Criminal activity at CAMS:

Henri Berestycki (CAMS, head of laboratory)

Marc Barthélémy (CEA and CAMS)

Jean-Pierre Nadal (LPS, ENS and CAMS)

Collaborations

Silvio Franz (ICTP, Trieste and Paris XI)

Mirta B. Gordon (TIMC-IMAG, Grenoble)

Roberto Iglesias (UFRGS, Porto Alegre)

Leila Kébir (INRIA, former postdoc at EHESS, Paris)

Viktoriya Semeshenko (postdoc, UFRGS Porto-Alegre)

Projects

- Paris: where do the offenders come from?
- Data from the Préfecture de Paris
- Years: 2000 and 2004
- Preliminary result:
- half of the illegal acts committed by people living in Paris, half from people living in the suburbs.
- Within Paris, offenders most often commits crime where they live (same arrondissement)

- Criminal behaviour and social influence

  see Mirta B. Gordon 's talk

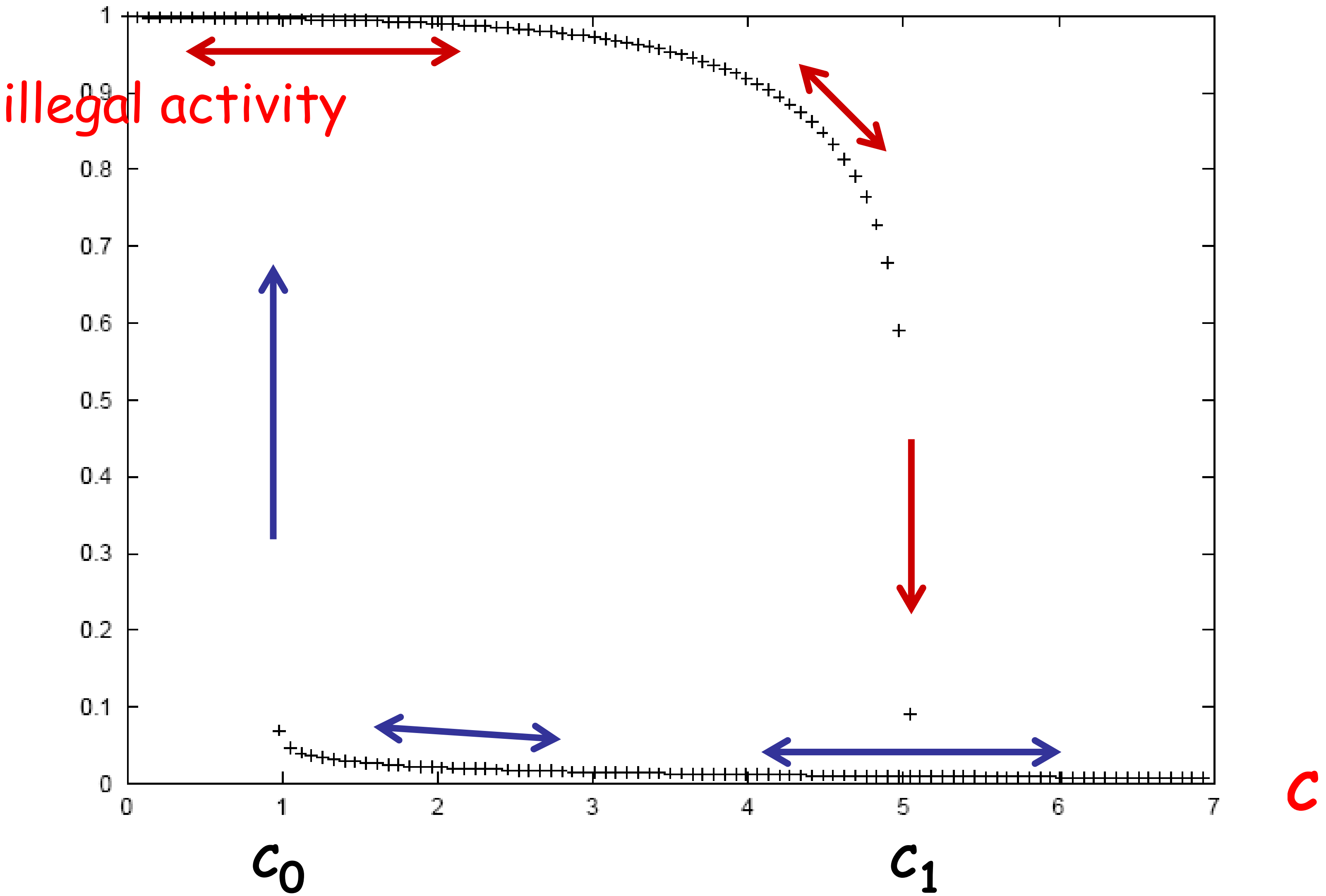
- Local social influence:

  based on T. C. Schelling's «dying seminar » (1973, 1978)
  and Granovetter analysis of riot formation (1978)

  both PDE and discrete choice approaches; « hot spots »

Level of illegal activity *vs.* Cost: hysteresis

multi equilibria



Global trend and local fluctuations

- Time series statistics:
  total number of illegal acts reported by the police

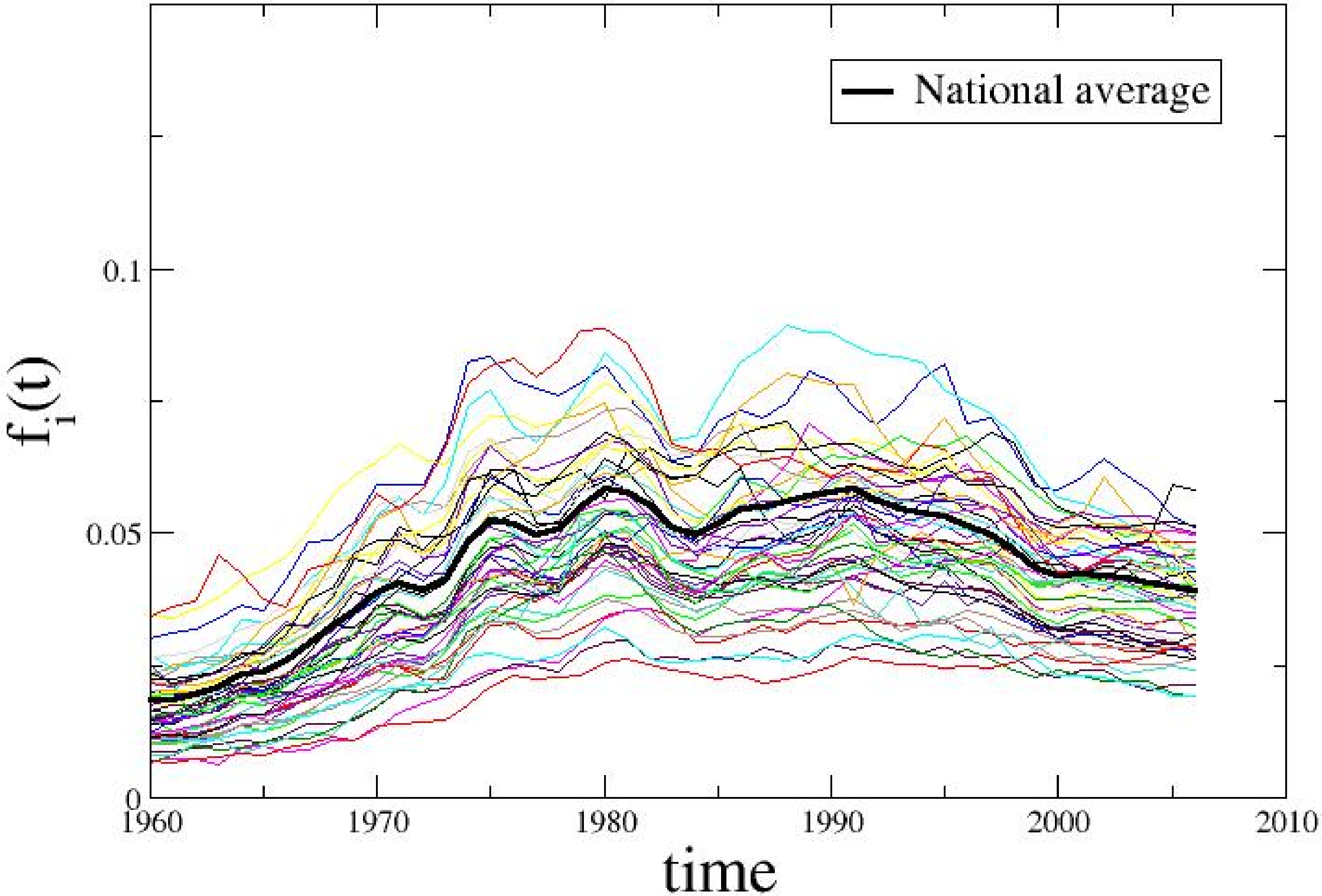
- National crime rate

- Local crime rates: States (USA), Départements (France)

