Parametric Inference and Dynamic State Recovery from Option Panels

Torben G. Andersen

Joint work with **Nicola Fusari** and **Viktor Todorov**

The Third International Conference
High-Frequency Data Analysis in Financial Markets

Hiroshima University of Economics

November 17, 2012

Motivation

- Under realistic assumptions: Derivatives Non-Redundant Assets.

- Contain important Information about Volatility and Jump risks and their **Pricing**:

Option Price $\Longleftrightarrow$ [Volatility & Jumps] $+$ [**Time-Varying Risk Premia**] $+$ [$Observation\ Error$]

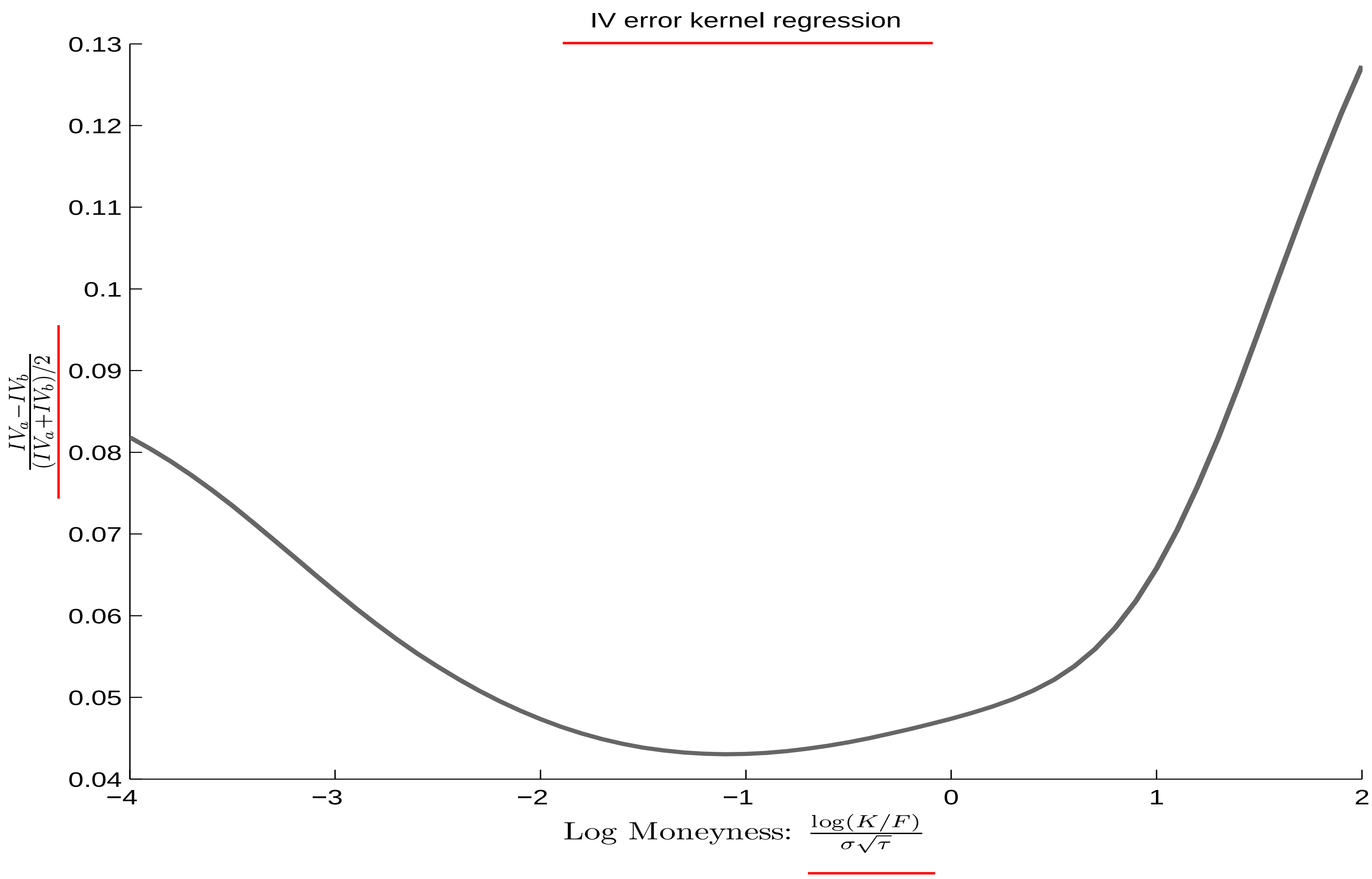
- Derivatives on Equity Indices actively Traded:
  - On Average, 200 plus SPX Option Quotes at Close of Trading,
  - Cover a wide range of Moneyness and Tenor (time-to-maturity).

Motivation

Most Parametric Option-based Estimation Methods of Risk Premia follow Two Steps:

(Bates (2000), Pan (2002), Eraker (2004), BCJ (2007), Christoffersen at al. (2006-08))

1. Identify Volatility and Jump Risks from underlying **Asset Data**,
2. Use **Price Levels in Option Panel** to estimate Risk Premia:
   - typically via restrictive Specification; i.e., small $\mathbb{P} - \mathbb{Q}$ wedge.
   - Bates (2000), Pan (2002) use aspects of Implied Volatility State Dynamics.
   - Option Price Error ignored or modeled with Normal Distribution.


IV error kernel regression
0.13
0.12
0.11
0.1
0.09
0.08
0.07
0.06
0.05
0.04
$\frac{IV_a - IV_b}{(IV_a + IV_b)/2}$
−4
−3
−2
−1
0
1
2
Log Moneyness: $\frac{\log(K/F)}{\sigma\sqrt{\tau}}$

Motivation

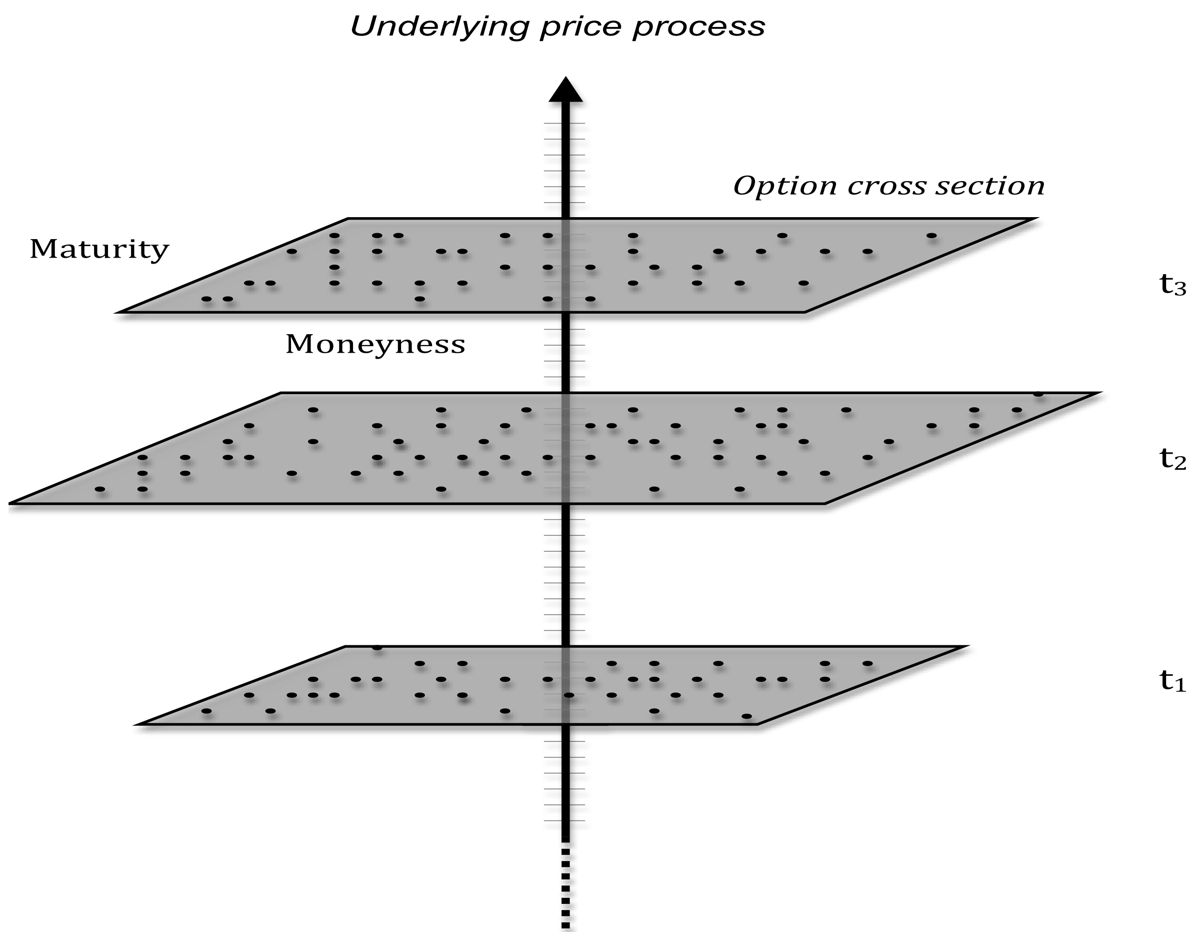
Goal to develop Estimation Technique that:

1. fully uses the State Dynamics implied by Option Prices;
2. is Robust with respect to Option Price Error Specification;
3. relies on In-Fill Asymptotics (Increasing number of Options each Trading Day);
4. Specifies only Risk-Neutral Dynamics (allows for flexible risk premia).

In Sum: Formal Estimation, Inference and Diagnostic Tests for Option Pricing

Obtain Path of State Vector Realizations solely from Option Panel

Set Stage for Risk Premia Estimation via Semi-Parametric $\mathbb{P}$ Estimation


Underlying price process
Option cross section
Maturity
Moneyness
t_3
t_2
t_1

Outline

- Information in Option Panels
- Inference in Presence of Noise
- Semiparametric Tests

Notation

Formally, underlying Price X_t has the following $\mathbb{P}$-Dynamics:

$$\frac{dX_t}{X_{t-}} = \alpha_t \, dt + \sqrt{V_t} \, dW_t^{\mathbb{P}} + \int_{x>-1} x \, \widetilde{\mu}^{\mathbb{P}}(dt, dx),$$

$W_t^{\mathbb{P}}$ is a Brownian motion; $\sqrt{V_t}$ is ***Spot*** Volatility (under both $\mathbb{P}$ and $\mathbb{Q}$);

$\mu^{\mathbb{P}}$ is an Integer-valued Random Measure; $\widetilde{\mu}^{\mathbb{P}} = \mu^{\mathbb{P}} - \nu^{\mathbb{P}}$

$\mu^{\mathbb{P}}$ Counts Jumps in X; Jump Compensator is $\nu^{\mathbb{P}}(ds, dx)$.

