Transaction Cost Analysis for Futures Trading

Robert Almgren

Oxford Frontiers in Quantitative Finance
Jan 2020

qb quantitative brokers

QB = trade execution

Execute order as agent for client
Goal: best final average execution price
What is a good price?
Evaluate relative to benchmark
- benchmark defines an "ideal" trade
- different benchmarks give different strategies

Slippage

Difference of final average execution price and benchmark

- execution - benchmark for buys
- benchmark - execution for sells

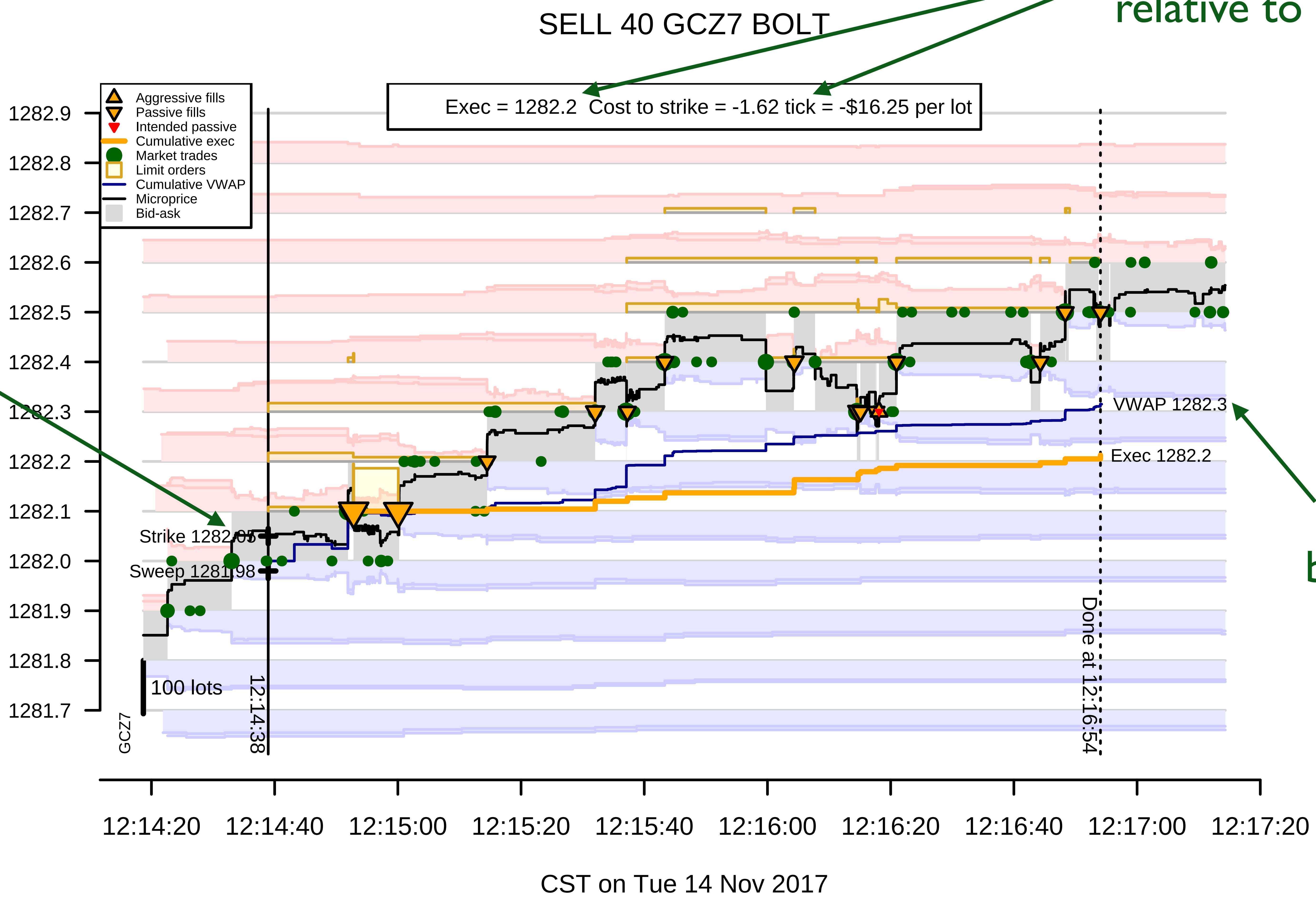
Positive slippage is bad, negative is good

