UNSTABLE VOLATILITY: THE BREAK
PRESERVING LOCAL LINEAR ESTIMATOR

Isabel Casas – CREATE, Aarhus University
joint work with
Irene Gijbels – K.U. Leuven

6th Bachelier Finance Society World Congress
Toronto, 2010
What is our work about?

- **Aim?** estimation of discontinuous volatility functions.
- **Discontinuities?** abrupt structural changes.
- **Method?** nonparametric kernel estimation.
- **Contribution?** Break preserving local linear.
Model

\[ Y_i = m(X_i) + \sigma(X_i)\epsilon_i \quad \epsilon \sim i.i.d(0, 1) \]

- Fixed design or random design.
- \( E(\epsilon|X) = 0, \; E(\epsilon^2|X) = 1 \) and \( E(\epsilon^4|X) < \infty \)
- \( E(Y|X = x) = m(x) \)
- \( E((Y - m(X))^2|X = x) = \sigma^2(x) \)
Previous work: drift estimator
Drift estimation

\[ Y_i = m(X_i) + 0.4\epsilon_i \text{ with } \epsilon \sim IID(0, 1) \]

Given a point \( x \) in the continuous part, estimator of \( m(x) \)?
Drift estimation: centred estimator

The **centred** estimator, $\hat{m}_c(x)$, is obtained as a regression using the points in a neighbourhood of $x$, Fan and Gijbels (1997).
Drift estimation: centred estimator

![Graph showing drift estimation with LL (centred) estimator and data points.]

- Drift estimation: centred estimator
- Point to estimate
- LL (centred) estimator
Motivation  Volatility estimation  Simulations  Summary

Drift estimation: What happens at discontinuities?

We expect the centred estimator to fall in the middle of the jump.
Drift estimation: What happens at discontinuities?

The asymmetric estimator: find two estimators, left and right, and choose appropriately, Qiu (2003).
Drift estimation: What happens at discontinuities?

We have three estimator, which is the best choice?
Contribution: volatility estimator
Estimation of a discontinuous volatility

\[ Y_i = m(X_i) + \sigma(X_i)\epsilon_i \quad \epsilon \sim i.i.d(0, 1) \]

Define \( \hat{r}_i = (Y_i - \hat{m}(X_i)) \). Then, \( E(\hat{r}^2|X = x) = \hat{\sigma}^2(x) \).

Fan and Yao (1998):

“While the bias of \( \hat{m} \) itself is of order \( O(h_1^2) \), its contribution to \( \hat{\sigma}^2(\cdot) \) is only of \( o(h_1^2) \).”

So, we expect to get a good estimate of the volatility even if the drift function is unknown.
Do you think that the **centred** estimator (Fan and Yao, 1998) is a good choice to estimate a discontinuous volatility function?
Do you think that the centred estimator (Fan and Yao, 1998) is a good choice to estimate a discontinuous volatility function?

- No, because it is not consistent at discontinuities.
- Solution: the break preserving local linear (BPLL) estimator.
Estimation of a discontinuous volatility

$$\hat{\sigma}^2_k(x) = \hat{a}_{0,k}(x) \quad \text{and} \quad \hat{\sigma}^2_k = \hat{a}_{1,k}$$

$$(\hat{a}_{0,k}(x), \hat{a}_{1,k}(x)) = \min_{(a_0, a_1)} \sum_{i=1}^{n} \left\{ \hat{r}_i^2 - a_{0,k} - a_{1,k}(X_i - x) \right\}^2 K_k \left( \frac{X_i - x}{h_2} \right)$$

left \(k=l\)  \hspace{2cm} centred \(k=c\)  \hspace{2cm} right \(k=r\)
Estimation of a discontinuous volatility

The expression of the three volatility estimators:

\[
\hat{\sigma}_k^2(x) = \sum_{i=1}^{n} \hat{r}_i^2 K_k \left( \frac{X_i - x}{h_2} \right) \frac{s_{k,2} - s_{k,1}(X_i - x)}{s_{k,0}s_{k,2} - s_{k,1}^2} \quad k = c, l, r
\]

where

\[
s_{k,j} = \sum (X_i - x)^j K_k \left( \frac{X_i - x}{h_2} \right)
\]