work with Marc Barthélémy and Henri Berestycki

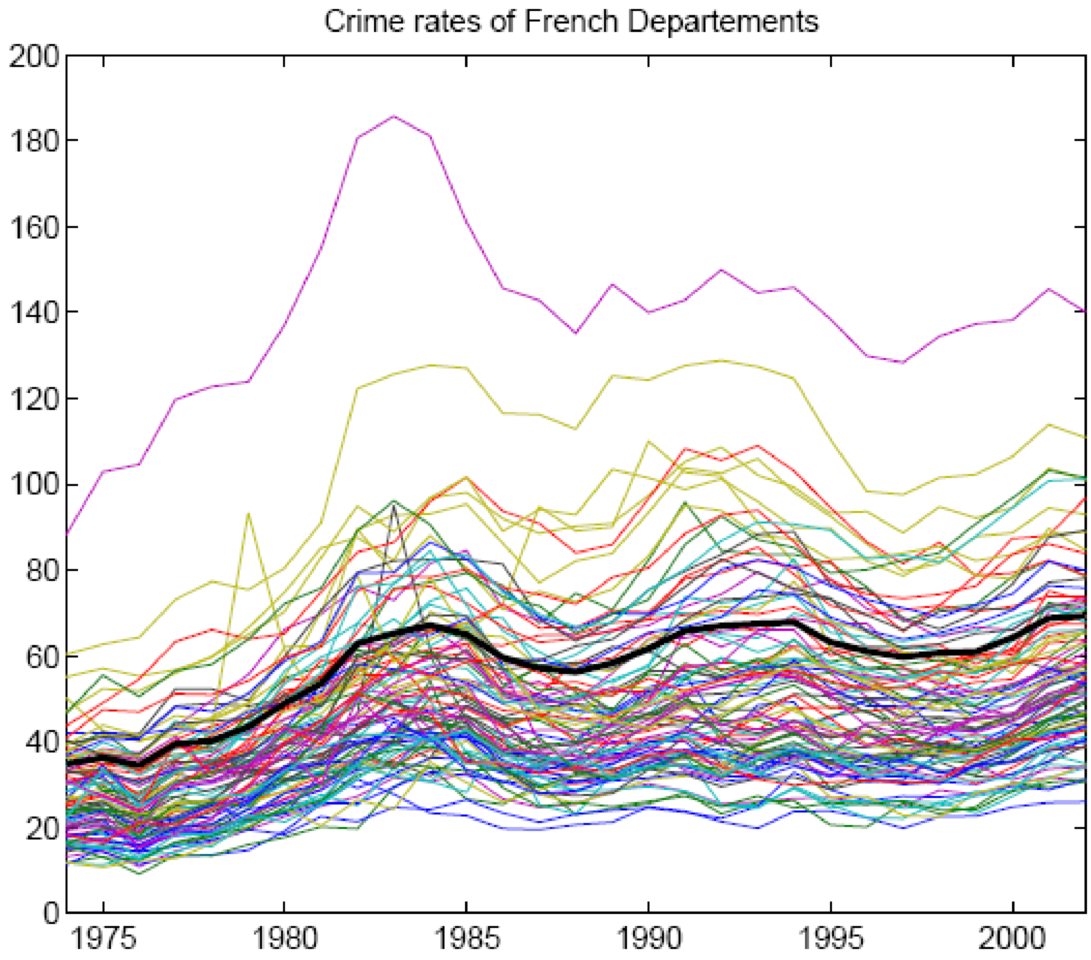
Data - USA: states crime rates

•



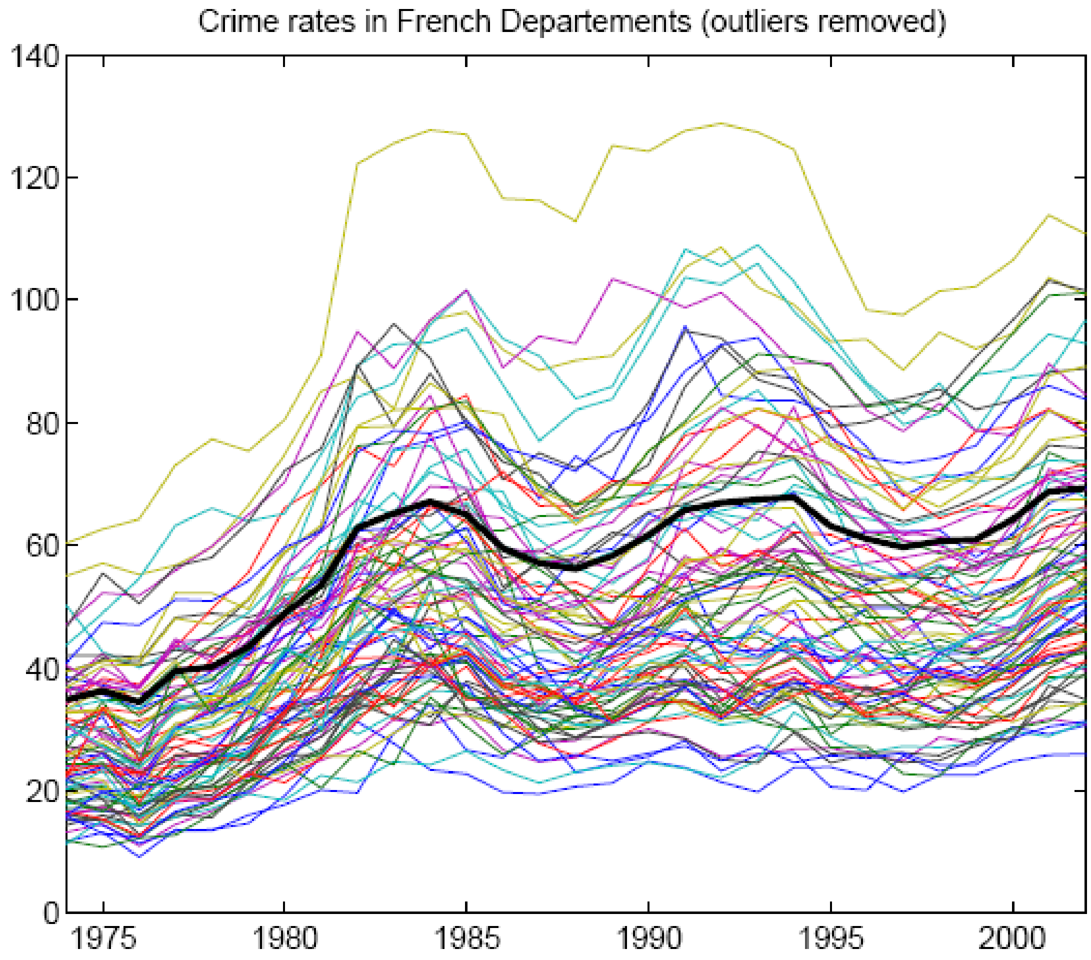
Data - French *Départements*

•



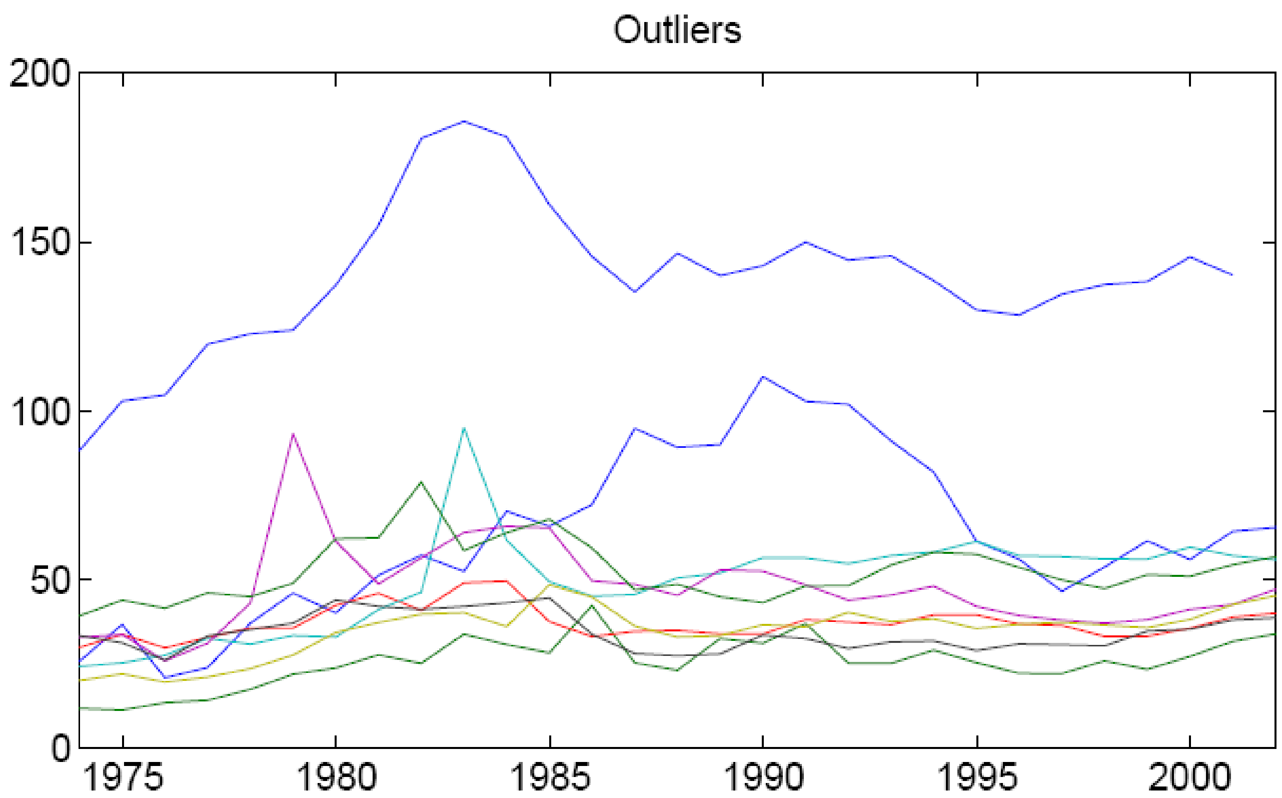
French *Départements* *(Paris excluded)*

•



Outliers: Paris, Corsica,…

•



Global trend and local fluctuations

- Issue: local crime rate = (global trend) @ (local fluctuations)
- proper separation of the global and local terms?

- General idea for such analysis (not specific to crime statistics):

M. Argollo de Menezes and A.-L. Barabasi, 2004
« Separating internal and external dynamics of complex systems »

$$f_i(t) = a_i\, w(t) + g_i(t)$$

local data = (local amplification x global trend) + local fluctuation

- Their method: (notation: $t=1,\dots,T$ $\langle u \rangle = (1/T)\, \Sigma_t\, u(t)$)
- assume $\langle g_i \rangle = 0$ hence $a_i = \langle f_i \rangle / \langle w \rangle$,
- compute an estimate of w as: $w(t) = \Sigma_i f_i(t) / \Sigma_i \langle f_i \rangle$
- hence $g_i(t) = f_i(t) - a_i\, w(t)$

Unfortunately, one can show that, whatever the data,
this is not a proper way to estimate the global trend.