For agency execution, minimize this

Bolt: arrival price

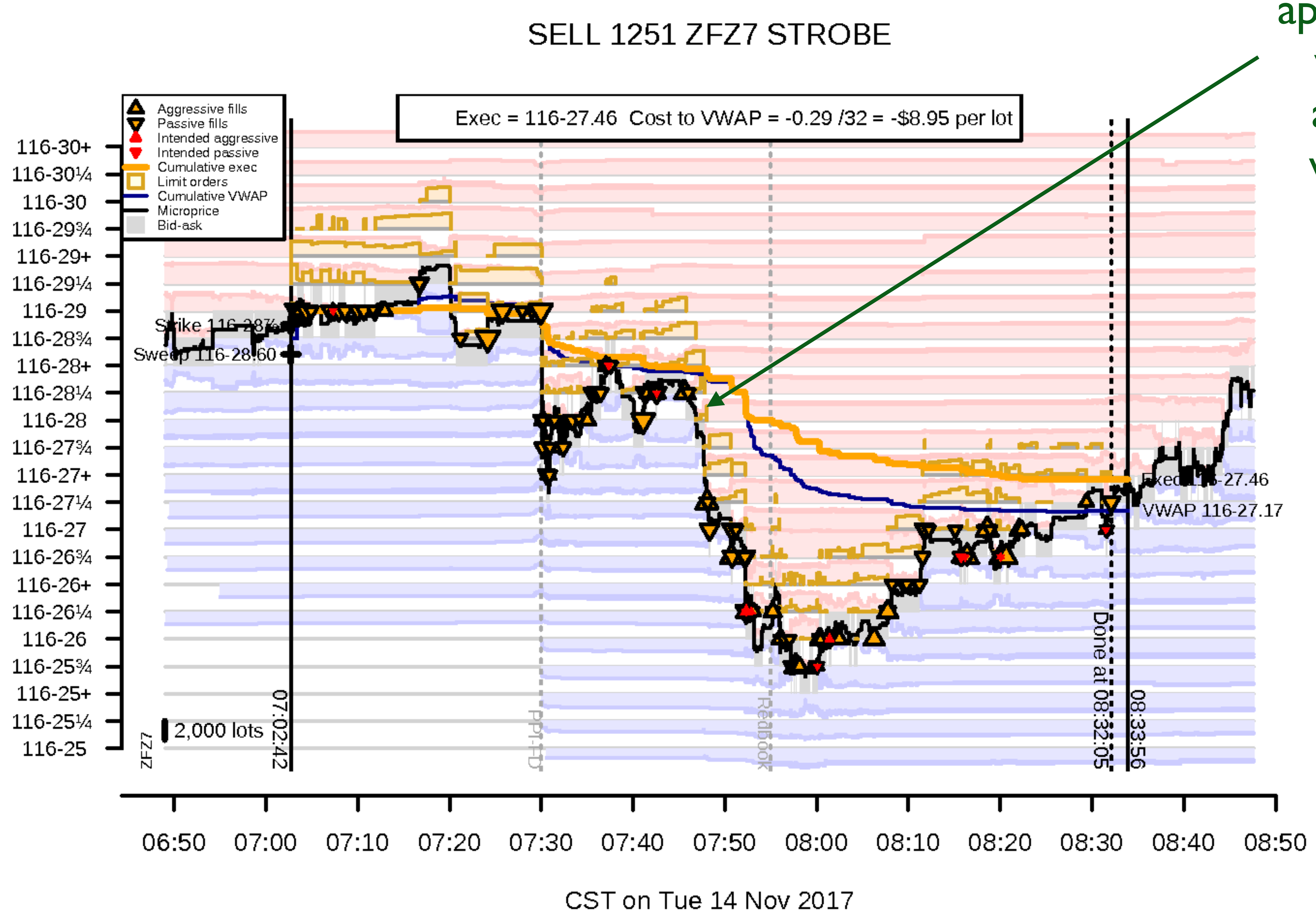
Report execution price and slippage relative to benchmark

Arrival price benchmark ("strike")



Also report other benchmarks for interest (but these are not targeted by this algo)

Strobe: average price on interval



For Strobe, execution approximately follows volume curve, but also opportunistic when can improve performance

We focus on Bolt and arrival price

Arrival price is most fundamental
- represents trade completed immediately at decision price

Arrival price is cleanest benchmark
- reference point is in past, not affected by trading

Arrival price is most challenging to model
- market direction is biggest contributor
- lots of statistical noise
- market impact and alpha are inextricable

Example time-dependent model

Equity market impact

The impact of large trades on prices is very important and widely discussed, but rarely measured. Using a large data set from a major bank and a simple but realistic theoretical model, Robert Almgren, Chee Thum, Emmanuel Hauptmann and Hong Li propose that impact is a 3/5 power law of block size, with specific dependence on trade duration, daily volume, volatility and shares outstanding. The results can be directly incorporated into an optimal trade scheduling algorithm and pre- and post-trade cost estimation

