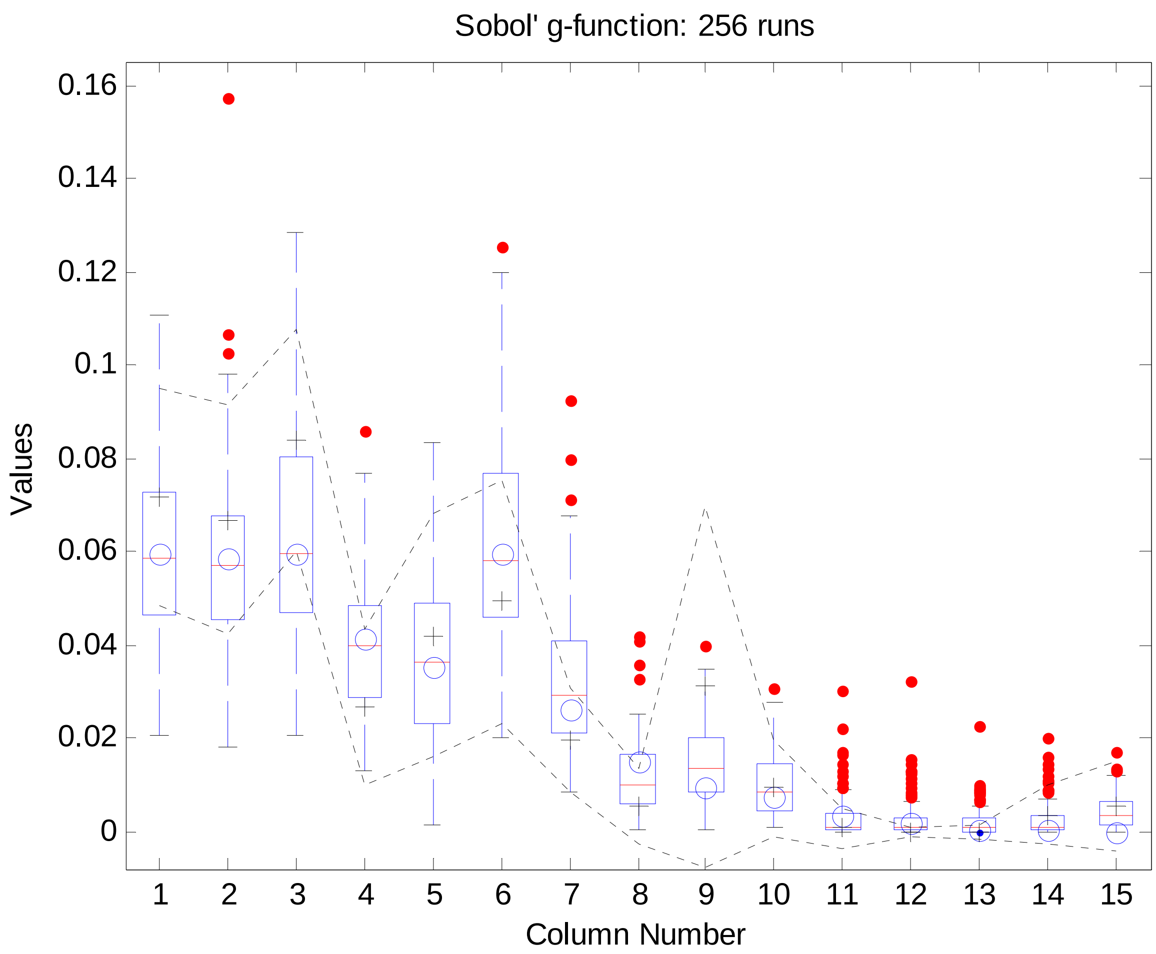
SDP properties

European Commission

1. Couched with optimal ML estimation;
2. Provides standard errors of all quantities involved: $E(Y|X_I)$, S_I
3. Flexible: Random Walk (RW) or Smoothed Random Walk (SRW) preferable in certain circumstances because they yield less smooth estimates. If any sharp changes or jumps seem possible, then these can be handled using `variance intervention' (see Young and Ng, 1989) of heteroscedastic noise assumptions (here).

