Microstructure under the Microscope

Tools to Survive and Thrive in the Age of (Too Much) Information

Ravi Kashyap

Seminar on AI, Machine Learning and Sentiment Analysis Applied to Finance, Hong Kong

March 2017
Topics for Discussion

- The Nature of Uncertainty
- Objectively Subjective
- The Circle of Investment
- The Miracle of Mathematics
- From Symbols to Numbers
THE NATURE OF UNCERTAINTY

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com

© 2016 Gain Knowledge Group. All Rights Reserved.
Sergei Bubka and the Regulators
The Uncertainty Principle of the Social Sciences

“Any generalization in the social sciences cannot be both popular and continue to yield accurate predictions or in other words, the more popular a particular generalization, the less accurate will be the predictions it yields.”
Simply Too Complex

All of finance, through time, has involved three simple outcomes -

“Buy, Sell or Hold”

The complications are mainly to get to these results.
OBJECTIVELY SUBJECTIVE

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
The problem of creating artificial intelligence can be a child's play, depending on which adult's brainpower acts as our gold standard.

Perhaps, the real challenge is to replicate the curiosity and learning an infant displays.

Intellect might be a byproduct of Inquisitiveness. Another instance of an unintended yet welcome (?) consequence.

If ignorance is bliss, intrusion might just be the opposite and bring misery. As the saying goes, Curiosity terminated the …
Yet to discover an objective measuring stick for comparison, a so called, True Comparison Theory.

Despite all the uncertainty, the one thing we can be almost certain about is the subjectivity in all decision making

Makes people react at varying degrees and at varying speeds

Decisions give rise to actions and subjectivity in the comparison means differing decisions and hence unpredictable actions
Merry-Go-Round …

… of comparisons, decisions and actions …

Actions of participants affects the state of the system

Explains the growing trend towards comprehending better and deciphering the decision process and the subsequent actions, by collecting more information

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
If Alice and Red Queen of the Wonderland fame were to visit Hedge-Fund-Land (or even Business-Land), the following modification of their popular conversation would aptly describe the situation today,

“My dear, here we must process as much data as we can, just to stay in business. And if you wish to make a profit you must process at least twice as much data.”

Applying this to HFT-Land: “… here we must trade as fast as we can …”, while reminiscing that the jury is still out on whether HFT is Good, Bad or Just Ugly and Unimportant.

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
THE CIRCLE OF INVESTMENT

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
Manager versus Machine (Man vs Mac)...

Man, Mac and the PC against Increasing Complexity

On the surface, it would seem that there is a repetitive nature to portfolio management

We currently lack a proper understanding of how, in some instances, our brains (or minds; and right now it seems we don’t know the difference!) make the leap of learning from information to knowledge to wisdom

This brings up the question of Art and Science in the practice of asset management (and everything else in life?); which are more related than we probably realize
Art and Science

Art is Science that we don’t know about; Science is Art restricted to a set of symbols governed by a growing number of rules”

Similarities between art and science, should give us hope.

We (including computers and intelligent machines?) can barely make the jump from the information to the knowledge stage; even with the use of cutting / (bleeding?) edge technology and tools.

Information is Hidden; Knowledge is Exchanged or Bartered; Wisdom is Dispersed.

Surely we are still in the Information Age since a disproportionate amount of our actions are geared towards accumulating unique data-sets for the sole benefits of the accumulators.

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
The more that shoppers shop, the more shops there will be and the more the shops will try to woo the shoppers.

If people are willing to pay (or bid) more than what is asked (or offered), then perhaps, we would not have specialized venues to trade, a small price to pay for, let us just say, peace on Earth.

Forgetting about Utopia - but keeping in mind that conceivably, the Bid-Offer spread can be a barometer to a civilization’s progress, till it becomes irrelevant, indicating that a society has transcended beyond mere material matters of accumulating and allocating wealth.
An exchange, as the word implies, is the process during which people give and take things of similar value.