We assume $V_t = \xi_1(\mathbf{S}_t)$, where $\mathbf{S}_t$ is Latent State Vector $(p \times 1)$.

Notation

Assumption A0. *The process* X*, defined over the fixed interval* $[0,\ T]$*, satisfies:*

(i) *For* $s, t \geq 0$*, exists* $K > 0$*:* $\mathbb{E}\left\{|V_t - V_s|^2 \wedge K\right\} \leq K|t-s|$.

(ii) $\int_{x>-1} (|x|^\beta \wedge 1)\, \nu^{\mathbb{P}}(dx) < \infty$*, for some* $\beta \in [0,2)$.

(iii) $\inf_{t\in[0,T]} V_t > 0$ *and the processes* α_t*,* V_t *and* a_t *are locally bounded.*

A0(i) satisfied if V_t is governed by (multivariate) Stochastic Differential Equation

A0(ii) restricts so-called Blumenthal-Getoor index of the jumps to be below β

A0(iii) implies, at each $t \in [0, T]$, the price process has Non-Vanishing BM Component

Assumption A0 does not involve Integrability or Stationarity Conditions for the Model

Notation

Likewise, X_t has $\mathbb{Q}$-Dynamics:

$$\frac{dX_t}{X_{t-}} = (\, r_t - \delta_t \,)\, dt + \sqrt{V_t}\, dW_t + \int_{x>-1} x\, \widetilde{\mu}(dt, dx),$$

$\nu(dt, dx) = \xi_2(\mathbf{S}_t) \otimes \nu(dx)$, where $\widetilde{\mu} = \mu - \nu$

We denote Options with Log-Moneyness $k = \log(K/X_t)$ and Tenor τ by

$$O_{t,k,\tau} = \mathbb{E}_t^{\mathbb{Q}} \left\{ e^{-\int_t^{t+\tau} (r_s - \delta_s) ds} \, (X_{t+\tau} - K)^+ \right\},$$

We denote associated Black-Scholes Implied Volatility by $\kappa(k, \tau, \mathbf{S}_t)$.

An Empirical Illustration

The "Double-Jump" Model of Duffie, Pan and Singleton (2001) has Risk-Neutral Dynamics:

$$\frac{dX_t}{X_{t-}} = (r - \delta)\, dt + \sqrt{V_t}\, dW_t + dL_{x,t},$$

$$dV_t = \kappa\, (\overline{v} - V_t)\; dt + \sigma_d \sqrt{V_t}\, dB_t + dL_{v,t},$$

$L_{x,t}$ and $L_{v,t}$ are Jump Martingales;

$(L_{x,t}, L_{v,t})$ Jump (simultaneously) with i.i.d. Probability λ_j, Jump Size $(Z_x\,, Z_v\,)$.

$Z_v \sim exp\,(\mu_v)$; $log\,(Z_x + 1)\,|Z_v \sim \mathcal{N}\,\left(\mu_x + \rho_j Z_v, \sigma_x^{\,2}\right)$; $Cor\,(dW_t, dB_t\,) = \rho_d$

State Vector (Realization): $\{V_t\,\}_{t=1}^{T}$

Risk-Neutral Parameters: $\theta = (\rho_d, \overline{v}, \kappa, \sigma_d, \lambda, \mu_x, \sigma_x, \mu_v, \rho_j)'\,.$

Monte Carlo Scenario

Inspired by Calibration/Empirical Estimates from BCJ (2007); $\rho_j = 0$.

Table 1: Parameter Setting for the Numerical Experiments

Under $\mathbb{P}$				Under $\mathbb{Q}$			
Parameter	Value	Parameter	Value	Parameter	Value	Parameter	Value
ρ_d	-0.4600	λ_j	1.0080	ρ_d	-0.4600	λ_j	1.0080
$\overline{v}$	0.0144	μ_x	-0.0284	$\overline{v}$	0.0144	μ_x	-0.0501
κ_d	4.0320	σ_x	0.0490	κ_d	4.0320	σ_x	0.0751
σ_d	0.2000	μ_v	0.0315	σ_d	0.2000	μ_v	0.0930

Observation Errors on Option Prices: $\epsilon_{t,k,\tau} = \sigma_{t,k,\tau} Z_{t,k,\tau}, \quad Z_{t,k,\tau} \sim \mathcal{N}(0,1).$

$\sigma_{t,k,\tau} = \frac{1}{2}\psi_k / Q_{0.995}, \quad \psi_k =$ Bid-Ask Spread "at k," $\quad Q_p = p^{\text{th}}$ Quantile of $\mathcal{N}(0,1)$.

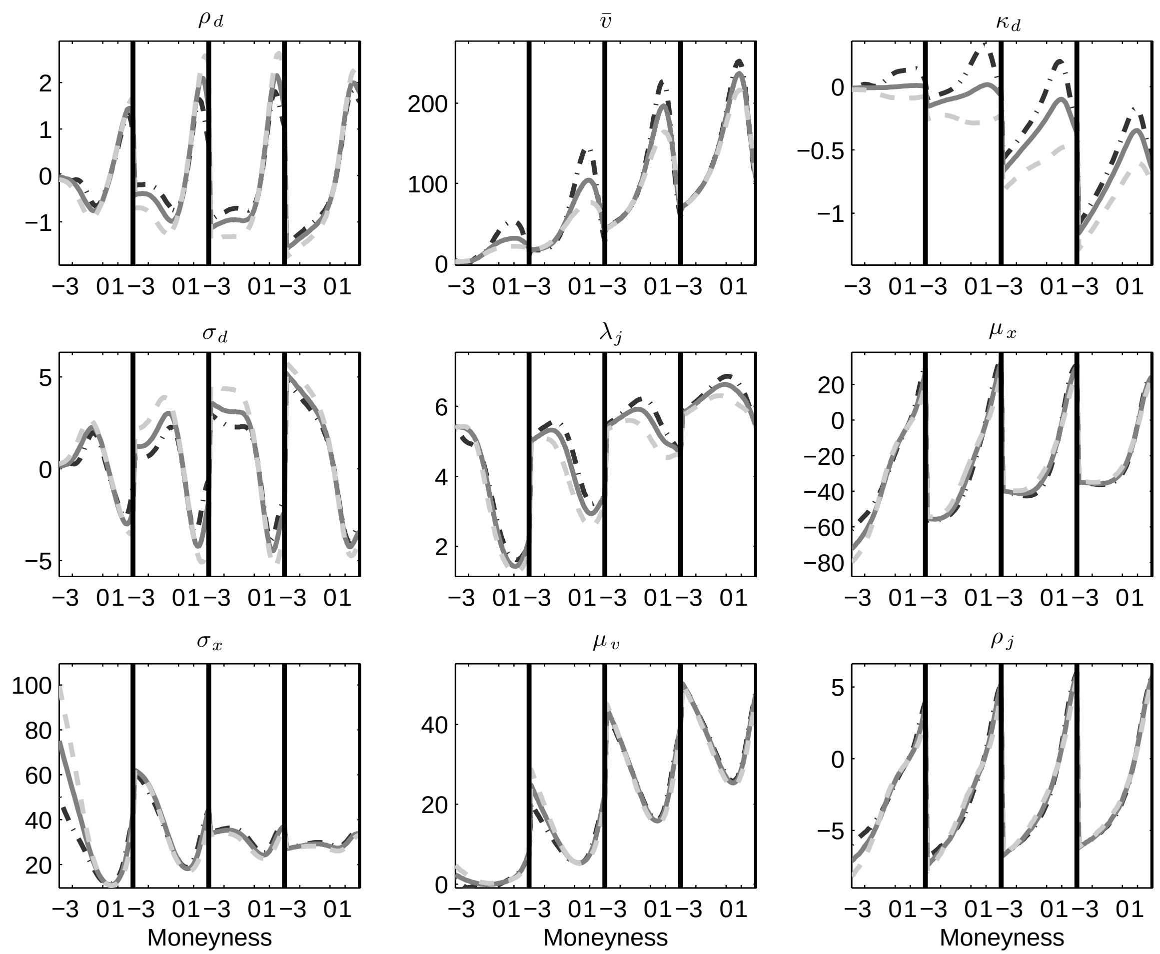
Information in Panels of Options

What can we Identify from Options?

- Different Parts of Volatility Surface Load differently on distinct Risks and their Pricing:
  - Short-Term OTM Options determined largely by Pricing of Jump Risks
  - Role of Volatility Risks more prominent for ATM Options
  - Different Maturities separate Persistent from Transient State Variables
  - Persistence of "Smirk" Identifies Sources of "Leverage" type Effects
  - ....

Can we identify the model?

Option Sensitivity to Parameters; Double-Jump Model



Information in Panels of Options

....

- $\Longrightarrow$ A Large Cross-Section of Option Prices observed without Error can Identify Risk-Neutral Parameters and the Current Value of the State Vector,

- Once Risk-Neutral Parameters are Known, Options are Known Transformations of the State Variables $\Longrightarrow$ Contain Same Information as observing directly the State Vector,

- $\Longrightarrow$ Options alone Contain Information to Estimate the Risk Premia!

Information in Panels of Options

Assumption A1. *Fix* $T > 0$. *For each Date* $t = 1, .., T$ *and Moneyness* τ,

 # options $N_t^\tau \uparrow \infty$ *with* $N_t^\tau / N_t \to \pi_t^\tau$ *and* $N_t / \sum_{t=1}^T N_t \to \varsigma_t$, *where* π_t^τ, $\varsigma_t > 0$.

Let $\underline{k}(t,\tau)$, $\overline{k}(t,\tau)$ *denote Min, Max Log-Moneyness on Day* t, *Maturity* τ.

Sequence of Grid Nested: $\underline{k}(t,\tau) = k_{t,\tau}(0) < k_{t,\tau}(1) ... < k_{t,\tau}(N_t^\tau) = \overline{k}(t,\tau)$.

$N_t \cdot (k_{t,\tau}(i) - k_{t,\tau}(i-1)) \to \psi_{t,\tau}(k)$ *Uniformly on* $(\underline{k}(t,\tau), \overline{k}(t,\tau))$.

