Dimensional analysis, scaling, and zero-intelligence modeling for financial markets

Eric Smith (SFI)

based on work with

Doyne Farmer (SFI)
Supriya Krishnamurthy (Swedish Inst. Comp. Sci)
Laci Gillemot (Budapest U. Technology and Economics)
Giulia Iori (City U. London)
Martin Shubik (Cowles Foundation, Yale)
Paolo Patelli (LANL)
Marcus Daniels (LANL)

Outline

- Where and why can a scientific point of view contribute to economic understanding?

- Dimensional analysis and scaling

- Zero-intelligence modeling

- A worked example: the continuous double auction of finance

- Compelling open problems

I. Relation of science to economics

- A little history and ideology of current mainstream economics

- Formal foundations and their flaws

- Entry points for "scientific" methodology

The intellectual homage of "neoclassical" economic theory

..every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, *neither intends to promote the public interest, nor knows how much he is promoting it**. By preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value,* *he intends only his own gain**, and he is in this, as in many other cases,* **led by an invisible hand** *to promote an end which was no part of his intention. Nor is it always the worse for the society that it was no part of it.* *By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it.* **I have never known much good done by those who affected to trade for the public good.**

Objectives of (much of) modern mathematical economics

- Formal concept of *optimal resource allocation*

- Identification of "final goals" of trade (~implies equilibrium)

- Proof that optimal equilibria are defined by the uncoordinated desires of self-regarding individuals (mathematical existence theorem)

Foundations of "neoclassical" economic theorizing

- ## Complete, costless contracts

  - All consequences known and bargained for   (No pig farms)

  - All contracts, costless, instantaneous, guaranteed   (No accountants, courts, gov't)

- ## "Rationality" := Global self-fulfilling prophecy

  - Strategic actions by everyone "as good as given"   (Mutual "best response")

  - Non-constructive existence proof for equilibria   (Multiple, uncomputable)

- ## "Utilitarian" preferences (usually self-regarding)

  - Partial order on all possible states of the world

    (Known false many ways)

  - All outcomes over all time have present value

Amounts to a process-free, institution-free world

Some cautions that go with criticism

- Beware the Laplacians:

  - People are not "simply" machines

  - No reason economic order will look like physical order

- Beware the "econophysicists"

  - Renaming mathematical models from physics would not count as good theory even in physics

  - Empirical validation of social science is tricky

Opportunities for science in economics

- Institutions are by nature "mechanistic"

    - Based on rules and constraints

    - Opportunity for structural and process analysis (analysis of market function, design of auctions, etc.)

- "Zero Intelligence" as a pure model of institutional constraint

    - Constraints can shape both action and reasoning

- Errors in ZI models can be among the best indicators of behavioral regularity

    - Provides better focus for proper behavioral science

What would a "scientific" economics look like?

- Richer conceptual substructure between empiricism and theory

- "Abstractions" should attach quantitative consequences to partial problem specifications

- Believe models when the paths that led to them have excluded everything else

- Look for notions of "universality" as the justification for simple models

II. Dimensional Analysis and scaling

- ## Two principles

  - Equations must be homogeneous in dimension

  - Quantities with the required dimensions control scaling

- ## Two consequences

  - Can guess the sizes of observables

  - Can relate differently-sized cases to a single model

Example: how fast do you walk?

- **A pure dimensional analysis**

  - You want to know a speed

  $$[\text{speed}] = \left[\frac{\text{length}}{\text{time}}\right]$$

  - You have a property of height

  $$[\text{height}] = [\text{length}]$$

  - You are walking in earth gravity

  $$[g] = [\text{acceleration}] = \left[\frac{\text{length}}{\text{time}^2}\right]$$

  - Only one combination has correct dimensions
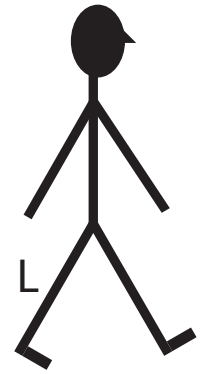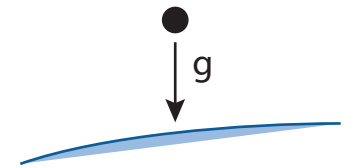
  $$[v] = \left[\sqrt{L\,g}\right]$$

- **Could break this down in more detail**

  - Your leg is a pendulum

  $$[\text{frequency}] = \left[\frac{1}{\text{time}}\right]$$

  - Pendulum is characterized by frequency

  - Only one combination works dimensionally

  $$[\omega] = \left[\sqrt{\frac{g}{L}}\right]$$

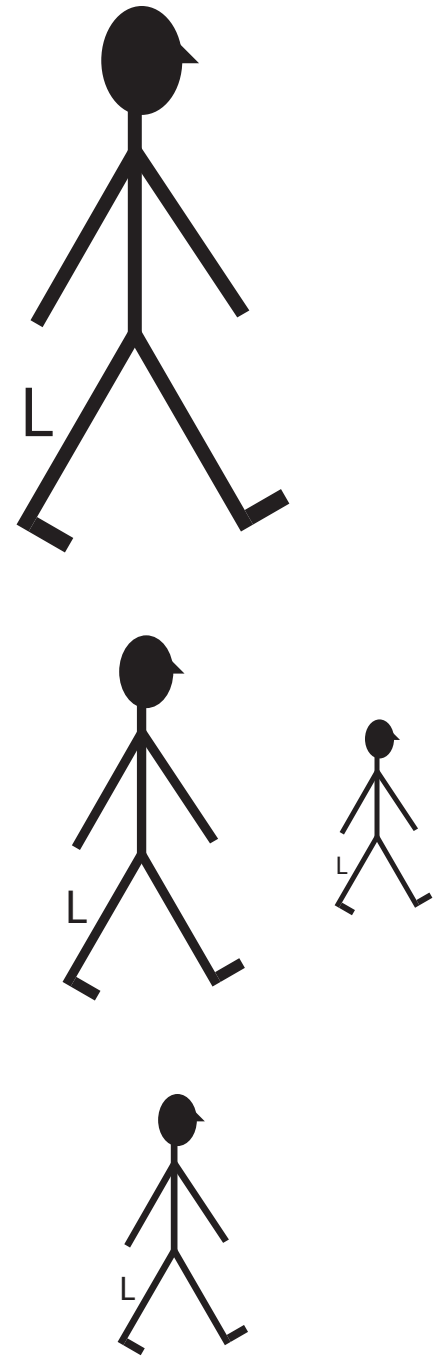 - Speed is "ticking" frequency times length
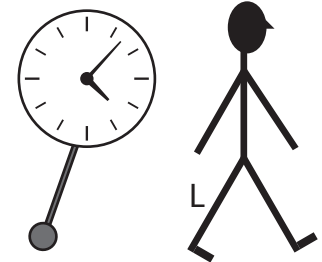
  $$[v] = [L\,\omega] = \left[L\sqrt{\frac{g}{L}}\right]$$

All walkers great and small

- Walking is a *scale-invariant* process

- All walkers share gravity, they differ mostly in their characteristic lengths

- Dimensional homogeneity can be used to generate scale factors relating different walkers to each other

$$[\omega_i] = \left[\sqrt{\frac{g}{L_i}} \right] = \left[\sqrt{\frac{L_0}{L_i}}\, \omega_0 \right]$$

$$[v_i] = \left[\sqrt{L_i\, g} \right] = \left[\sqrt{\frac{L_i}{L_0}}\, v_0 \right]$$

III. Zero-Intelligence (ZI) modeling: a pure formulation of institutional constraint

- We don't know what "intelligence" or "rationality" are; but we do know how to exclude them

- An example: the minority game

- ZI modeling can be a foundation to which more formal models of learning are added

The minimal Minority Game: the original zero-intelligence model

- A population of N players (usually odd) choose one move from the set (0,1), independently and simultaneously.

- Each player whose choice was the minority in the population is awarded a payoff ("wins").

- Moves have no intrinsic value

- The number of constellations of winners is enormous (frustration)

- A model of the non-rational component of purely speculative stock trading

Minority Games: a basis for learning models

- Each agent remembers outcomes of M prior rounds of play

- A set of random lookup-tables (particular to each agent) provides "next moves" in response to each possible history (2^M)