Global trend and local fluctuations

- Issue: local crime rate = (global trend) @ (local fluctuations)
- proper separation of the global and local terms?

$$f_i(t) = a_i \, w(t) + g_i(t)$$

local data = (local amplification x global trend) + local fluctuation

- Our approach: an Independent Component Analysis approach

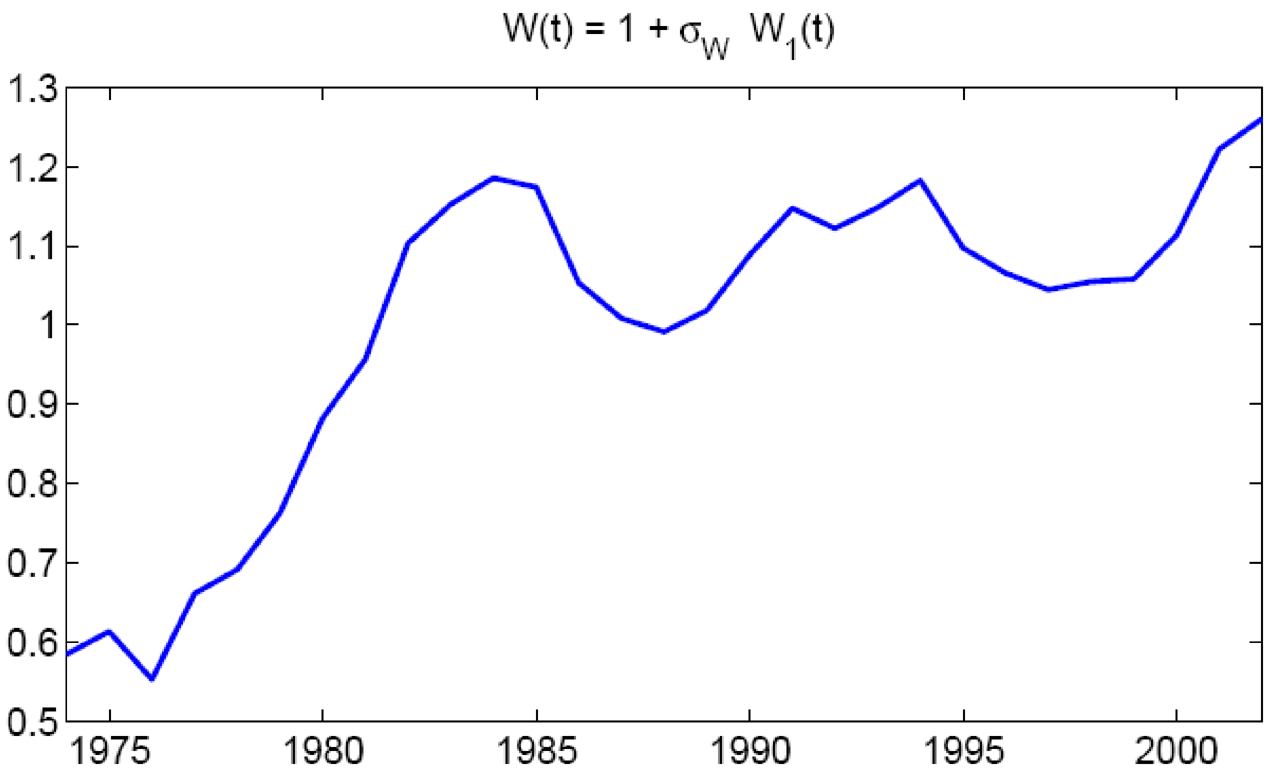
- Method: (notation: $\langle u \rangle = (1/T) \, \Sigma_t \, u(t)$)
- (do **not** assume $\langle g_i \rangle = 0$)
- compute an estimate of the a_is and w from the hyopthesis:

$$\langle w \, g_i \rangle - \langle w \rangle \langle g_i \rangle = 0$$

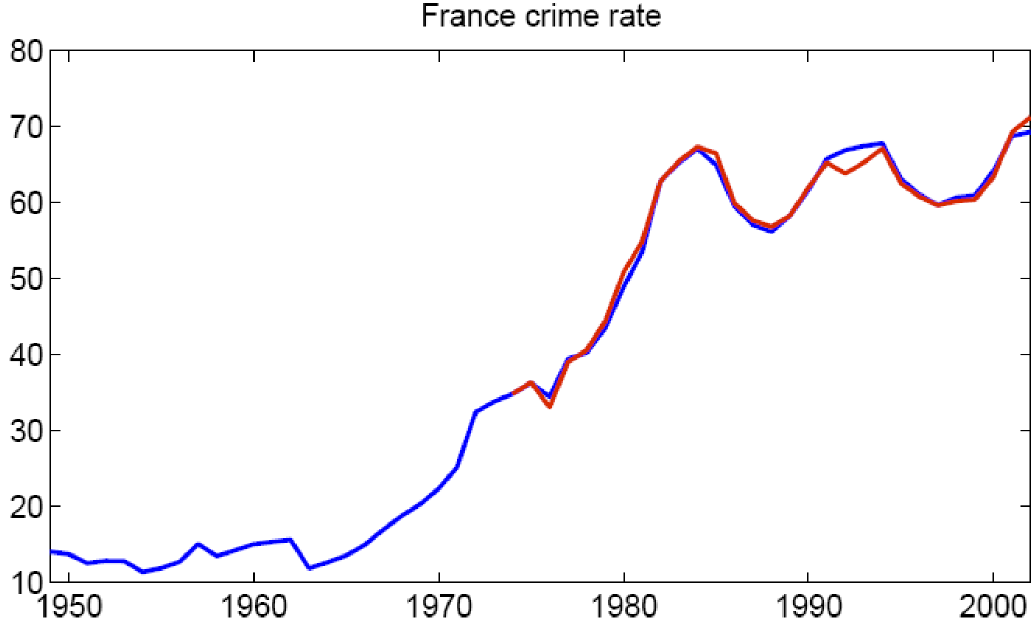
$$\langle g_i \, g_k \rangle - \langle g_i \rangle \langle g_k \rangle = 0 \quad \text{for } i \neq k$$

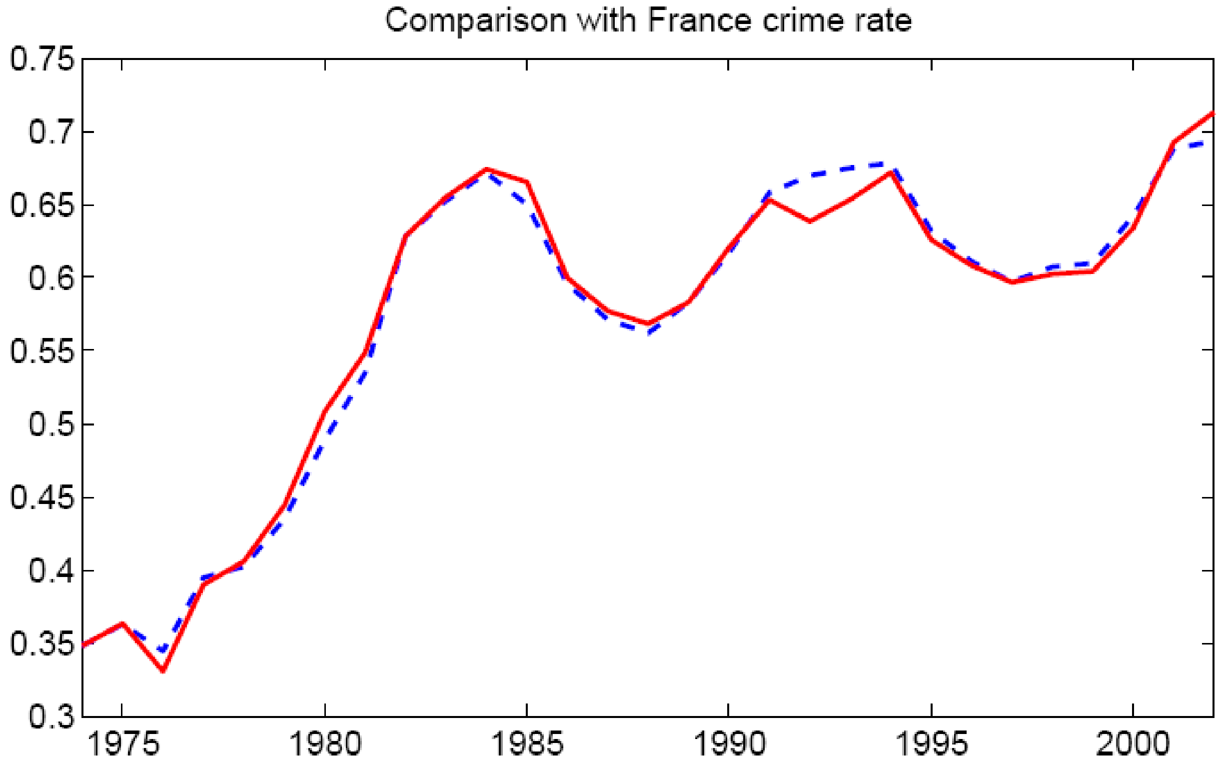
Estimated trend - France

•



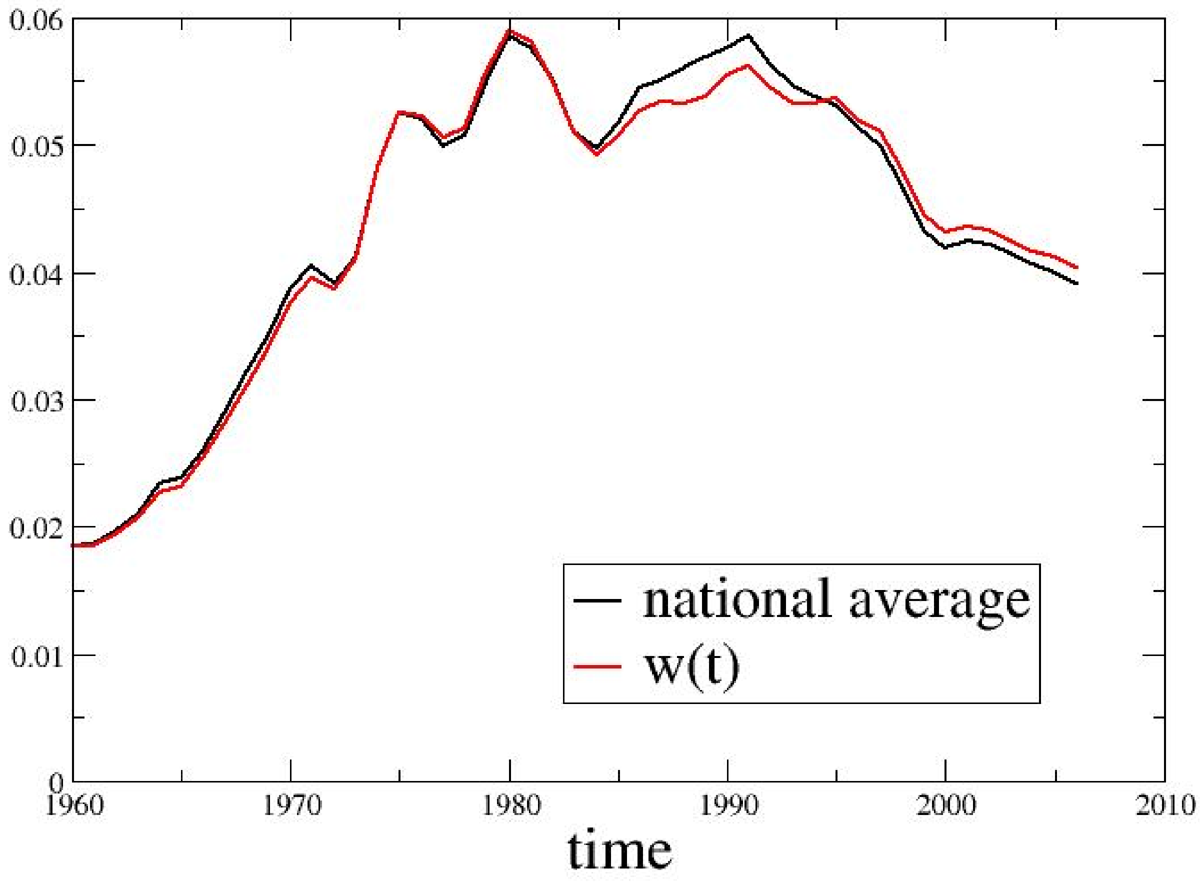
France


France crime rate
80
70
60
50
40
30
20
10
1950
1960
1970
1980
1990
2000


Comparison with France crime rate
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
1975
1980
1985
1990
1995
2000