- Easy to compute.
- No numerical minimisation.
Estimation of a discontinuous volatility

How well are the estimators fitted to the data set? **Weighted Residuals Mean Square.**

\[ \text{WRMSE}_k(x) = \sum_{i=1}^{n} \left\{ \hat{r}_i^2 - \hat{a}_{0,c} - \hat{a}_{1,c}(X_i - x) \right\}^2 K_k \left( \frac{X_i - x}{h_2} \right) \]

\[ \sum_{i=1}^{n} K_k \left( \frac{X_i - x}{h_2} \right) \]
The break preserving local linear estimator:

\[
\hat{\sigma}_{BPPL}^2(x) = \begin{cases} 
\hat{\sigma}_c^2(x) & \text{diff}(x) < u \\
\hat{\sigma}_l^2(x) & \text{diff}(x) \geq u \text{ and } \text{WRMS}_l(x) < \text{WRMS}_r(x) \\
\hat{\sigma}_r^2(x) & \text{diff}(x) \geq u \text{ and } \text{WRMS}_l(x) > \text{WRMS}_r(x) \\
\frac{\hat{\sigma}_l^2(x) + \hat{\sigma}_r^2(x)}{2} & \text{diff}(x) \geq u \text{ and } \text{WRMS}_l(x) = \text{WRMS}_r(x) 
\end{cases}
\]

where \(\text{diff}(x) = \max(\text{WRMS}_c(x) - \text{WRMS}_l(x), \text{WRMS}_c(x) - \text{WRMS}_r(x))\), and \(0 \leq u \leq Q\) for all \(x\) and \(Q\) a constant.
How is the WRMS for each estimator?

Let \([a, b]\) be the support of \(X\) and \(\{x_q\}\) for \(q = 1, \ldots, m\) be the finite set of points where the volatility function is discontinuous. Then, two regions can be differentiated:

- \(D_1\) is the region where the volatility function is continuous,

\[
D_1 = \left[ a + \frac{h_2}{2}, b - \frac{h_2}{2} \right] \setminus D_2
\]

- \(D_2\) contains the points of discontinuity and their neighbourhoods:

\[
D_2 = \bigcup_{q=1}^{m} \left[ x_q - \frac{h_2}{2}, x_q + \frac{h_2}{2} \right]
\]
How is the WRMS for each estimator?

Under certain regularity conditions:
For \( x \in D_1 \),

\[
WRMS_k(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + R_{k,1}(x)
\]
How is the WRMS for each estimator?

Under certain regularity conditions:
For $x \in D_1$, 

$$WRMS_k(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + R_{k,1}(x)$$

For $x \in D_2$ such that $x = x_q + \tau h_2$ with $\tau \in [0, \frac{1}{2}]$ and a jump of magnitude $d$,

$$WRMS_l(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + d^2C_{l,\tau}^2 + R_{l,2}(x)$$

$$WRMS_r(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + R_{r,2}(x)$$

$$WRMS_c(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + d^2C_{c,\tau}^2 + R_{c,2}(x)$$
How is the WRMS for each estimator?

Under certain regularity conditions:
For $x \in D_1$,

$$W R M S_k(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + R_{k,1}(x)$$

For $x \in D_2$ such that $x = x_q + \tau h_2$ with $\tau \in [-\frac{1}{2}, 0]$ and a jump of magnitude $d$,

$$W R M S_l(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + R_{l,3}(x)$$

$$W R M S_r(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + d^2 C^2_{r,\tau} + R_{r,3}(x)$$

$$W R M S_c(x) = \sigma^4(x)(E(\epsilon^4|X) - 1) + d^2 C^2_{c,\tau} + R_{c,3}(x)$$
MSE (continuous points)