- Agents "learn" by choosing the table with best performance so far

- A phase transition occurs in $z = 2^M / N$

- Most satisfaction at critical

Figure 6: s=2

Variance in the number of winners

○	N=11
×	25
+	101
∗	1001
—	slope=−1

Symmetric Unpredictable

Ordered non-symmetric Predictable

From R. Savit, Y. Li, R. Riolo, *Physica A* 276 (2000): 265
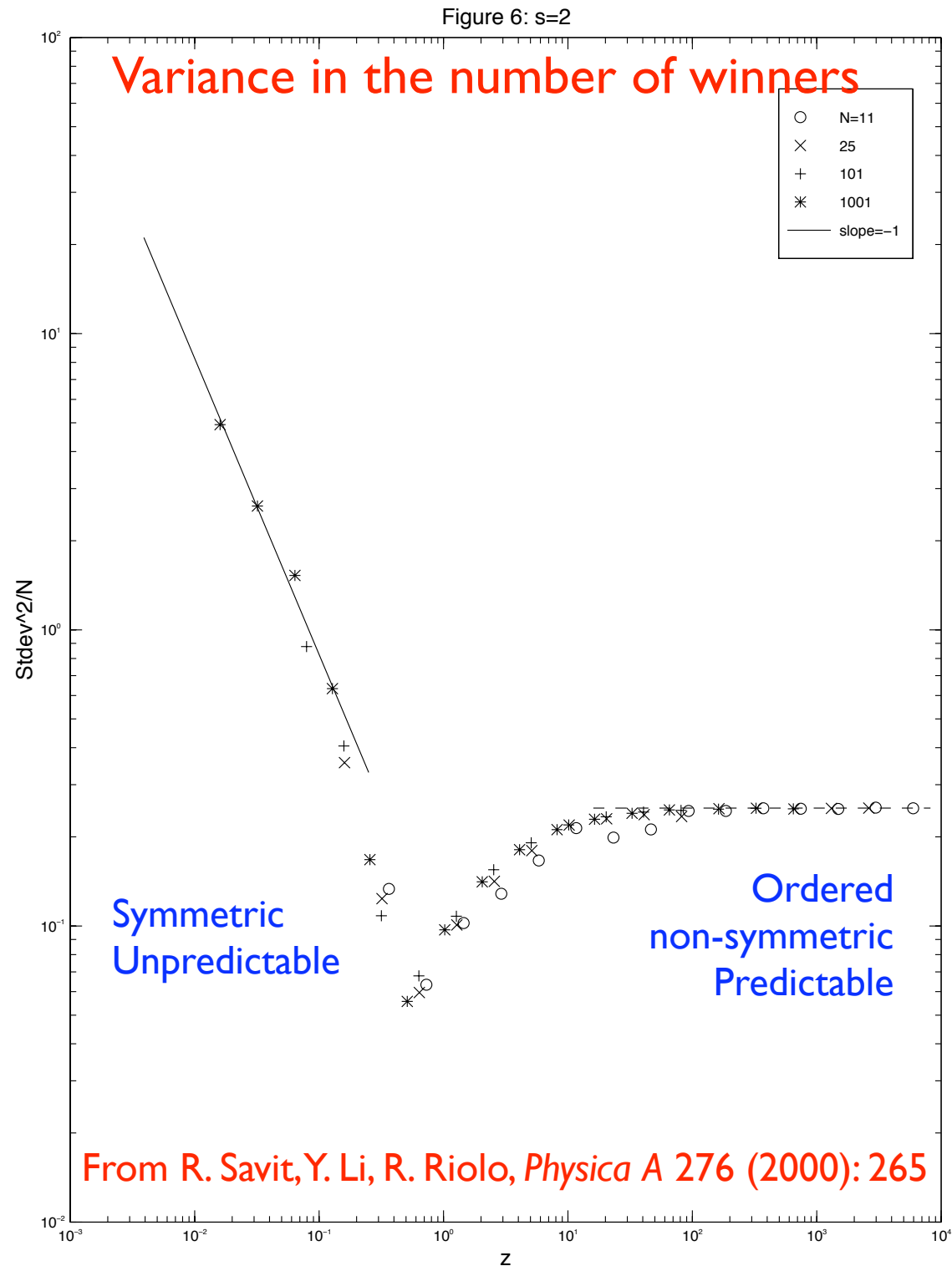
IV. Worked example of dimensional analysis and ZI: the continuous double auction of financial markets

- The basis of all continuous-trading financial markets today

- Solves problem of matching asynchronously-arriving orders to buy and sell

- Heavily institutionalized and high volume (~5-7 billion dollars / day in 1999)

- "Complete", good-quality data

Mostly from Smith, E., J. D. Farmer, L. Gillemot, and S. Krishnamurthy. "Statistical Theory of the Continuous Double Auction." *Quant. Fin.* 3(6) (2003): 481-514

The Continuous Double Auction (CDA) mechanism

- Two kinds of orders

  - Market (v)

  - Limit (p,v)

- Limit orders (LO) accumulate

- Market orders (MO) clear immediately

- LO at single price clear in order of arrival
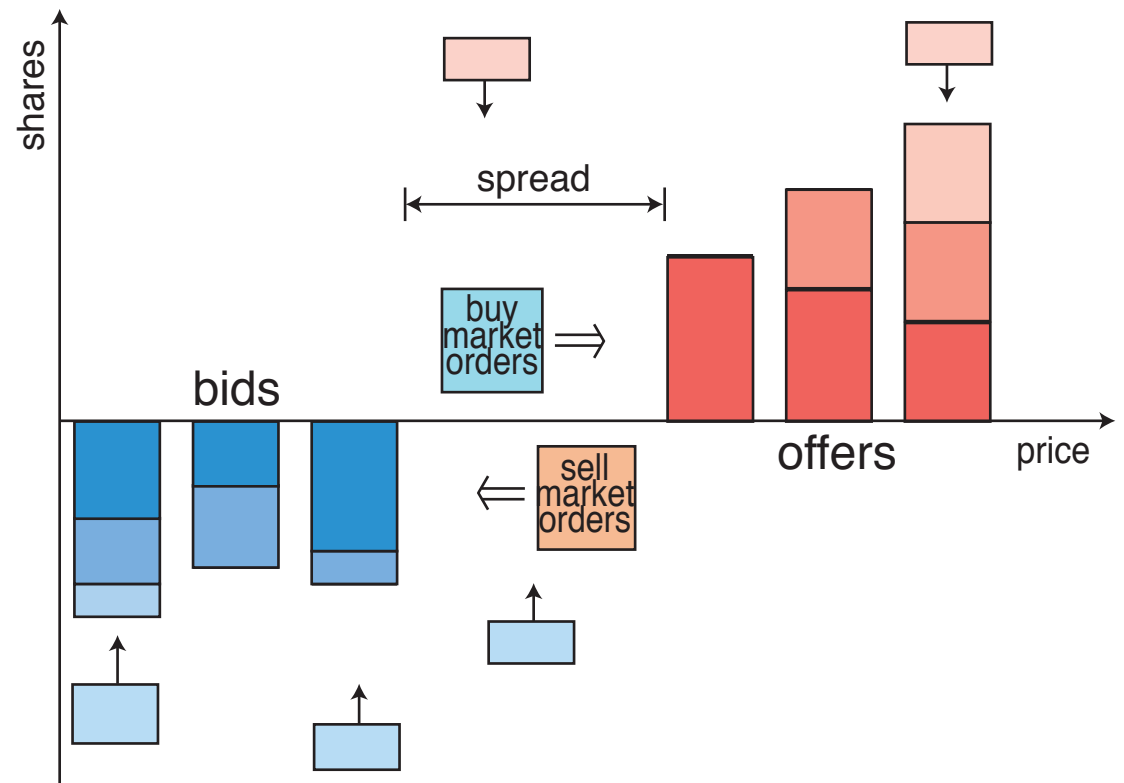
Observables of the CDA

- Bid-ask spread (responsible for costs)

- Volatility (responsible for risk and profit)

- Market impact (responsible for costs and risk)

- Liquidity (resistance to market impact)

- Depth (responsible for stability)

The bid-ask spread

- **Bid** = best offered buying price

- **Ask** = best offered selling price

- **Spread** = ask - bid

- Rapid small buy-sell alterations will pay the average spread per pair of transactions

- Nonzero spread is the leading source of transaction costs

- Nonzero spread is the regulator against divergent volume of trading
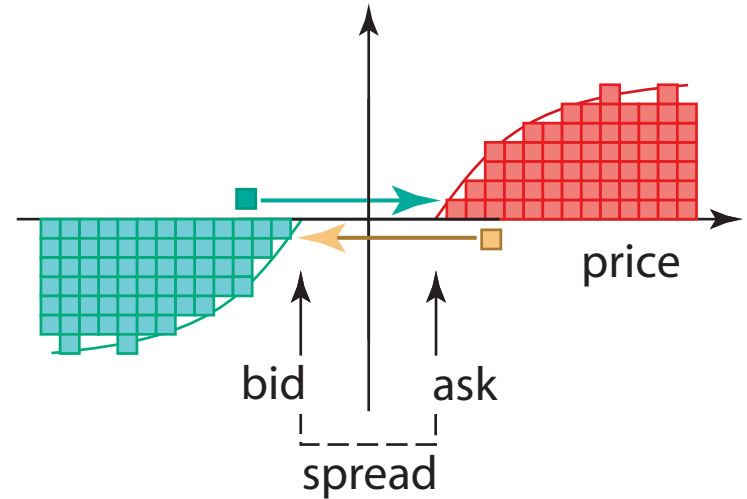
The volatility

- Midprice:
  m = (ask + bid) / 2

- Only midprice motion creates risk + profit opportunities in the presence of a spread

- Has traditionally been modeled as a random walk (Bachelier 1900)