At a place where this transfer happens, also an exchange, shares or holdings can be liquidated and hence the primary mission of an exchange is to provide liquidity.

We do not differentiate the exchange of OTC (Over The Counter) securities, which are traded wherever, whenever and however one can trade them; but we provide the analogy that if Exchange Trading is similar to collecting tolls on a road; OTC Trading is like highway robbery.
Market Structure, The Big Kahuna!!

- To Automate or Not to Automate
- Auction versus Dealer Markets
- Pick a Size for the Perfect Tick
- The Price of Connections to High (and Faraway) Places
- Speed Thrills but Kills (?) / The HFT Conundrum!

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com

© 2016 Gain Knowledge Group. All Rights Reserved.
Market microstructure is the investigation of the process and protocols that govern the exchange of assets with the objective of reducing frictions that can impede the transfer.

In financial markets, where there is an abundance of recorded information, this translates to the study of the dynamic relationships between observed variables, such as price, volume and spread, and hidden constituents, such as transaction costs and volatility, that hold sway over the efficient functioning of the system.
A Swimsuit Model Wearing a Burkha

Mathematics is like a swimsuit model wearing a Burkha; we need to see beyond the symbols and the surface to appreciate the beauty.

Mathematics is built on one simple operation, addition, making it a fractal with addition as its starting point.

Mathematics has become complex because of the confusion that different notations, assumptions not made explicit and missed steps can create.

Mathematics without the steps is like a treasure hunt without the clues.

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
The Bhattacharyya distance is defined as the negative logarithm of the Bhattacharyya coefficient.

\[
\rho (p_i, p'_i) = \sum_i \sqrt{p_i p'_i} \quad \text{and} \quad D_{BC} (p_i, p'_i) = -\ln [\rho (p_i, p'_i)]
\]

\[
\rho (p_i, p'_i) = \int \sqrt{p_i (x) p'_i (x)} \, dx
\]

\[
\rho (p_i, p'_i) = \cos \theta = \sum_i \sqrt{p_i p'_i}
\]

The original paper on the Bhattacharyya distance (Bhattacharyya 1943) mentions a natural extension to the case of more than two populations. For an \( M \) population system, each with \( k \) random variates, the definition of the coefficient becomes,

\[
\rho (p_1, p_2, \ldots, p_M) = \int \cdots \int [p_1 (x) p_2 (x) \ldots p_M (x)]^{\frac{1}{k}} \, dx_1 \cdots \, dx_k
\]
Nothing Lost in (Dimension) Reduction

JL Lemma

Lemma 1. For any $0 < \epsilon < 1$ and any integer $n$, let $k$ be a positive integer such that

$$k \geq 4 \left( \frac{\epsilon^2}{2} - \frac{\epsilon^3}{3} \right)^{-1} \ln n$$

Then for any set $V$ of $n$ points in $\mathbb{R}^d$, there is a map $f : \mathbb{R}^d \rightarrow \mathbb{R}^k$ such that for all $u, v \in V$,

$$(1 - \epsilon) \|u - v\|^2 \leq \|f(u) - f(v)\|^2 \leq (1 + \epsilon) \|u - v\|^2$$

Furthermore, this map can be found in randomized polynomial time and one such map is $f = \frac{1}{\sqrt{k}} Ax$ where, $x \in \mathbb{R}^d$ and $A$ is a $k \times d$ matrix in which each entry is sampled i.i.d from a Gaussian $N(0, 1)$ distribution.
Normally (not so) Normal

To transform a column vector with \( d \) observations of a random variable into a lower dimension of order, \( k < d \), we can multiply the column vector with a matrix, \( A \sim N(0; \frac{1}{k}) \) of dimension \( k \times d \).