Under certain regularity conditions and with

\[ \mu_{k,j} = \int w^j K_k(u) \, du \quad \text{and} \quad V_k = \int K_k^2(u) \left[ \frac{\mu_{k,2} - \mu_{k,1} u}{\mu_{k,0} \mu_{k,2} - \mu_{k,1}^2} \right]^2 \, du : \]

For \( x \in D_1 \) (Continuous points),

\[ \text{Bias}(\hat{\sigma}_k^2(x)) = \frac{h_2^2 \hat{o}^2(x)}{2} \left( \frac{\mu_{k,2}^2 - \mu_{k,1} \mu_{k,3}}{\mu_{k,2} \mu_{k,0} - \mu_{k,1}^2} \right) + o_p(h_1^2 + h_2^2 + \frac{1}{nh_2}) \]

\[ \text{Variance}(\hat{\sigma}_k^2(x)) = \frac{(E(\varepsilon^4|X) - 1) \sigma^4(x)}{nh_2 f_X(x)} V_k + o_p \left( \frac{1}{nh_2} \right) \]

\[ \text{MSE}(\hat{\sigma}_k^2(x)) = \text{Bias}^2 + \text{Variance} \]
Under certain regularity conditions and with
\[ \mu_{k,j} = \int u^j K_k(u) \, du \quad \text{and} \quad V_k = \int K_k^2(u) \left[ \frac{\mu_{k,2} - \mu_{k,1}u}{\mu_{k,0}\mu_{k,2} - \mu_{k,1}^2} \right]^2 \, du : \]

For \( x \in D_1 \) (Continuous points),
\[ \text{Bias}(\hat{\sigma}_k^2(x)) = \frac{h_2^2 \sigma^2(x)}{2} \frac{\mu_{k,2}^2 - \mu_{k,1}\mu_{k,3}}{\mu_{k,2}\mu_{k,0} - \mu_{k,1}^2} + o_p(h_1^2 + h_2^2 + \frac{1}{nh_2}) \]
\[ \text{Variance}(\hat{\sigma}_k^2(x)) = \frac{(E(\epsilon^4|X)-1)\sigma^4(x)}{nh_2 f_X(x)} V_k + o_p \left( \frac{1}{nh_2} \right) \]

If \( h_1, h_2 \to 0, n \to \infty \) and \( nh_2 \to \infty \)
\[ \text{MSE}(\hat{\sigma}_k^2(x)) = \text{Bias}^2 + \text{Variance} \]
Under certain regularity conditions and with

\[ \mu_{k,j} = \int u^j K_k(u) \, du \quad \text{and} \quad V_k = \int K_k^2(u) \left[ \frac{\mu_{k,2} - \mu_{k,1} u}{\mu_{k,0} \mu_{k,2} - \mu_{k,1}^2} \right]^2 \, du: \]

For \( x \in D_1 \) (Continuous points),

\[ \text{Bias} \left( \hat{\sigma}_k^2(x) \right) = \]

\[ \text{Variance} \left( \hat{\sigma}_k^2(x) \right) = \frac{(E(\epsilon^4 | X) - 1) \sigma^4(x)}{nh_2 f_X(x)} V_k + o_p \left( \frac{1}{nh_2} \right) \]

If \( h_1, h_2 \to 0, n \to \infty \) and \( nh_2 \to \infty \)

\[ \text{MSE} \left( \hat{\sigma}_k^2(x) \right) = \text{Bias}^2 + \text{Variance} \]
MSE (continuous points)

Under certain regularity conditions and with

$$\mu_{k,j} = \int w^j K_k(u) du \quad \text{and} \quad V_k = \int K_k^2(u) \left[ \frac{\mu_{k,2} - \mu_{k,1}u}{\mu_{k,0}\mu_{k,2} - \mu_{k,1}^2} \right]^2 du,$$

For $x \in D_1$ (Continuous points),

Bias($\hat{\sigma}_k^2(x)$) =

Variance($\hat{\sigma}_k^2(x)$) =

If $h_1, h_2 \to 0$, $n \to \infty$ and $nh_2 \to \infty$

$$\text{MSE}(\hat{\sigma}_k^2(x)) = \text{Bias}^2 + \text{Variance}$$
MSE (continuous points)