- Diffusion constant of the random walk is called the *volatility*

Midprice motion is a good approximation to diffusion

Vodafone, August 4 1998
D=1.498e! 06, R^2=0.998
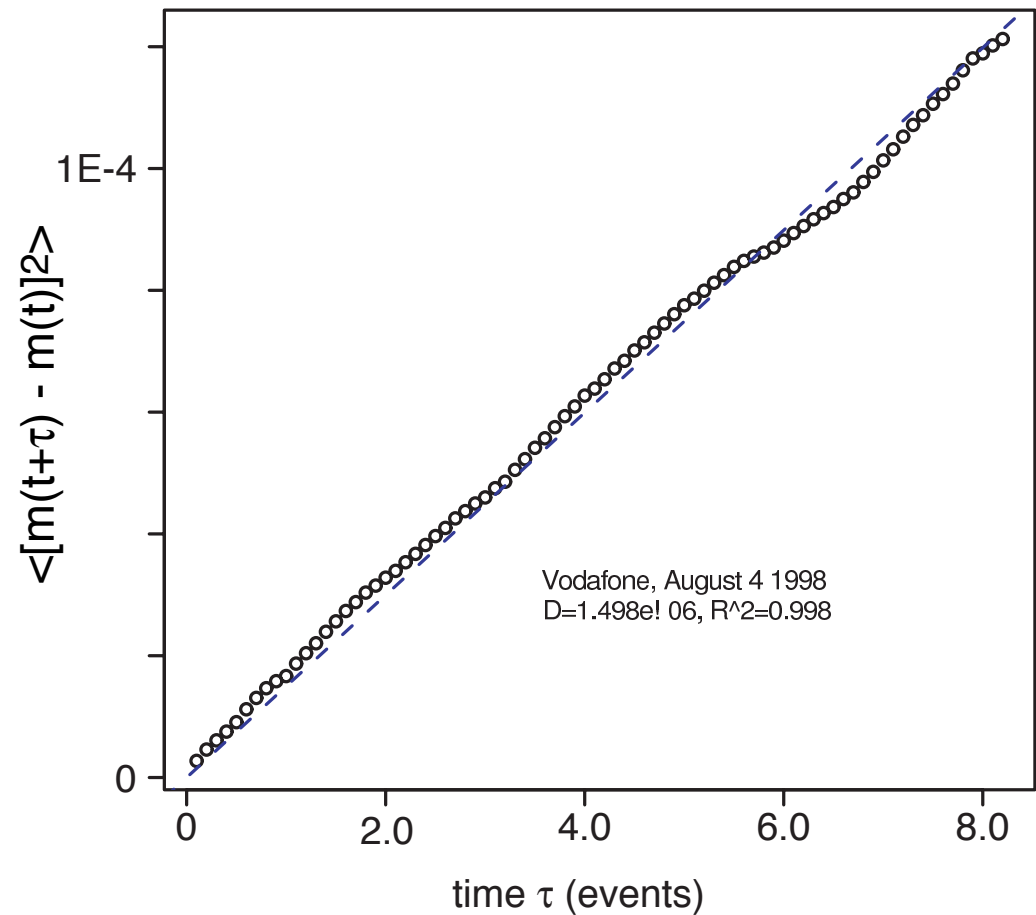
Courtesy Paolo Patelli

Market Impact

- Average price of an order always worse than the starting best price

- Return depends only on price ratios
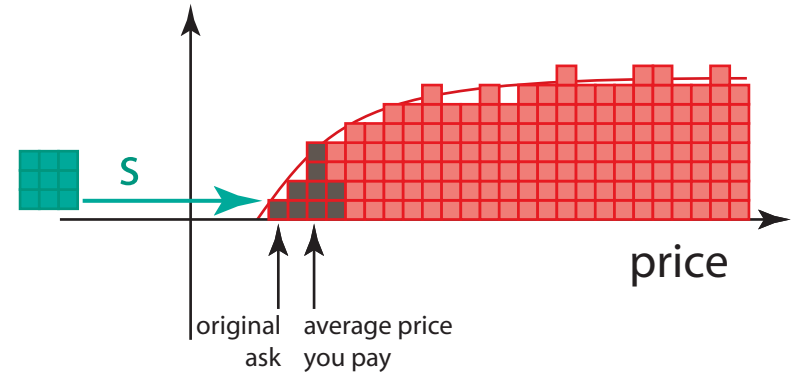
$$\text{return} = \frac{p_{\text{sell}}}{p_{\text{buy}}}$$

- No profit if impact > return

$$(\text{tolerated impact}) \leq \langle \text{return} \rangle$$

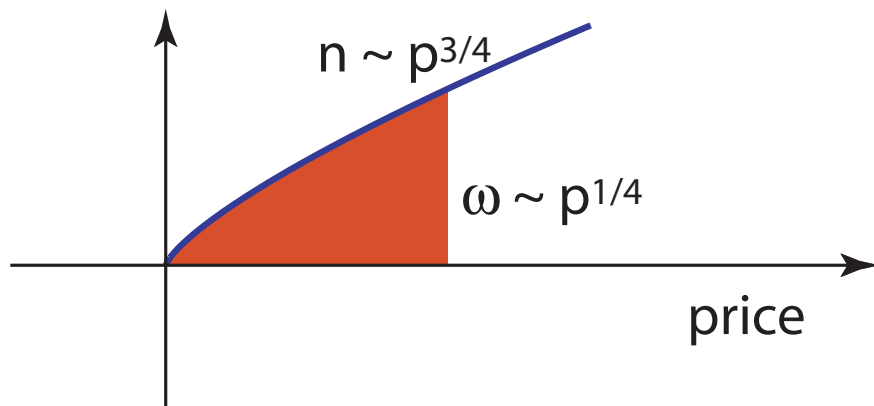- Expect impact to scale exponentially, or else there is an incentive to split or join orders
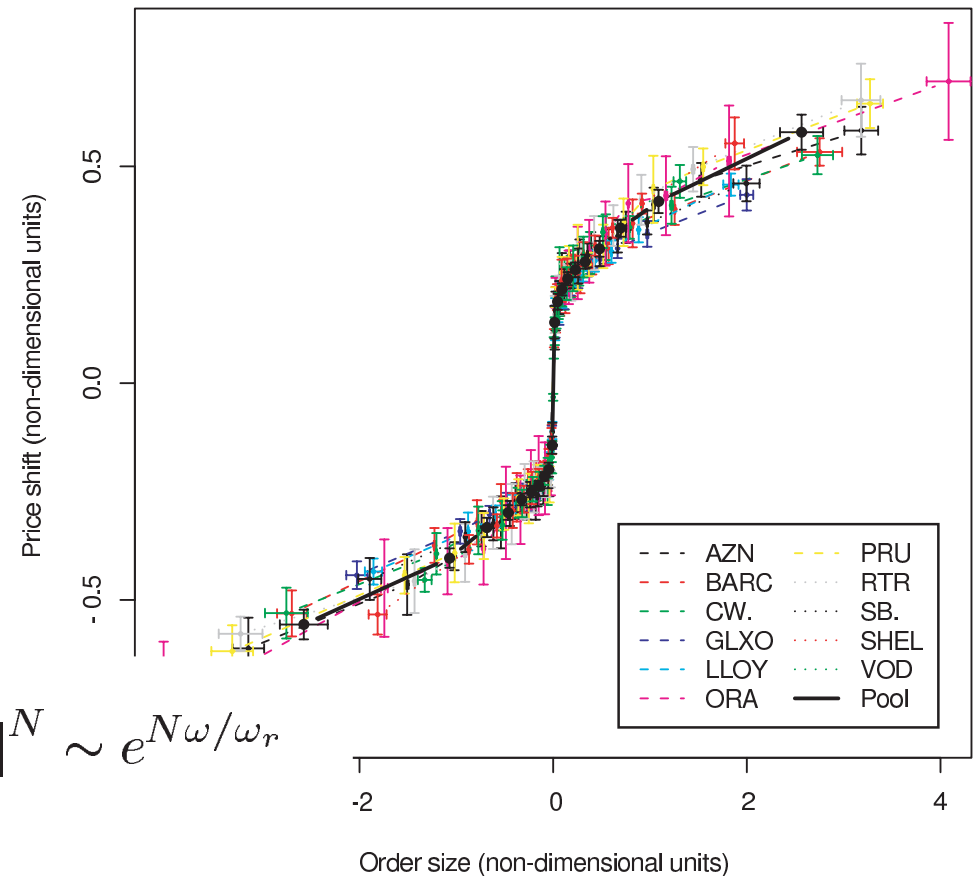
$$(\text{tolerated impact})_{N\omega} \sim [(\text{tolerated impact})_{\omega}]^N \sim e^{N\omega/\omega_r}$$

- Yet it doesn't

Farmer, J. D., P. Patelli, and I. I. Zovko. "The Predicitive Power of Zero Intelligence in Financial Markets" *PNAS USA* 102(11) (2005): 2254-2259
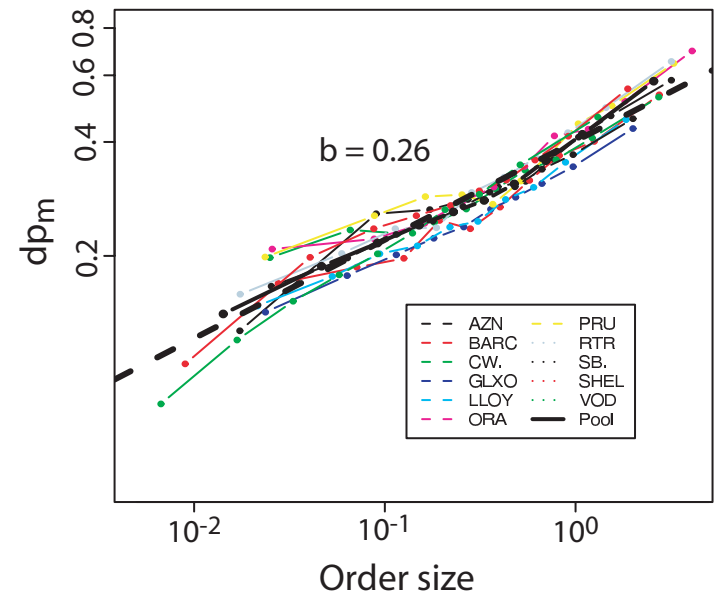
Observed regularities of the market impact

- Power laws with exponents 0.25 - 0.5 are common

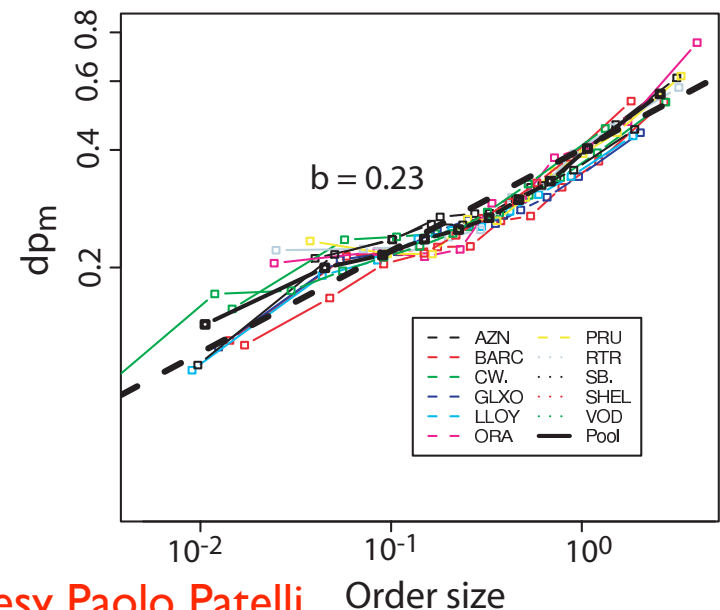- Corresponds to a power-law distribution of limit orders