SDP: st. errors

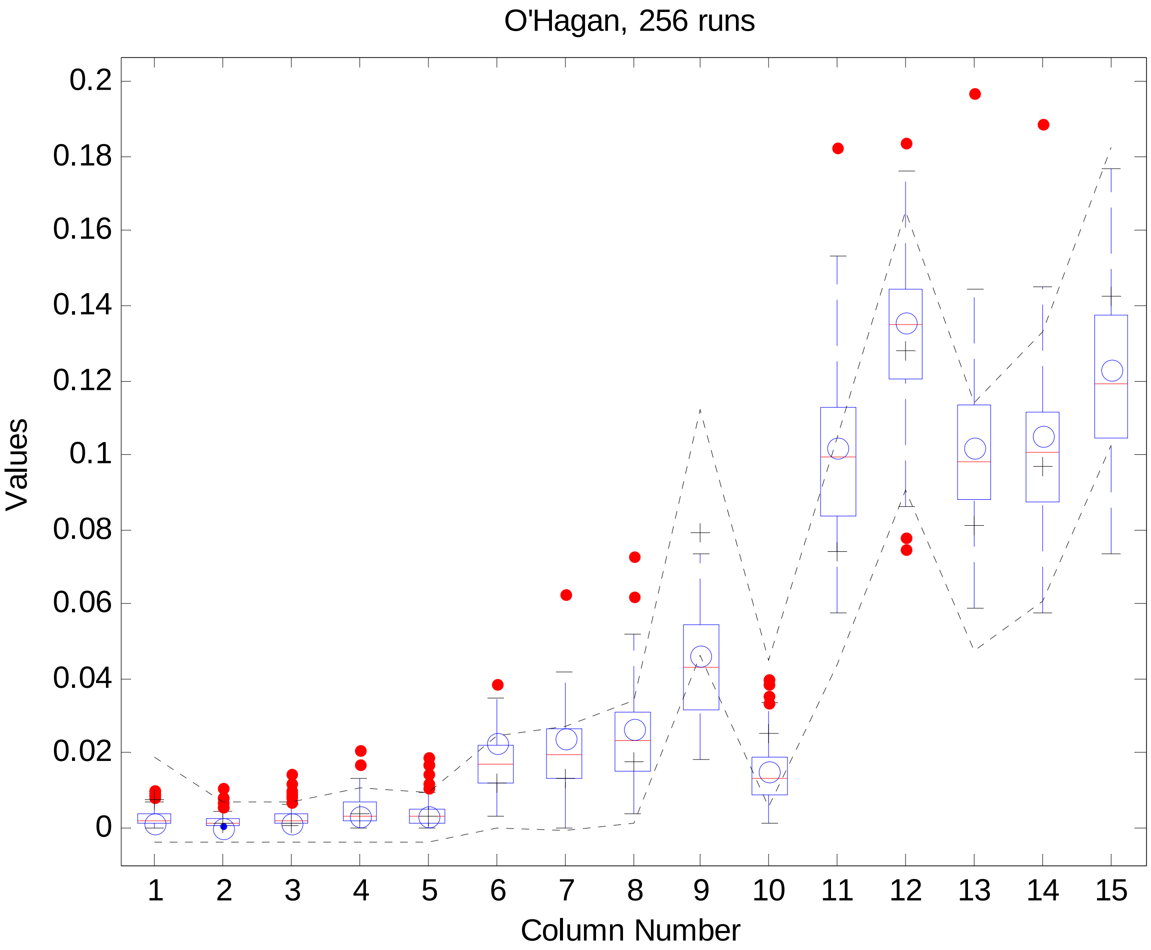
European Commission



http://eemc.jrc.ec.europa.eu

SDP: st. errors

European Commission



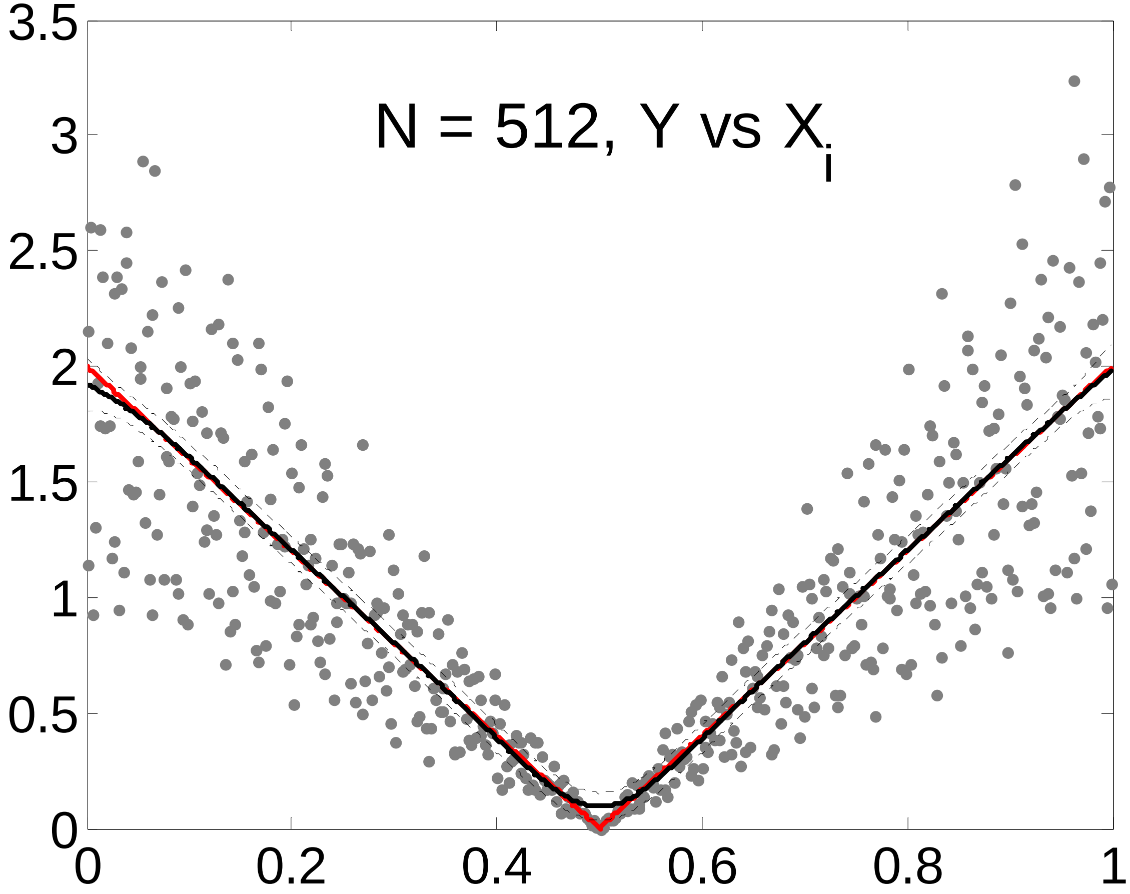
SDP flexibility

Let us consider those effects that cannot attributed to shifts in the mean:

they are not accounted for by $E(Y|X_i)$ and the related variance-based sensitivity indices