Under certain regularity conditions and with

$$\mu_{k,j} = \int u^j K_k(u) du \quad \text{and} \quad V_k = \int K_k^2(u) \left[ \frac{\mu_{k,2} - \mu_{k,1} u}{\mu_{k,0} \mu_{k,2} - \mu_{k,1}^2} \right]^2 du :$$

For $x \in D_1$ (Continuous points),

$$\text{Bias} (\hat{\sigma}_k^2(x)) =$$

$$\text{Variance} (\hat{\sigma}_k^2(x)) =$$

If $h_1, h_2 \to 0, n \to \infty$ and $nh_2 \to \infty$

$$\text{MSE} (\hat{\sigma}_k^2(x)) =$$
MSE (right side of discontinuity)

For $x \in D_2$ such that $x = x_q + \tau h_2$ with $\tau \in [0, \frac{1}{2}]$ and a jump of magnitude $d$,

\[
\text{MSE}(\hat{\sigma}^2_l(x)) = \left[ d \int_{-\frac{1}{2}}^{\tau} K_l(u) \frac{\mu_{l,2} - \mu_{l,1}u}{\mu_{l,0}\mu_{l,2} - \mu_{l,1}^2} du \right]^2 + \frac{(E(\epsilon^4|X) - 1)\sigma^4(x)}{nh_2f_X(x)}V_l + o_p(1)
\]

\[
\text{MSE}(\hat{\sigma}^2_c(x)) = \left[ d \int_{-\frac{1}{2}}^{\tau} K_c(u) du \right]^2 + \frac{(E(\epsilon^4|X) - 1)\sigma^4(x)}{nh_2f_X(x)}V_c + o_p(1)
\]
MSE (right side of discontinuity)

For $x \in D_2$ such that $x = x_q + \tau h_2$ with $\tau \in [0, \frac{1}{2}]$ and a jump of magnitude $d$,

$$\text{MSE}(\hat{\sigma}_l^2(x)) = \left[ d \int_{-\frac{1}{2}}^{\tau} K_l(u) \frac{\mu_{l,2} - \mu_{l,1} u}{\mu_{l,0} \mu_{l,2} - \mu_{l,1}^2} du \right]^2 + \frac{(E(\epsilon^4|X) - 1)\sigma^4(x)}{nh_2 f_X(x)} V_l + o_p(1)$$

$$\text{MSE}(\hat{\sigma}_c^2(x)) = \left[ d \int_{-\frac{1}{2}}^{\tau} K_c(u) du \right]^2 + \frac{(E(\epsilon^4|X) - 1)\sigma^4(x)}{nh_2 f_X(x)} V_c + o_p(1)$$

If $h_1, h_2 \to 0$, $n \to \infty$ and $nh_2 \to \infty$
MSE (right side of discontinuity)

For \( x \in D_2 \) such that \( x = x_q + \tau h_2 \) with \( \tau \in [0, \frac{1}{2}] \) and a jump of magnitude \( d \),

\[
\text{MSE}(\hat{\sigma}_l^2(x)) = \left[ d \int_{-\frac{1}{2}}^{\tau} K_l(u) \frac{\mu_{l,2} - \mu_{l,1} u}{\mu_{l,0} \mu_{l,2} - \mu_{l,1}^2} du \right]^2 + \]

\[
\text{MSE}(\hat{\sigma}_c^2(x)) = \left[ d \int_{-\frac{1}{2}}^{\tau} K_c(u) du \right]^2 + \]

If \( h_1, h_2 \to 0, \ n \to \infty \) and \( nh_2 \to \infty \)
Consistency

- At points of continuity: all the estimators are consistent.
- At the right of the discontinuity: only the right estimator is consistent.
- At the left of the discontinuity: only the left estimator is consistent.
- The BPLL is consistent everywhere.
Theorem