Proposition 1. A dimension transformation of \( d \) observations of a log-normal variable into a lower dimension, \( k \), using Lemma 4, yields a probability density function which is the sum of random variables with a normal log-normal mixture, given by the convolution,

\[
fs (a) = f_{U_1} (u_1) \ast f_{U_2} (u_2) \ast \ldots \ast f_{U_k} (u_k)
\]

Here, \( f_{U_k} (u_k) = \frac{\sqrt{k}}{2\pi\sigma_k} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma_k^2}} \frac{|y|^k}{\sigma_k} dy \)

\[
U_k = X_k e^{Y_k}
\]

\[
\begin{bmatrix} X_t \\ Y_t \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ \mu_{Y_t} \end{bmatrix} , \begin{bmatrix} \frac{1}{\mu_{Y_t}} & 0 \\ 0 & \sigma_{Y_t}^2 \end{bmatrix} \right)
\]

The convolution of two probability densities arises when we have the sum of two independent random variables, \( Z = X + Y \). The density of \( Z \), \( h_Z (z) \), is given by,

\[
h_Z (z) = (f_X f_Y) (z) = \int_{-\infty}^{\infty} f_X (z - y) \ast f_Y (y) dy = \int_{-\infty}^{\infty} f_X (z) \ast f_Y (z - x) dx
\]

When the number of independent random variables being added is more than two, or the reduced dimension after the Lemma 4 transformation is more than two, \( k > 2 \), then we can take the convolution of the density resulting after the convolution of the first two random variables, with the density of the third variable and so on in a pair wise manner, till we have the final density.
Proposition 2. A dimension transformation of $d$ observations of a normal variable into a lower dimension, $k$, using Lemma 1, yields a probability density function which is the sum of random variables with a normal normal product distribution, given by the convolution,

$$f_S(s) = f_U(u_1) \ast f_U(u_2) \ast \ldots \ast f_U(u_k)$$

Here, $f_U(u_k) = \int_{-\infty}^{\infty} \left( \frac{1}{\sqrt{2\pi}} \right) \frac{1}{\sigma_Y \sqrt{2\pi}} e^{-\frac{(x-\mu_Y)^2}{2\sigma_Y^2}} \sqrt{\frac{k}{2\pi}} e^{-\frac{s(x)^2}{2}} dx$

$$U_k = X_kY_k$$

$$\begin{bmatrix} X_k \\ Y_k \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ \mu_{Y_k} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \sigma_Y^2 \end{bmatrix} \right)$$
Lemma 4. The Bhattacharyya coefficient when we have truncated multivariate normal distributions \( p, q \) and all the \( k \) dimensions have some overlap, is given by

\[
D_{BC-TMN}(p, q) = \frac{1}{8} (\mu_p - \mu_q)^T \Sigma^{-1} (\mu_p - \mu_q) + \frac{1}{2} \ln \left( \frac{\det \Sigma}{\sqrt{\det \Sigma_p \det \Sigma_q}} \right) + \frac{1}{2} \ln \left( \frac{1}{\sqrt{(2\pi)^k (|\Sigma_p|)}} \int_b^a \exp \left( -\frac{1}{2} (x - \mu_p)^T \Sigma_p^{-1} (x - \mu_p) \right) \, dx; x \in \mathbb{R}^k_{a \leq x \leq b} \right)
\]

\[
+ \frac{1}{2} \ln \left( \frac{1}{\sqrt{(2\pi)^k (|\Sigma_q|)}} \int_c^d \exp \left( -\frac{1}{2} (x - \mu_q)^T \Sigma_q^{-1} (x - \mu_q) \right) \, dx; x \in \mathbb{R}^k_{c \leq x \leq d} \right)
\]

\[
- \ln \left( \frac{1}{\sqrt{(2\pi)^k \det (\Sigma_p \Sigma_q^{-1} \Sigma_q)}} \right)
\]

\[
\int_a^b \exp \left( -\frac{1}{2} \left\{ (x - m)^T (\Sigma_q^{-1} | \Sigma | \Sigma_p^{-1}) (x - m) \right\} \right) \, dx; x \in \mathbb{R}^k_{\min(a,c) \leq x \leq \min(b,d)}
\]