SDP: conditional expectation

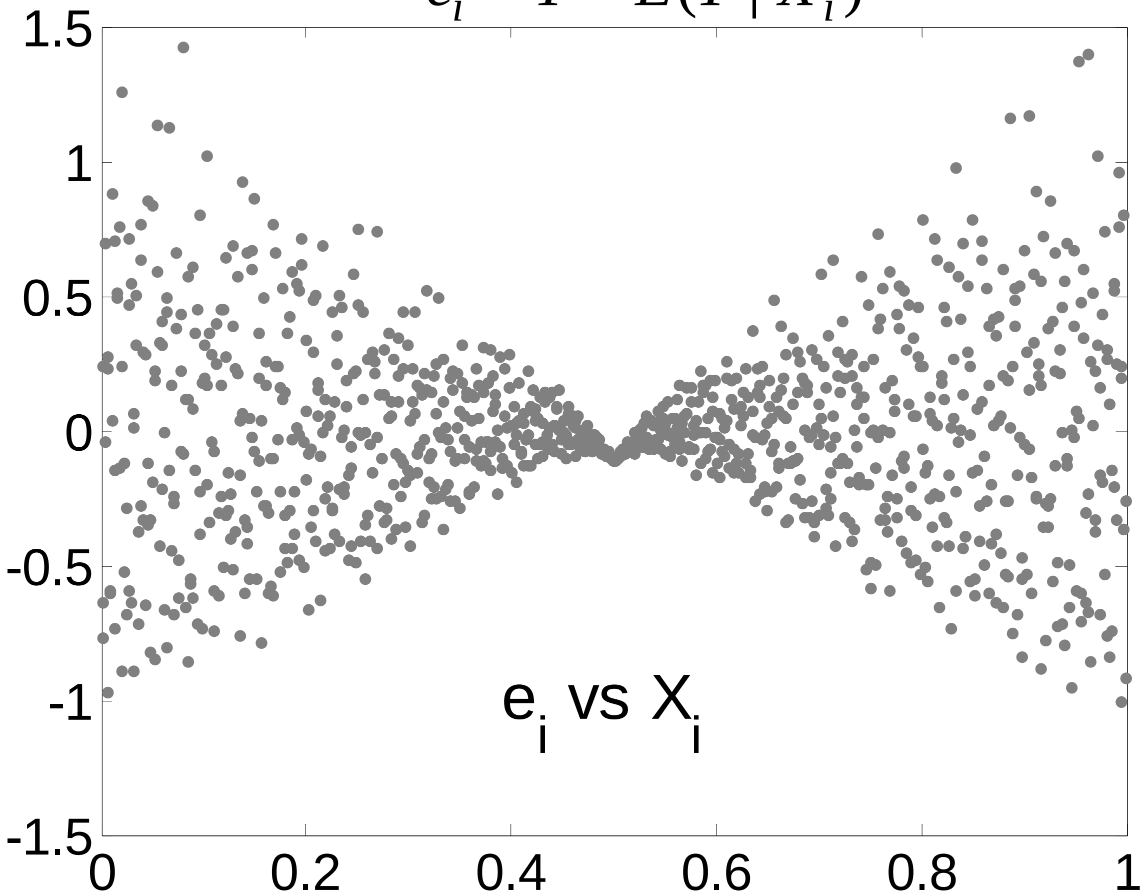
European Commission



SDP: residuals

European Commission

$$e_i = Y - E(Y \mid X_i)$$



SDP: conditional variance (heteroscedastic noise)