$n \sim p^{3/4}$

$\omega \sim p^{1/4}$

price

- Market impact is integral of the depth

Buy orders

dp_m

b = 0.26

AZN
BARC
CW.
GLXO
LLOY
ORA
PRU
RTR
SB.
SHEL
VOD
Pool

Order size

Sell orders

dp_m

b = 0.23

AZN
BARC
CW.
GLXO
LLOY
ORA
PRU
RTR
SB.
SHEL
VOD
Pool

Order size

Liquidity

- Liquidity = resistance of price to change in response to market orders

- More standing orders should mean more liquidity

- Coefficient of the power law of market impact is a natural liquidity measure

- For impact power ~0.5 this is just the *slope* of the distribution
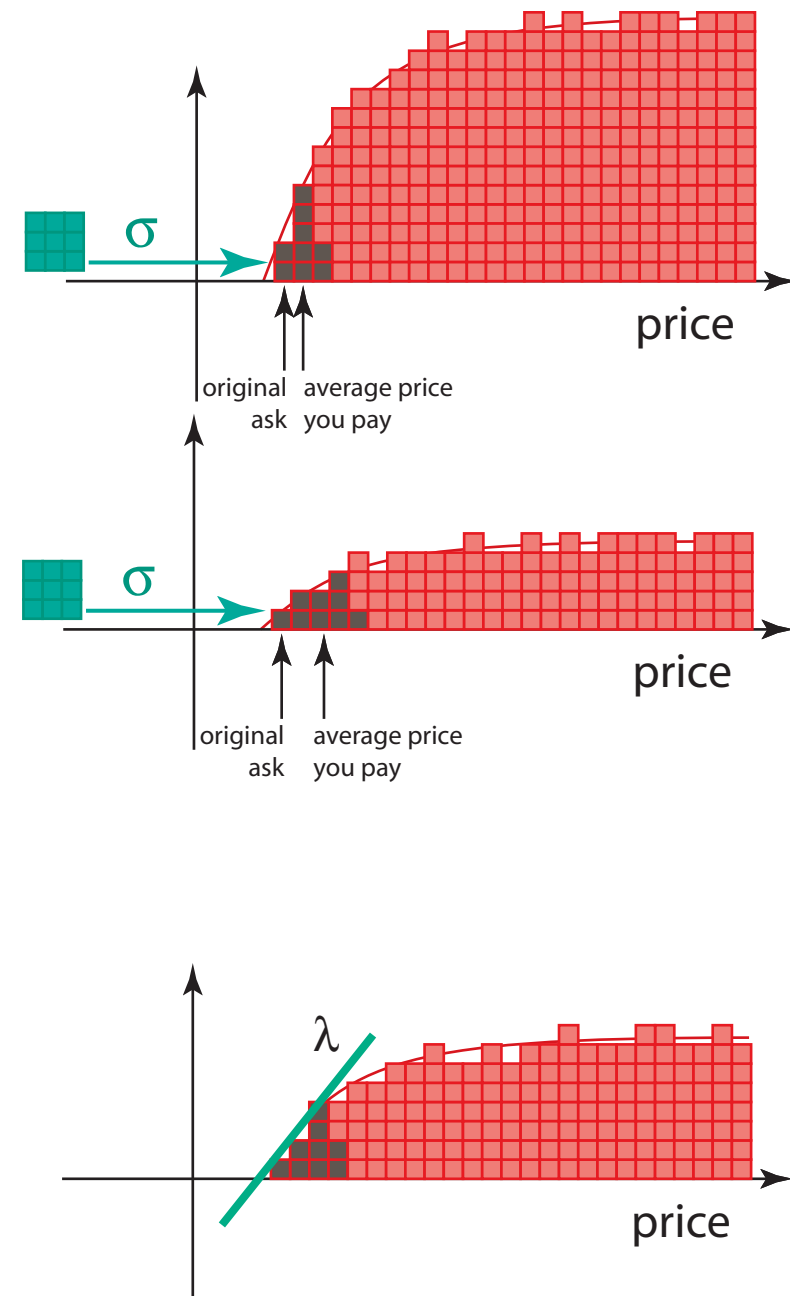
Depth of the limit-order "book"

- Clearly asymptotic order density determines where (concave) power-law scaling ends and linear scaling begins (and gives the coefficient)

- Order density (shares / price) is termed "depth" (n)
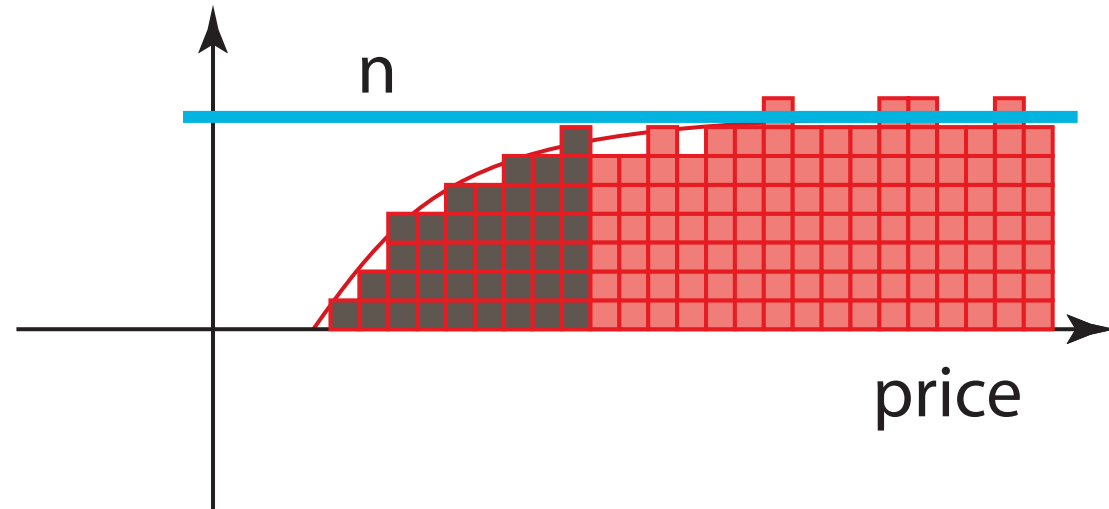
Dimensional analysis of the continuous flows

- Summarize the order-placement processes by a collection of continuous rates and rate densities

- Limit order placement α

- Limit order deletion δ

- Market order arrival μ
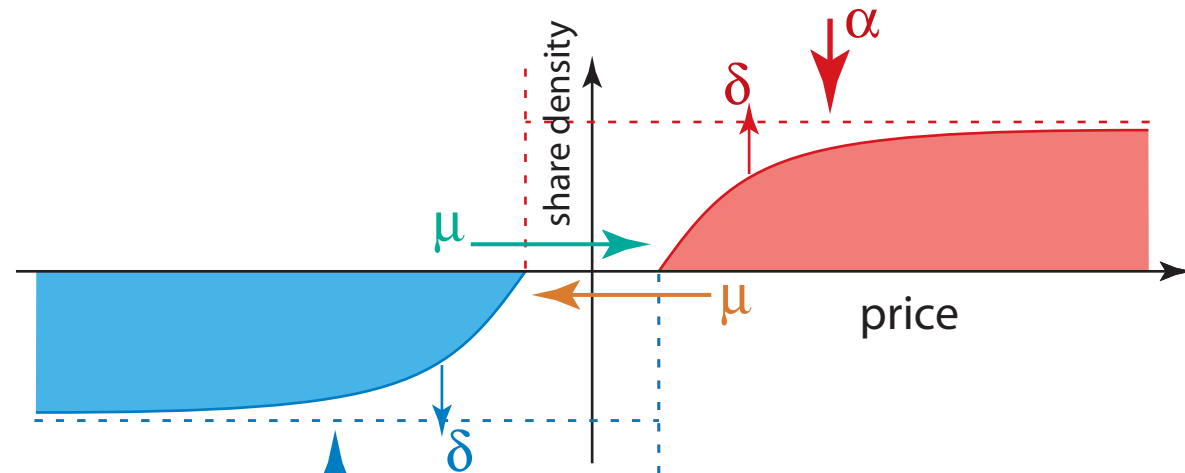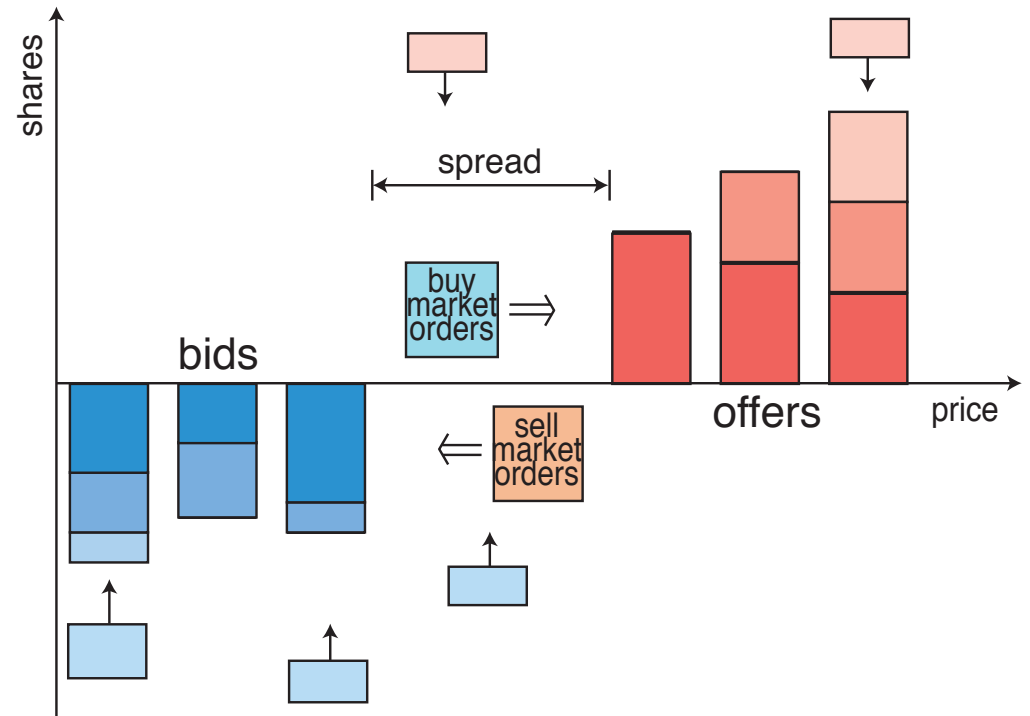
$$[\,\alpha\,] = \left[\frac{\text{shares}}{\text{price}\cdot\text{time}}\right] \quad [\,\mu\,] = \left[\frac{\text{shares}}{\text{time}}\right]$$