Here,

\[
p \sim N(\mu_p, \Sigma_p, a, b)
\]

\[
q \sim N(\mu_q, \Sigma_q, c, d)
\]

\[
l = \min(a, c) \\; ; \\; m = \min(b, d)
\]

\[
m = \left( \mu_p^T \Sigma_p^{-1} + \mu_q^T \Sigma_q^{-1} \right) (\Sigma_p^{-1} + \Sigma_q^{-1})^{-1}
\]

\[
\gamma = \frac{\Sigma_p + \Sigma_q}{2}
\]

Lemma 3. The Bhattacharyya distance, which is zero and when they overlap, is given by

\[
D_{BC-TN}(p, q) = \frac{1}{2} \left( \frac{(\mu_p - \mu_q)^2}{\sigma_p^2 + \sigma_q^2} \right) + \frac{1}{2} \ln \left( \frac{1}{\sqrt{\sigma_p^2 + \sigma_q^2 + 2}} \right)
\]

\[
+ \frac{1}{2} \ln \left( \frac{\Phi \left( \frac{b - \mu_p}{\sigma_p} \right) - \Phi \left( \frac{a - \mu_p}{\sigma_p} \right)}{\Phi \left( \frac{d - \mu_q}{\sigma_q} \right) - \Phi \left( \frac{c - \mu_q}{\sigma_q} \right)} \right)
\]

\[
- \ln \left( \Phi \left( \frac{u - \nu}{\zeta} \right) \right)
\]

Here,

\[
p \sim N(\mu_p, \sigma_p^2, a, b)
\]

\[
q \sim N(\mu_q, \sigma_q^2, c, d)
\]

\[
l = \min(a, c) \; ; \; u = \min(b, d)
\]

\[
\nu = \frac{(\mu_p \sigma_q^2 + \mu_q \sigma_p^2)}{(\sigma_p^2 + \sigma_q^2)} \; ; \; \zeta = \sqrt{\frac{2 \sigma_p^2 \sigma_q^2}{(\sigma_p^2 + \sigma_q^2)}}
\]
The Kryptonite for Infoman

Combining Covariance and Distance

**Lemma 5.** The following equations govern the relationship between the Bhattacharyya distance, $\rho(f_X, f_Y)$, and the covariance between any two distributions with joint density function, $f_{XY}(t,u)$, means, $\mu_X$ and $\mu_Y$ and density functions $f_X(t)$ and $f_Y(t)$.

\[
\text{Cov}[c(X), Y] = \text{Cov}(X,Y) - E \left[ \sqrt{\frac{f_Y(t)}{f_X(t)}} \right] + \mu_Y \rho(f_X, f_Y)
\]

\[
\text{Cov}(X,Y) + \mu_Y \rho(f_X, f_Y) = E[e' (X) g(X,Y)] + E \left[ \sqrt{\frac{f_Y(t)}{f_X(t)}} Y \right]
\]

Here,

\[
c(t) = t - \sqrt{\frac{f_Y(t)}{f_X(t)}}
\]

and $g(t,u)$ is a non-vanishing function such that,

\[
\frac{f'_{XY}(t,u)}{f_{XY}(t,u)} = \frac{g'(t,u)}{g(t,u)} + \frac{[\mu_Y - u]}{g(t,u)}, \quad t,u \in (a,b)
\]
FROM SYMBOLS TO NUMBERS

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
**Figure 2: Volume Distance Measures over Randomly Chosen Sub Universe**

(a) Volume PCA Dimension Reduction

| Iteration=1 | AUS   | GBR   | HKG   | IND   | JPN   | SGP   |
|-------------|-------|-------|-------|-------|-------|-------|
| AUS         | 84.81 | 76.82 | 85.01 | 88.46 | 85.64 | -     |
| GBR         | 93.91 | 94.34 | 82.85 | 82.30 | 91.94 | -     |
| HKG         | 74.70 | 87.92 | 82.37 | 93.98 | 70.89 | -     |
| IND         | 91.10 | 81.72 | 82.37 | 62.41 | 89.73 | -     |
| JPN         | 90.30 | 87.39 | 98.24 | 61.81 | 97.81 | -     |
| SGP         | 69.21 | 89.96 | 69.95 | 85.73 | 98.17 | -     |