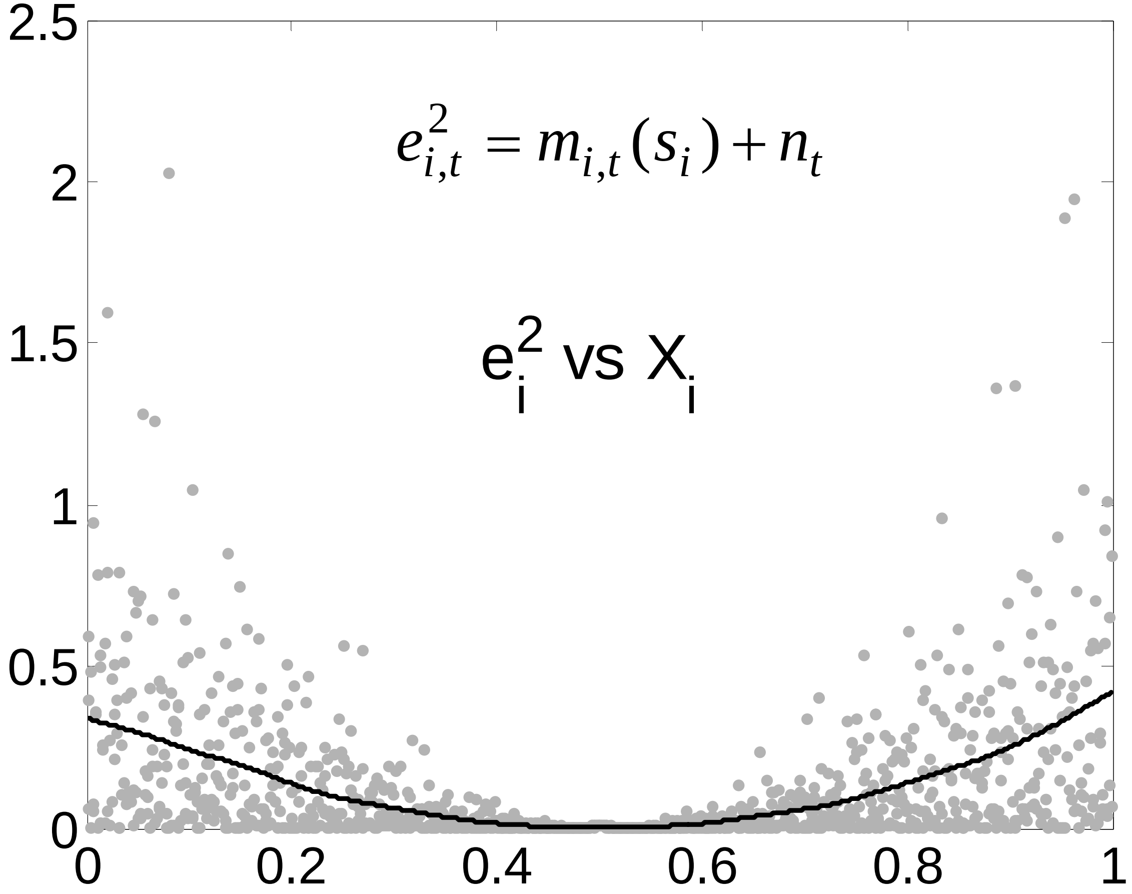
European Commission

$$V(Y \mid X_i) = E(Y^2 \mid X_i) - (E(Y \mid X_i))^2 =$$
$$E((Y - E(Y \mid X_i))^2 \mid X_i) = E(e_i^2 \mid X_i)$$

e_i is exactly the residual of the estimated first order HDMR term

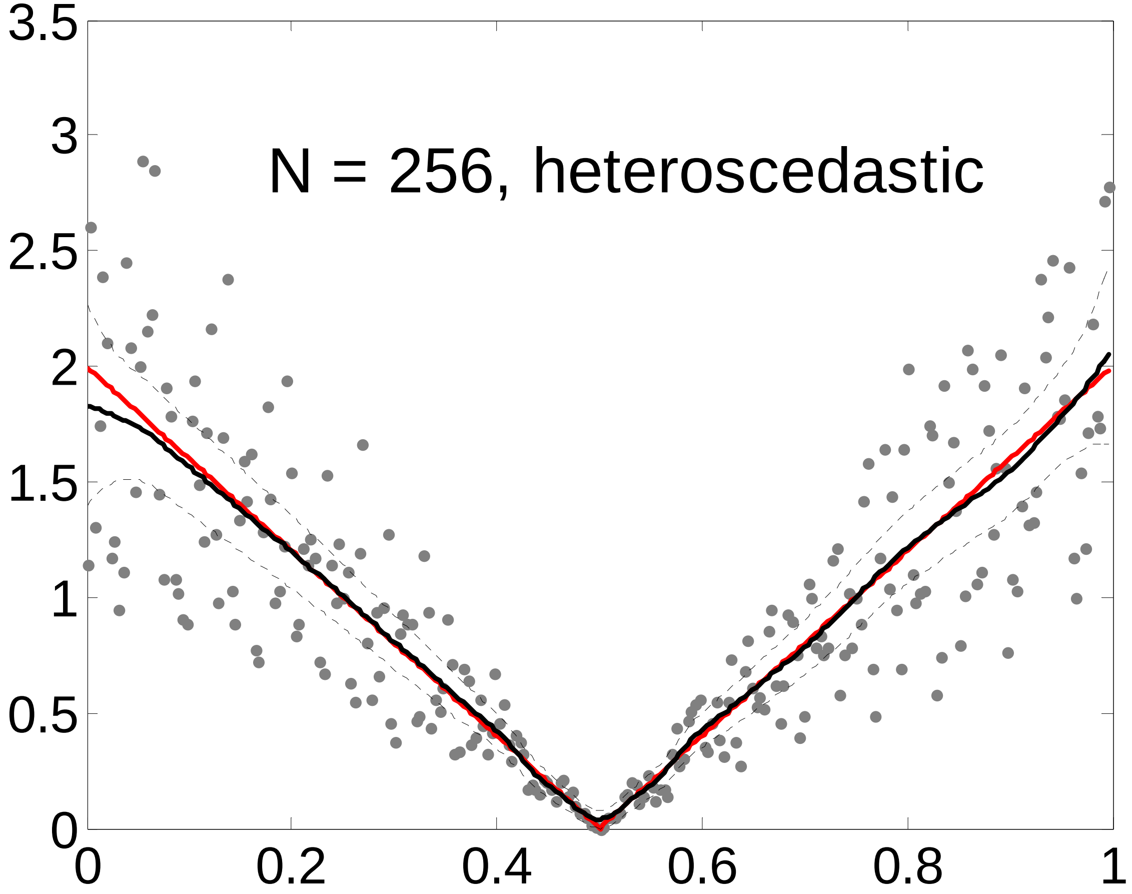
SDP: conditional variance (heteroscedastic noise)

