Empirical Bounds of Log-Returns Characteristics

Yoshihiro Shirai

Applied Mathematics Dept.
University of Washington

June 27, 2023
Introduction
Folly and Fantasy in Finance

- Valuations should be based on log-returns dynamics explaining both
  1. **historical prices**, reflecting **risks** deemed acceptable by operators/regulators;
  2. **option prices**, reflecting **market expectations**.

- It is possible to consistently model empirical/option-implied **finite dimensional distribution** of asset prices (Madan, [2022]):
  - Bid and ask defined by a set of equivalent laws distorting a measure $\mathcal{C}$, chosen by the market, reflecting options’ mid prices;
  - Historical (mid) price dynamics are specified by a measure $\mathbb{P}$;

- Inconsistencies however arise over path sets of **probability zero**, i.e. $\mathbb{P}$ and $\mathcal{C}$ are typically not equivalent;

Possible Solution: introducing Statistical Model Uncertainty

⇒ Dynamics specified by **nondominated** set $\mathcal{P}$ of laws;
⇒ For each such law market chooses an EMM $\mathcal{Q}$;
⇒ Bid and ask are inf and sup over prices generated by each $\mathcal{Q}$;
⇒ If $\mathcal{P}$ is singleton, we go back to a classical set up;
Volatility Uncertainty

- $\Omega = \mathcal{D}(\mathbb{R}_+, \mathbb{R})$ denotes the Skorohod space of real valued cadlag paths $\omega = \{\omega_t\}_{t \in \mathbb{R}_+}$ on $\mathbb{R}_+$ satisfying $\omega_0 = 0$;
- $\mathcal{F}$ is the Borel $\sigma$-algebra generated by the Skorohod topology on $\Omega$;
- $X$ is the canonical process on $(\Omega, \mathcal{F})$ defined by $X_t(\omega) = \omega_t$;
- $\mathcal{P}_{[\sigma, \bar{\sigma}]}$ is the set of laws on $(\Omega, \mathcal{F})$ under which $X$ is a semimartingale with differential characteristics $(\mu, \sigma, 0)$ where the process $\sigma$ evolves in $[\sigma, \bar{\sigma}]$.

"Escalator up and elevator down"

Hard to capture local asymmetries in log-returns due e.g. to panic/immediate selloff.
Speed Uncertainty

- \((\Omega, \mathcal{F})\) and \(X\) as before;
- \(\kappa_k(x)\) is defined, for \(k = (b_p, c_p, b_n, c_n)\), by
  \[
  \kappa_k(x) = c_p \frac{e^{-x/b_p}}{x} \mathbb{1}\{x>0\} + c_n \frac{e^{-|x|/b_n}}{|x|} \mathbb{1}\{x<0\}
  \]

Given \(K \subset \mathbb{R}^4_+\), \(\Theta = \{(\mu, 0, \kappa_k dx)\}_{k \in K}\) is the set of Levy triplets corresponding to the Bilateral Gamma processes with parameters \((b_p, c_p, b_n, c_n) \in K\);
- \(\mathfrak{P}_\Theta\) is the corresponding set of BG laws on \((\Omega, \mathcal{F})\).

Note:

i. BG law capable to fit options mid prices;
ii. Two BG laws are equivalent iff their speed is the same;
\(\Rightarrow\) no need to include local BG laws in \(\mathfrak{P}_\Theta\);
\(\Rightarrow\) uncertainty in statistical parameters \((c_p, c_n)\).
Goals

To construct $\Psi_\Theta$, we
- estimate $\tilde{K}$ from historical prices;
- estimate $\hat{K}$ from risk neutral prices;
- combine them to form $K$

Question:
How well can we match bid-ask spreads?