Assumption A2. *For every* $\epsilon > 0$ *and* $T > 0$ *finite, we have a.s.*

$$\inf_{t=1,..,T:\ \cup||\mathbf{Z}_t - \mathbf{S}_t|| > \epsilon\ \cup ||\theta - \theta_0|| > \epsilon} \sum_{t=1}^T \sum_\tau \int_{\underline{k}(t,\tau)}^{\overline{k}(t,\tau)} \left(\kappa(k,\tau,\mathbf{S}_t,\theta_0) - \kappa(k,\tau,\mathbf{Z}_t,\theta)\right)^2 dk > 0,$$

where θ *is the risk-neutral parameter vector.*

Inference in the Presence of Noise

Options are Observed with Error, i.e., we observe $\widehat{\kappa}_{t,k,\tau}$ for

$$\widehat{\kappa}_{t,k,\tau} = \kappa_{t,k,\tau} + \epsilon_{t,k,\tau},$$

where the errors, $\epsilon_{t,k,\tau}$, are defined on an extension of the original probability space.

We assume the Error can be averaged out by Pooling Options across Moneyness:

Assumption A3. *For every* $\epsilon > 0$ *and* $T > 0$ *finite, we have*

$$\sup_{t=1,..,T:\ \cup||\mathbf{Z}_t - \mathbf{S}_t|| > \epsilon\ \cup||\theta-\theta_0||>\epsilon} \frac{\sum_{t=1}^{T} \frac{1}{N_t} \sum_{j=1}^{N_t} \Big(\kappa(k_j,\tau_j,\mathbf{S}_t,\theta_0) - \kappa(k_j,\tau_j,\mathbf{Z}_t,\theta)\Big)\,\epsilon_{t,k,\tau}}{\sum_{t=1}^{T} \frac{1}{N_t} \sum_{j=1}^{N_t} \Big(\kappa(k_j,\tau_j,\mathbf{S}_t,\theta_0) - \kappa(k_j,\tau_j,\mathbf{Z}_t,\theta)\Big)^2} \xrightarrow{\mathbb{P}} 0,$$

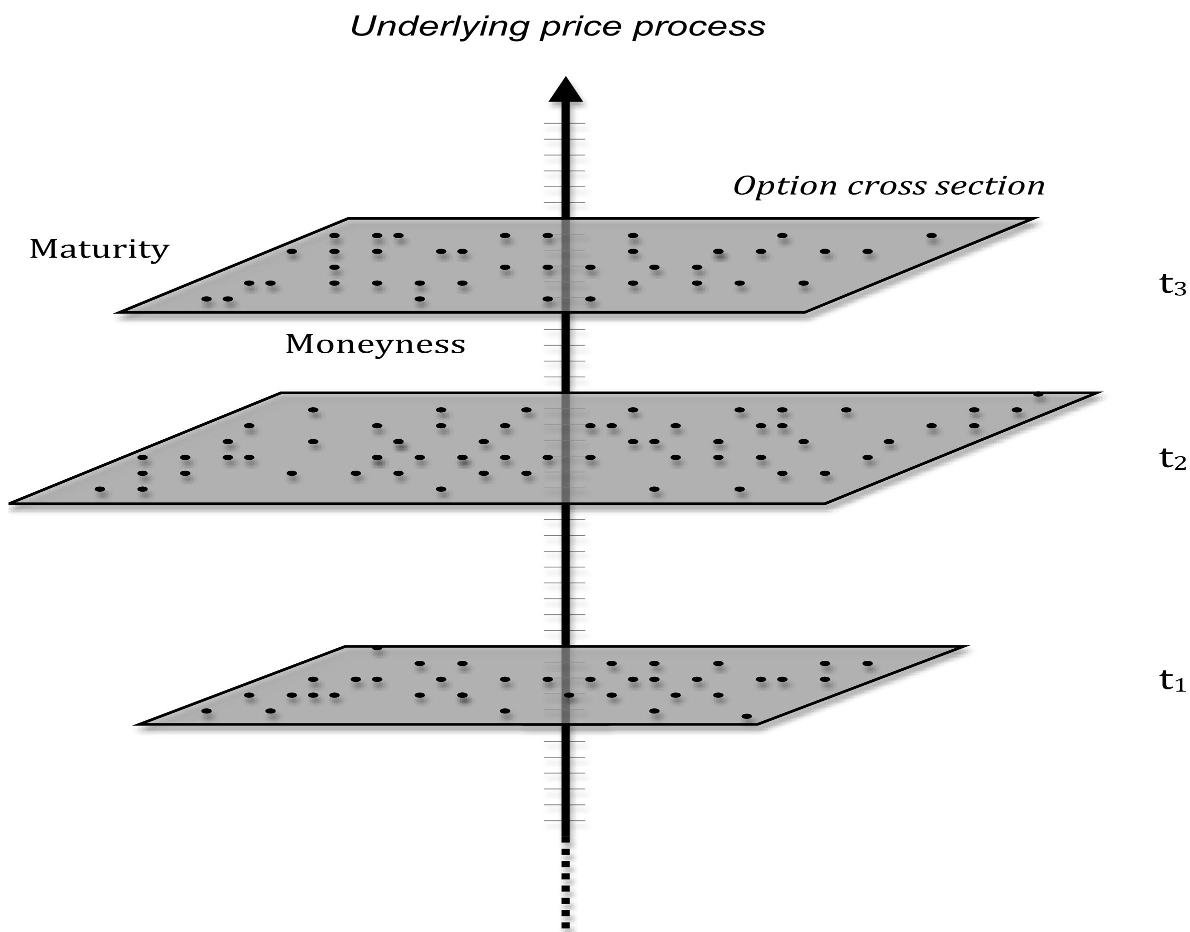
when $\min_{t=1,\ldots,T} N_t \to \infty$ *for all* $\theta \in \Theta$*.*

Estimation

We define our estimator of risk-neutral parameters and state variables as

$$\left(\{\widehat{\mathbf{S}}_t^n\}_{t=1,\dots,T}, \widehat{\theta}^n\right) = \underset{\{\mathbf{Z}_t\}_{t=1,\dots,T},\, \theta\in\Theta}{\operatorname{argmin}} \sum_{t=1}^{T} \left\{ \frac{1}{N_t} \sum_{j=1}^{N_t} \left(\widehat{\kappa}_{t,k,\tau} - \kappa(k_j, \tau_j, \mathbf{Z}_t, \theta)\right)^2 + \lambda_n \left(\widehat{V}_t^n - \xi_1(\mathbf{Z}_t)\right)^2 \right\},$$

$\lambda_n \geq 0$, $\widehat{V}_t^n$ is Nonparametric Estimator of Volatility from High-Frequency Data.


Underlying price process
Option cross section
Maturity
Moneyness
t3
t2
t1

Estimation

Theorem 1. *Suppose Assumptions A1-A3 Hold for some* $T \in \mathbb{N}$ *fixed,*

and $\{\widehat{V}_t^n\}_{t=1,\ldots,T}$ *is Consistent for* $\{V_t\}_{t=1,\ldots,T}$, *as* $n \to \infty$.

Then, if $\min_{t=1,\ldots,T} N_t \to \infty$ *and* $\lambda_n \to \lambda$ *for some finite* $\lambda \geq 0$, *as* $n \to \infty$,

we have that $\left(\widehat{\mathbf{S}}_t^n, \widehat{\theta}_t^n\right)$ *exists with probability approaching* 1, *and*

$$||\widehat{\mathbf{S}}_t^n - \mathbf{S}_t|| \xrightarrow{\mathbb{P}} 0, \quad ||\widehat{\theta}^n - \theta_0|| \xrightarrow{\mathbb{P}} 0, \quad t = 1, \ldots, T.$$

Estimation

To quantify precision of estimation we need slightly stronger assumption on errors:

Assumption A4. *For the error process, $\epsilon_{t,k,\tau}$, we have,*

(i) $\mathbb{E}\left(\epsilon_{t,k,\tau} | \mathcal{F}^{(0)}\right) = 0$

(ii) $\mathbb{E}\left(\epsilon^2_{t,k,\tau} | \mathcal{F}^{(0)}\right) = \phi_{t,k,\tau}$, *for* $\phi_{t,k,\tau}$ *continuous in its second argument*

(iii) $\epsilon_{t,k,\tau}$, $\epsilon_{t',k',\tau'}$ *are independent, conditional on* $\mathcal{F}^{(0)}$, *for* $(t, k, \tau) \neq (t', k', \tau')$

(iv) $\mathbb{E}\left(|\epsilon_{t,k,\tau}|^4 | \mathcal{F}^{(0)}\right) < \infty$, *almost surely*

where $\mathcal{F}^{(0)}$ is the σ-algebra associated with X.

Estimation

Theorem 2. *Assume Assumptions A1-A4 Satisfied for $T \in \mathbb{N}$ Fixed, and $\kappa(t, \tau, \mathbf{Z}, \theta)$ is twice Continuously-Differentiable in its arguments.*