European Commission



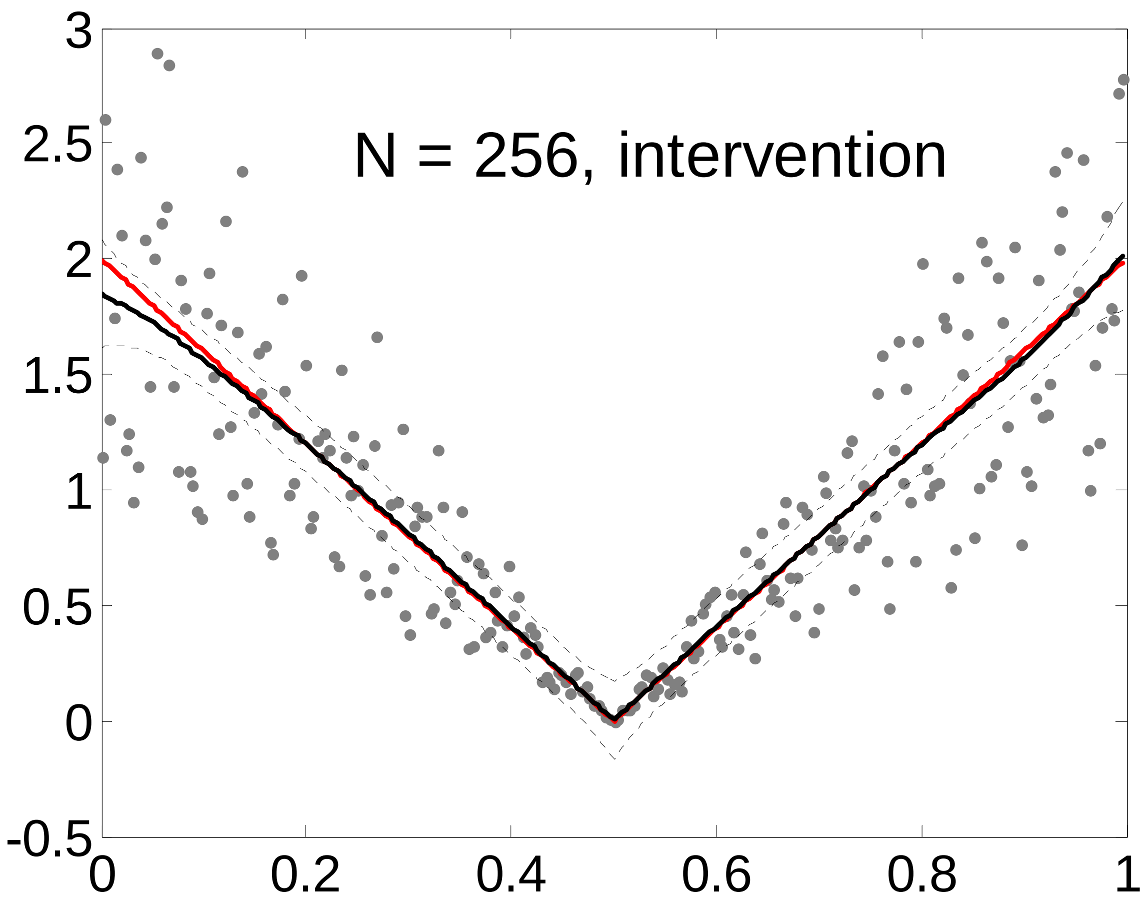
SDP: conditional expectation (heteroscedastic noise)