Potential Applications:
Good and fast quotes for reversals and combos.
BG Process for Log Returns
Observation: prices exhibit exponential growth

**Black-Scholes**: log-returns are Gaussian (maximal entropy law on \( \mathbb{R} \))

**Issues**:

i. **Risk Aversion**
   - Days with intense selloff alternate with lower activity ones
   - Daily *realized variance/quadratic variation* is not constant
   - OTM puts priced higher than OTM calls (*volatility smile*)

ii. Prices exhibit leptokurtic features and often *jump*
   - need to look at *higher moments* than just variance

**Possible Solution**

Randomize time to track periods with higher/slower activity
From Black-Scholes to Bilateral Gamma

- **VG**: Quadratic variation is **gamma process** (maximum entropy law on $\mathbb{R}_+$)

$$\Rightarrow S(t) = S(0)e^{\mu g(t) + B(g(t))}$$

- **Characteristic exponent**: by conditioning on the random time $g(t)$,

$$IE[e^{i\theta B(g(t))}] = IE[e^{-\theta^2\sigma^2/2 g(t)}] = \left(1 + \frac{\sigma^2 v \theta^2}{2}\right)^{-\frac{t}{v}}$$

$$= (1 - ia\theta)^{-\frac{t}{v}} (1 + ia\theta)^{-\frac{t}{v}}$$

where $v = \nabla[g(1)]$, $a^2 = \frac{\sigma^2 v}{2}$ and we assumed wlog $IE[g(1)] = 1$.

**Excess Return**

$$\Rightarrow$$ is difference of two gamma processes

$$\Rightarrow$$ has **finite variation**
From Black-Scholes to Bilateral Gamma

- VG is sum of processes of gains and losses with same mean and variance
- **Issue**: downward jumps > upward jumps (escalator up and elevator down)
- BG process defined as difference of independent gamma processes, i.e.

\[
\mathbb{E}[e^{i\theta X(t)}] = (1 - i\theta b_p)^{-tc_p} (1 + i\theta b_n)^{-tc_n}
\]

with Levy density \( \kappa(x) = \left( \frac{c_p}{x} e^{-b_p x} 1_{(0,\infty)}(x) + \frac{c_n}{|x|} e^{-b_n |x|} 1_{(-\infty,0)}(x) \right) \)

\( \Rightarrow \) BG is a finite variation Levy process
\( \Rightarrow \) BG is self-decomposable \( \Rightarrow \) sum of “independent news”

- Moments
  - Gains: \( \mu_p = c_p b_p, \sigma_p = \sqrt{c_p b_p} \)
  - Losses: \( \mu_n = c_n b_n, \sigma_n = \sqrt{c_n b_n} \)
Estimating $\tilde{K}$
Some Intuition

- Estimating $\tilde{K}$ is equivalent to identifying relationships between the parameters $(b_p, c_p, b_n, c_n)$ of a bilateral gamma density.
- Equivalently, we can estimate relationships between $(\mu_p, \sigma_p, \mu_n, \sigma_n)$;
- In a symmetric, Black-Scholes, setting, it is reasonable to expect that a positive relationship exists between reward, $\mu$, and risk, $\sigma$, and several studies confirm such relationships;

**Question:**

⇒ Can we identify $(\sigma_p, \mu_n, \sigma_n)$ as risks and $\mu_p$ as their compensation?
⇒ If so, one would expect to see bounds for $\mu_p$ given $(\sigma_p, \mu_n, \sigma_n)$, to be increasing in each of the risks.
Risks and Compensation

**Theorem 1.**

Let $X^+, X^-, Y^+, Y^-$ have gamma distribution. If

$$\mathbb{E}[X^+] \geq \mathbb{E}[Y^+], \quad \mathbb{E}[X^-] \leq \mathbb{E}[Y^-], \quad V[X^+] \leq V[Y^+], \quad V[X^-] \leq V[Y^-],$$

then $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for every concave function $u$.

**Theorem 2.**

A strictly increasing and concave function $v \in C^2((0, \infty))$ has local CRRA coefficient $\epsilon$ bounded below by 1 if and only if there is a strictly increasing and concave function $u \in C^2(\mathbb{R})$ such that $v(x) = u(\log(x))$ for every $x \in (0, \infty)$.

**Kelly’s Criterion**

LT investors maximize log utility. ST ones are more risk averse $\Rightarrow \epsilon \geq 1$

$\Rightarrow$ For BG log-returns, 3D risks vector $(\sigma_p, \mu_n, \sigma_n)$ compensated by $\mu_p$
Estimation of Boundaries of the set $\tilde{K}$

- **Dataset:** $(\mu_p, \sigma_p, \mu_n, \sigma_n)$ daily estimated for 184 stocks for the period between 1/1/2008 to 31/12/2020.