A. Distinguishing features of our model

Advantages
- Calibrated from real data
- Includes time component
- Incorporates intra-day profiles
- Uses non-linear impact functions
- Confidence levels for coefficients

Disadvantages
- Based only on Citigroup data
- Little data for small-cap stocks
- Little data for very large trades

WWW.RISK.NET • JULY 2005 RISK

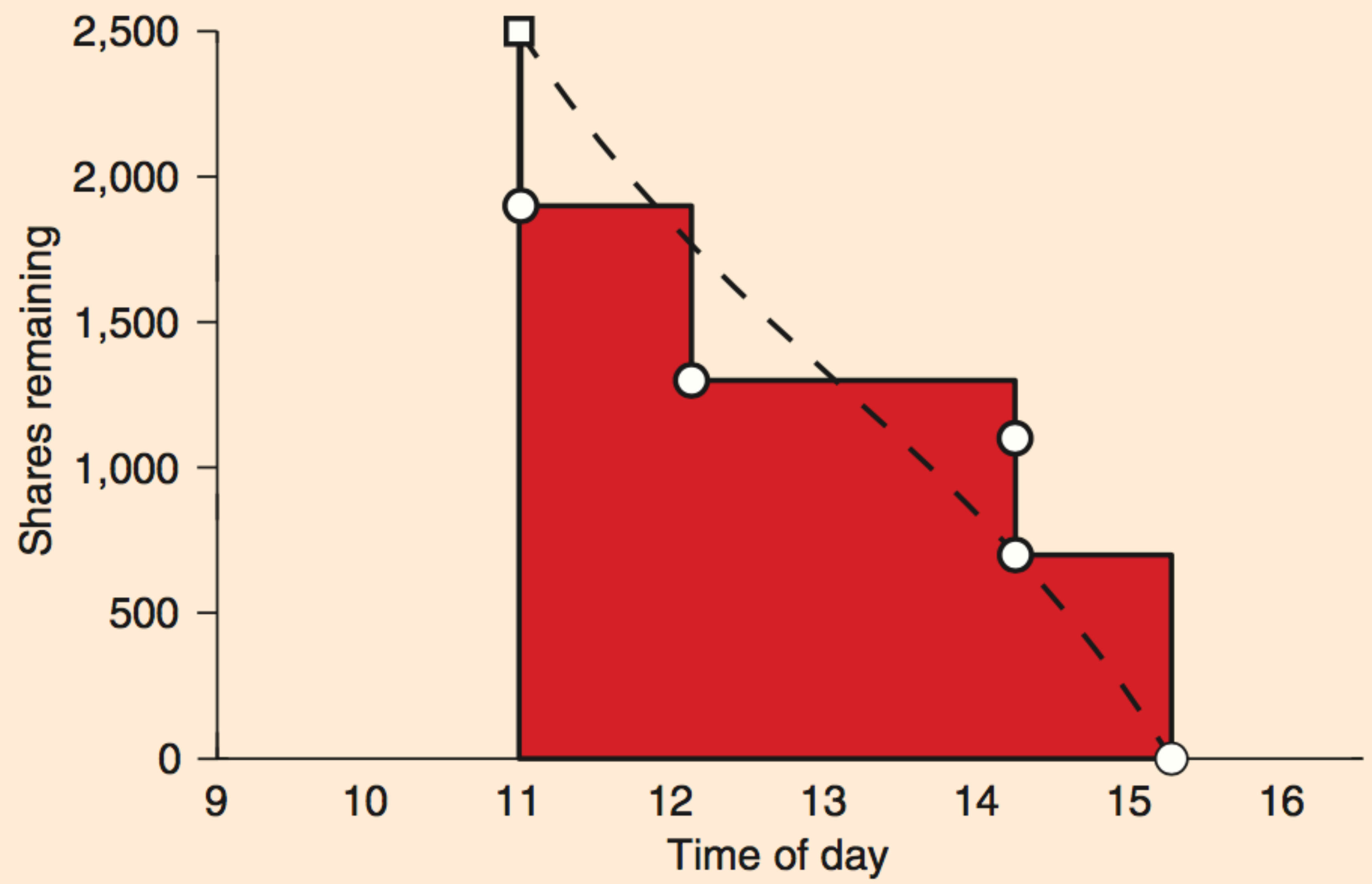
Available data is typical of institutional trade records

- The stock symbol, requested order size (number of shares) and sign (buy or sell) of the entire order. Client identification is removed.
- The times and methods by which transactions were submitted by the Citigroup trader to the market. We take the time t_0 of the first transaction to be the start of the order. Some of these transactions are sent as market orders, some are sent as limit orders, and some are submitted to Citigroup's automated VWAP server. Except for the starting time t_0, and except to exclude VWAP orders, we make no use of this transaction information.
- The times, sizes, and prices of execution corresponding to each transaction. Some transactions are cancelled or only partially executed; we use only the completed price and size. We denote execution times by $t_1, \ldots, t_n$, sizes by $x_1, \ldots, x_n$, and prices by $S_1, \ldots, S_n$.