If \( h_1, h_2 \to 0, \ n \to \infty \) and \( nh_1, nh_2 \to \infty \) and under certain regularity conditions, \( \sqrt{nh_2} (\sigma^2(x) - \hat{\sigma}^2_{BPLL}(x) - \beta_n(x)) \) is asymptotically normal with mean 0 and variance

\[
\frac{(E(\epsilon^4|X) - 1)\sigma^4(x)}{nh_2 f_X(x)} \int K_k^2(u) \left[ \frac{\mu_{k,2} - \mu_{k,1}u}{\mu_{k,0}\mu_{k,2} - \mu_{k,1}^2} \right]^2 du + o_p \left( \frac{1}{nh_2} \right),
\]

and bias

\[
\beta_n = \frac{h_2^2 \bar{\sigma}^2(x) \mu_{k,2}^2 - \mu_{k,1} \mu_{k,3}}{2 \mu_{k,2}\mu_{k,0} - \mu_{k,1}^2}
\]

for \( k = c, l, r \) as appropriate.
Alternative to the plug-in bandwidth estimator:

1. The leave–one–out cross validation:

\[
(h_{cv}^2, u_{cv}) = \arg \min_h \sum_{i=1}^n [\hat{r}_i^2 - \hat{\sigma}_i^2]^2
\]

where \(\hat{\sigma}_i^2\) is calculated without using the pair \((X_i, \hat{r}_i^2)\).

2. The leave a b–block–out cross validation for dependent data (Patton, Politis and White (2009) shows how to find the size of the block):

\[
(h_{b}^2, u_b) = \arg \min_h \sum_{i=1}^n [\hat{r}_i^2 - \hat{\sigma}_{-bi}^2]^2
\]

where \(\hat{\sigma}_{-bi}^2\) is calculated without using the \(2b+1\) pairs \((X_{i-b}, \hat{r}_{i-b}^2), \ldots, (X_i, \hat{r}_i^2), \ldots, (X_{i+b}, \hat{r}_{i+b}^2)\).
Ensuring positivity

The LL estimator, and therefore the BPLL estimator, is sometime negative for finite samples.

Solutions:

- Discard negative values.
- The re–weighted Nadaraya–Watson estimator (see Hall et al., 1999; Cai, 2002; and Phillips and Xu, 2007). It cannot be extended to estimate discontinuous volatility functions.
- The exponential local linear (ELL) (see Ziegelmann, 2002). Computationally heavy and theoretically obscure.
- Substitute any negative values of $\hat{\sigma}^2_k(x)$ by $\hat{\sigma}^2_{k,\text{ELL}}(x)$ for $k = c, l, r$. 
Experiment 1: iid variables

\[ Y_i = m(X_i) + \sigma(X_i)\epsilon_i \]

- \( \epsilon \sim IIDN(0, 1) \).
- \( X_i = IIDU(-2, 2) \), random design.
- \( x \) are \( T = 250 \) equidistant values in \([-1.8, 1.8]\).
- \( n = 500, 1000, 2000 \), number of simulations \( N=200 \).
- \( \epsilon_i \) and \( X_i \) are independent.
- Leave–one–out cross validation.
- \( \sigma(x) \) has two discontinuities at \( x = -1, 1 \).
- Four scenarios depending on \( m(x) \):
  - Scenario I: \( m \equiv 0 \)
  - Scenarios II, III, IV:
# Comparison LL vs. BPLL (MISE)

<table>
<thead>
<tr>
<th>Method</th>
<th>LL</th>
<th>BPLL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\widehat{\text{MISE}}$</td>
<td>$\widehat{\text{MISE}}_q$</td>
</tr>
<tr>
<td>$n = 500$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Scenario I | 0.0089  
0.0060 | 0.0098  
0.0039 |
| Scenario II | 0.0098  
0.0062 | 0.0120  
0.0039 |
| Scenario III | 0.0093  
0.0061 | 0.0109  
0.0040 |
| Scenario IV | 0.0107  
0.0067 | 0.0123  
0.0042 |
| $n = 1000$|
| Scenario I | 0.0047  
0.0034 | 0.0037  
0.0015 |
| Scenario II | 0.0044  
0.0032 | 0.0043  
0.0018 |
| Scenario III | 0.0048  
0.0034 | 0.0045  
0.0017 |
| Scenario IV | 0.0044  
0.0032 | 0.0040  
0.0016 |
| $n = 2000$|
| Scenario I | 0.0021  
0.0016 | 0.0012  
0.0005 |
| Scenario II | 0.0020  
0.0016 | 0.0012  
0.0006 |
| Scenario III | 0.0020  
0.0016 | 0.0012  
0.0006 |
| Scenario IV | 0.0022  
0.0016 | 0.0013  
0.0005 |
Comparison LL vs. BPLL