$$[\,\delta\,] = \left[\frac{1}{\text{time}}\right]$$

These are complete for shares, price, and time

Dimensional analysis of granularities

- Orders arrive and are removed in typical-sized chunks σ

- Prices are delimited in "ticks" Δp

- Both units are redundant with dimension provided by flow variables

- Redundancy creates the possibility of functions of non-dimensional variables
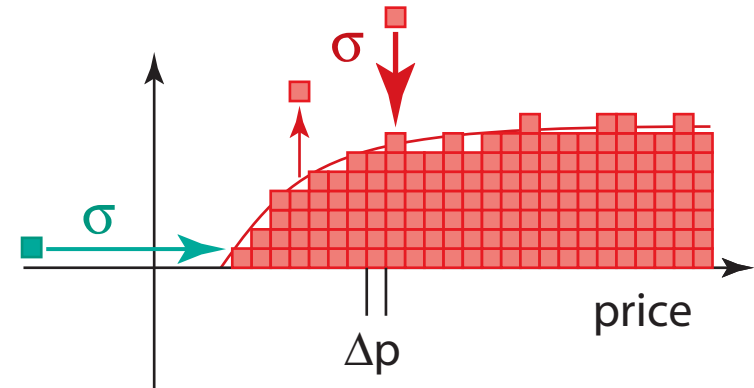
$$[\,\sigma\,]=[\mathrm{shares}] \qquad [\Delta p]=[\mathrm{price}]$$

Here I will take the flow variables to define the "classical" scaling dimensions, and treat the discreteness parameters as the source of nondimensional corrections

Classical scaling and dimensional analysis

- Flow variables define characteristic scales for fundamental properties

$$[\text{shares}]: \quad N_c \equiv \frac{\mu}{2\delta}$$

$$[\text{price}]: \quad p_c \equiv \frac{\mu}{2\alpha}$$

- Create "guesses" for observables based on their dimensions

$$[\text{depth}] = \left[\frac{\text{shares}}{\text{price}}\right] \leftrightarrow \frac{N_c}{p_c} = \frac{\alpha}{\delta}$$

$$[\text{spread}] = [\text{price}] \leftrightarrow 2p_c = \frac{\mu}{\alpha}$$

$$[\text{liquidity (slope)}] = \left[\frac{\text{shares}}{\text{price}^2}\right] \leftrightarrow \frac{\alpha^2}{\mu\delta} = \left[\frac{\text{depth}}{\text{spread}}\right]$$

$$[\text{midprice diffusion}] = \left[\frac{\text{price}^2}{\text{time}}\right] \leftrightarrow \frac{\mu^2\delta}{\alpha^2}$$

Scaling "collapse" of different instances

- Non-dimensionalized variables relate cases differing only by scale to a universal model

  - Price (coordinate)
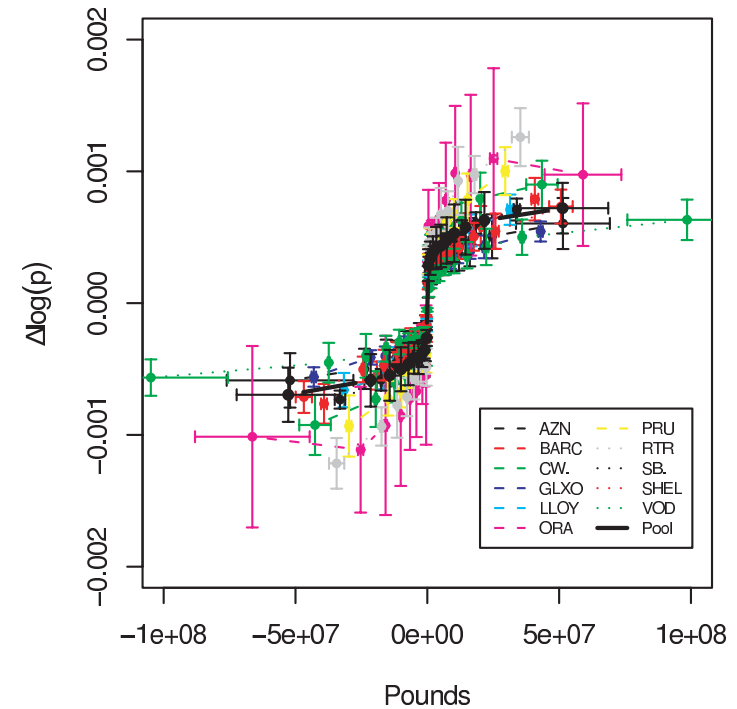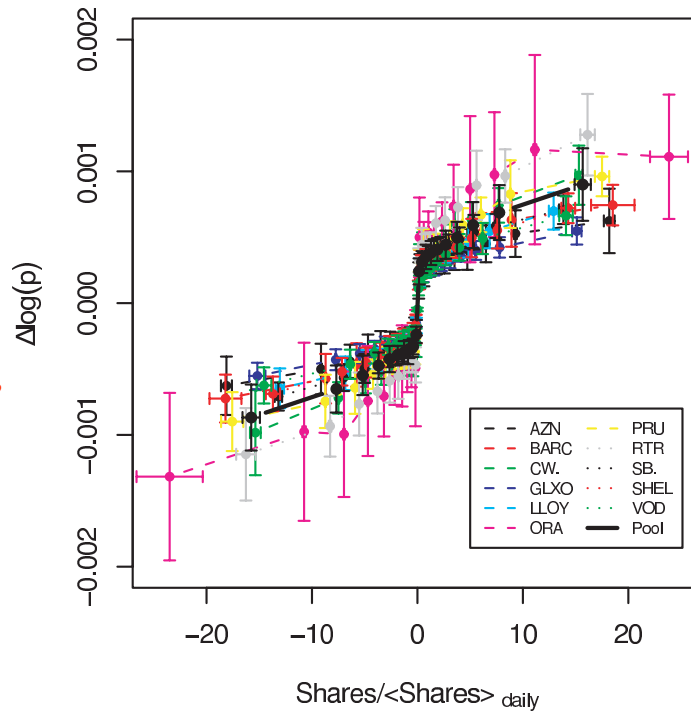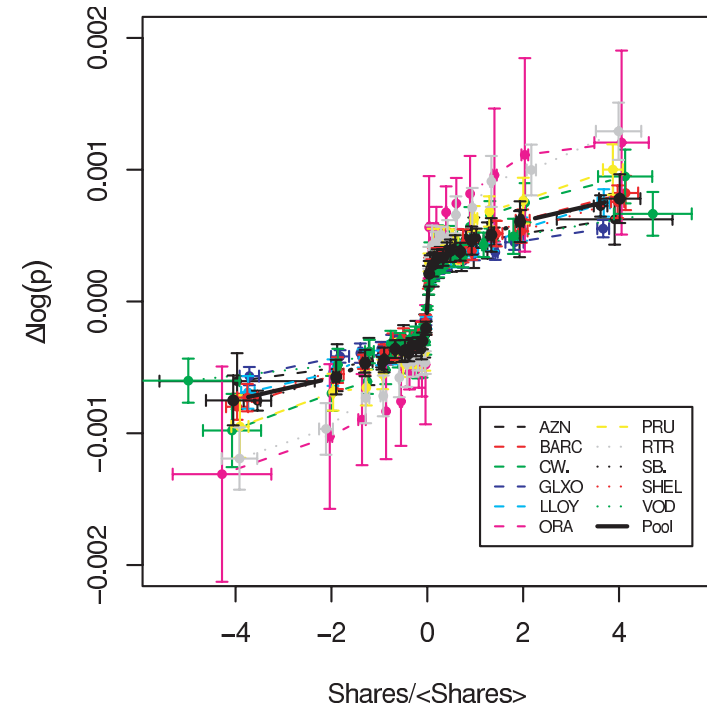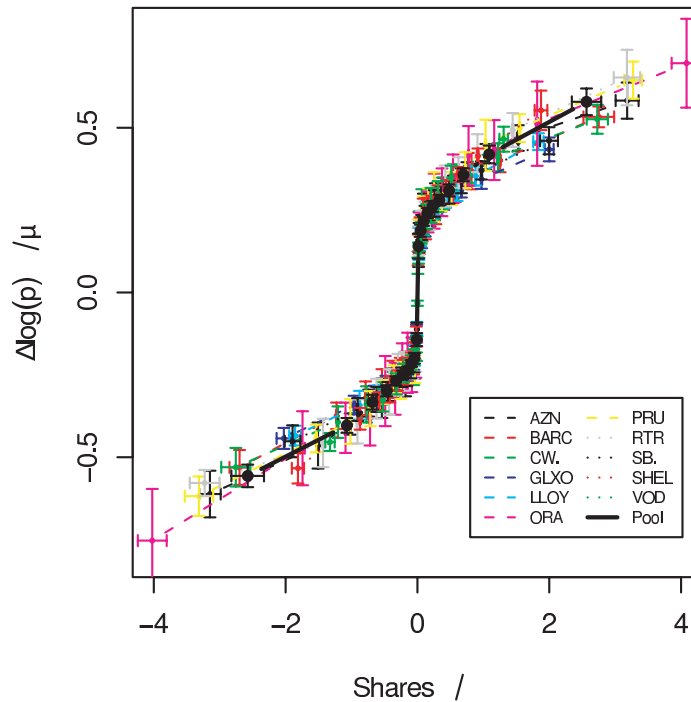 $$\hat{p} \equiv \frac{p}{p_c} = p\frac{2\alpha}{\mu}$$