(b) Volume JL Lemma Dimension Reduction

| Iteration=2 | AUS   | GBR   | HKG   | IND   | JPN   | SGP   |
|-------------|-------|-------|-------|-------|-------|-------|
| AUS         | 81.86 | 78.23 | 96.07 | 90.37 | 70.70 | -     |
| GBR         | 90.42 | 92.18 | 78.88 | 87.85 | 90.18 | -     |
| HKG         | 67.18 | 92.73 | 82.37 | 95.56 | 72.50 | -     |
| IND         | 85.44 | 81.43 | 82.37 | 62.35 | 86.19 | -     |
| JPN         | 82.29 | 85.22 | 92.64 | 61.67 | 97.58 | -     |
| SGP         | 71.25 | 89.07 | 76.25 | 88.51 | 96.56 | -     |

| Iteration=3 | AUS   | GBR   | HKG   | IND   | JPN   | SGP   |
|-------------|-------|-------|-------|-------|-------|-------|
| AUS         | 91.83 | 67.99 | 87.39 | 86.02 | 54.20 | -     |
| GBR         | 91.75 | 93.97 | 80.84 | 84.74 | 92.59 | -     |
| HKG         | 73.46 | 95.92 | -     | 82.37 | 96.44 | 72.98 |
| IND         | 86.44 | 80.28 | 82.37 | 62.12 | 85.65 | -     |
| JPN         | 92.56 | 84.56 | 95.61 | 63.02 | 102.12 | -     |
| SGP         | 63.40 | 89.66 | 69.56 | 86.12 | 96.98 | -     |

| Iteration=4 | AUS   | GBR   | HKG   | IND   | JPN   | SGP   |
|-------------|-------|-------|-------|-------|-------|-------|
| AUS         | 78.91 | 72.97 | 83.66 | 90.74 | 64.95 | -     |
| GBR         | 81.57 | 92.61 | 80.52 | 84.93 | 93.73 | -     |
| HKG         | 69.17 | 91.17 | 82.37 | 96.88 | 67.78 | -     |
| IND         | 86.65 | 82.47 | 82.37 | 62.77 | 91.47 | -     |
| JPN         | 83.34 | 85.57 | 96.62 | 63.14 | 97.85 | -     |
| SGP         | 66.94 | 92.89 | 74.65 | 84.17 | 96.33 | -     |

| Iteration=5 | AUS   | GBR   | HKG   | IND   | JPN   | SGP   |
|-------------|-------|-------|-------|-------|-------|-------|
| AUS         | 84.48 | 71.55 | 83.71 | 90.53 | 72.69 | -     |
| GBR         | 80.39 | 89.52 | 78.57 | 86.37 | 91.60 | -     |
| HKG         | 79.32 | 92.58 | 82.37 | 92.38 | 71.09 | -     |
| IND         | 91.27 | 83.53 | 82.37 | 63.25 | 85.50 | -     |
| JPN         | 91.05 | 86.06 | 98.11 | 62.77 | 95.25 | -     |
| SGP         | 85.91 | 91.14 | 68.00 | 89.46 | 95.75 | -     |
## A Pricey Prescription (?)