European Commission

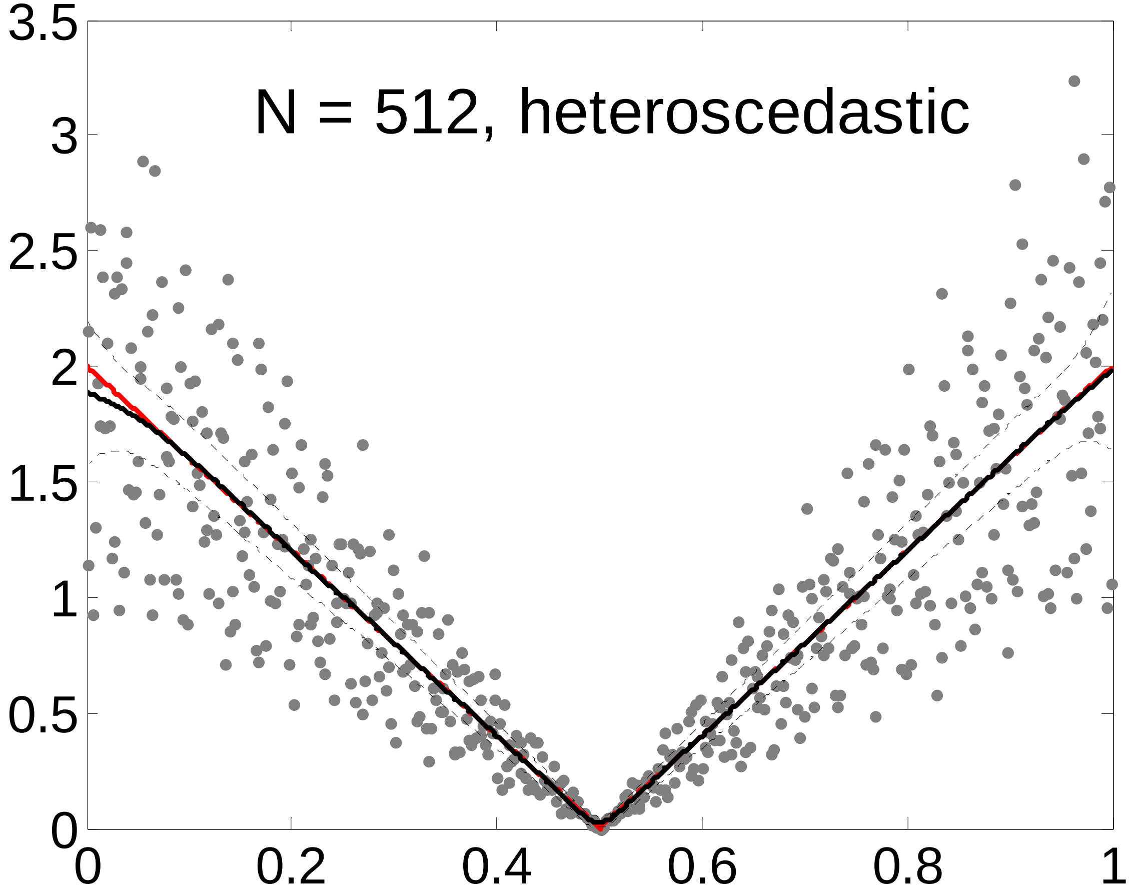


SDP: conditional expectation (intervention)

European Commission

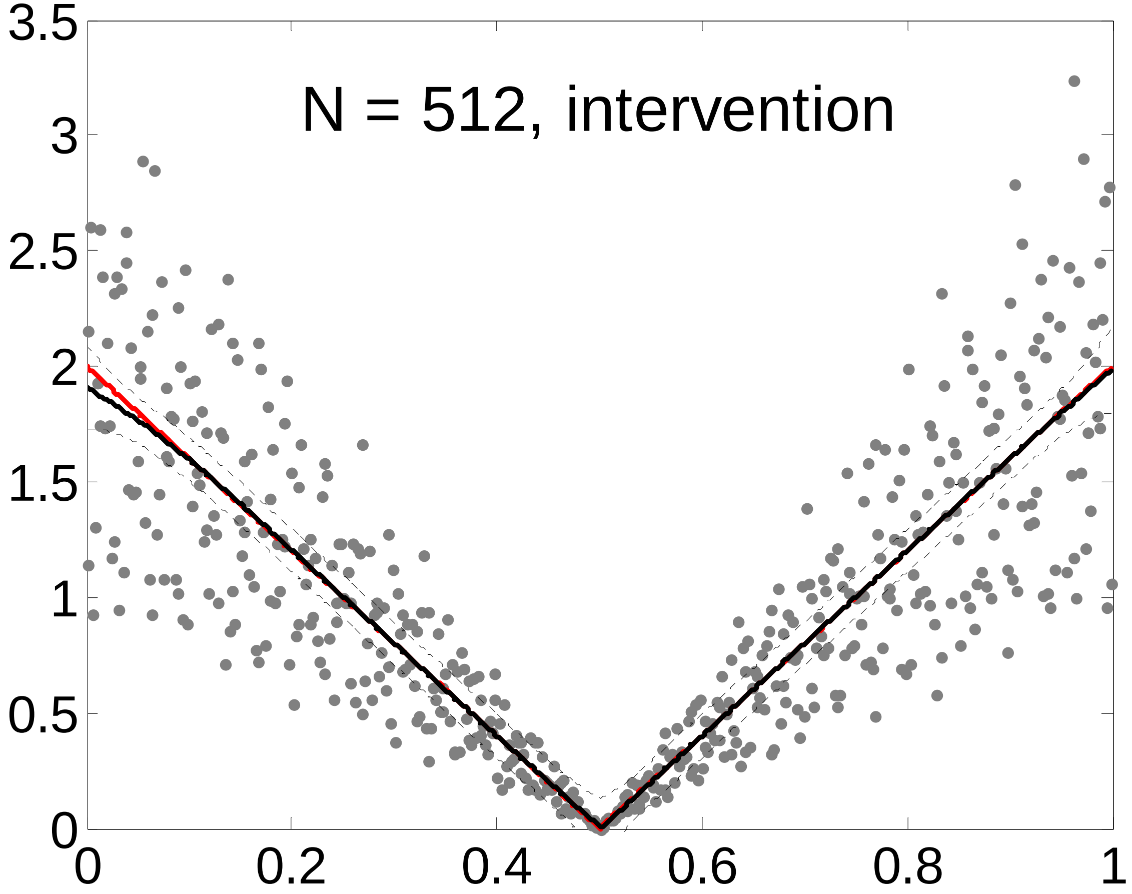


SDP: conditional expectation (heteroscedastic noise)