  - Share depth (variable)
  $$\hat{n} \equiv n\frac{p_c}{N_c} = n\frac{\delta}{\alpha}$$

  - Spread (observable)
  $$\hat{s} \equiv \frac{s}{p_c} = s\frac{2\alpha}{\mu}$$

  - Diffusivity (observable)
  $$\hat{D} \equiv \frac{D}{\delta p_c^2} = D\frac{4\alpha^2}{\mu^2\delta}$$

How well does classical scaling work?

Try collapsing the market impact four different ways

Farmer, J. D., P. Patelli, and I. I. Zovko. "The Predicitive Power of Zero Intelligence in Financial Markets" *PNAS USA* 102(11) (2005): 2254-2259

Classical scaling predictions for the spread

- Prediction for the spread from the flow variables is

$$[\text{spread}] = [\text{price}] \leftrightarrow 2p_c = \frac{\mu}{\alpha}$$

- Broad selection of stocks from the London Stock Exchange agree with this prediction to overall scale

Farmer, J. D., P. Patelli, and I. I. Zovko. "The Predicitive Power of Zero Intelligence in Financial Markets" *PNAS USA* 102(11) (2005): 2254-2259
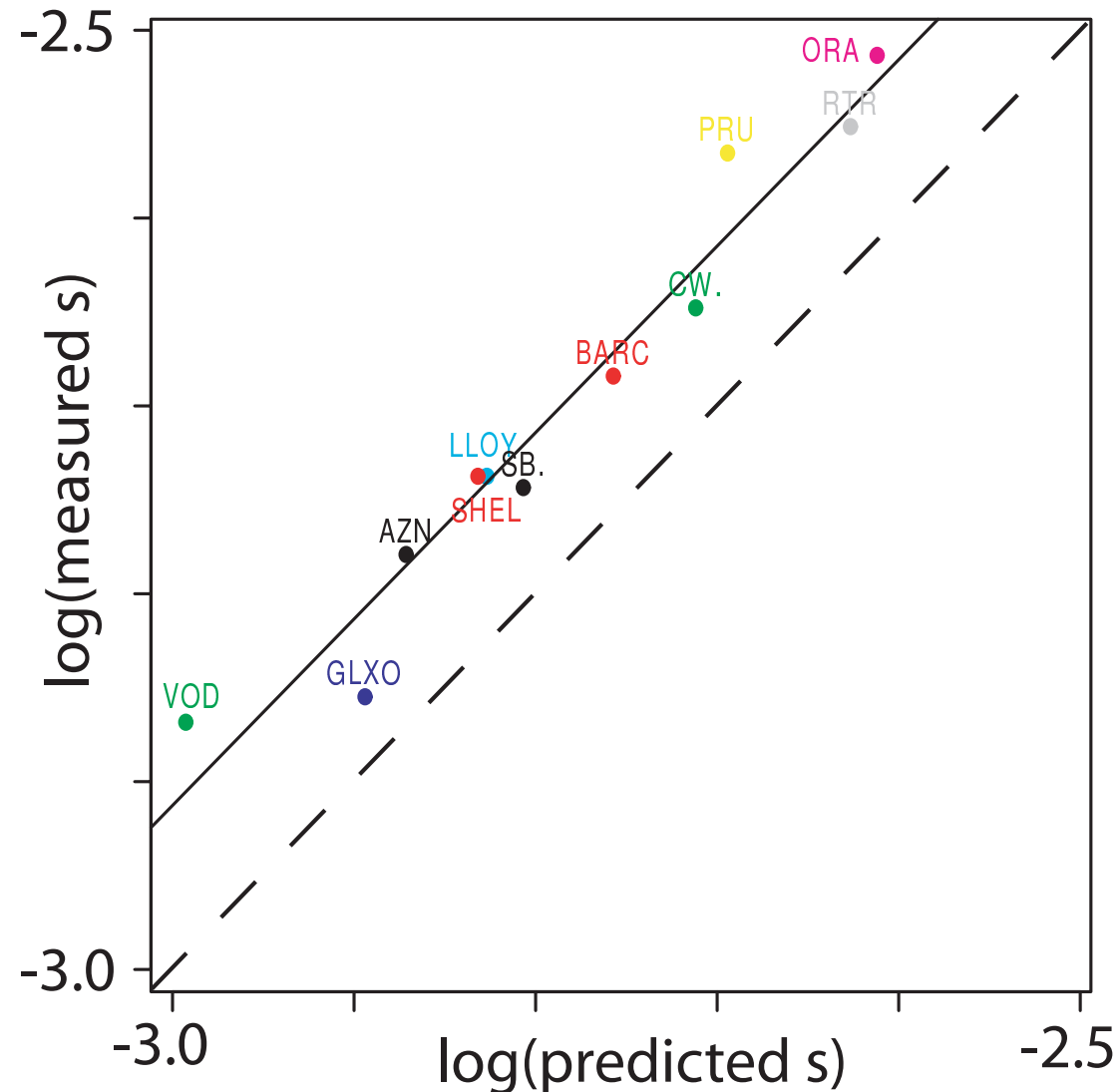
Discreteness parameters determine possible importance of dimensionless ratios

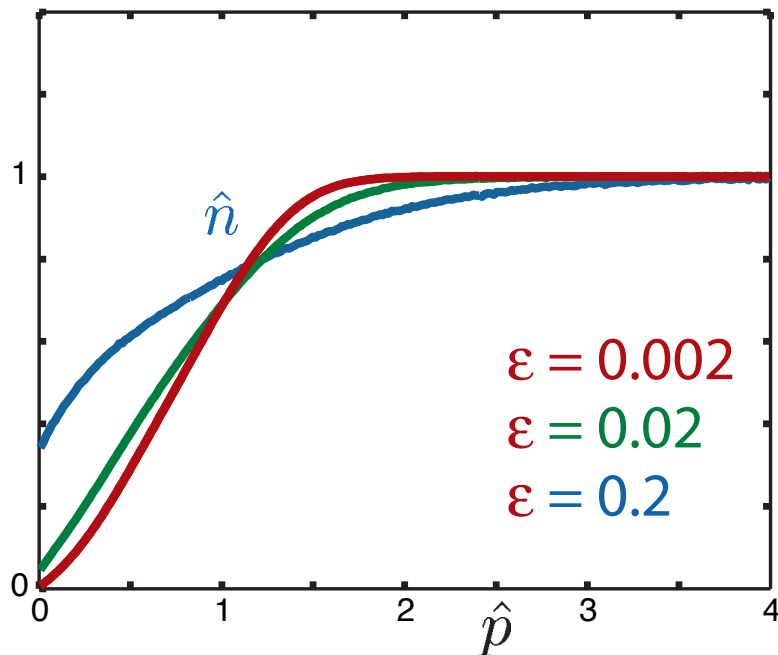- "Tick" size

- Order chunk size

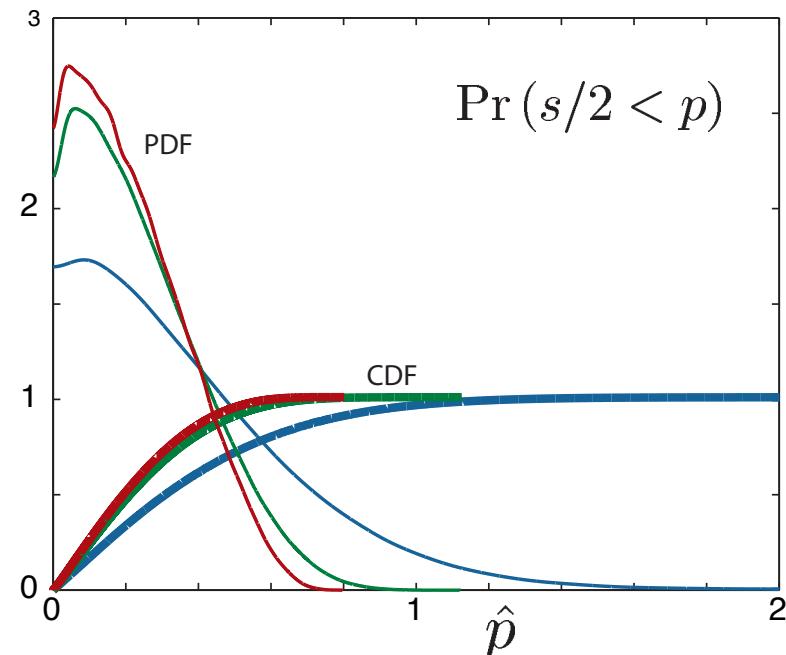$$\Delta \hat{p} \equiv \frac{\Delta p}{p_c} \to 0$$

$$\hat{\sigma} \equiv \epsilon \equiv \frac{\sigma}{N_c} = \sigma \frac{2\delta}{\mu}$$