December 2001 through June 2003.

29,509 orders in our data set.

1. A typical trading trajectory



The vertical axis is shares remaining and each step downwards is one execution. The trajectory starts at the first transaction recorded in the system; the program ends when the last execution has been completed. The dashed line is our continuous-time approximation

Subtlety: Impact vs alpha

You anticipate price increase

↓

You enter a buy order to profit from increase

↓

Price goes up

Impact or alpha?

No action in market is independent
of what came before, nor
of what is expected to come after.

Impossible to separate these two: take an empirical point of view and only summarize combined result

Goal of market impact modeling

Predict

- Slippage for a particular contemplated order
  - buy 500 10-year Treasury over the next 2 hours
- Dependence on variables that can be adjusted
  - what happens if we trade 200 or 1000, or take 3 hours?

Uses

- Trade decision-making
  - how should we choose execution parameters?
- Post-trade analysis
  - what products / brokers / traders were good or bad relative to model?

Data resources

Large variety of orders executed in past

- thousands per day
- many different products and market conditions

'What happened the last time that we did something "like" this?'

Classic problem in statistics / machine learning

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longrightarrow y$$

One output variable: slippage

Many input variables:

- Order parameters:
  - symbol, side, size, start time, duration, etc
- Market parameters known before trading:
  - forecast volume, volatility, spread, quote size
  - real-time volume, volatility, spread, quote size
- Market parameters discovered during trading:
  - price direction (most important)
  - evolution of volume, volatility, etc

Analytical techniques

$$y = f(x_1, \ldots, x_n)$$

$$y = \alpha + \beta_1 x_1 + \cdots + \beta_n x_n$$

- Regression
  - specific function depending on parameters
  - easy interpretation, not always accurate
- Supervised learning
  - many powerful modern techniques
    - neural nets, trees, support vector machines, etc
  - not always easy to interpret

Challenge in market impact modeling

Very large noise relative to signal
Criterion of minimum discrepancy is hard to apply
Main criterion: residuals, etc, should not depend other variables

Strategy

1. Single asset fitting
2. Multi-asset fitting across all universe of futures products

Example:

ES (SP500 futures) trades from Sep 2018 through Dec 2019

- all clients merged together (except private data)