- Assume that, for given risks $(\sigma_p, \mu_n, \sigma_n)$, acceptable compensation ranges between $f_m(\sigma_p, \mu_n, \sigma_n)$ and $f_M(\sigma_p, \mu_n, \sigma_n)$.

- $f_m$ and $f_M$ estimated via **quantile regression**, i.e. we solve

  $$\min_{f \in \mathcal{F}} (1-\tau) \sum_i [\mu_p(i) - f_M(\sigma_p(i), \mu_n(i), \sigma_n(i))]^+$$

  $$- \tau \sum_i [\mu_p(i) - f_M(\sigma_p(i), \mu_n(i), \sigma_n(i))]^-,$$

- We set $\tau = 0.05$ for $f_m$ and $\tau = 0.95$ for $f_M$.

- For $\mathcal{F}$ we considered the class of linear and Gaussian process regressors.
Results

- **Linear Regression:**
  \[
  f_m(\sigma_p, \mu_n, \sigma_n) = 0.0017 + 0.2029\sigma_p + 0.9951\mu_n - 0.3711\sigma_n,
  \]
  \[
  f_M(\sigma_p, \mu_n, \sigma_n) = 0.0017 + 0.2710\sigma_p + 1.0102\mu_n - 0.2311\sigma_n.
  \]

- **Gaussian process regression:**
  - \(\partial f_m / \partial \sigma_n\) always negative;
  - \(\partial f_M / \partial \sigma_n\) negative at all but two of 16 representative points

**Observation**

⇒ Risk seeking behavior in the pure loss prospect
### Estimation of Boundaries of the set $\tilde{K}$

<table>
<thead>
<tr>
<th>$\frac{\partial f_M}{\partial \sigma_p}$</th>
<th>$\frac{\partial f_M}{\partial \mu_n}$</th>
<th>$\frac{\partial f_M}{\partial \sigma_n}$</th>
<th>$\frac{\partial f_m}{\partial \sigma_p}$</th>
<th>$\frac{\partial f_m}{\partial \mu_n}$</th>
<th>$\frac{\partial f_m}{\partial \sigma_n}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2667</td>
<td>2.4704</td>
<td>0.7577</td>
<td>-0.0130</td>
<td>2.0042</td>
<td>-0.2421</td>
</tr>
<tr>
<td>0.8691</td>
<td>1.9402</td>
<td>-1.3539</td>
<td>1.1140</td>
<td>1.8974</td>
<td>-0.8361</td>
</tr>
<tr>
<td>1.5243</td>
<td>1.9553</td>
<td>-1.1134</td>
<td>1.4108</td>
<td>1.9274</td>
<td>-1.2346</td>
</tr>
<tr>
<td>1.0459</td>
<td>2.0254</td>
<td>-0.4887</td>
<td>0.5666</td>
<td>1.9927</td>
<td>-1.2635</td>
</tr>
<tr>
<td>1.0867</td>
<td>1.9956</td>
<td>-1.0836</td>
<td>0.8823</td>
<td>2.0199</td>
<td>-1.2220</td>
</tr>
<tr>
<td>0.4639</td>
<td>2.0065</td>
<td>-1.4194</td>
<td>0.6053</td>
<td>2.0648</td>
<td>-1.1568</td>
</tr>
<tr>
<td>1.3013</td>
<td>2.0509</td>
<td>-1.4681</td>
<td>1.2715</td>
<td>2.0128</td>
<td>-1.4149</td>
</tr>
<tr>
<td>0.9669</td>
<td>2.0019</td>
<td>-0.2462</td>
<td>0.4477</td>
<td>1.9760</td>
<td>-1.0806</td>
</tr>
<tr>
<td>1.4434</td>
<td>2.2522</td>
<td>0.3978</td>
<td>0.5052</td>
<td>2.0026</td>
<td>-0.5761</td>
</tr>
<tr>
<td>0.9710</td>
<td>1.9465</td>
<td>-0.9840</td>
<td>0.9472</td>
<td>1.8900</td>
<td>-0.8995</td>
</tr>
<tr>
<td>1.0653</td>
<td>1.9423</td>
<td>-1.4702</td>
<td>1.3075</td>
<td>1.9230</td>
<td>-0.9990</td>
</tr>
<tr>
<td>0.9307</td>
<td>1.9594</td>
<td>-0.5941</td>
<td>0.6044</td>
<td>1.9087</td>
<td>-1.0416</td>
</tr>
<tr>
<td>1.3390</td>
<td>2.0444</td>
<td>-1.8664</td>
<td>1.4529</td>
<td>2.0287</td>
<td>-1.5394</td>
</tr>
<tr>
<td>0.8652</td>
<td>1.9872</td>
<td>-1.3001</td>
<td>0.9281</td>
<td>2.0499</td>
<td>-1.0898</td>
</tr>
<tr>
<td>1.1957</td>
<td>2.0398</td>
<td>-0.9586</td>
<td>0.9027</td>
<td>1.9967</td>
<td>-1.3931</td>
</tr>
<tr>
<td>0.9283</td>
<td>1.9956</td>
<td>-0.0906</td>
<td>0.3913</td>
<td>1.9830</td>
<td>-0.9539</td>
</tr>
</tbody>
</table>
## Implied Boundaries of Measure Performance