USA

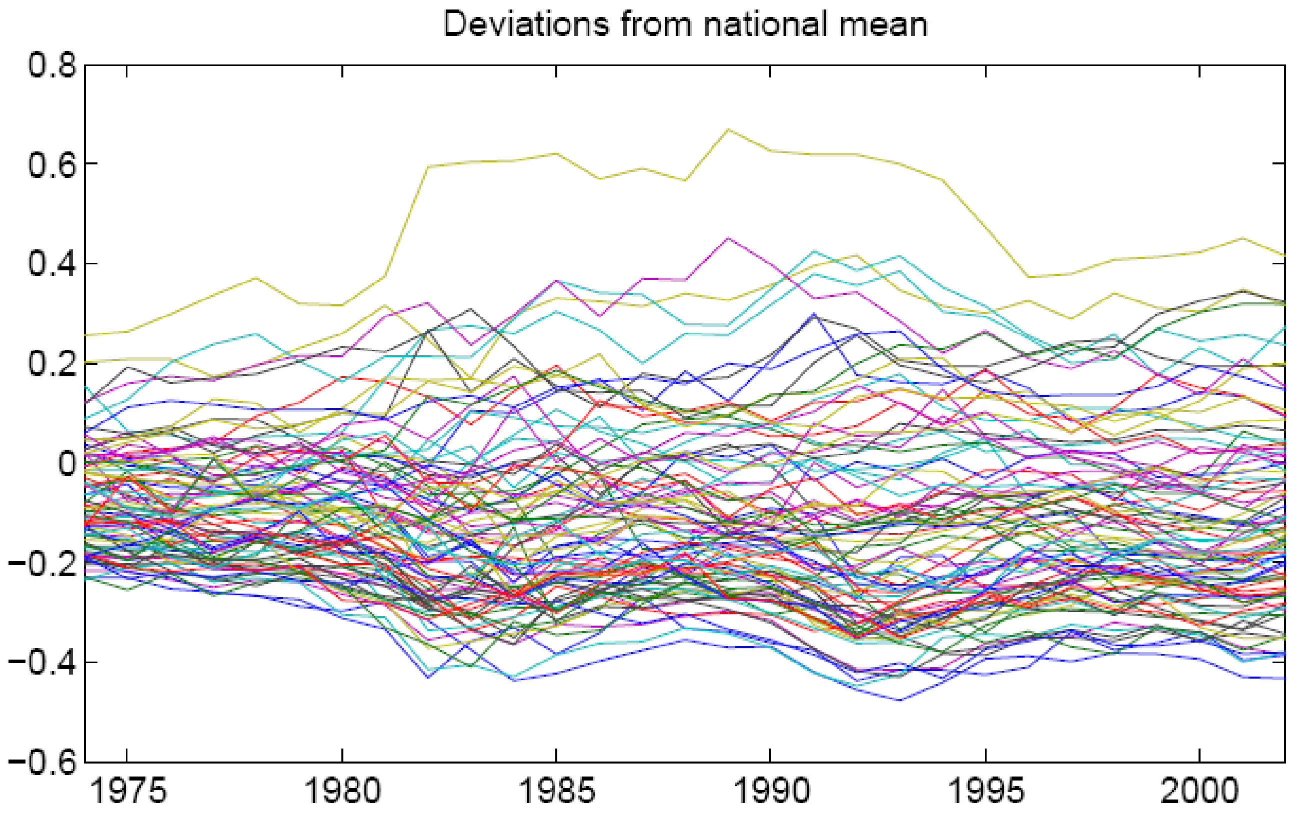
•



0.06
0.05
0.04
0.03
0.02
0.01
0
1960
1970
1980
1990
2000
2010
time
national average
w(t)