Outright contracts only, mostly front month

Exclude orders with limit price

Exclude clients who cancel more than 5% of orders

For "market impact", most important variable is size

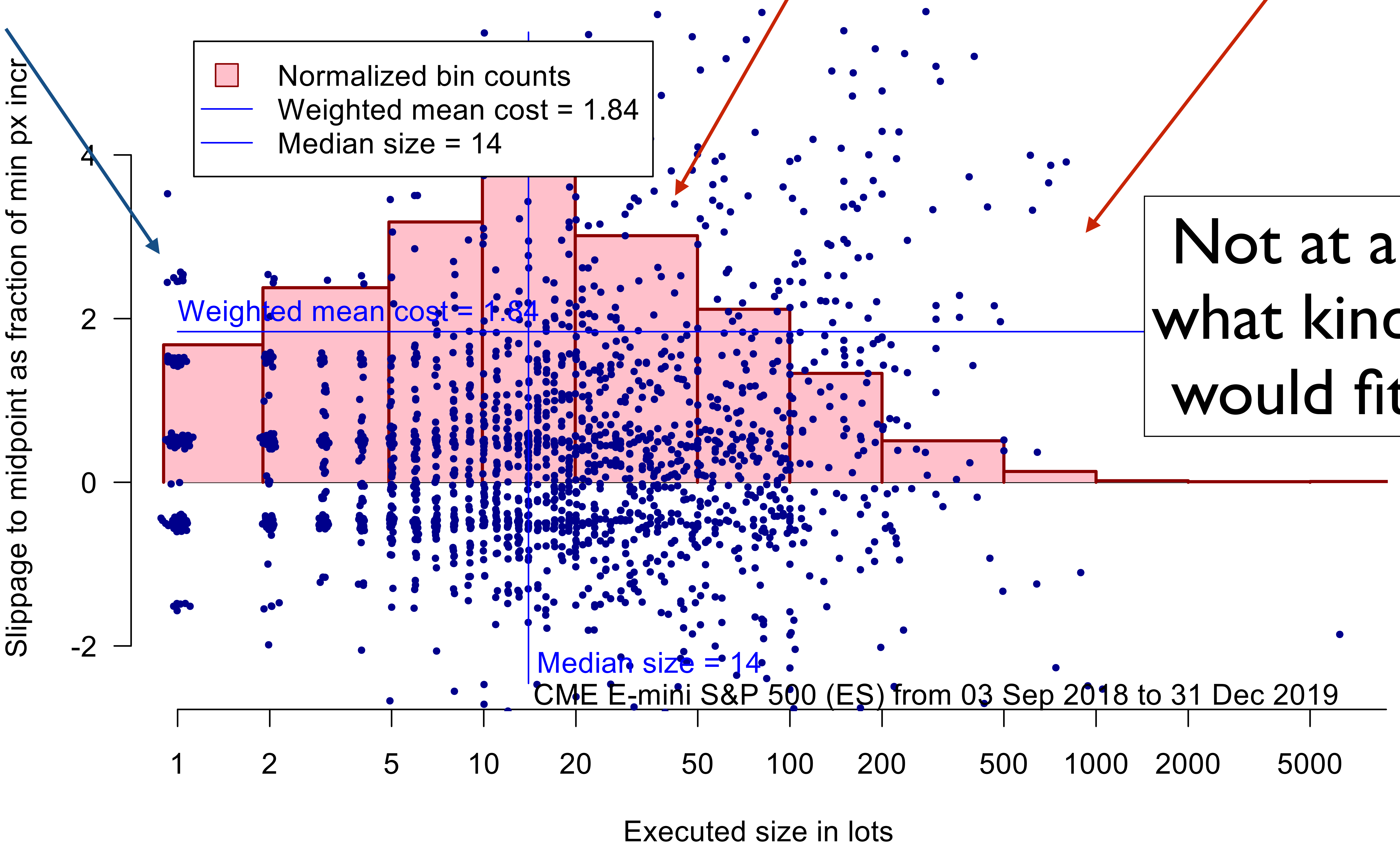
- hypothesis: large trades have higher slippage

Raw data

Slight "jitter" to show overlapping points

Majority of orders are small

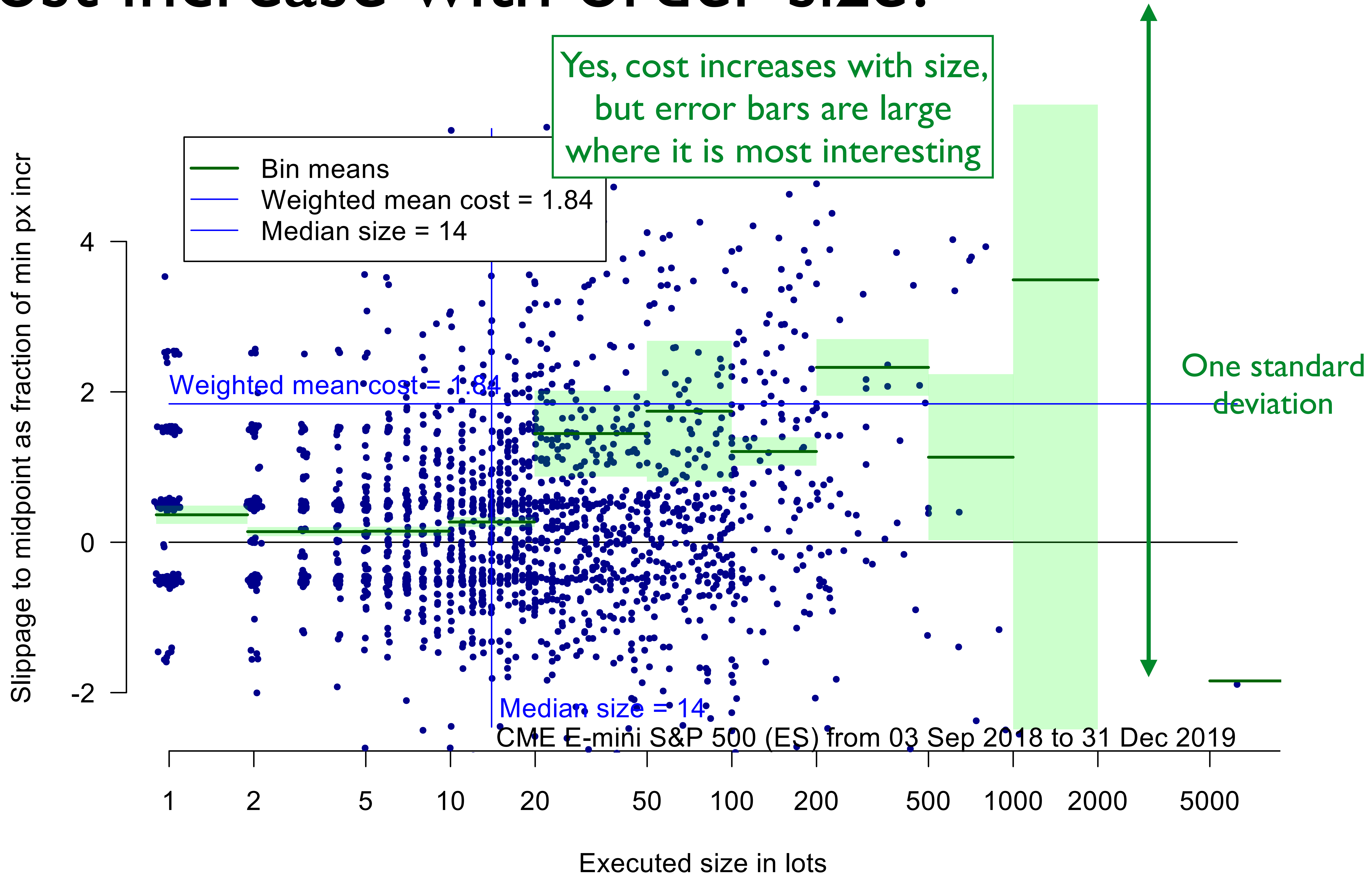
But cost is experienced on these few large orders