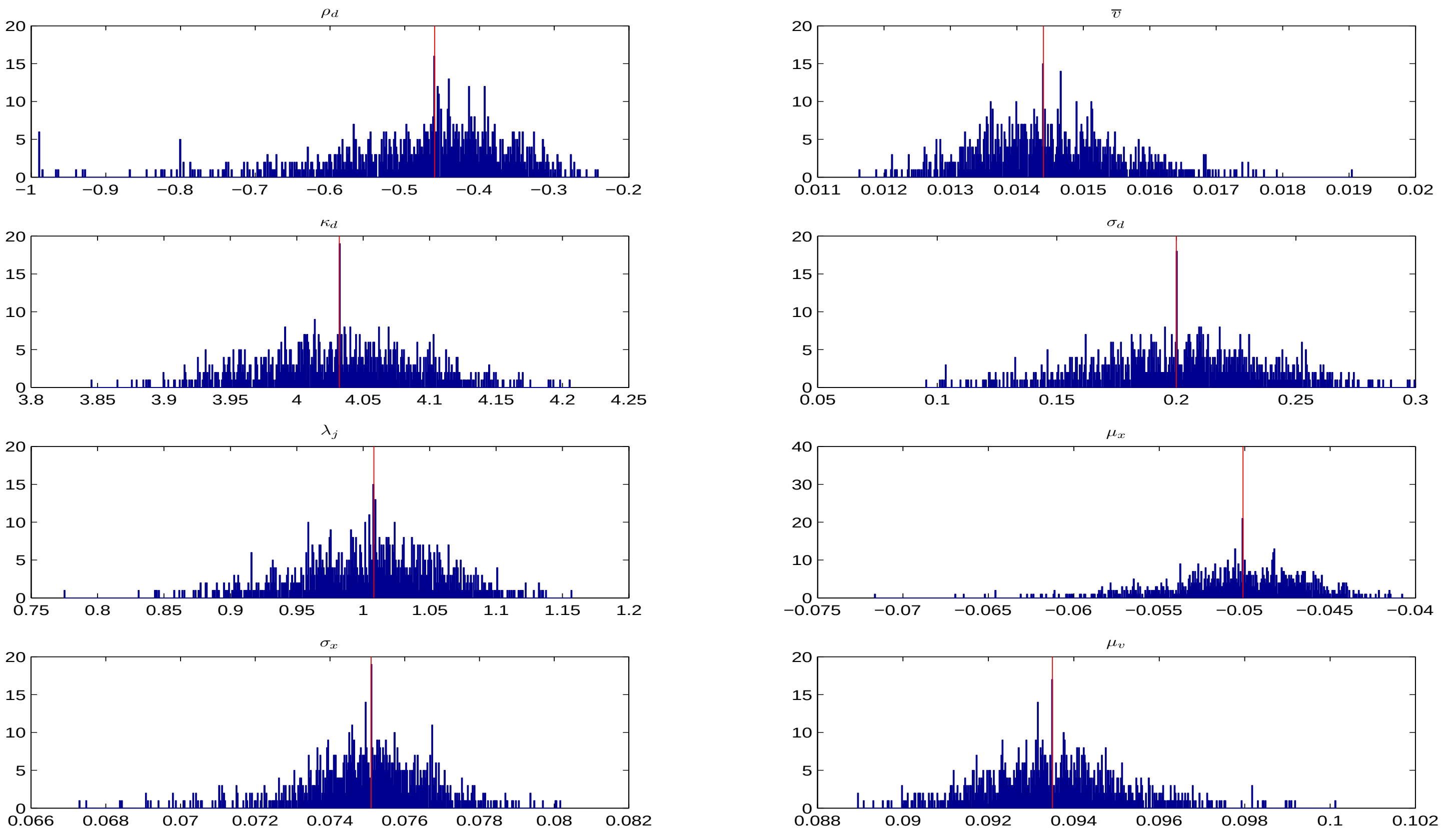
Then, if $\min_{t=1,\dots,T} N_t \to \infty$ *and* $\lambda_n^2 \min_{t=1,\dots,T} N_t \to 0$, *for* $n \to \infty$:

$$\begin{pmatrix} \sqrt{N_1}(\widehat{\mathbf{S}}_1^n - \mathbf{S}_1) \\ \vdots \\ \sqrt{N_T}(\widehat{\mathbf{S}}_T^n - \mathbf{S}_T) \\ \sqrt{\frac{N_1+\dots+N_T}{T}}(\widehat{\theta}^n - \theta_0) \end{pmatrix} \xrightarrow{\mathcal{L}-s} \mathbf{H}_T^{-1} \left(\boldsymbol{\Omega}_T\right)^{1/2} \begin{pmatrix} \mathbf{E}_1 \\ \vdots \\ \mathbf{E}_T \\ \mathbf{E}' \end{pmatrix},$$

$\mathbf{E}_1$,...,$\mathbf{E}_T$ *are* $p \times 1$ *vectors,* $\mathbf{E}'$ *is* $q \times 1$ *vector, all are i.i.d. Standard Normal and Defined on an Extension of the original Probability Space,*

$\mathbf{H}_T$, $\boldsymbol{\Omega}_T$ *are* $\mathcal{F}_T^{(0)}$*-adapted Random Matrices, for which Consistent Estimates can be Constructed from Options Data.*

MC Estimation; Double-Jump Model; 1,000 Replications



Empirical Application

We use the following Data Set in the Application

- CBOE European-style (SPX) Options on the S&P 500 index,
- The Options have Maturity Ranging from 8 Days to 1 Year,
- The Data Covers Period $1996 - 2010$ for a Total of $3,500$ Days,
- We Apply Standard Filters; Retain only OTM and ATM Options; Wide Strike Range,
- For Semi-Parametric Tests: 5-minute S&P 500 Futures, Same Sample Period.

Model-free vs Option-Implied Volatility

- We specify and estimate the risk-neutral distribution of the underlying process X and we do not impose any parametric structure for the dynamics under the true statistical measure $\mathbb{P}$.

- Absence of arbitrage implies that recovered volatility from options should be the same with that "observed" in the underlying asset X.

- This is a semiparametric restriction: it is based on a parametric specification for the risk-neutral distribution as well as nonparametric estimate for the stochastic volatility.

An Empirical Illustration

The "Double-Jump" Model of Duffie, Pan and Singleton (2001) has Risk-Neutral Dynamics:

$$\frac{dX_t}{X_{t-}} = (r - \delta)\, dt + \sqrt{V_t}\, dW_t + dL_{x,t},$$

$$dV_t = \kappa\, (\overline{v} - V_t)\, dt + \sigma_d \sqrt{V_t}\, dB_t + dL_{v,t},$$

$L_{x,t}$ and $L_{v,t}$ are Jump Martingales;

$(L_{x,t}, L_{v,t})$ Jump (simultaneously) with i.i.d. Probability λ_j, Jump Size (Z_x, Z_v).

$Z_v \sim exp\,(\mu_v)$; $log\,(Z_x + 1)\,|Z_v \sim \mathcal{N}\left(\mu_x + \rho_j Z_v, \sigma_x^2\right)$; $Cor\,(dW_t, dB_t) = \rho_d$

State Vector (Realization): $\{V_t\}_{t=1}^{T}$

Risk-Neutral Parameters: $\theta = (\rho_d, \overline{v}, \kappa, \sigma_d, \lambda, \mu_x, \sigma_x, \mu_v, \rho_j)'$.

Empirical Application

Qualitative Features of the Double-Jump Model

- Both Stochastic Volatility and Price Jumps Present,
- Volatility can Move through Jumps,
- Price and Volatility may be Correlated through Small and Big Moves,
- Jump Intensity is Constant.
- (Only) One Volatility Factor.
- Slightly more General than BCJ "State-of-Art"

Empirical Application

Table 2: Parameter Estimates of One-Factor Model

Parameter	Estimate	Standard Error	Parameter	Estimate	Standard Error
ρ_d	-0.9586	0.0059	λ	0.0150	0.0006
$\overline{v}$	0.0356	0.0004	μ_x	0.5833	0.0417
κ	1.4350	0.0198	σ_x	0.3220	0.0430
σ_d	0.2965	0.0032	μ_v	1.6162	0.1912
			ρ_j	-0.8579	0.1572

Risk Neutral Mean of Volatility: 24.3% vs Sample RV Estimate: 21.35%

Mean Risk-Neutral (Log-Return) Jump: -80% Annual Jump Probability: 1.5%

Mean Risk-Neutral Volatility (Level) Jump: $\sqrt{1.62} = 127\%$

Diagnostic Tests

We design the following Diagnostic Tests of Model Performance:

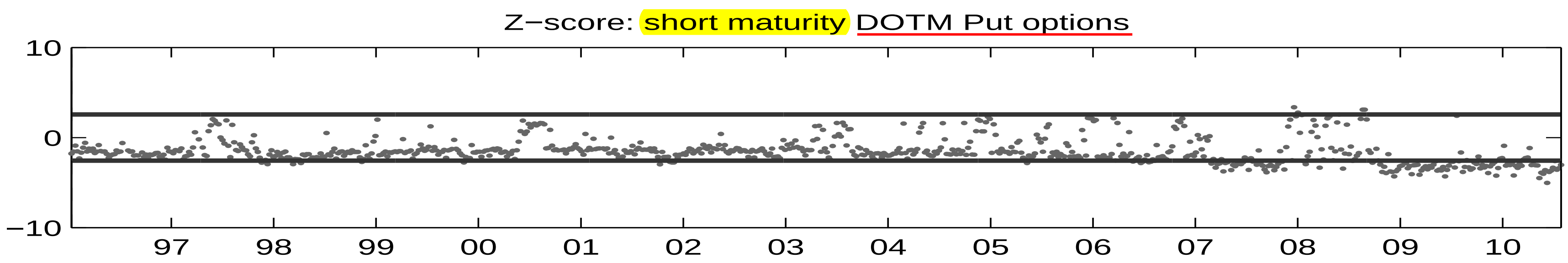
- Fit to the Volatility Surface over some Period of Time
- Parameter Stability across Time
- Distance between Model-Free and Option-Model Implied Volatility

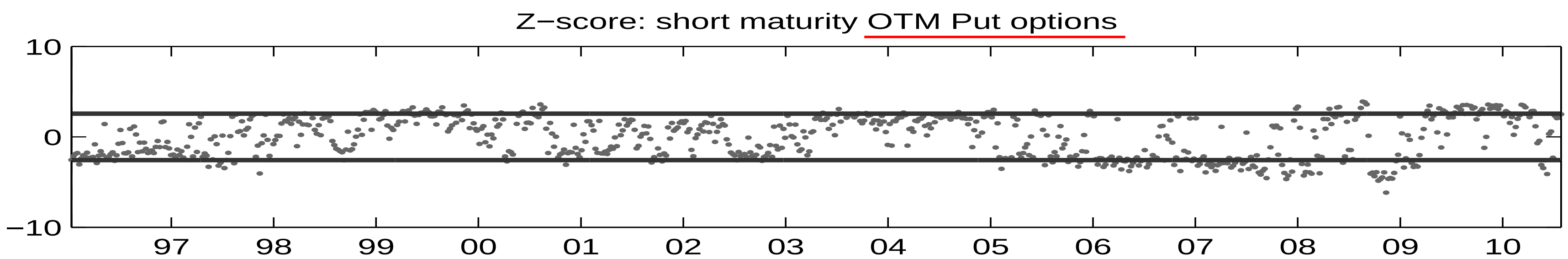
Diagnostic Test I: Fit to Volatility Surface