Local fluctuations: the « naive » measure

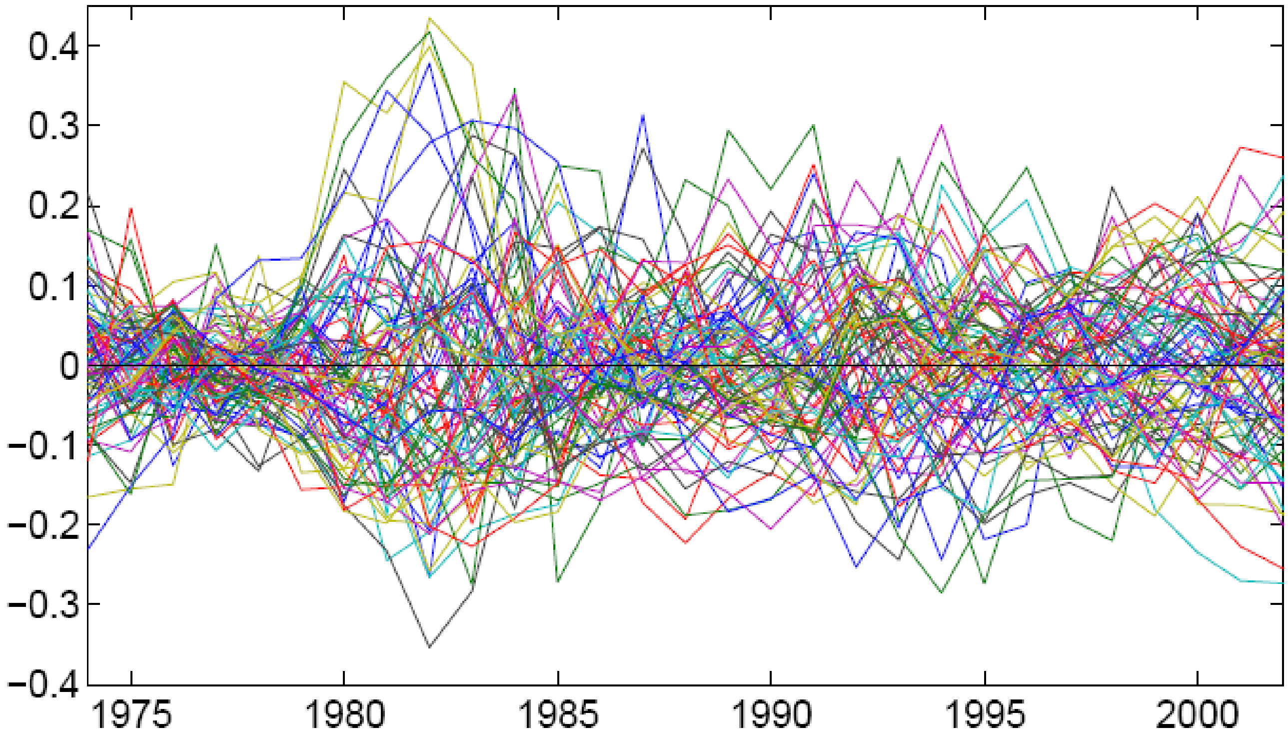
$h_i(t) = f_i(t) - f_{national}(t)$



Local fluctuations - France

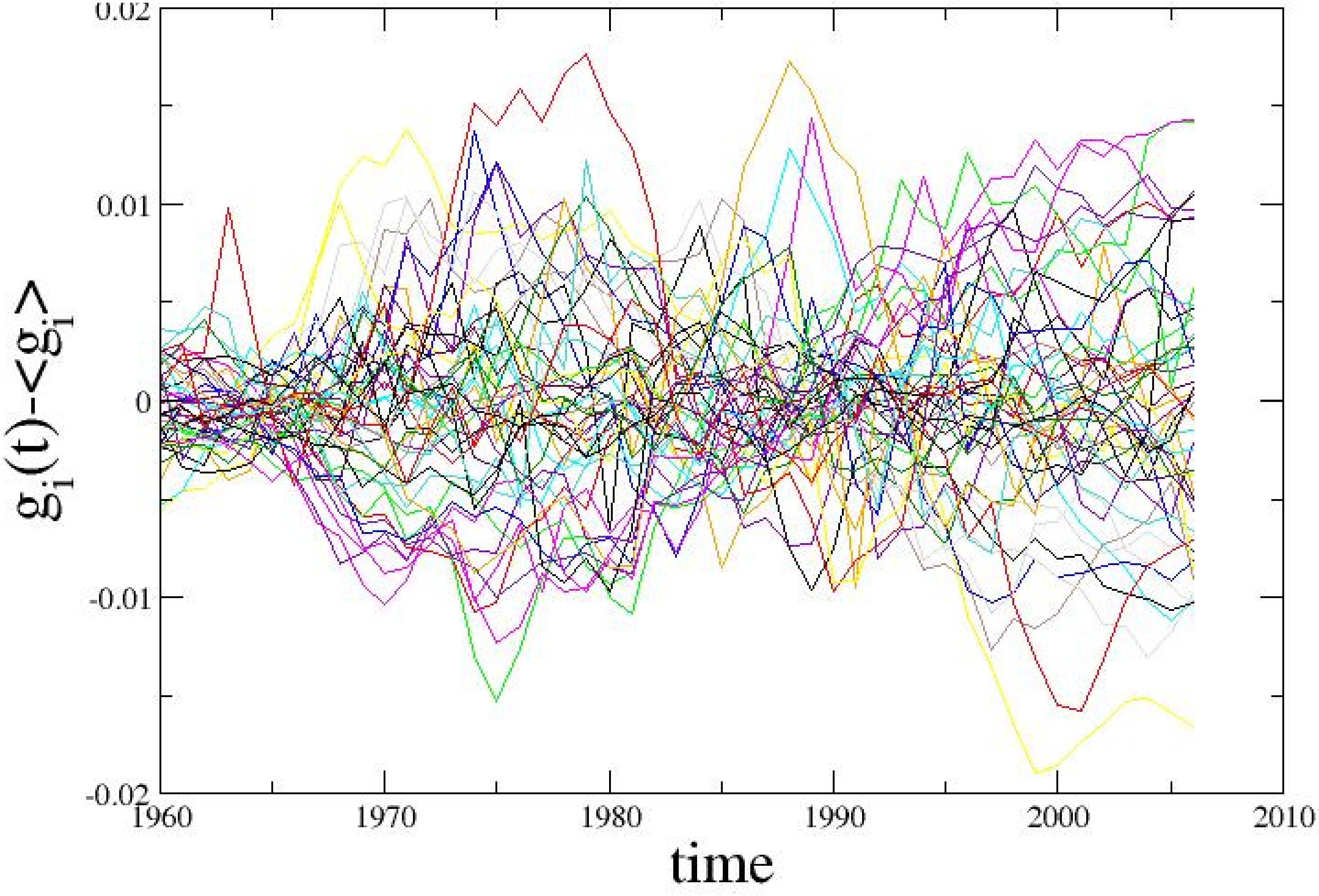
$g_i(t) = f_i(t) - a_i\, w(t)$

$g_i(t) - \langle g_i \rangle$



Local fluctuations - USA

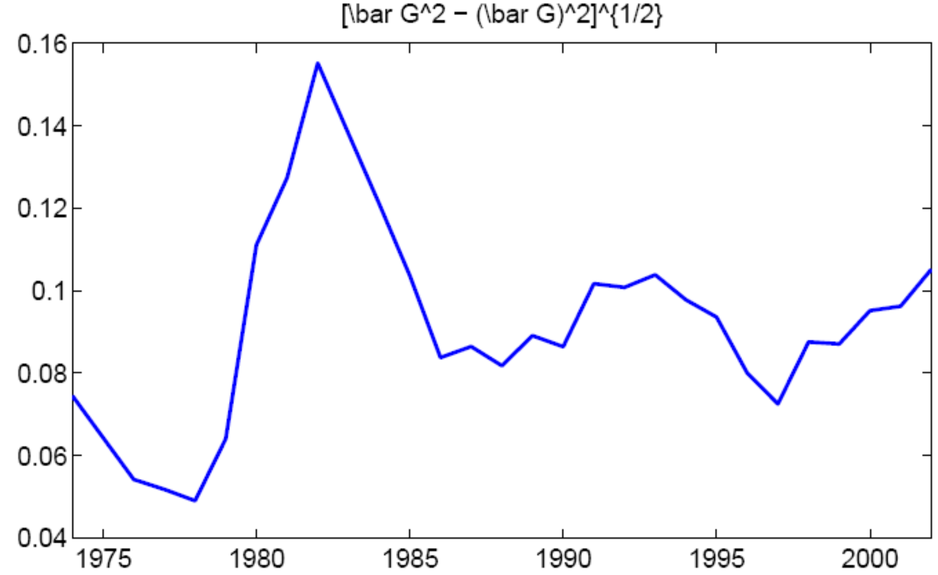
$g_i(t) = f_i(t) - a_i\, w(t)$



state-to-state fluctuations (volatility) - France

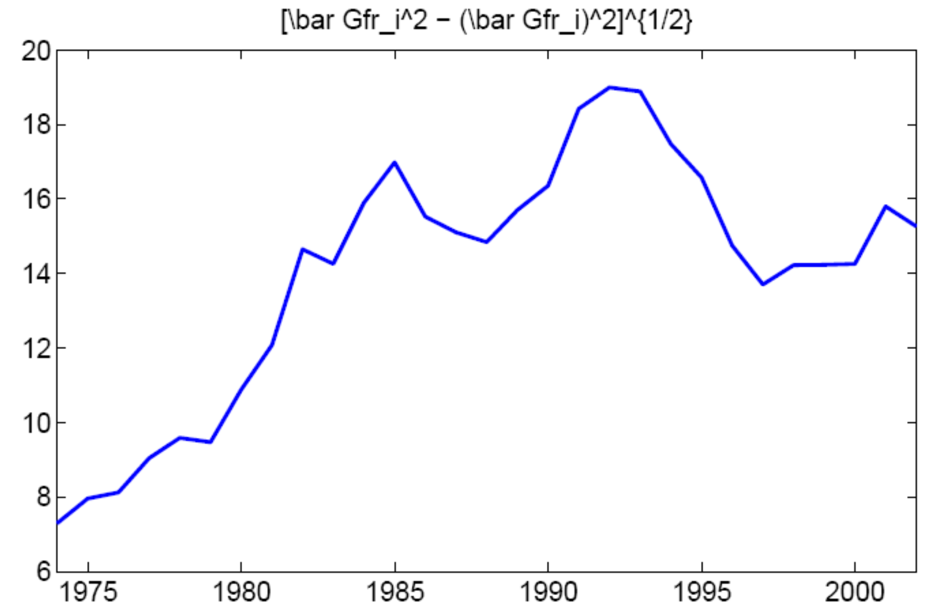
CAMS

$$(1/N)\ \Sigma_i\ g_i(t)^2 - [(1/N)\ \Sigma_i\ g_i(t)]^2$$



$$(1/N)\ \Sigma_i\ h_i(t)^2 - [(1/N)\ \Sigma_i\ h_i(t)]^2$$