Not at all obvious what kind of model would fit this data

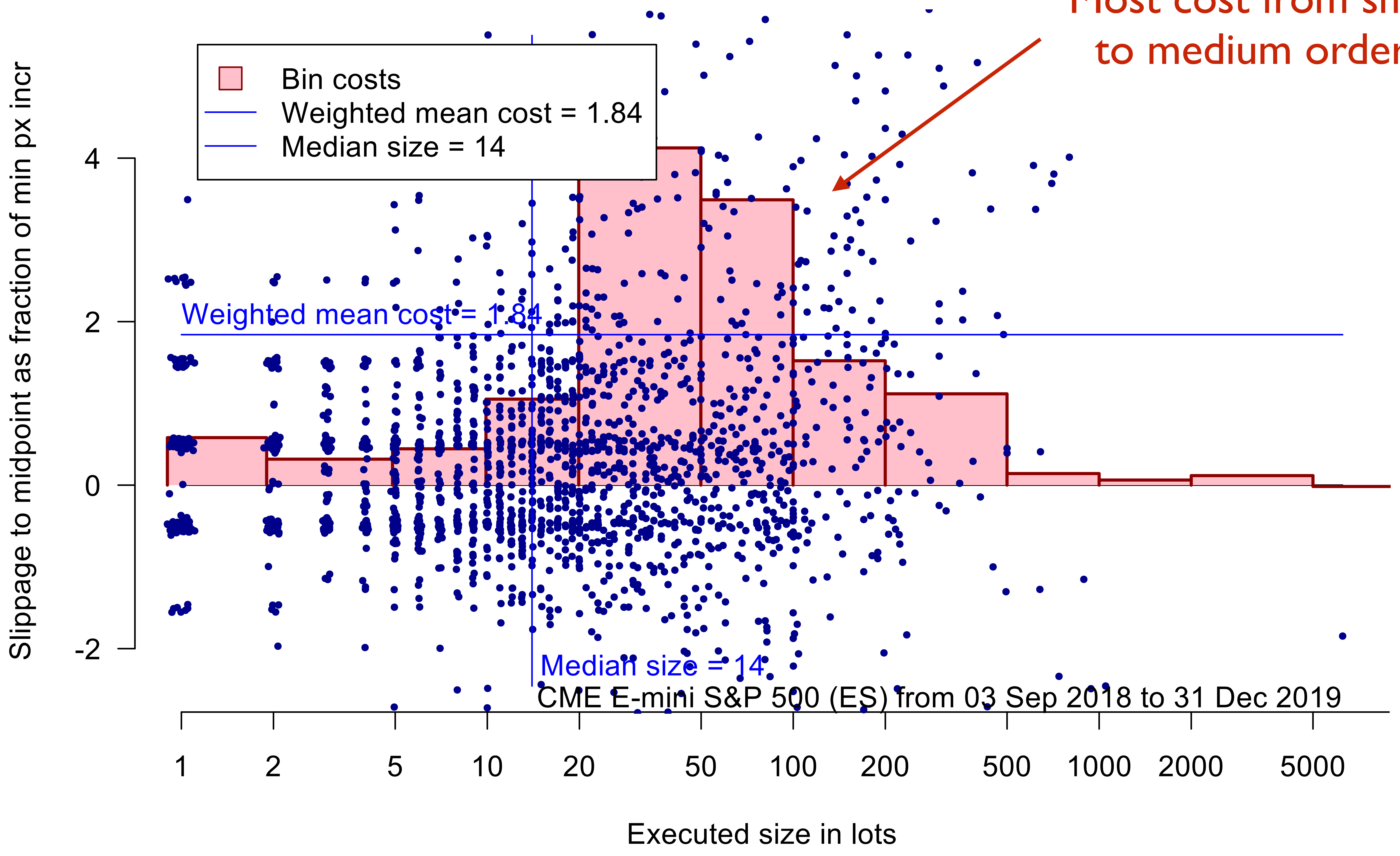
Choice of size as unique independent variable (for now)