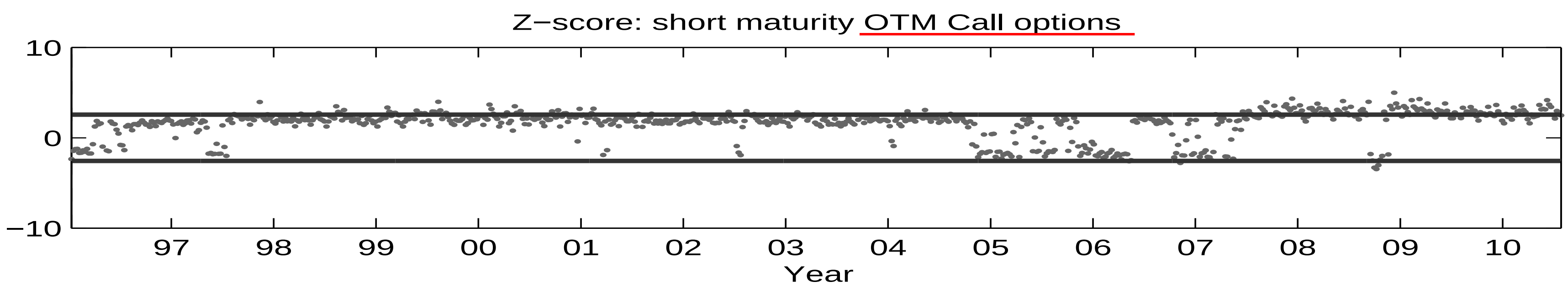
Corollary 1. *Let $\mathcal{K} \subset \left(\underline{k}(t,\tau^*),\ \overline{k}(t,\tau^*)\right)$ be a set with positive Lebesgue measure and $N_t^{\mathcal{K}}$ be the number of options on day t with tenor τ^* and log-moneyness in $\mathcal{K}$. Then, given our Assumptions, we have,*

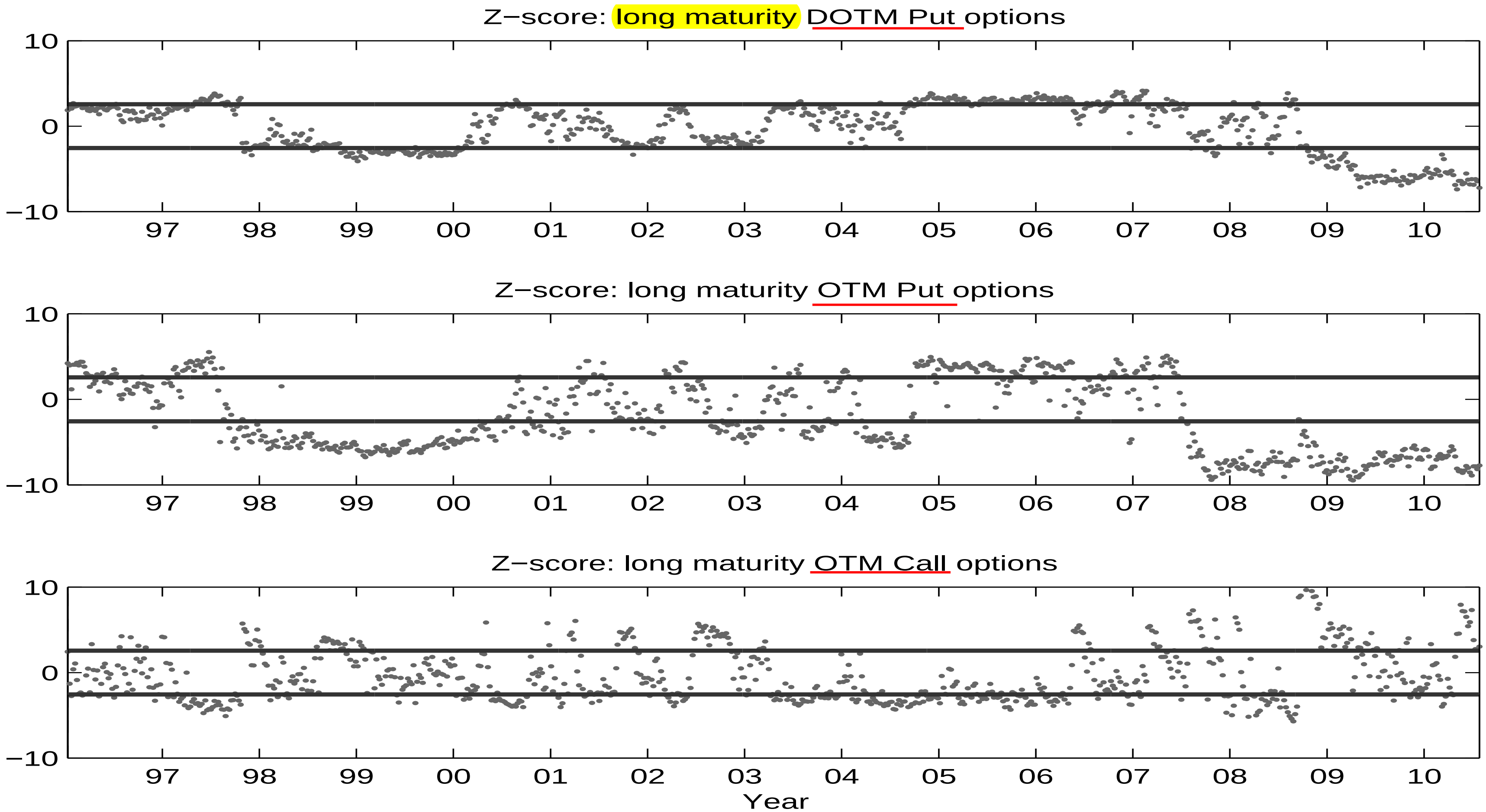
$$\frac{\sum_{j:k_j\in\mathcal{K}}\left(\widehat{\kappa}_{t,k_j,\tau^*} - \kappa(k_j,\tau^*,\widehat{\mathbf{S}}_t^n,\widehat{\theta}^n)\right)}{\sqrt{\widehat{\Pi}_T'\widehat{\Xi}_T\widehat{\Pi}_T}} \quad \xrightarrow{\mathcal{L}-s} \quad \mathcal{N}(0,1),$$

where $\widehat{\Pi}_T$ and $\widehat{\Xi}_T$ are some $\mathcal{F}_T^{(0)}$-adapted random matrices.


Z−score: short maturity DOTM Put options
10
0
−10
97
98
99
00
01
02
03
04
05
06
07
08
09
10


Z−score: short maturity OTM Put options
10
0
−10
97
98
99
00
01
02
03
04
05
06
07
08
09
10


Z−score: short maturity OTM Call options
10
0
−10
97
98
99
00
01
02
03
04
05
06
07
08
09
10
Year


Z−score: long maturity DOTM Put options
10
0
−10
97 98 99 00 01 02 03 04 05 06 07 08 09 10
Z−score: long maturity OTM Put options
10
0
−10
97 98 99 00 01 02 03 04 05 06 07 08 09 10
Z−score: long maturity OTM Call options
10
0
−10
97 98 99 00 01 02 03 04 05 06 07 08 09 10
Year

Diagnostic Test II: Parameter Stability

Parameters Estimated over Non-Overlapping Periods should, up to Statistical Error, be Identical. Thus,

$$\left(\widehat{\theta}_1^n - \widehat{\theta}_2^n\right)' \left(\widehat{\text{Avar}}(\widehat{\theta}_1^n) + \widehat{\text{Avar}}(\widehat{\theta}_2^n)\right)^{-1} \left(\widehat{\theta}_1^n - \widehat{\theta}_2^n\right) \quad \xrightarrow{\mathcal{L}-s} \quad \chi^2(q),$$

where $\widehat{\text{Avar}}(\widehat{\theta}_j^n)$ is Consistent Asymptotic Variance Estimate for $\widehat{\theta}_j^n$.

Note:

- Under Model Misspecification, Parameter Estimates Converge to Pseudo-True Values.
- However, as State Vector changes over Time $\Longrightarrow$ Pseudo-True Values Change as well.

SPX Options – Parameter Stability Test

Table 3: Parameter Stability; S&P 500 Options Data

Parameter	Nominal size of test		Parameter	Nominal size of test	
	1%	5%		1%	5%
Panel A: One-Factor Model					
ρ_d	62.86%	70.48%	λ_j	56.19%	67.62%
$\overline{v}$	71.43%	73.33%	μ_x	20.00%	25.71%
κ	91.43%	93.33%	σ_x	49.52%	61.91%
σ_d	77.14%	80.95%	μ_v	31.43%	36.19%
			ρ_j	13.33%	17.14%
Panel B: Two-Factor Model					
$\rho_{d,1}$	8.57%	16.19%	$\lambda_{j,0}$	37.14%	49.52%
$\overline{v}_1$	7.62%	12.38%	$\lambda_{j,1}$	21.91%	33.33%
$\kappa_{d,1}$	72.38%	70.05%	μ_x	0.00%	4.76%
$\sigma_{d,1}$	42.86%	53.33%	σ_x	9.52%	17.14%
$\rho_{d,2}$	7.62%	16.19%	μ_v	23.81%	34.29%
$\overline{v}_2$	76.19%	80.95%	ρ_j	1.91%	2.86%
$\kappa_{d,2}$	69.52%	77.14%			
$\sigma_{d,2}$	0.95%	0.95%			

Model-Free vs Option-Implied Volatility

Two Nonparametric Estimators for Spot Volatility from High-Frequency Data:

$$\widehat{V}_t^{\pm,n} = \frac{n}{k_n} \sum_{i \in I^{\pm,n}} (\Delta_i^{t,n} X)^2 1\left(|\Delta_i^{t,n} X| \leq \alpha n^{-\varpi}\right), \quad \Delta_i^{t,n} X = \log\left(X_{t+\frac{i}{n}}\right) - \log\left(X_{t+\frac{i-1}{n}}\right),$$

where $\alpha > 0$, $\varpi \in (0, 1/2)$, k_n is Deterministic sequence, $k_n/n \to 0$, and

$$I^{-,n} = \{-k_n + 1, \ldots, 0\} \quad \text{and} \quad I^{+,n} = \{1, \ldots, k_n\}.$$