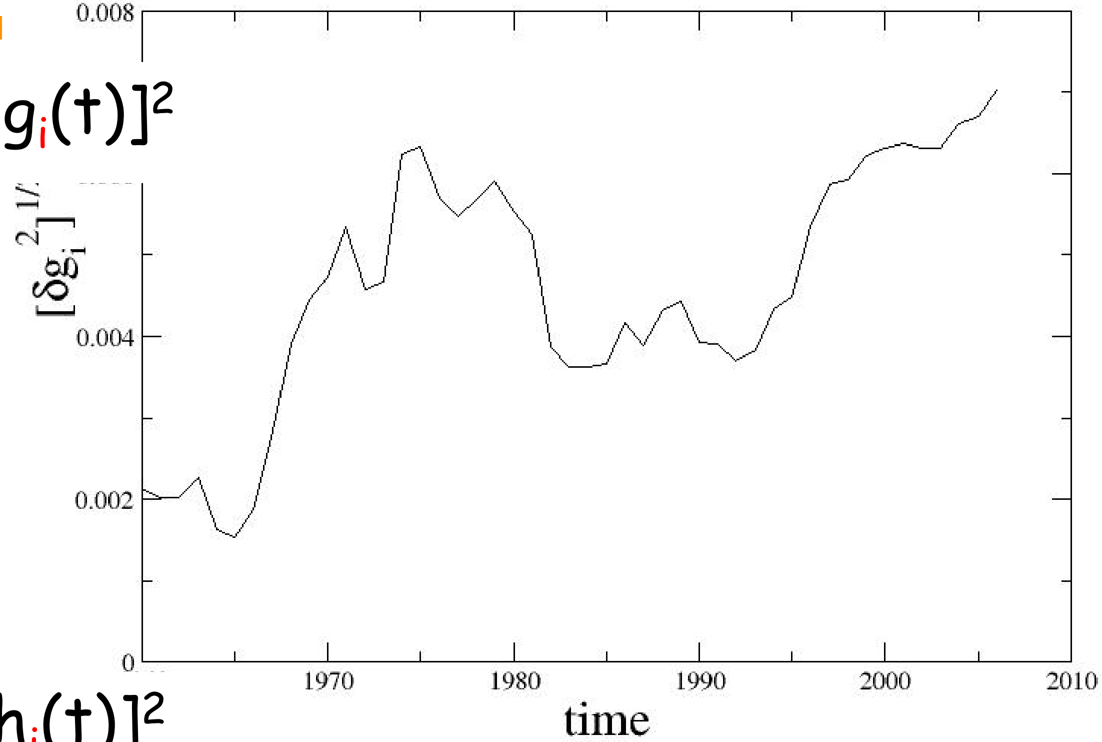
$$h_i(t) = f_i(t) - f_{national}(t)$$



state-to-state fluctuations - USA

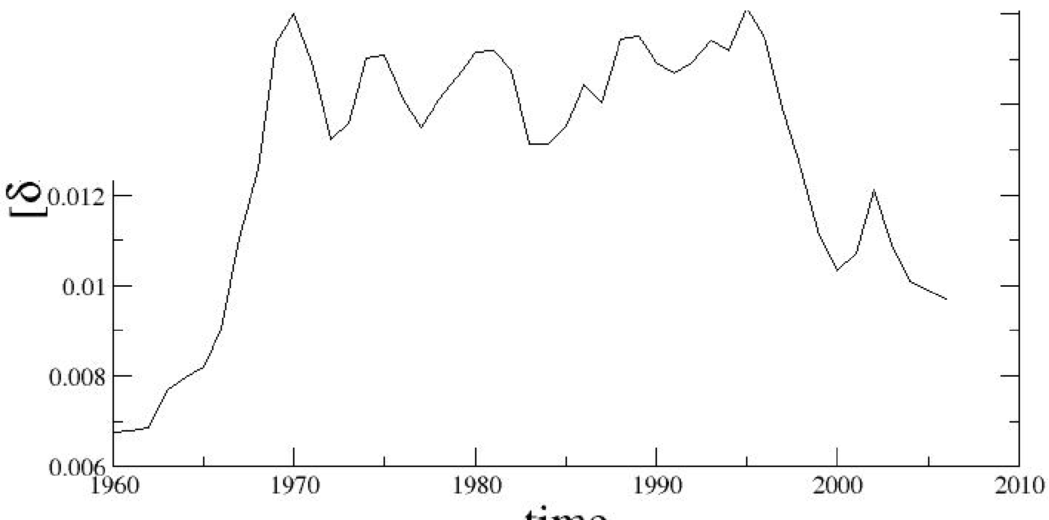
CAMS

$(1/N)\, \Sigma_i\, g_i(t)^2 - [(1/N)\, \Sigma_i\, g_i(t)]^2$



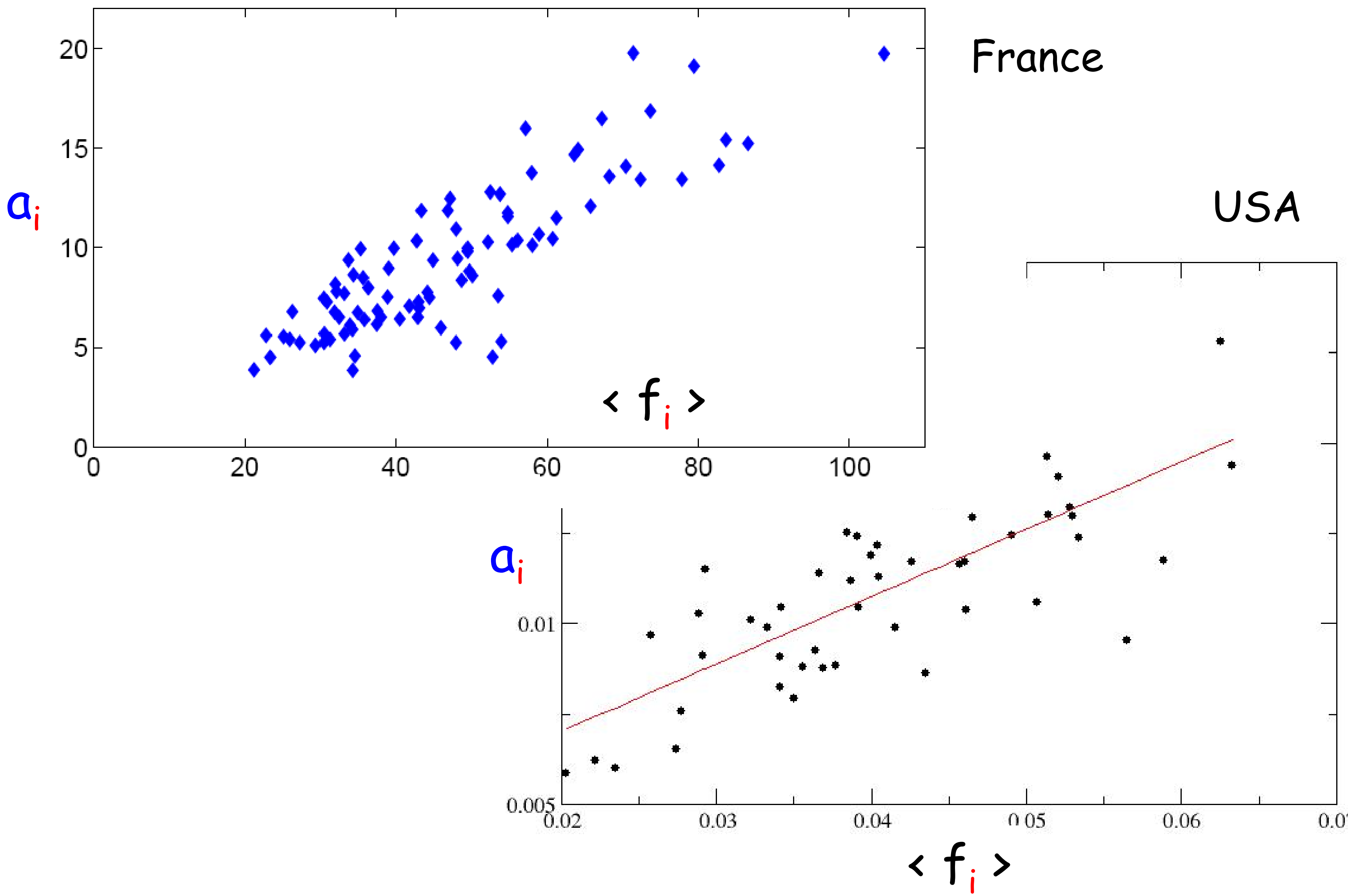
$(1/N)\, \Sigma_i\, h_i(t)^2 - [(1/N)\, \Sigma_i\, h_i(t)]^2$

$h_i(t) = f_i(t) - f_{national}(t)$



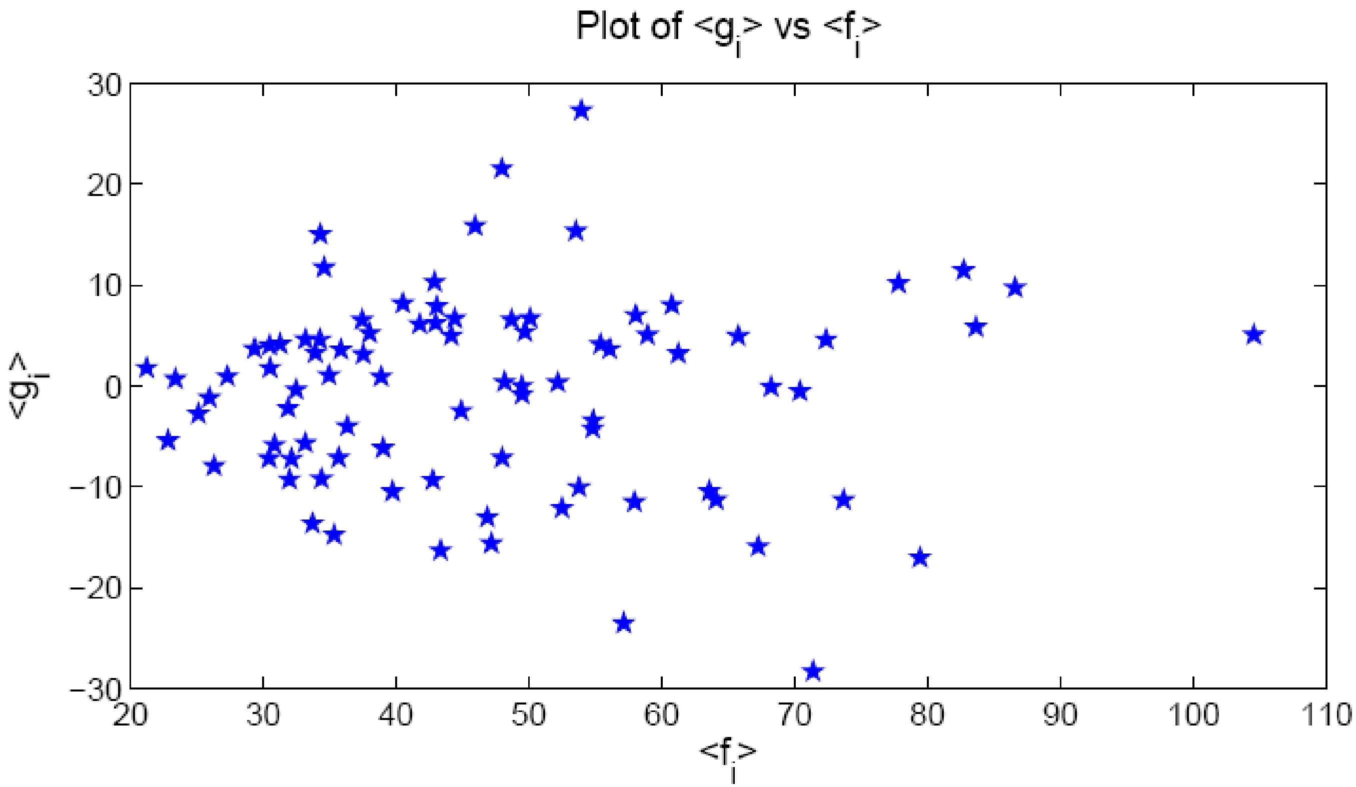
ai vs. <fi>

$f_i = a_i W + g_i$



$< g_i >$

•



Plot of <g_i> vs <f_i>
30
20
10
0
−10
−20
−30
20
30
40
50
60
70
80
90
100
110
<f_i>
<g_i>