Does cost increase with order size?



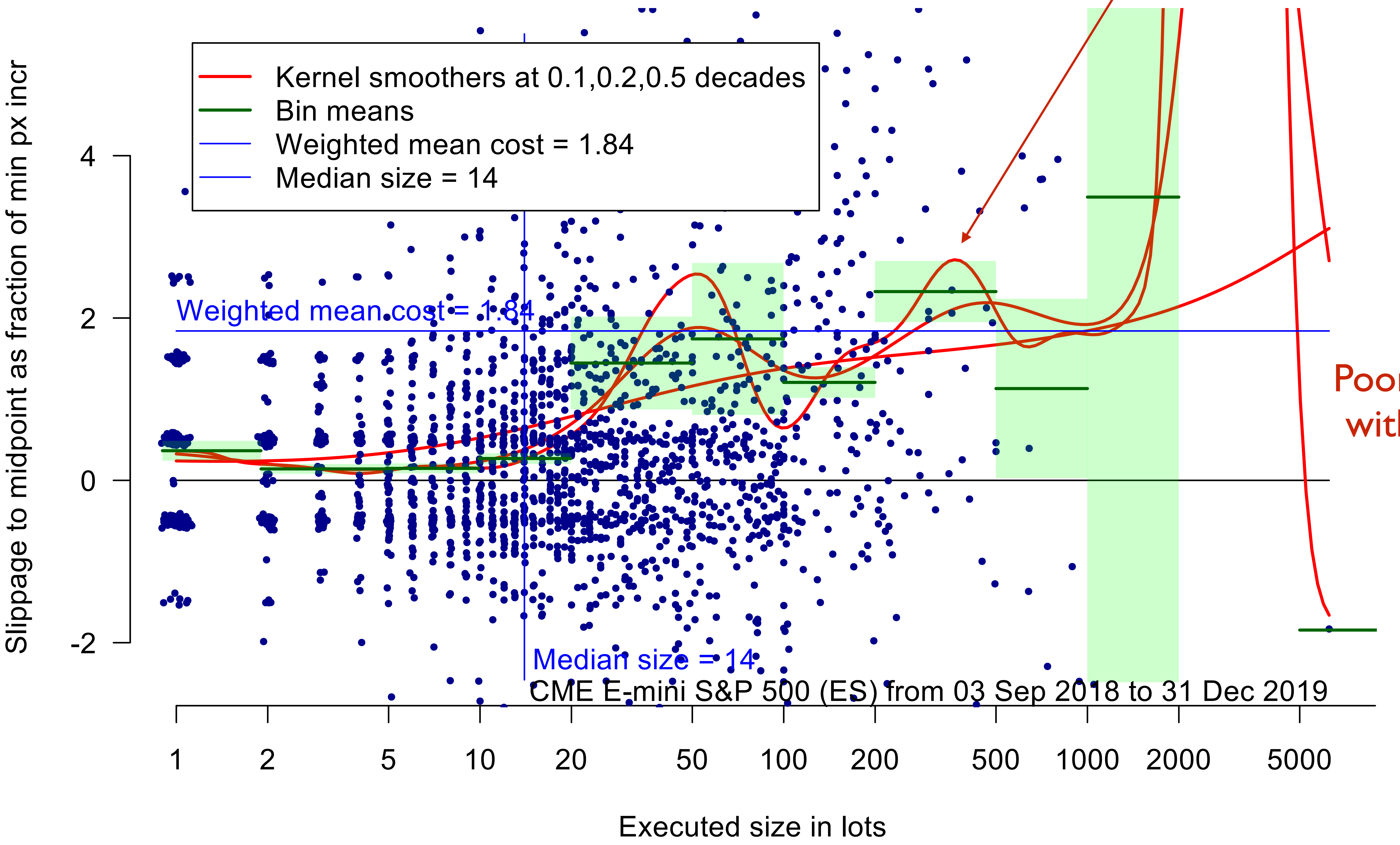
Where does the cost come from?