Table 1: Close PCA Dimension Reduction

| Sig Level = 2 | AUS | GBR | HKG | IND | JPN | SGP |
|---------------|-----|-----|-----|-----|-----|-----|
| AUS           | 0.76 | 0.61 | 0.58 | 0.61 | 0.60 | 10.96 |
| GBR           | 1.19 | 0.52 | 0.70 | 0.52 | 0.59 | 22.54 |
| HKG           | 0.33 | 0.13 | 0.60 | 2.58 | 12.81 |
| IND           | 13.14 | 10.78 | 16.64 | 22.84 | 16.05 |

Table 2: Close JL Lemma Dimension Reduction

| Iteration 1 | AUS | GBR | HKG | IND | JPN | SGP |
|-------------|-----|-----|-----|-----|-----|-----|
| AUS         | 0.76 | 0.61 | 0.58 | 0.61 | 0.60 | 10.96 |
| GBR         | 1.19 | 0.52 | 0.70 | 0.52 | 0.59 | 22.54 |
| HKG         | 0.33 | 0.13 | 0.60 | 2.58 | 12.81 |
| IND         | 13.14 | 10.78 | 16.64 | 22.84 | 16.05 |

Figure 5: Close Distance Measures over Randomly Chosen Sub Universe

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com

© 2016 Gain Knowledge Group. All Rights Reserved.
## Taming The (Volatility) Skew

### Sig Level 1:

|          | AUS | GBR | HRC | IND | JPN | SGP |
|----------|-----|-----|-----|-----|-----|-----|
| Iteration 1 | 0.48 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 2 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 3 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 4 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |

### Sig Level 2:

|          | AUS | GBR | HRC | IND | JPN | SGP |
|----------|-----|-----|-----|-----|-----|-----|
| Iteration 1 | 0.48 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 2 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 3 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 4 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |

### Sig Level 3:

|          | AUS | GBR | HRC | IND | JPN | SGP |
|----------|-----|-----|-----|-----|-----|-----|
| Iteration 1 | 0.48 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 2 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 3 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |
| Iteration 4 | 0.38 | 0.70 | 0.89 | 0.93 | 0.94 | 0.95 |

**Figure 10:** Close Volatility PCA Dimension Reduction

**Figure 11:** Close Volatility JL Lemma Dimension Reduction

---

Gain Knowledge Group  
First & Last Stop For Your Research Problems  
www.gainknowledge.com

© 2016 Gain Knowledge Group. All Rights Reserved.
Visions of Bubbles and Busts?
Story of Beauty and the Beast!

- Bubbles are seductive but scary when they burst.

- Just as Beauty and the Beast must Coexist; Bubbles and Busts must be close to one another.

- Trading costs could be a signal of the building up of a bubble and a later bust.

- If microstructure variables, especially implicit trading costs, are showing steady movement, the change in transaction costs could be a signal of a potential building up of a bubble and a later bust.

- Our study will allow the comparison of trading costs across aggregations of individual securities, allowing inferences to be drawn across sectors or markets, enabling us to find early indications of bubbles building up in corners of the economy.
Sleeping Like A Koala

- It is said that the Universe is but the Brahma’s (Creator’s) dream.
- Research (Effort / Struggle) can help us understand this world; Sleep (Ease / Peace of Mind) can help us create our own world
- A lesson from close by and down under; We need to “Do Some Yoga and Sleep Like A Koala”.

Gain Knowledge Group
First & Last Stop For Your Research Problems
www.gainknowledge.com
Some Sleeping Aids

Carroll, L. (1865). (2012 Reprint) Alice’s adventures in wonderland. Random House.

Carroll, L. (1871). (2009 Reprint) Through the looking glass: And what Alice found there. Lulu.com.

Bhattacharyya, A. (1943). On a Measure of Divergence Between Two Statistical Populations Defined by their Probability Distributions, Bull. Calcutta Math. Soc., 35, pp. 99-110.

Bhattacharyya, A. (1946). On a measure of divergence between two multinomial populations. Sankhya: The Indian Journal of Statistics, 401-406.

Eco, U., & Chilton, N. (1972). The myth of Superman.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. Econometrica: journal of the Econometric Society, 135-155.

Smith, M. (1979). Behaviour of the Koala, Phascolarctos Cinereus Goldfuss, in Captivity. 1. Non-Social Behaviour. Wildlife Research, 6(2), 117-129.