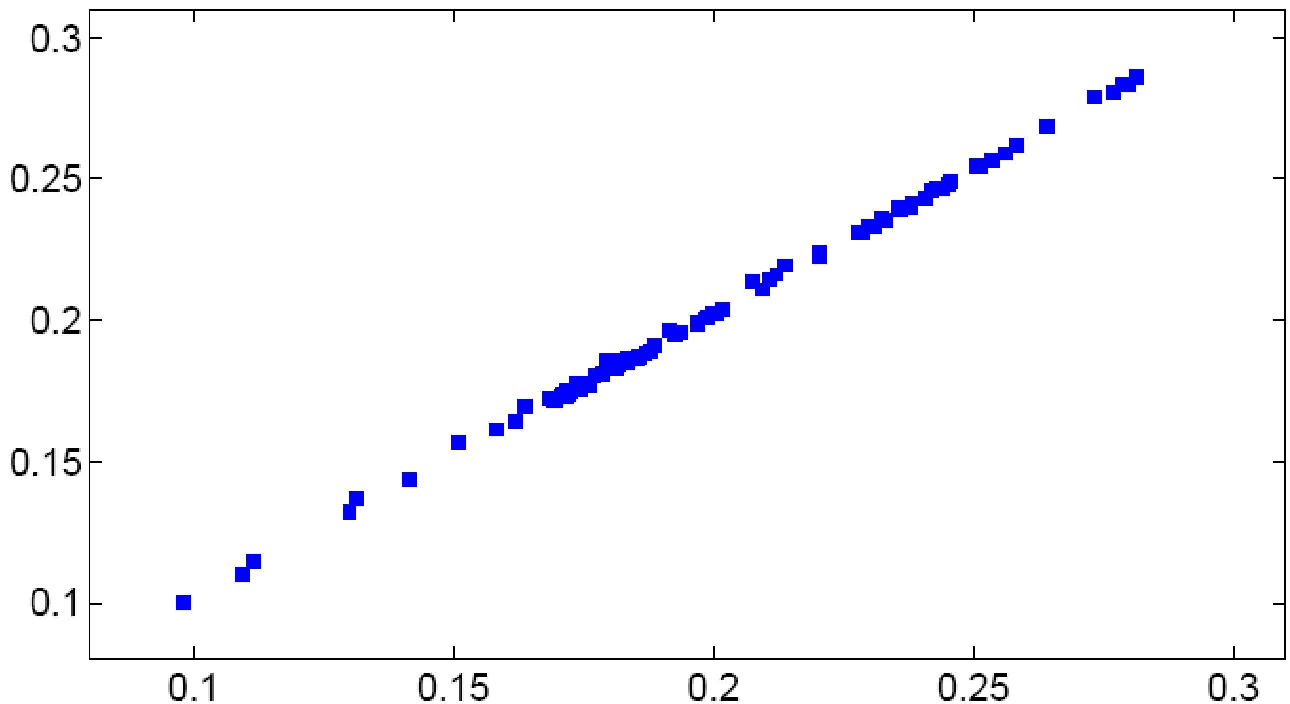
Consistency check

check of: $\langle W g_i \rangle - \langle W \rangle \langle g_i \rangle = 0$

$W = 1 + \sigma_W W_1$

$\langle W \rangle = 0, \ \langle W^2 \rangle = 1$

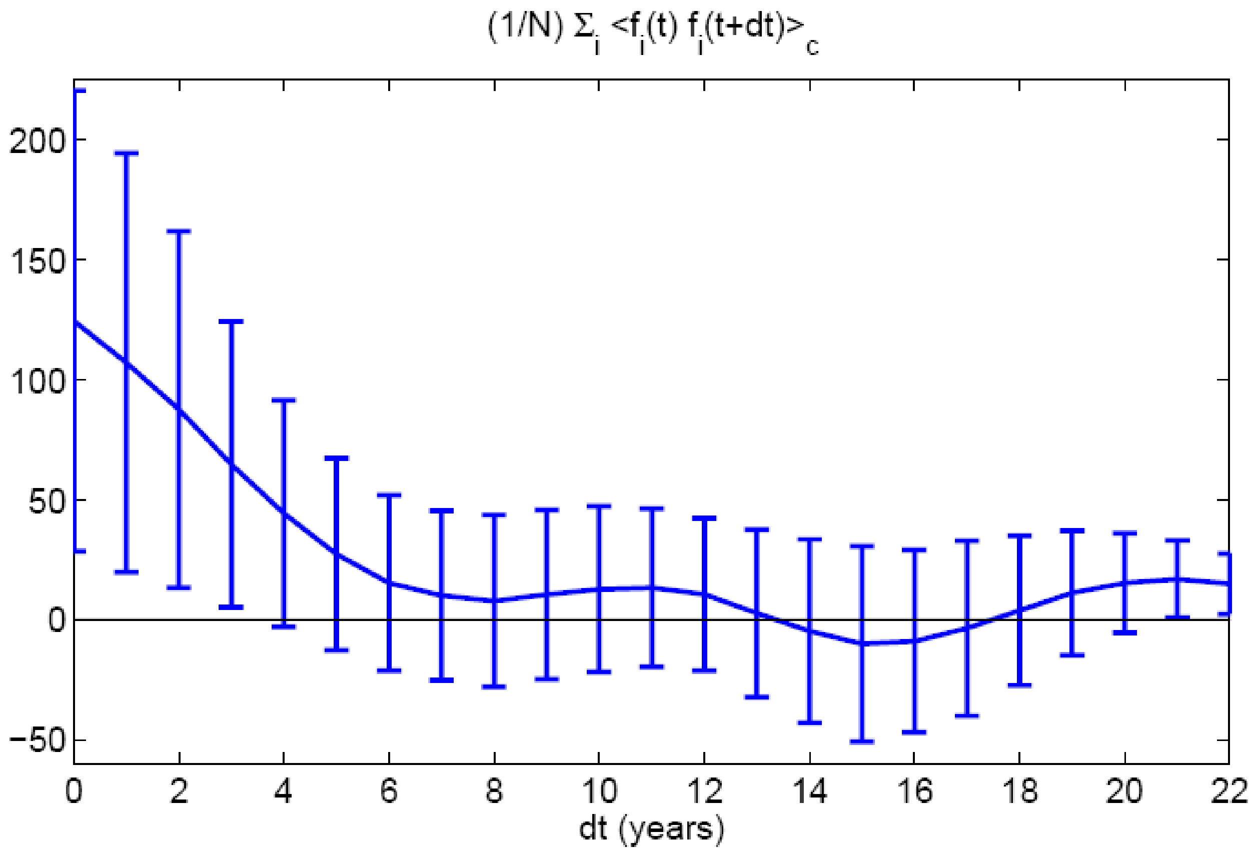
$\langle W (f_i - \langle f_i \rangle) \rangle$



$a_i \sigma_W$

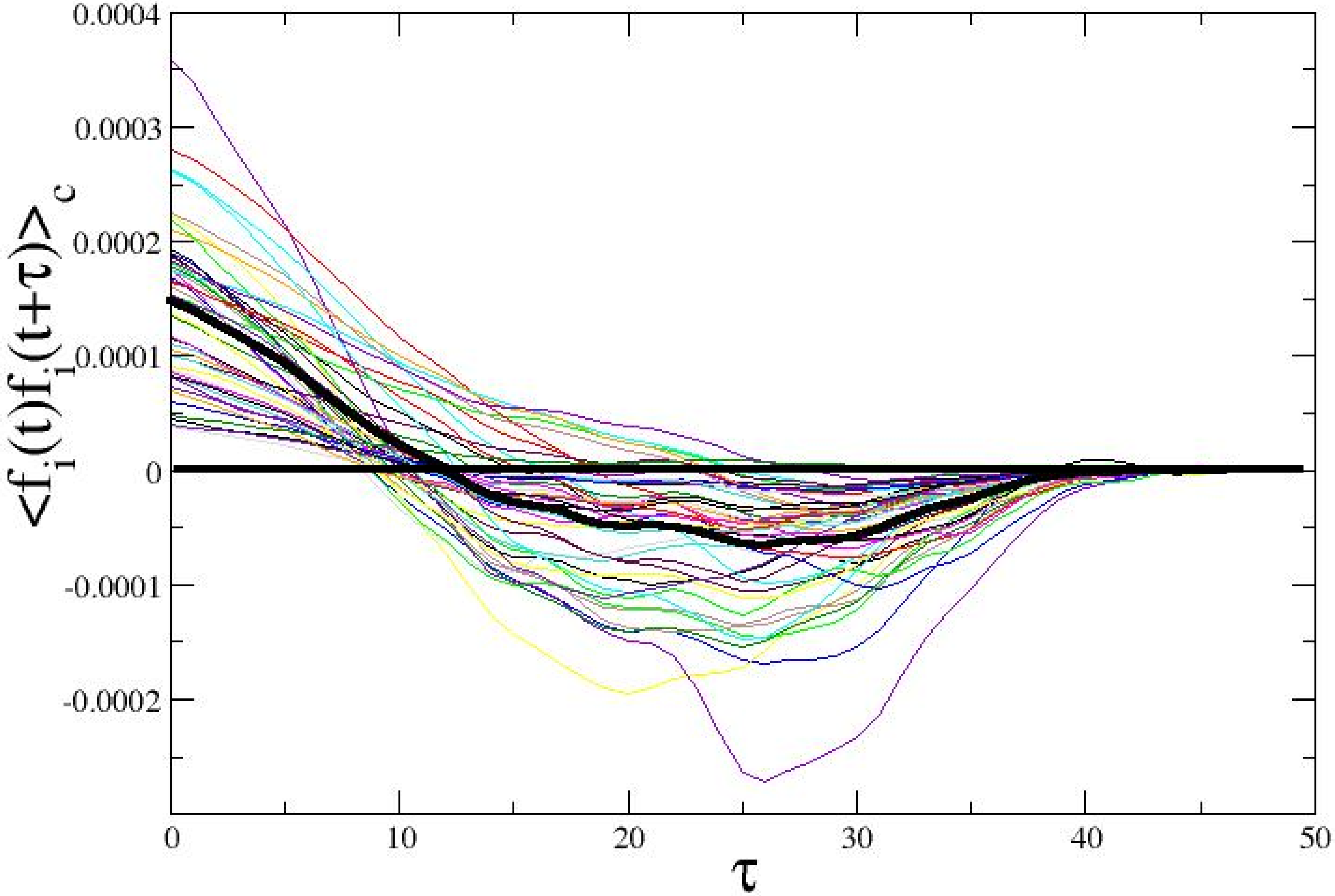
Time correlations: raw data (France)

$$(1/N)\ \Sigma_i\ [\ \langle f_i(t)\, f_i(t+dt) \rangle - \langle f_i(t) \rangle^2\ ]$$



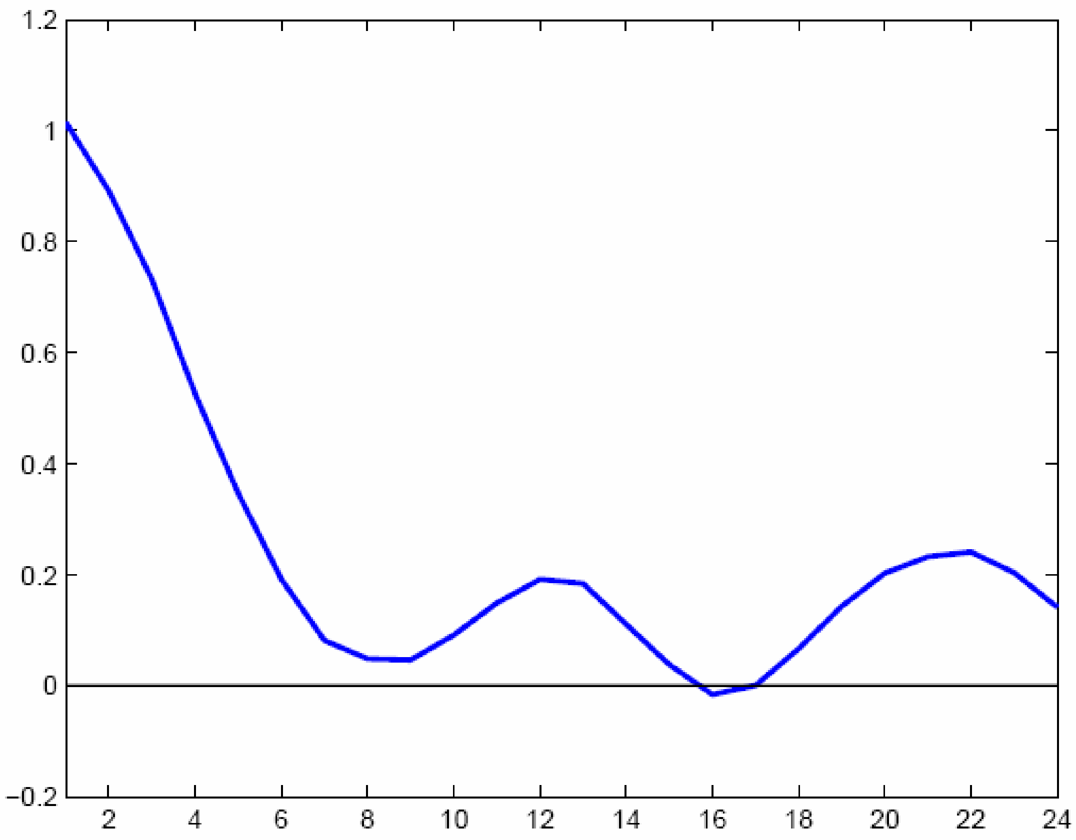
Time correlations: raw data (USA)