(a) LL with $n = 500$

(b) LL with $n = 2000$

(c) BPLL with $n = 500$

(d) BPLL with $n = 2000$
Comparison LL vs. BPLL (Error boxplot)

(a) $n = 2000$ in $D_1$

(b) $n = 2000$ in $D_2$
Experiment 2: a square root diffusion

The process is of the form:

\[ dX_t = \kappa(\theta - X_t)dt + \sigma \sqrt{X_t}dB_t \]

The process was generated following the algorithm in Section 3.4 of Glasserman (2004).

- \( x \) are \( T = 250 \) equidistant values in \([0.03,0.12]\).
- \( n = 500, 1000, 2000 \), number of simulations \( N = 400 \).
- \( B_t \) and \( X_t \) are independent.
- Leave–b–block–out cross validation to obtain the bandwidth.
- The drift and diffusion are discontinuous at \( x = 0.1 \).
Comparison LL vs. BPLL (MISE)

<table>
<thead>
<tr>
<th>Method</th>
<th>LL MISE</th>
<th>LL MISE&lt;sub&gt;q&lt;/sub&gt;</th>
<th>BPLL MISE</th>
<th>BPLL MISE&lt;sub&gt;q&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 500</td>
<td>0.0241</td>
<td>0.0083</td>
<td>0.0114</td>
<td>0.0057</td>
</tr>
<tr>
<td>n = 1000</td>
<td>0.0120</td>
<td>0.0041</td>
<td>0.0039</td>
<td>0.0022</td>
</tr>
<tr>
<td>n = 2000</td>
<td>0.0059</td>
<td>0.0021</td>
<td>0.0013</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

**Table:** MISE of LL and BPLL comparison for Experiment 2.
Comparison LL vs. BPLL

(a) LL with $n = 500$

(b) LL with $n = 2000$

(c) BPLL with $n = 500$

(d) BPLL with $n = 2000$
Comparison LL vs. BPLL (Boxplots)

(a) $n = 500$ in $D_1$

(b) $n = 500$ in $D_2$

(a) $n = 2000$ in $D_1$

(b) $n = 2000$ in $D_2$
Conclusions

- The break preserving estimator is consistent in the presence of discontinuities.
- It is always positive.
- It keeps some of the smooth properties of the LL in the continuous parts.
Further interest

- Application to the spot volatility of intra–day data (SPDR).
  - ...
Further interest

- Application to the spot volatility of intra–day data (SPDR).
  - ...

- Application to the estimation of interest rates: changes of structure in the drift and volatility.
  - ...

...
Simulated volatility function

Volatility function

\[ \sigma(x) \]

Graph showing the volatility function \( \sigma(x) \) with values ranging from -1.5 to 1.5 on the x-axis and from 0.2 to 0.8 on the y-axis.
Scenario II

Continuous function.

Scenario II: Drift function

\[ m(x) \]
Scenario III

One discontinuity at $x = 0$. 

Scenario III: drift function

$m(x)$

$x$
Two discontinuities at the same points than the volatility function $x = -1$ and $x = 1$. 

**Scenario IV: drift function**

\[ m(x) \]

\[ -1.5 \quad -1.0 \quad -0.5 \quad 0.0 \quad 0.5 \quad 1.0 \quad 1.5 \]

\[ -1.5 \quad -1.0 \quad -0.5 \quad 0.0 \quad 0.5 \quad 1.0 \quad 1.5 \]