Overall scale is respected, but the discreteness parameter matters

Depth profile versus ε

Nondimensional spread versus ε

$\hat{n}$

$\epsilon = 0.002$
$\epsilon = 0.02$
$\epsilon = 0.2$

$\hat{p}$

$\Pr(s/2 < p)$

PDF

CDF

$\hat{p}$

Correction of the spread from the classical guess

- Recall scale factor and non-dimensionalization of prices

- Classically we would expect non-dimensionalized spread to take value
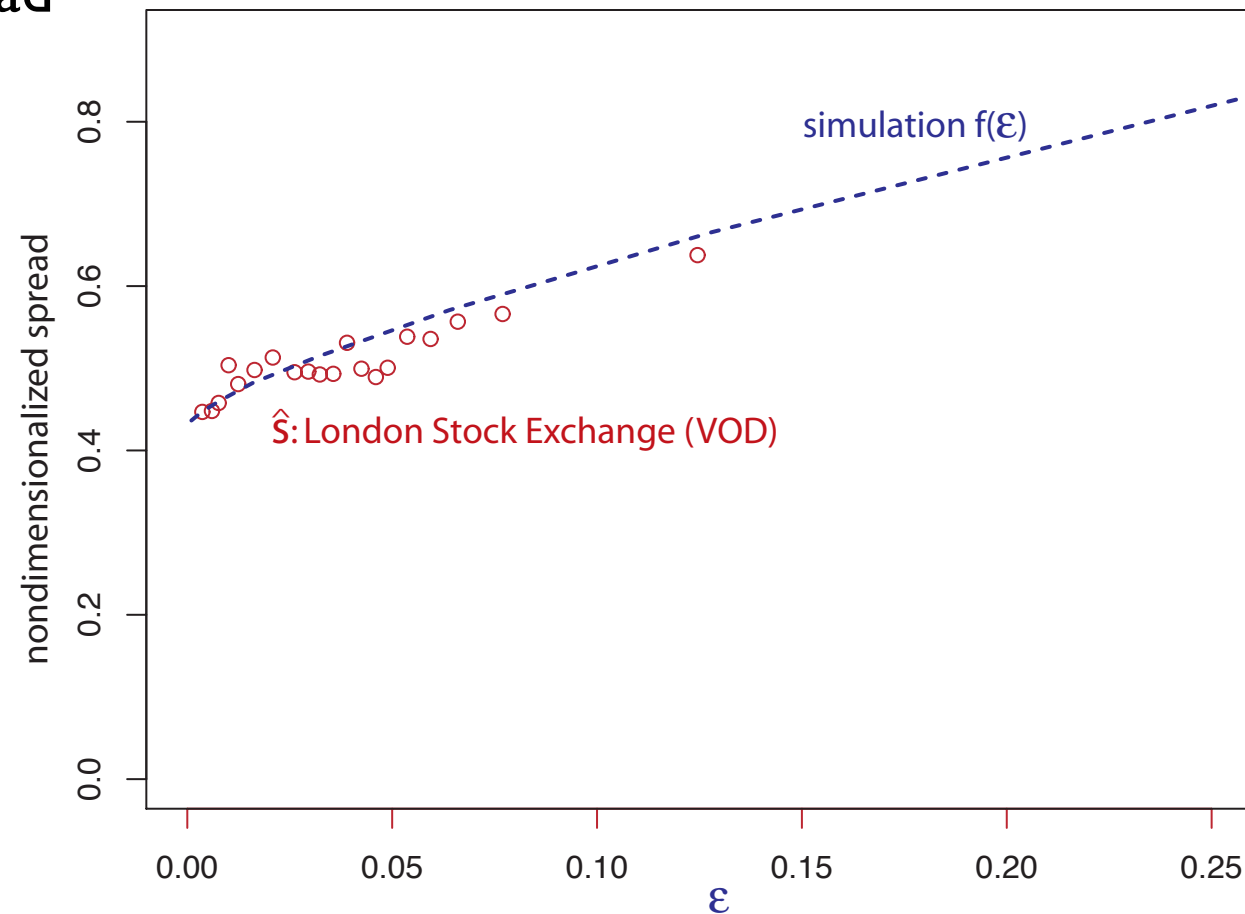
$$\hat{s} \rightarrow 1$$

- From simulations get small correction with epsilon

$$\hat{s} \rightarrow f(\varepsilon)$$

- Comparisons to data are remarkably good

$$[\text{spread}] = [\text{price}] \leftrightarrow 2p_c = \frac{\mu}{\alpha}$$

$$\hat{s} \equiv \frac{s}{p_c} = s\frac{2\alpha}{\mu}$$

simulation f(ε)

ŝ: London Stock Exchange (VOD)
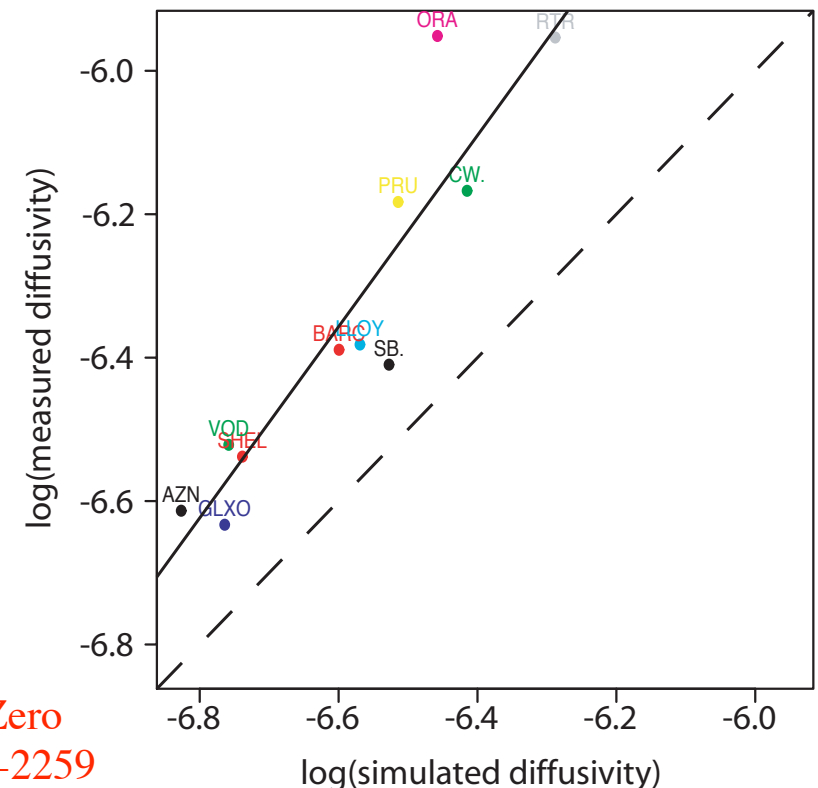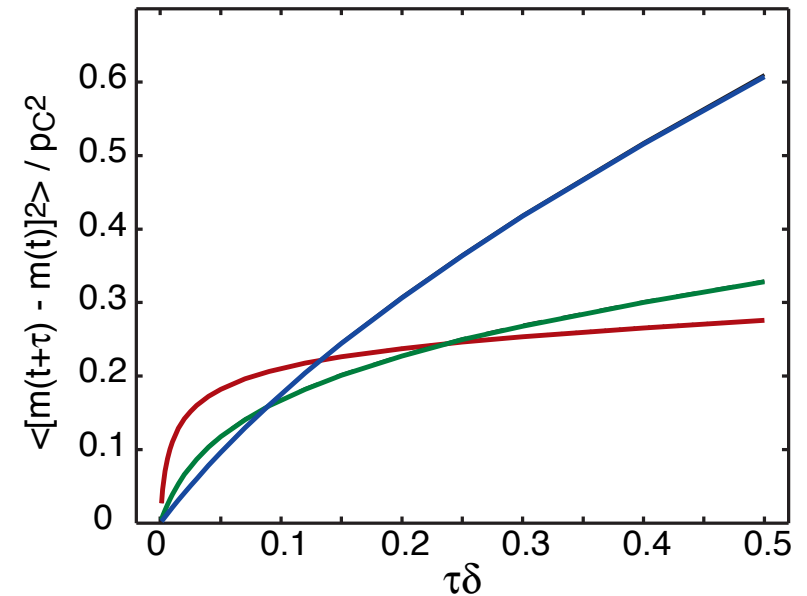
nondimensionalized spread

ε

The volatility

- Classical scaling of the diffusivity

$$\hat{D} \equiv \frac{D}{\delta p_c^2} = D \frac{4\alpha^2}{\mu^2 \delta}$$

- Simulations show different short-term and long-term corrections

$$\hat{D}_0 \sim \epsilon^{-1/2} \qquad \hat{D}_\infty \sim \epsilon^{1/2}$$

- Data scale with approximately with the short-term correction, with no apparent qualitative change

Farmer, J. D., P. Patelli, and I. I. Zovko. "The Predicitive Power of Zero Intelligence in Financial Markets" *PNAS USA* 102(11) (2005): 2254-2259

Learning from the limitations of ZI models

- Simplest ZI model leads to infinite range of limit prices and number of orders (clearly silly)

- Real order distributions are still very regular, but not by any institutional rule