<table>
<thead>
<tr>
<th>Upper Boundary</th>
<th>Observation</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Observation</th>
<th>Lower Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0856</td>
<td>0.0694</td>
<td>0.0697</td>
<td>0.0536</td>
<td>0.0467</td>
<td>0.0469</td>
</tr>
<tr>
<td>0.0224</td>
<td>0.0208</td>
<td>0.0190</td>
<td>0.0184</td>
<td>0.0165</td>
<td>0.0143</td>
</tr>
<tr>
<td>0.0348</td>
<td>0.0343</td>
<td>0.0339</td>
<td>0.0269</td>
<td>0.0260</td>
<td>0.0252</td>
</tr>
<tr>
<td>0.0180</td>
<td>0.0167</td>
<td>0.0153</td>
<td>0.0147</td>
<td>0.0130</td>
<td>0.0112</td>
</tr>
<tr>
<td>0.0706</td>
<td>0.0685</td>
<td>0.0661</td>
<td>0.0473</td>
<td>0.0453</td>
<td>0.0428</td>
</tr>
<tr>
<td>0.1440</td>
<td>0.1428</td>
<td>0.1421</td>
<td>0.1024</td>
<td>0.1002</td>
<td>0.0986</td>
</tr>
<tr>
<td>0.0329</td>
<td>0.0308</td>
<td>0.0284</td>
<td>0.0243</td>
<td>0.0225</td>
<td>0.0204</td>
</tr>
<tr>
<td>0.0127</td>
<td>0.0119</td>
<td>0.0107</td>
<td>0.0092</td>
<td>0.0088</td>
<td>0.0081</td>
</tr>
</tbody>
</table>

**Table:** $\mu_p$ boundaries (estimated via Quantile GPR), at 16 representative points.
Estimation of Boundaries of the set $\tilde{K}$

<table>
<thead>
<tr>
<th>Upper Boundary</th>
<th>Observation</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Observation</th>
<th>Lower Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0960</td>
<td>-0.1253</td>
<td>-0.0500</td>
<td>1.7966</td>
<td>-0.2803</td>
<td>-0.2231</td>
</tr>
<tr>
<td>1.4038</td>
<td>0.3108</td>
<td>-0.8727</td>
<td>2.2708</td>
<td>0.6168</td>
<td>-1.1672</td>
</tr>
<tr>
<td>-0.0615</td>
<td>-0.2441</td>
<td>-0.4023</td>
<td>0.4234</td>
<td>-0.0286</td>
<td>-0.4702</td>
</tr>
<tr>
<td>3.2179</td>
<td>1.6361</td>
<td>-0.1370</td>
<td>2.7574</td>
<td>0.9385</td>
<td>-0.9590</td>
</tr>
<tr>
<td>2.7685</td>
<td>0.8818</td>
<td>-1.3389</td>
<td>2.1867</td>
<td>0.7031</td>
<td>-1.2120</td>
</tr>
<tr>
<td>1.2525</td>
<td>0.5067</td>
<td>0.0483</td>
<td>2.3867</td>
<td>0.6779</td>
<td>-0.5442</td>
</tr>
<tr>
<td>2.4903</td>
<td>0.7203</td>
<td>-1.2574</td>
<td>2.9818</td>
<td>1.1174</td>
<td>-0.9577</td>
</tr>
<tr>
<td>3.0642</td>
<td>1.8490</td>
<td>0.2862</td>
<td>2.7892</td>
<td>2.1576</td>
<td>0.9789</td>
</tr>
</tbody>
</table>