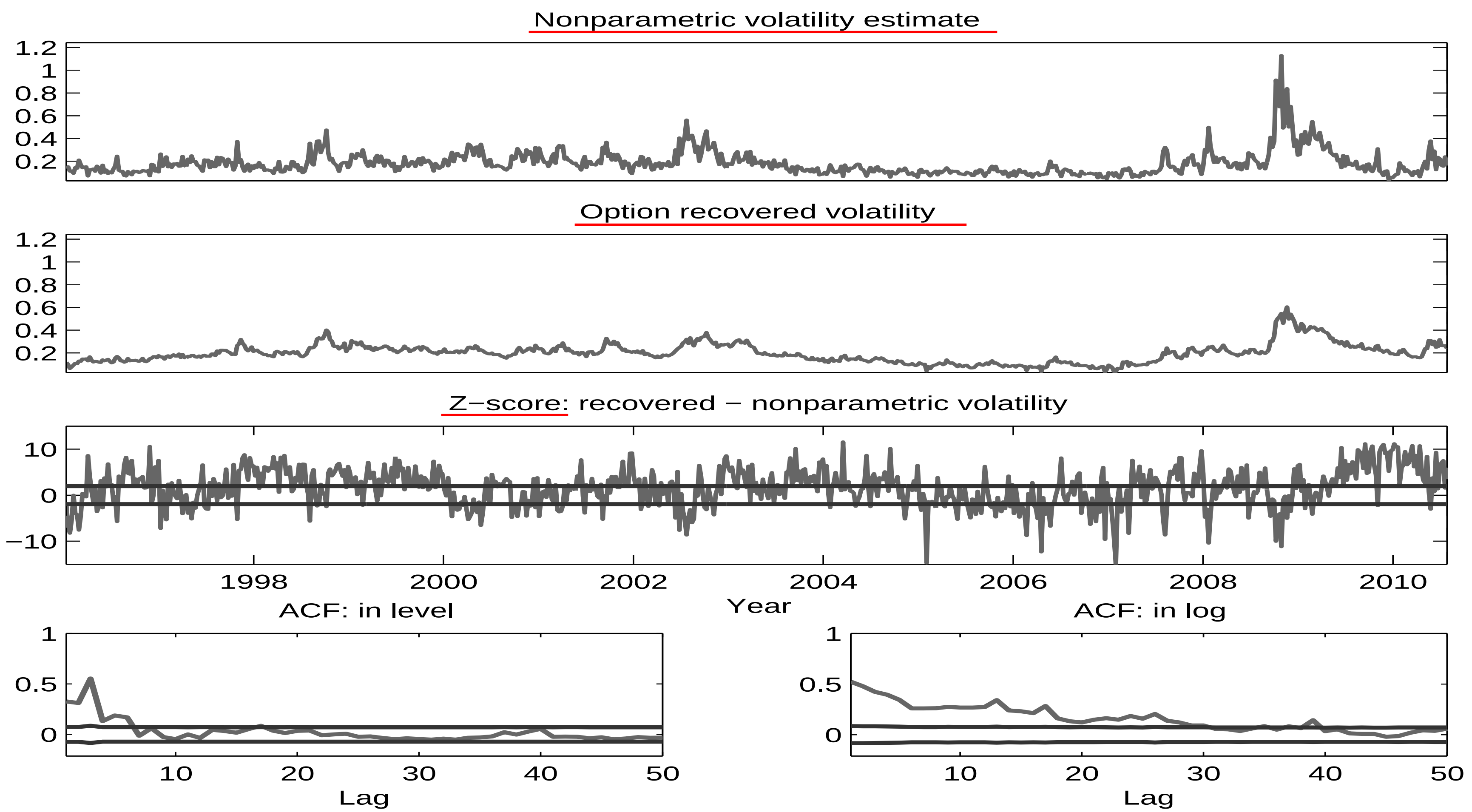
- $V_t^{-,n}$ is Estimator for Spot Variance from Left; $V_t^{+,n}$ Estimator from Right.
- $V_t^{-,n}$ and $V_t^{+,n}$ Differ only if Volatility Jumps at t (Probability Zero Event).

Diagnostic Test III: Model-free vs Option-Implied Volatility

Corollary 2. *Under the same conditions as in Theorem 3, we have for* $k_n \to \infty$, $\min_{t=1,\ldots,T} N_t \to \infty$ *and* $\lambda_n^2 \min_{t=1,\ldots,T} N_t \to 0$,

$$\left\{ \frac{\xi_1(\widehat{\mathbf{S}}_t^n) - \widehat{V}_t^{+,n}}{\sqrt{\frac{\nabla_{\mathbf{S}}\xi_1(\widehat{\mathbf{S}}_t^n)' \widehat{\chi}_t \nabla_{\mathbf{S}}\xi_1(\widehat{\mathbf{S}}_t^n)}{N_t} + \frac{2(\widehat{V}_t^{+,n})^2}{k_n}}} \right\}_{t=1,\ldots,T} \stackrel{\mathcal{L}-s}{\longrightarrow} \begin{pmatrix} \breve{E}_1 \\ \vdots \\ \breve{E}_T \end{pmatrix},$$

where $\widehat{\chi}_t$ *is the part of* $\widehat{\mathbf{H}}_T^{-1}\widehat{\mathbf{\Omega}}_T(\widehat{\mathbf{H}}_T^{-1})'$ *corresponding to the variance-covariance of* $\widehat{\mathbf{S}}_t^n$ *and* $(\breve{E}_1, \ldots, \breve{E}_T)'$ *is a vector of standard normals independent of each other and of* $\mathcal{F}$.


Nonparametric volatility estimate
1.2
1
0.8
0.6
0.4
0.2
Option recovered volatility
1.2
1
0.8
0.6
0.4
0.2
Z−score: recovered − nonparametric volatility
10
0
−10
1998
2000
2002
2004
2006
2008
2010
Year
ACF: in level
1
0.5
0
10
20
30
40
50
Lag
ACF: in log
1
0.5
0
10
20
30
40
50
Lag

Empirical Application – Two-Factor Model

We now extend the Model to Include Two SV Factors:

$$\frac{dX_t}{X_{t-}} = (r - \delta)dt + \sqrt{V_{1,t}}dW_{1,t} + \sqrt{V_{2,t}}dW_{2,t} + dL_{x,t},$$
$$dV_{1,t} = \kappa_1(\overline{v}_1 - V_{1,t})dt + \sigma_{1,d}\sqrt{V_{1,t}}dB_{1,t} + dL_{v,t},$$
$$dV_{2,t} = \kappa_2(\overline{v}_2 - V_{2,t})dt + \sigma_{2,d}\sqrt{V_{2,t}}dB_{2,t},$$

Now, Jump Intensity is $\lambda_0 + \lambda_1 V_{1,t}$: Jumps Self-Exciting, as Jumps Impact Volatility.

Note:

- Extension allows “Fears” to be Time-Varying.
- Breaks Tight Link between Pricing of Risk and its Level.

Empirical Application

Table 4: Parameter Estimates of Two-Factor Model

Parameter	Estimate	Standard Error	Parameter	Estimate	Standard Error
$\rho_{1,d}$	-0.9800	0.0253	λ_0	0.0217	0.0027
$\overline{v}_1$	0.0331	0.0019	λ_1	6.0683	0.8749
κ_1	1.2327	0.0638	μ_x	-0.0145	0.0136
$\sigma_{1,d}$	0.2640	0.0113	σ_x	0.0877	0.0082
$\rho_{2,d}$	-0.1824	0.0388	μ_v	0.1501	0.0124
$\overline{v}_2$	0.0066	0.0001	ρ_j	-0.7756	0.0718
κ_2	29.8797	0.5951			
$\sigma_{2,d}$	0.2341	0.0569			

Risk Neutral Mean of Volatility: 22.8% vs Sample RV Estimate: 21.35%

Mean Risk-Neutral (Log-Return) Jump: -13% Annual Jump Probability: 22.25%

Mean Risk-Neutral Volatility (Level) Jump: $\sqrt{0.1501} = 38.7\%$

SPX Options – Parameter Stability Test

Table 3: Parameter Stability; S&P 500 Options Data

Parameter	Nominal size of test		Parameter	Nominal size of test	
	1%	5%		1%	5%
		Panel A: One-Factor Model			
ρ_d	62.86%	70.48%	λ_j	56.19%	67.62%
$\overline{v}$	71.43%	73.33%	μ_x	20.00%	25.71%
κ	91.43%	93.33%	σ_x	49.52%	61.91%
σ_d	77.14%	80.95%	μ_v	31.43%	36.19%
			ρ_j	13.33%	17.14%
		Panel B: Two-Factor Model			
$\rho_{d,1}$	8.57%	16.19%	$\lambda_{j,0}$	37.14%	49.52%
$\overline{v}_1$	7.62%	12.38%	$\lambda_{j,1}$	21.91%	33.33%
$\kappa_{d,1}$	72.38%	70.05%	μ_x	0.00%	4.76%
$\sigma_{d,1}$	42.86%	53.33%	σ_x	9.52%	17.14%
$\rho_{d,2}$	7.62%	16.19%	μ_v	23.81%	34.29%
$\overline{v}_2$	76.19%	80.95%	ρ_j	1.91%	2.86%
$\kappa_{d,2}$	69.52%	77.14%			
$\sigma_{d,2}$	0.95%	0.95%			