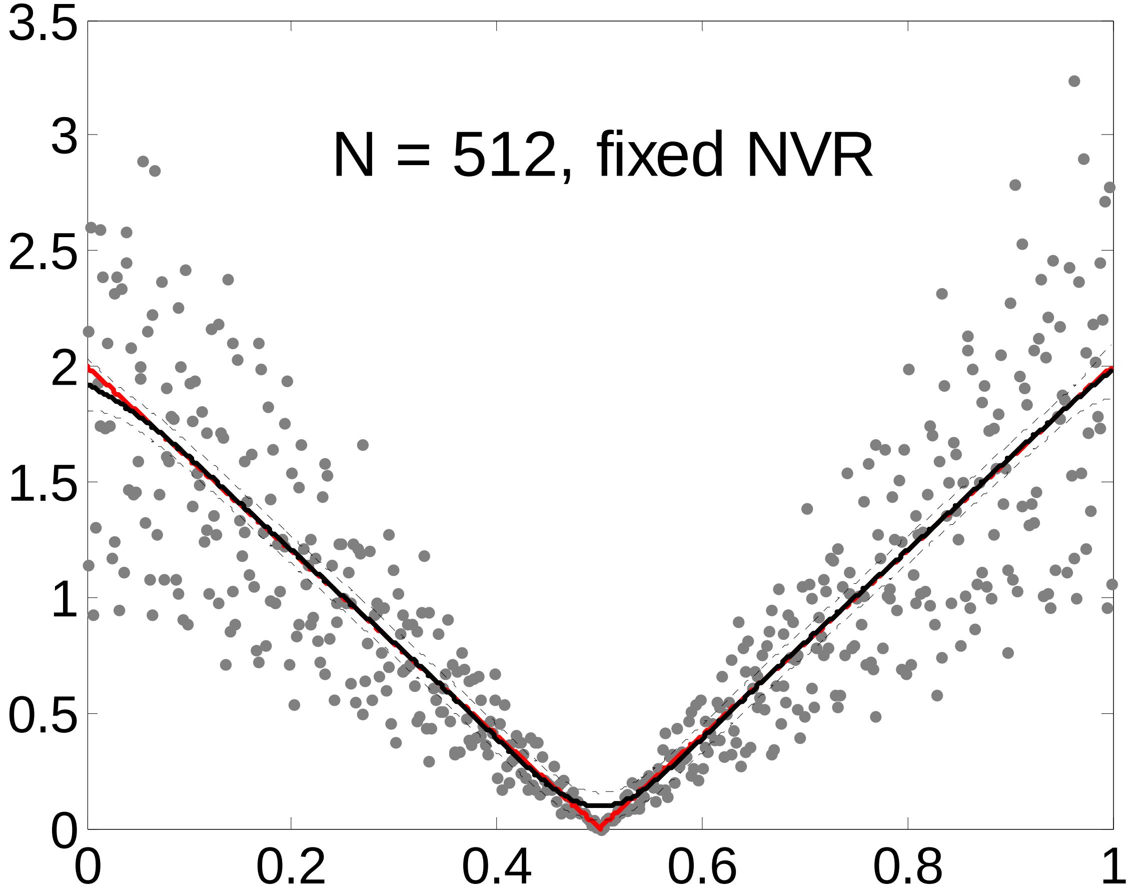
SDP: conditional expectation (intervention)

European Commission



SDP

European Commission



SDP: conditional variance (heteroscedastic noise)

European Commission

With no heteroscedastic noise, the sharp change at X1=0.5 is not within the SE's even at 1024 runs,

while it is already within bounds at 256 runs with heteroscedastic noise.