Stein, C. M. (1981). Estimation of the mean of a multivariate normal distribution. The annals of Statistics, 1135-1151.

Barnett, L. D. (1984). The Brahma Knowledge (Vol. 1). Library of Alexandria.

Lawson, T. (1985). Uncertainty and economic analysis. The Economic Journal, 95(380), 909-927.

Frankl, P., & Machara, H. (1988). The Johnson-Lindenstrauss lemma and the sphericity of some graphs. Journal of Combinatorial Theory, Series B, 44(3), 355-362.

Frankl, P., & Machara, H. (1990). Some geometric applications of the beta distribution. Annals of the Institute of Statistical Mathematics, 42(3), 463-474.

Ramamurthi, B. (1995). The fourth state of consciousness: The Thuriya Avastha. Psychiatry and clinical neurosciences, 49(2), 107-110.

Dasgupta, S., & Gupta, A. (1999). An elementary proof of the Johnson-Lindenstrauss lemma. International Computer Science Institute, Technical Report, 99-006.

Chiani, M., Dardani, D., & Simon, M. K. (2003). New exponential bounds and approximations for the computation of error probability in fading channels. Wireless Communications, IEEE Transactions on, 2(4), 840-845.

Horrace, W. C. (2005). Some results on the multivariate truncated normal distribution. Journal of Multivariate Analysis, 94(1), 209-221.

Burges, C. J. (2009). Dimension reduction: A guided tour. Machine Learning, 2(4), 275-365.

Kattumannil, S. K. (2009). On Stein’s identity and its applications. Statistics & Probability Letters, 79(12), 1444-1449.

Ghatage, S. (2010). Brahma’s Dream. Anchor Canada.

Kashyap, R. (2014). Dynamic Multi-Factor Bid-Offer Adjustment Model: A Feedback Mechanism for Dealers (Market Makers) to Deal (Grapple) with the Uncertainty Principle of the Social Sciences. The Journal of Trading, 9(3), 42-55.

Kashyap, R. (2015). A Tale of Two Consequences. The Journal of Trading, 10(4), 51-95.

Kashyap, R. (2016a). Notes on Uncertainty, Unintended Consequences and Everything Else. Working Paper.

Kashyap, R. (2016b). Combining Dimension Reduction, Distance Measures and Covariance. Working Paper.

Kashyap, R. (2017). Microstructure under the Microscope: Tools to Survive and Thrive in The Age of (Too Much) Information. Accepted and Forthcoming.
Now that the US elections are over, not just America, but perhaps, the entire world might need the creator’s blessings.

Thank You.

The views and opinions expressed in this presentation, along with any mistakes, are mine alone and do not necessarily reflect the official policy or position of either of my affiliations or any other agency.
Partial List of Acknowledgements (Halle Berry Style?)

- The following individuals have been a constant source of inputs (sometimes without their realization) and encouragement, more continuous than the flow of orders in an extremely liquid venue:

  - Brad Hunt, Sam Zou, Ross Allen, Dan Campion, Karen King, Julian Chesser, Richard Earl, Yaacov Mutnikas, Dave Weisberger, Mike Aldridge, Rob Flatley, Alex Gillula, Henry Yegerman, Ron Ang, Tom Conigliaro, Mike Richter, Chris Hammond, Rob Beauregard, Patrick Fang, Melissa Allison, Kent Bevel, Laura Davis, Jerrine Chia, Leanna Lim, Randolf Tantzscher, Sophie Cortes, Lange Uggl, Adam Kansler, Shane Akeroyd, Kevin Gould, Jen Theiss, Philippe Rivet, Jiyi Kim, Vincent Huang / Wang, Tracy Man, JP Chan, Mehmet Gurtin and Nic Berner at IHS Markit.