Table: Sharpe ratio boundaries (estimated via Quantile GPR), at 16 representative points.
### Estimation of Boundaries of the set $\tilde{K}$

<table>
<thead>
<tr>
<th>Upper Boundary</th>
<th>Observation</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Observation</th>
<th>Lower Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.151</td>
<td>-0.003</td>
<td>-0.003</td>
<td>0.065</td>
<td>-0.009</td>
<td>-0.009</td>
</tr>
<tr>
<td>0.051</td>
<td>0.011</td>
<td>-0.031</td>
<td>0.082</td>
<td>0.022</td>
<td>-0.041</td>
</tr>
<tr>
<td>-0.015</td>
<td>-0.008</td>
<td>-0.002</td>
<td>0.015</td>
<td>-0.001</td>
<td>-0.017</td>
</tr>
<tr>
<td>0.117</td>
<td>0.059</td>
<td>-0.006</td>
<td>0.100</td>
<td>0.034</td>
<td>-0.034</td>
</tr>
<tr>
<td>0.101</td>
<td>0.032</td>
<td>-0.049</td>
<td>0.079</td>
<td>0.026</td>
<td>-0.044</td>
</tr>
<tr>
<td>0.045</td>
<td>0.019</td>
<td>0.001</td>
<td>0.087</td>
<td>0.025</td>
<td>-0.020</td>
</tr>
<tr>
<td>0.090</td>
<td>0.026</td>
<td>0.046</td>
<td>0.109</td>
<td>0.040</td>
<td>-0.035</td>
</tr>
<tr>
<td>0.112</td>
<td>0.067</td>
<td>0.009</td>
<td>0.102</td>
<td>0.078</td>
<td>0.034</td>
</tr>
</tbody>
</table>

**Table:** Acceptability index boundaries (estimated via Quantile GPR), at 16 representative points. Negative signs represent acceptability indices of short positions.

⇒ Upper and lower performances consistent with empirical observations.
Uncertainty Quantification

The uncertainty around $\mu_p$ given $(\sigma_p, \mu_n, \sigma_n)$ can be quantified by a dimensional analysis of the manifold $\tilde{K}$:

<table>
<thead>
<tr>
<th></th>
<th>PCA cumulative weight (in %)</th>
<th>Diffusion Map cumulative weight (in %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>2.7529</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>68.82</td>
<td>70.27</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>1.1778</td>
<td>0.0045</td>
</tr>
<tr>
<td></td>
<td>98.27</td>
<td>98.58</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.0685</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>99.98</td>
<td>99.64</td>
</tr>
<tr>
<td>$\lambda_4$</td>
<td>0.0009</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table: Eigenvalues’s weights for PCA and diffusion map on the quantized dataset.

Observations

⇒ Upper and lower boundaries for $\mu_p$ are relatively close;
⇒ Boundaries are well approximated by linear functions
A Modified Lucas Tree Economy

Is prospects theory consistent with the risk seeking behaviors in losses?

Consider the following variation of a Lucas tree economy:

- Two periods $i = 0, 1$
- Each agent endowed with a single risky asset (a tree) with payoff $S_i$, $i = 0, 1$
- Assume: $S_0$ known, $S_1 = S_0 e^{G - L}$, with $G$ and $L$ independent gamma variates
- There is a risk free asset in zero net supply with risk free rate $r_f$
- Consumption is determined by borrowing/lending $\ell$ at time zero:

$$C_0 = S_0 + \ell, \quad C_1 = S_0 e^{G - L} - \ell e^{r_f}.$$

- Preferences: let $X = s_0 + G - L$, $0 < \beta, \rho < 1$, and (logarithms in lower case)

$$U(C_0, C_1) = u(c_0) + e^{-\beta} \mathbb{E} \left[ u(c_1) \mathbb{1}_{\{X \geq 0\}} - u(-c_1) \mathbb{1}_{\{X \leq 0\}} \right].$$