SDP + polynomials

European Commission

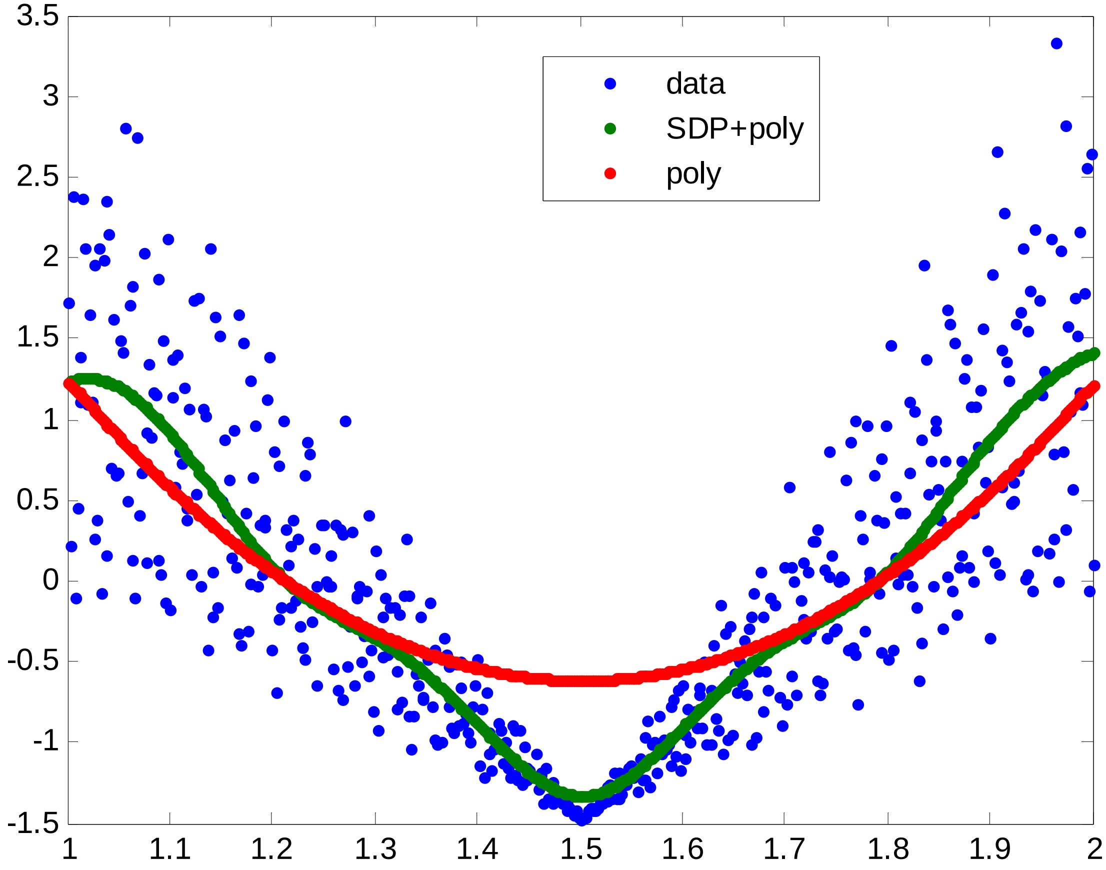
Alternative/generalised specification

$$Y_t = E(Y \mid \mathbf{X}_{\mathbf{I},t}) + e_t = p_{\mathbf{I},t}(s_{\mathbf{I}})\phi(\mathbf{X}_{\mathbf{I},t}) + e_t$$

Assume $\phi(X_I)$ = QUADRATIC POLYNOMIAL

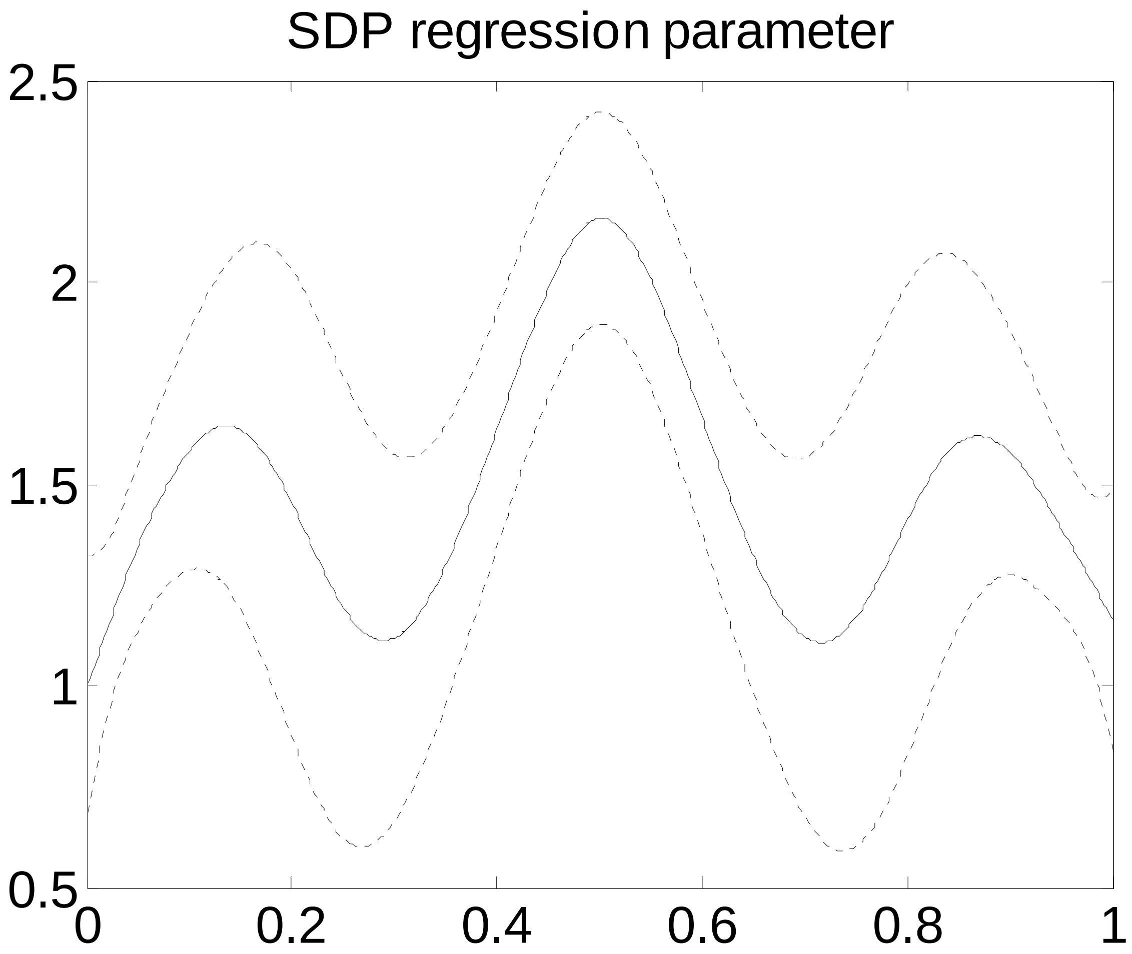
SDP + polynomials

European Commission



SDP + polynomials

European Commission



SDP: CPU cost

1 output, 1024 runs, 22 parameters:

1. no ML estimate of NVR's: 1.53 s;
2. ML estimate of NVR's: 98 s;
3. ML estimate of NVR's with heteroscedastic noise: 143 s.

Conclusions

European Commission

1. SDP modelling allows for effectively treat smoothing processes with heteroscedastic noise (i.e. linked to higher order interactions) and allows to 'fix' one main criticism to variance-based methods;
2. SDP can be used as a generalisation of explicit RS-HDMR;
3. No particular computational needs;
4. Software available soon.