  - Patrick Lawlor, Joanna Wong, Eugene Kanevsky, Michael Yau, Jerry Ing, Andy Maynard, John Mackie, Nigel Lucas, Joakim Axelsson, Ricca Pang, Kaushal Shah, Kunj Gandhi, Rafal Czerniawski, Jason Kwok, Roshan Khan, and Stella Kong at CLSA.

  - Several unnamed clients at undisclosed firms.

  - Dr. Yong Wang, Dr. Isabel Yan, Dr. Vikas Kakkar, Dr. Fred Kwan, Dr. William Case, Dr. Srikant Marakani, Dr. Qiang Zhang, Dr. Costel Andonie, Dr. Jeff Hong, Dr. Guangwu Liu, Dr. Humphrey Tung, Dr. Xu Han at the City University of Hong Kong; Dr. Richard Sylla, Dr. Adam Brandenburger, Dr. Richard Freedman, Dr. Robert Engle, Prof. Larry Zicklin, Prof. Seth Freeman, Dr. Laura Veldkamp, Dr. Ignacio Esponda at New York University; Dr. Liam Lenten at La Trobe University; and Dr. Paul Joseph at National Institute of Technology Calicut.

  - Brian Bruce, Melinda Estelle, Deborah Trask, Harry Katz, Christine Kemper, Courtney Adams and the anonymous many at the Institutional Investor Journals, Journal of Trading.

  - Aqeela Rahman, Julie Valentine, Gautam Mitra and the Organizing Team at UNICOM Seminars.

- The views and opinions expressed in this presentation, along with any mistakes, are mine alone and do not necessarily reflect the official policy or position of either of my affiliations or any other agency.
(1) Good Morning Hong Kong. I am Ravi Kashyap. When someone asks me what I do for a living; I say I work with models and then if they ask, what type of models, I say exotic models. Thankfully the models I work with are not as sophisticated as the models that come to the minds of most people when they hear the word model. Perhaps, my models are not even the models that most people wish to work with.

(2) Today we are going to talk about information explosion and ways to combat it. This requires us to understand Uncertainty as it relates to people. Our empirical study is about prices and volumes, observed on a venue. Hence to set the stage, we will briefly look at the motivations of venues, the connection between portfolio management and the need for smarter computing, before reviewing the methodology used. We will see how these techniques can be useful for decision making and for building intelligence artificially.

(4) Sergei Bubka is our Icon of Uncertainty. To the younger people in the room, he broke the pole vault world record 35 times. We can think of regulatory change as raising the bar. Each time the bar is raised, the spirit of Sergei Bubka, in all of us, will find a way over it.

(5) There is an Uncertainty Principle in Physics called the Heisenberg Principle. I only have a high school physics background and am not qualified to talk about it. But in the social sciences, we have something similar. What it says is that, the more popular a generalization, the less accurate are the predictions it yields. A simple example of this: if there is a prediction that the price of the stock will go up tomorrow, then if more people become aware of it; the less certain we can be of the result or what will happen to the price tomorrow.

(6) Specific to finance, where we have many active participants, throughout time, we have had only three simple outcomes - "Buy, Sell or Hold". The complications are mostly in obtaining these results.

(8) We are yet to discover an objective measuring stick for comparison, a so called, True Comparison Theory, which can be an aid for arriving at objective decisions. Despite all the uncertainty, the one thing we can be almost certain about is the subjectivity in all decision making. This lack of an objective measure for comparisons makes people react at varying degrees and at varying speeds, as they make their subjective decisions. A decision gives rise to an action and subjectivity in the comparison means differing decisions and hence unpredictable actions. This inability to make consistent predictions in the social sciences explains the growing trend towards comprehending better and deciphering the decision process and the subsequent actions, by collecting more information across the entire cycle of comparisons, decisions and actions.

(10) If Alice and Red Queen of Wonderland fame were to visit Hedge-Fund-Land (or even Business-Land), the following modification of their popular conversation would aptly describe the situation today, “My dear, here we must process as much data as we can, just to stay in business. And if you wish to make a profit you must process at least twice as much data.”