Equilibrium condition $\ell = 0$ gives

$$r^e_f = \beta - \rho \log(s_0) - \log \left( \mathbb{E}[(X)^{-\rho} e^{-X} \mathbb{1}_{\{X \geq 0\}}] - \mathbb{E}[(-X)^{-\rho} e^{-X} \mathbb{1}_{\{X \leq 0\}}] \right).$$
A Modified Lucas Three Economy

Figure: Equilibrium rate as function of $\sigma_p, \mu_p, \sigma_n, \mu_n$.
Estimating $\hat{K}$
Dataset and Methods

- **Dataset:** \((b_p, c_p, b_n, c_n)\) calibrated every 10 days for 10 sector ETFs for the period between 1/1/2015 to 31/12/2020 for each of the four middle maturities traded \(\Rightarrow 4812\) observations;

- **Bounds for** \(c_p\) **are estimated utilizing:**
  
  - **quantile regression:** Quantile loss function replaced with
    \[
    S(x) = \tau x + \alpha \log(1 - e^{-x/\alpha}), \quad \alpha = 10^{-4}
    \]
  
  - **distorted least squares:** Objective function:
    \[
    \min_{f \in \mathcal{F}} \sum_i r_i^2 \left( \Psi(q_i) - \Psi(q_i - \frac{1}{n}) \right),
    \]
    where \(r_i\) is residual, \(\Psi\) is MINMAXVAR distortion, \(q_i\) is the \(i\)-th empirical quantile of the residual’s empirical distribution
Visualization of Speed Bounds

Figure: Visualization of quantile (upper pictures) and distorted (lower pictures) GPR boundaries around a randomly selected point (in red).
### Table: Eigenvalues's weights for PCA and diffusion map on the risk neutral dataset.

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>Diffusion Map</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cumulative weight (in %)</td>
<td>cumulative weight (in %)</td>
<td></td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>1.2234</td>
<td>30.58</td>
<td>0.9999</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.9949</td>
<td>55.46</td>
<td>0.9155</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0.9574</td>
<td>79.39</td>
<td>0.0575</td>
</tr>
<tr>
<td>$\lambda_4$</td>
<td>0.8244</td>
<td>100.0</td>
<td>0.0092</td>
</tr>
</tbody>
</table>

- Embedding and boundaries are nonlinear
- Variance of speed well explained by that of scale
## Options Implied Bid-Ask Prices to Unwind a 1$ Position

<table>
<thead>
<tr>
<th>Upper Valuation</th>
<th>Lower Valuation</th>
<th>% of Points Represented</th>
<th>Upper Valuation</th>
<th>Lower Valuation</th>
<th>% of Points Represented</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9806</td>
<td>1.0364</td>
<td>0.1641</td>
<td>0.9113</td>
<td>1.1205</td>
<td>0.0347</td>
</tr>
<tr>
<td>0.9712</td>
<td>1.0210</td>
<td>0.1610</td>
<td>0.9319</td>
<td>1.0358</td>
<td>0.0339</td>
</tr>
<tr>
<td>0.9630</td>
<td>1.0359</td>
<td>0.1240</td>
<td>0.9250</td>
<td>1.1468</td>
<td>0.0265</td>
</tr>
<tr>
<td>0.9646</td>
<td>1.0306</td>
<td>0.0943</td>
<td>0.9759</td>
<td>1.2024</td>
<td>0.0253</td>
</tr>
<tr>
<td>0.9538</td>
<td>1.0321</td>
<td>0.0920</td>
<td>0.9497</td>
<td>1.0631</td>
<td>0.0214</td>
</tr>
<tr>
<td>0.9586</td>
<td>1.0586</td>
<td>0.0799</td>
<td>0.9089</td>
<td>1.1389</td>
<td>0.0164</td>
</tr>
<tr>
<td>0.9638</td>
<td>1.0666</td>
<td>0.0608</td>
<td>0.8442</td>
<td>1.1997</td>
<td>0.0109</td>
</tr>
<tr>
<td>0.9286</td>
<td>1.0911</td>
<td>0.0452</td>
<td>0.8946</td>
<td>1.2626</td>
<td>0.0094</td>
</tr>
</tbody>
</table>

### Work in Progress

Comparison with Model-Free Options Implied Prices
The set $K$
Constructing the Pricing Measures