- Candidate for a regularity of *behavior*
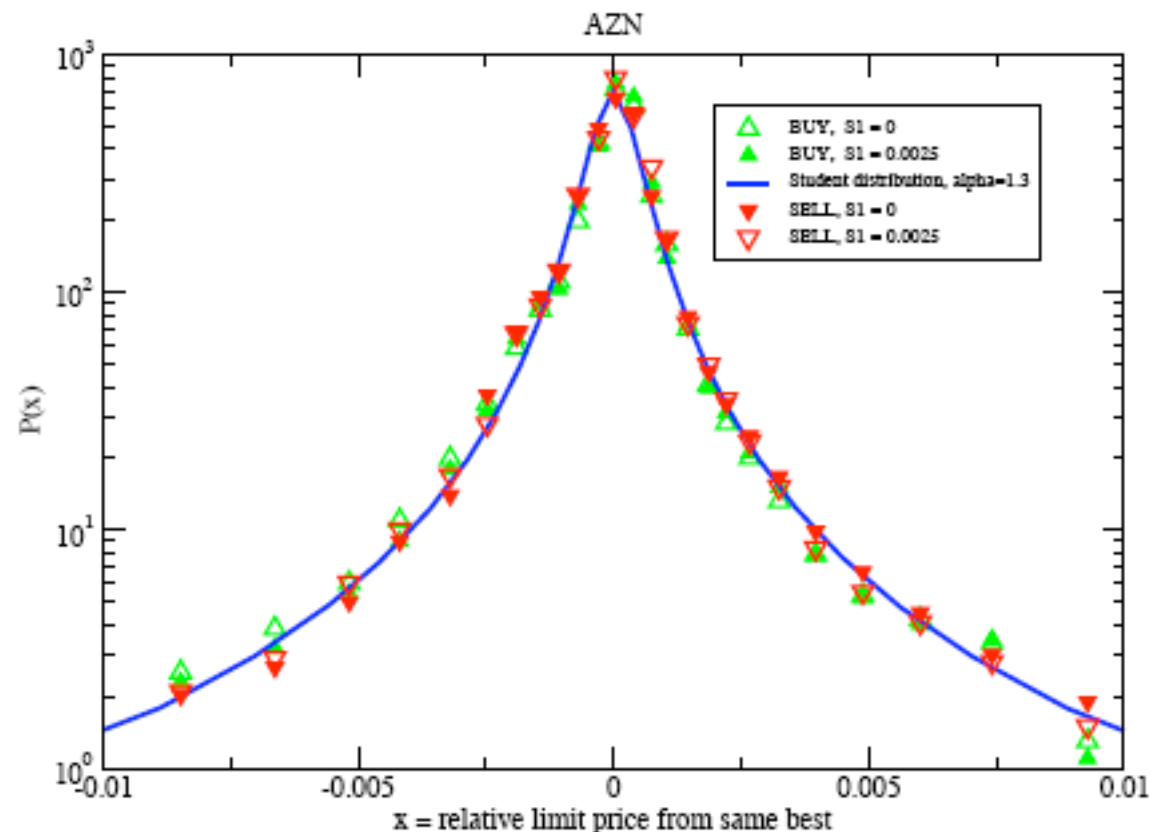
From Mike, S., and J. D. Farmer. "An Empirical Behavior Model of Price Formation." Santa Fe Institute Working Paper 05-10-039

V. Some compelling problems in income and wealth distribution

- An economically central issue

  - At the heart of questions of inequality and social welfare

  - Sets wealth from capital ownership apart from wealth from wage labor

  - Concerns whole-nation real productivity

  - Reflects relation of finance to the "real economy"

- The regularities are stunning and durable

- Nobody even knows whether they are institutional or behavioral

Income distribution in modern nations

- Two scaling regimes

- Income in "wage range" is lognormal or exponential

- Income of "wealthy" is power-law

- Both distributions are forms of maximum-entropy subject to different constraints

- Laws, customs, etc. all get folded into just four or five constants
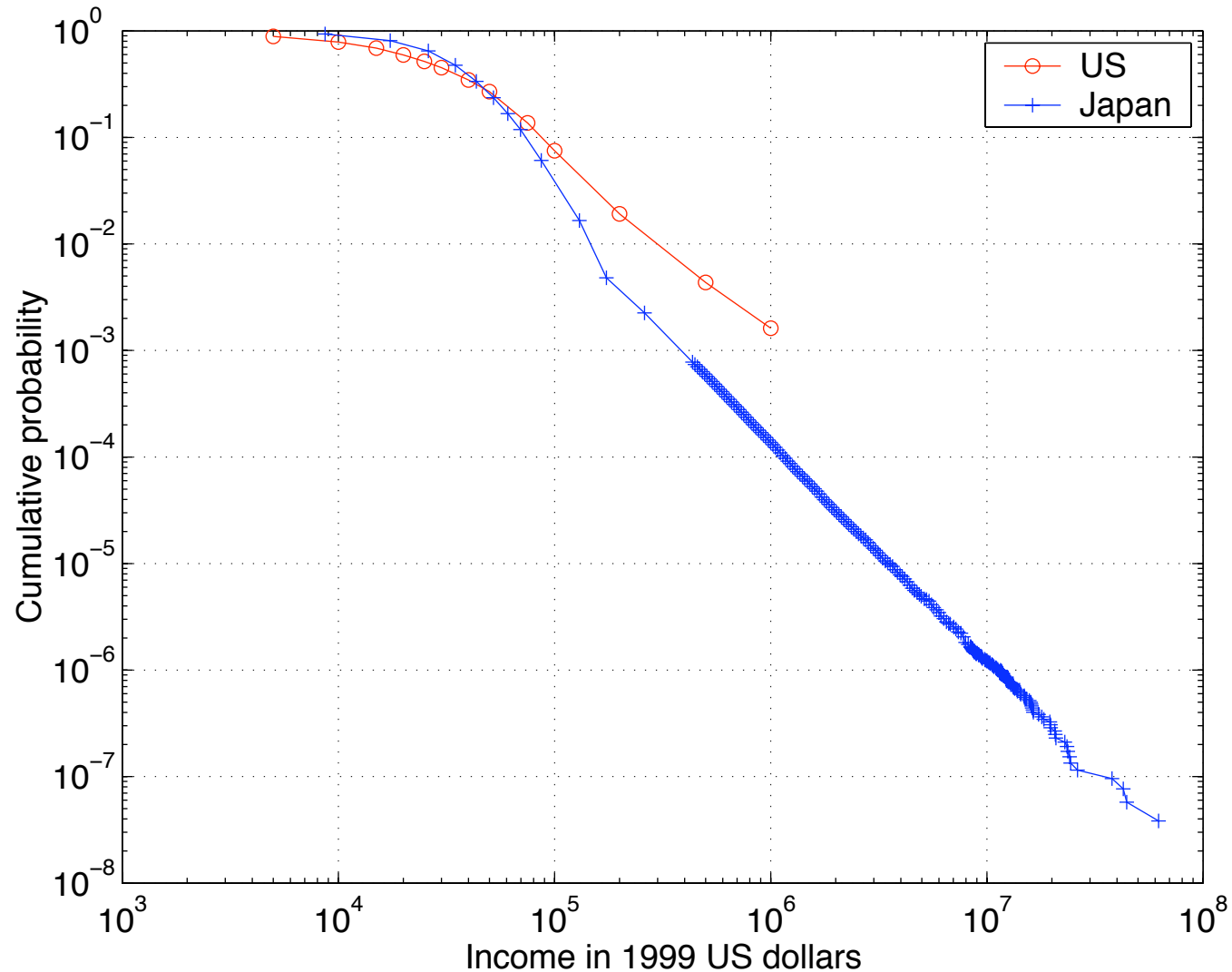
Courtesy Makoto Nirei and Wataru Souma

Concluding thoughts

- Theory is best constrained by empiricism

  - Institutions, behavior, the physical world, etc. are real and accessible to experiment

- Empiricism is best constrained by theory

  - Numbers don't become meaningful until you understand (something about) the measurement system that generates them

- Good methods tie falsifiable consequences to approximate descriptions

- Dimensional and scaling analysis and ZI have proved useful for some institutional process analysis

Some further reading

- Farmer, J. D., D. E. Smith, and M. Shubik. "Is Economics the Next Physical Science?" *Physics Today* 58 (9) (2005): 37-42.

- Farmer, J. D., P. Patelli, and I. I. Zovko. "The Predicitive Power of Zero Intelligence in Financial Markets" *PNAS USA* 102(11) (2005): 2254-2259.

- Lillo, F., S. Mike, and J. D. Farmer. "Theory for Long Memory in Supply and Demand." *Phys. Rev. E* 7106 (6 pt 2) (2005): 287-297.

- Lillo, F., and J. D. Farmer. "The Key Role of Liquidity Fluctuations in Determining Large Price Fluctuations." *Fluctuations and Noise Lett.* 5 (2005): L209-L216.

- Challet, D., Marsili, M., and Zhang, Y.-C. *Minority Games,* Oxford U. Press, New York (2005)

- Farmer, J. D., L. Gillemot, F. Lillo, S. Mike, and A. Sen. "What Really Causes Large Price Changes?" *Quant. Fin.* 4(4) (2004): 383-397.

- Smith, E., J. D. Farmer, L. Gillemot, and S. Krishnamurthy. "Statistical Theory of the Continuous Double Auction." *Quant. Fin.* 3(6) (2003): 481-514.

- Lillo, F., J. D. Farmer, and R. Mantegna. "Master curve for price impact function." *Nature* 421 (2003): 129.

- Bouchaud, J.-P., and Potters, M. *Theory of Financial Risk: From Statistical Physics to Risk Management*, Cambridge U. Press, New York (2000)