- $(1/N) \sum_i [< f_i(t) f_i(t+\tau) > - < f_i(t) >^2]$



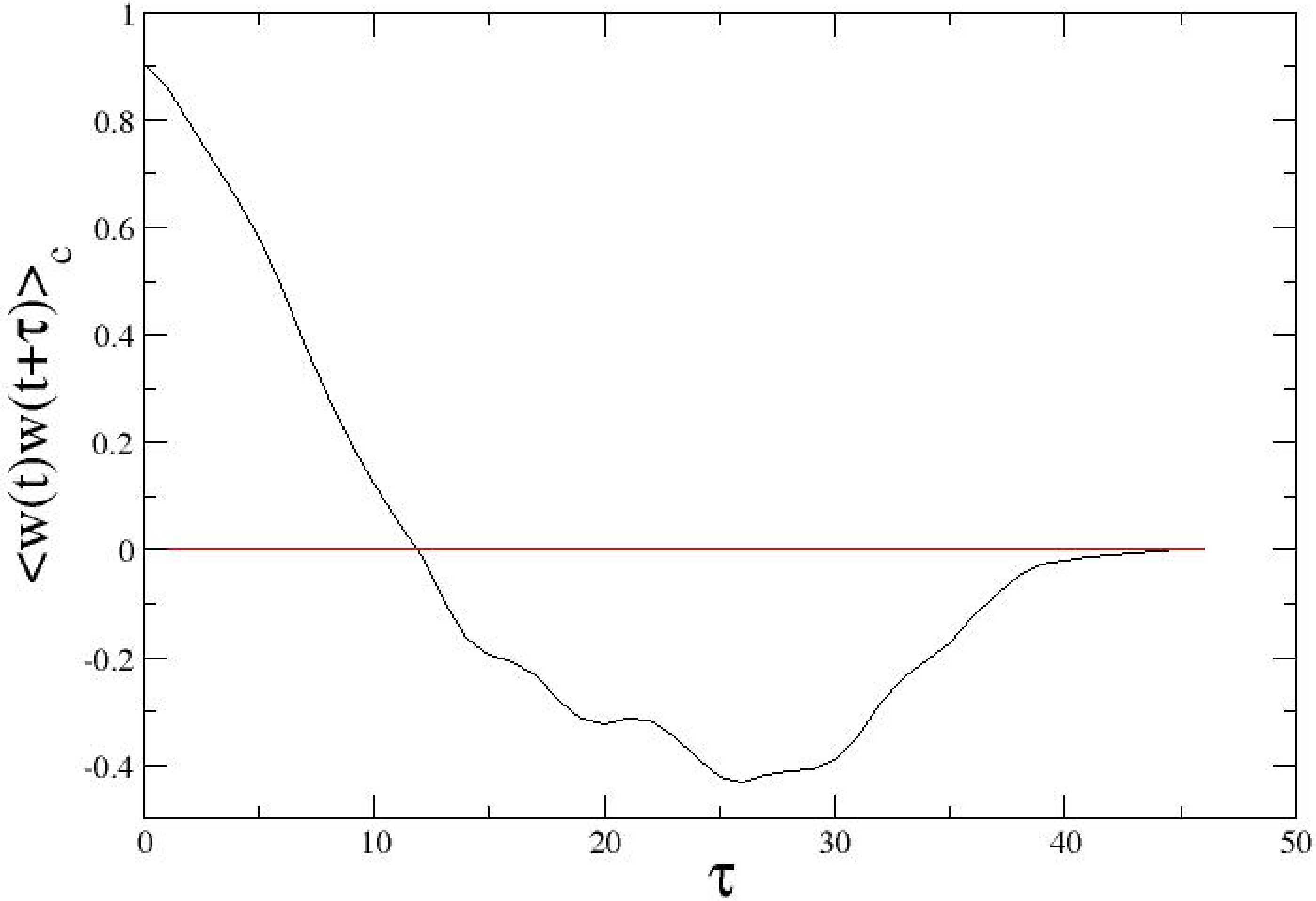
Trend: time correlations - France

$$< W(t)\ W(t+dt) > - < W(t) >^2$$



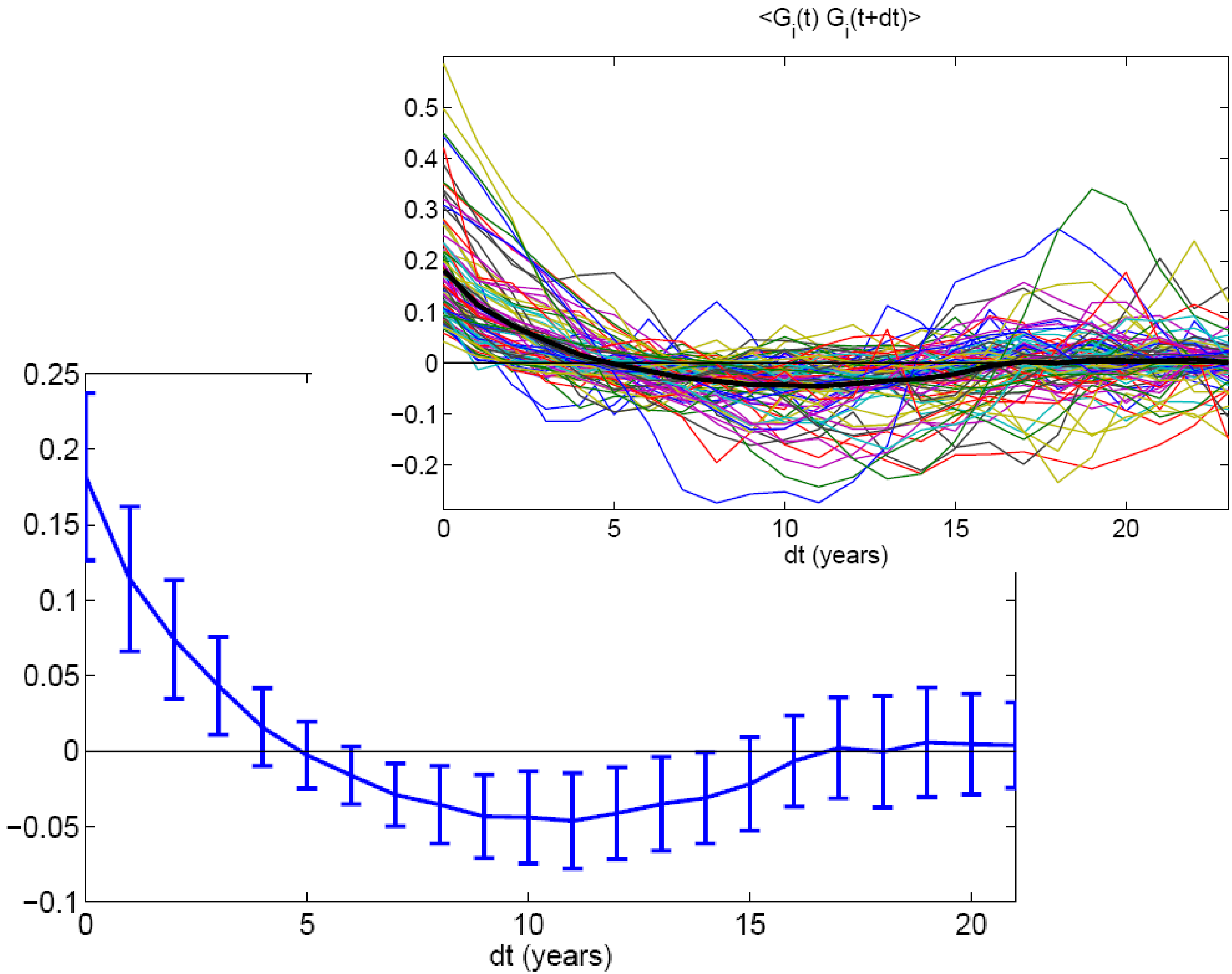
Trend: time correlations – USA

$$< W(t)\, W(t+\tau) > - < W(t) >^2$$

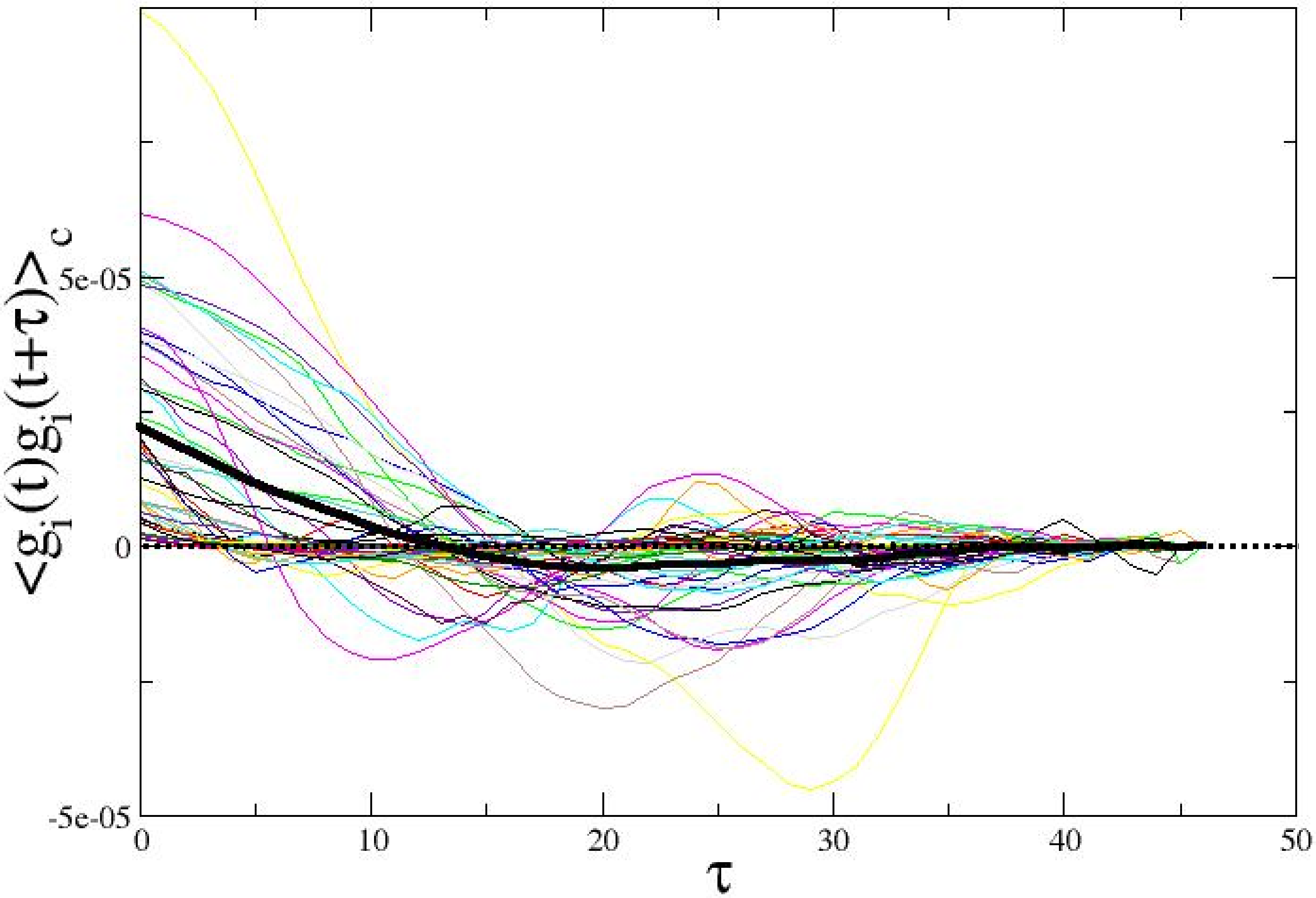


Time correlations: local fluctuations - France

•



Time correlations: local fluctuations - USA



to be continued…

- work in progress

  econometric analysis

  multiscale analysis

  other states

  origine of fluctuations

  …

  similar analysis on other/better data