- From options mid prices, estimate:
  - risk neutral $(\hat{b}_p, \hat{c}_p, \hat{b}_n, \hat{c}_n)$;
  - range $\tilde{C}_p = (\hat{c}_{p,m}, \hat{c}_{p,M})$ for $c_p$ given $(\hat{b}_p, \hat{b}_n, \hat{c}_n)$;

- From equity prices, estimate
  - statistical $(\tilde{b}_p, \tilde{c}_p, \tilde{b}_n, \tilde{c}_n)$;
  - range $\tilde{C}_p = (\hat{c}_{p,m}, \hat{c}_{p,M})$ for $c_p$ given $(\tilde{b}_p, \tilde{b}_n, \tilde{c}_n)$;

- Set $C = \tilde{C}_p \times \tilde{c}_n \cup \hat{C}_p \times \hat{c}_n$;

- For each pair $(c_p, c_n) \in C$, estimate $(b_p, b_n)$ that match best option prices.

- The pricing measures consists of resulting BG laws.
Constructing the Pricing Measures

Example with data on SPY as of October 8 2020.¹

\[(\hat{b}_p, \hat{c}_p, \hat{b}_n, \hat{c}_n) = (0.0175, 24.1090, 0.0262, 42.9922)\]
\[\Rightarrow \hat{C}_p = (22.1135, 27.5439)\]

\[(\tilde{b}_p, \tilde{c}_p, \tilde{b}_n, \tilde{c}_n) = (0.0082, 0.1802, 0.0224, 0.4165)\]
\[\Rightarrow \tilde{C}_p = (0.6950, 1.3105)\]

Set \(C = \hat{C}_p \cup \tilde{C}_p\), and

\[\Rightarrow \text{for each } (c_p, c_n) \in C, \text{ compute } (b_p, b_n) \text{ by matching first and second moment of options implied risk neutral distribution:}\]

\[\varphi(-i; c_p, c_n, b_p, b_n) = \varphi(-i; \hat{c}_p, \hat{c}_n, \hat{b}_p, \hat{b}_n)\]
\[\varphi(-2i; c_p, c_n, b_p, b_n) = \varphi(-2i; \hat{c}_p, \hat{c}_n, \hat{b}_p, \hat{b}_n)\]

\[\Rightarrow \text{set ask price operator } a(\cdot) = \sup_{c_p, c_n \in C} IE^{Q_{\kappa}(c_p, c_n)}[\cdot]\]

¹Upper and lower bounds on \(c_p\) are estimated via quantile GPR and quantile regression.
Valuation of Long Term Options

- Speed bounds control **short term** uncertainty;
- For long maturities, evolution of risk neutral/statistical BG parameters specified by 4D Markov chain \( \{K_{t,j}\}_{j=1,...,N} \):
  - Quantize the dataset of options implied BG parameters into \( S \) representative points \( \{k_s\}_{s=1,...,S} \), each representing a fraction \( p_s \) of points;
  - Define
    \[
    \hat{q}_{s,r} := \frac{1}{\|F_{k_s} - F_{k_r}\|_W}, \quad Q_{s,r} := \arg\min_{Q \in S, pQ=0} \sum_{s \neq r} \|\hat{q}_{s,r} - Q_{s,r}\|^2
    \]
    \[
    (P)^j := e^{Q_{t,j}}
    \]
  - Stationary distribution \( p \) and transition probability to close states maximized;
  - Simulate \( K^t \) and along each path compute bid ask prices backward.
Conclusions
Summary of Results

- Speed uncertainty allows to **consistently** model equity and option prices;
- This seems fundamental as quotes must depend on both **market expectations** and range of **risks** deemed acceptable by operators/regulators;
- Focus on Speed is needed to reflect **biases** in the financial markets introduced by risk averse-seeking behaviors;
- Quantile/Distorted GPR show certain promise in capturing risk neutral and statistical features of equity and option prices, such as:
  - Increasing utility of **variance of losses**;
  - Sharpe ratios and other **performance** measures;
  - **Forward** prices to unwind 1$ valuations.

- Potential Application: pricing **combos and reversals**, for which fast and good quotes need to be provided not to lose market shares.

- Future work include:
  - Additional statistical studies to compare forward prices with **model-free** prices;
  - Development of **statistical methods** on implied volatility surface to generate prices of portfolios of options.
Thank you!