Diagnostic Tests I and III

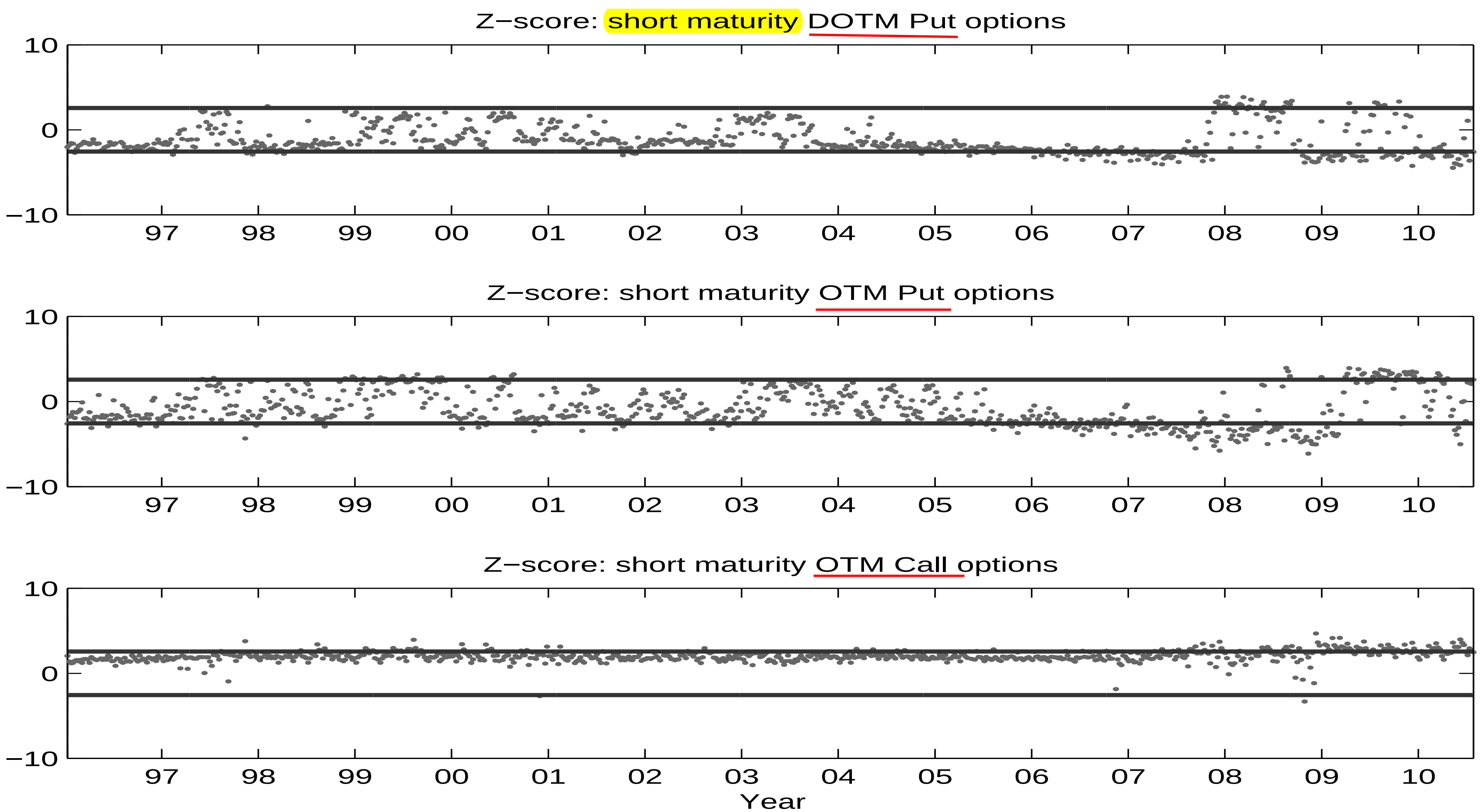
Table 5: Tests on S&P 500 options data

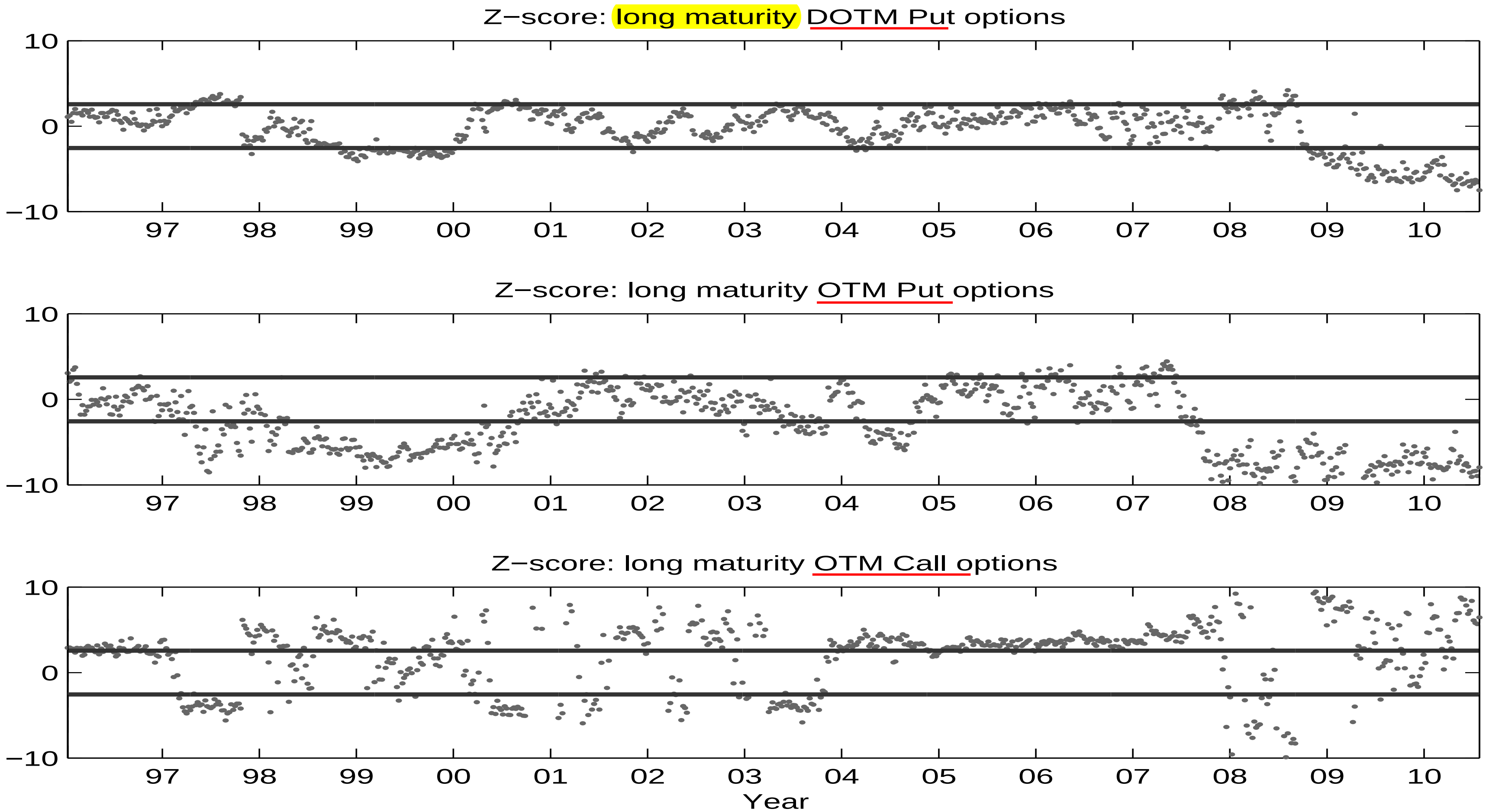
Test	One-factor Model Nominal size of test		Two-factor Model Nominal size of test	
	1%	5%	1%	5%
Panel A: Option Fit Tests				
DOTM, short-maturity puts	18.16%	39.74%	21.84%	45.53%
OTM, short-maturity puts	24.87%	53.03%	27.89%	50.39%
OTM, short-maturity calls	20.53%	55.00%	16.58%	48.55%
DOTM, long-maturity puts	41.45%	62.11%	27.50%	41.97%
OTM, long-maturity puts	72.63%	80.79%	52.89%	60.53%
OTM, long-maturity calls	53.16%	65.53%	79.47%	86.45%
Panel B: Distance implied-nonparametric volatility				
	54.08%	65.66%	49.74%	61.58%
Panel C: Root-mean squared error of fit				
	3.09%		2.32%	

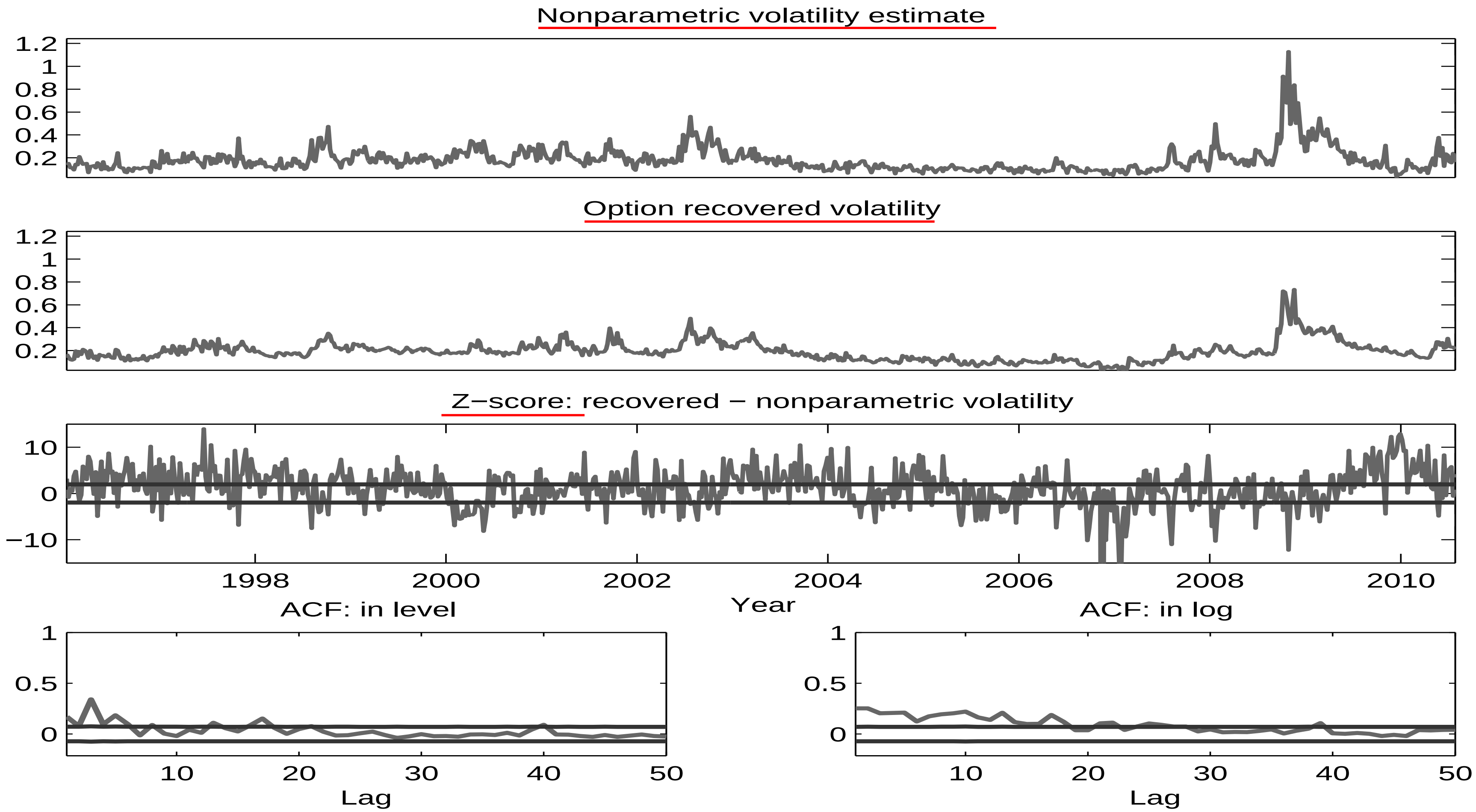
Empirical Application

Things to Note for Two-Factor Model:

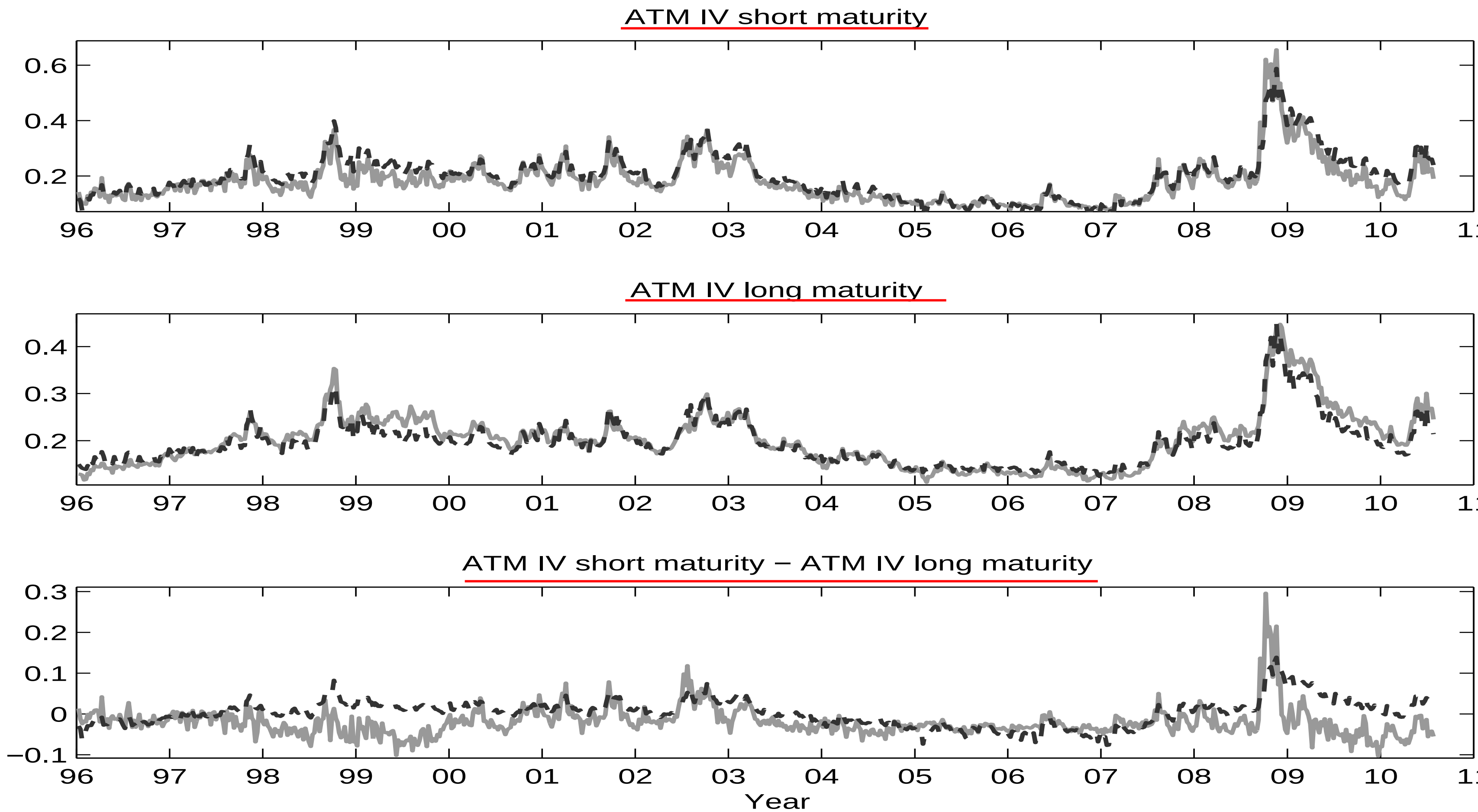
- Fit Improves Significantly (RMSE Drops about 25%).
- Constant Part of Jump Intensity Small $\Longrightarrow$ Jump Risk Premia Time-Varying.
- First Return-Volatility Correlation (remains) Extremely Negative.
- Second Volatility Factor Smaller, much Less Persistent.
- Model still Struggles with Short-Maturity OTM Calls, Long-Term OTM Options.
- Parameters Vary over time, particularly the ones driving Jump Distribution.
- In Quiet Period Jump Intensity Near Zero – Jump Parameters not Identified.
- Time-Varying Parameters $\Longrightarrow$ Missing State Variables?
- Period 2006-2010 very Hard to Fit Reasonably.
- Model still Misspecified, even on Stretches of One Year.


Z−score: short maturity DOTM Put options
10
0
−10
97 98 99 00 01 02 03 04 05 06 07 08 09 10
Z−score: short maturity OTM Put options
10
0
−10
97 98 99 00 01 02 03 04 05 06 07 08 09 10
Z−score: short maturity OTM Call options
10
0
−10
97 98 99 00 01 02 03 04 05 06 07 08 09 10
Year

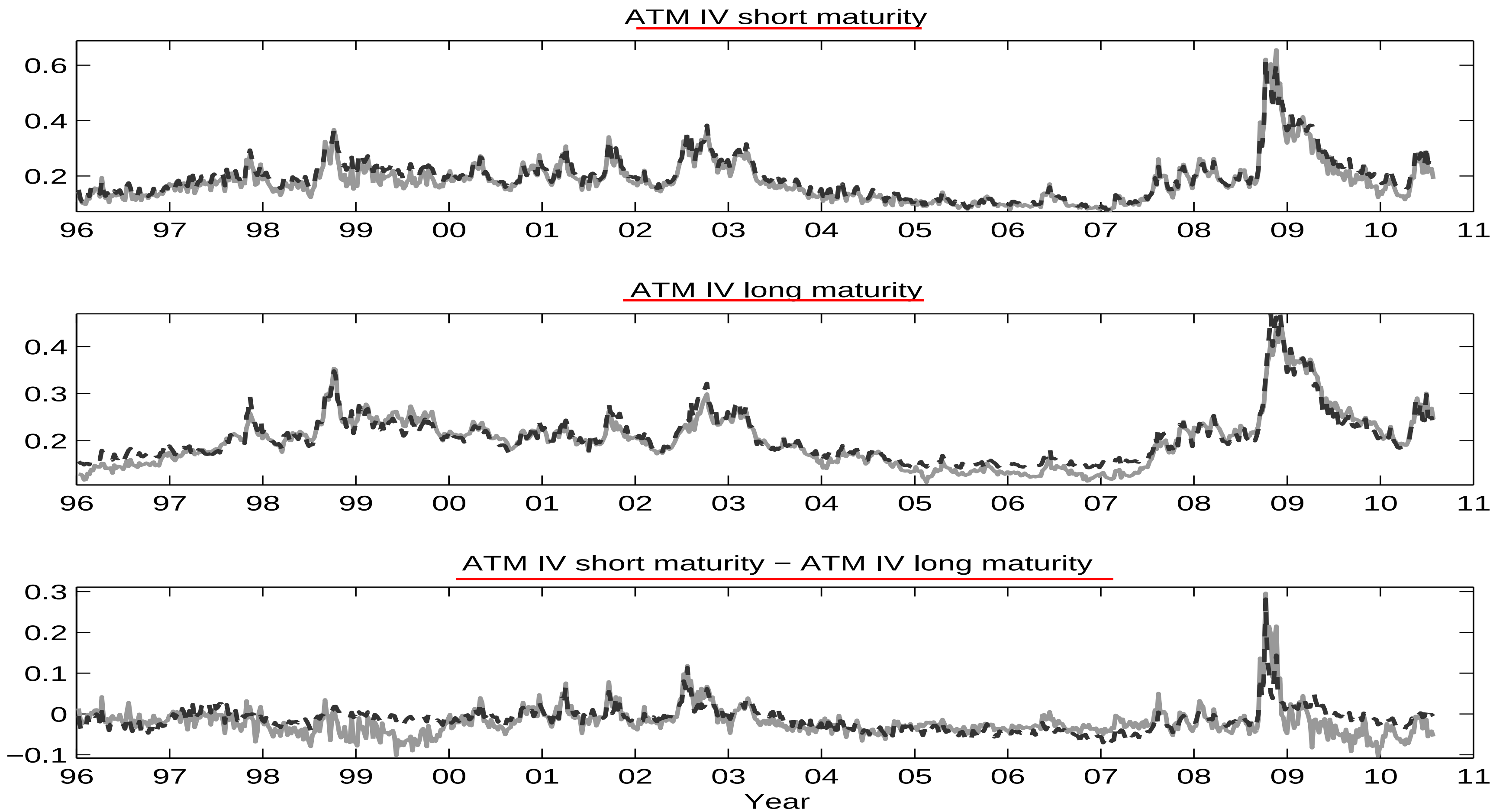

Z−score: long maturity DOTM Put options
10
0
−10
97
98
99
00
01
02
03
04
05
06
07
08
09
10
Z−score: long maturity OTM Put options
10
0
−10
97
98
99
00
01
02
03
04
05
06
07
08
09
10
Z−score: long maturity OTM Call options
10
0
−10
97
98
99
00
01
02
03
04
05
06
07
08
09
10
Year


Nonparametric volatility estimate
1.2
1
0.8
0.6
0.4
0.2
Option recovered volatility
1.2
1
0.8
0.6
0.4
0.2
Z−score: recovered − nonparametric volatility
10
0
−10
1998
2000
2002
2004
2006
2008
2010
Year
ACF: in level
1
0.5
0
10
20
30
40
50
Lag
ACF: in log
1
0.5
0
10
20
30
40
50
Lag

One-Factor Model – Fit to ATM Term Structure of IV



Two-Factor Model – Fit to ATM Term Structure of IV



Conclusions

- We Propose and Derive Asymptotic Properties of Estimation in Large Option Panels with Fixed Time Span and Increasing Cross-Section.

- Method requires Risk-Neutral Model only, is Nonparametric about Option Pricing Errors, and Allows for Heteroscedasticity in the latter.

- Battery of Statistics to Detect Sources of Model Misspecification:
  - Testing Model Fit Over Time and different Parts of Volatility Surface,
  - Testing Model Stability,
  - Testing Consistency between Model Option-Implied Volatility and Nonparametric Estimate from High-Frequency Data on underlying Asset.