Most cost from small to medium orders



Kernel estimator

Roughly agrees with bin averages



Poor behavior with outliers

Parametric model

$$y = a + b\, x^{\gamma}$$

$$\min_{a,b,\gamma} \sum_{j=1}^{n} \left(a + b x_j^{\gamma} - y_j\right)^2$$

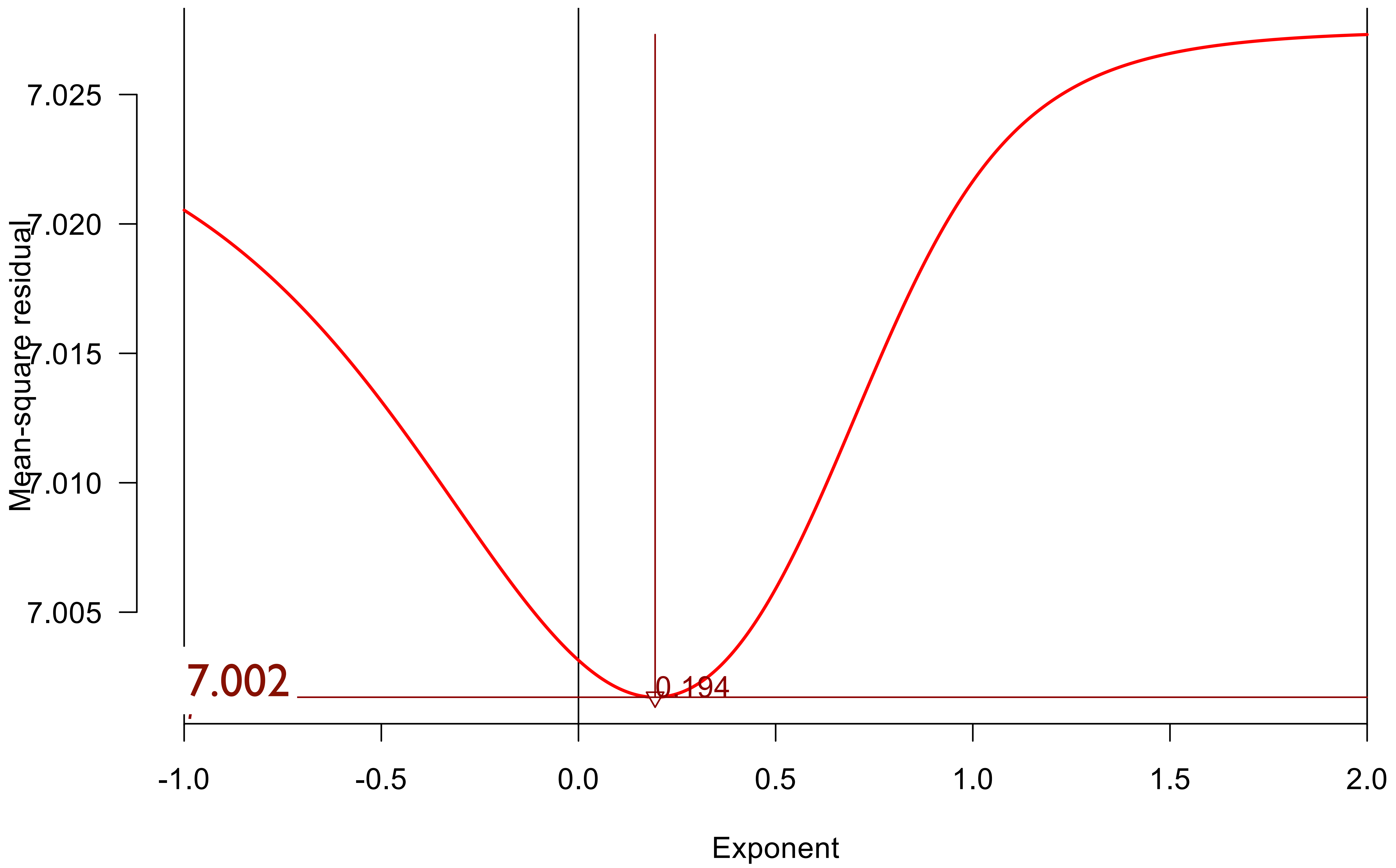
$$x = \text{order size}$$
$$y = \text{slippage}$$

Two linear coefficients a,b
One exponent γ

Coefficients a, b determined linearly for each exponent γ

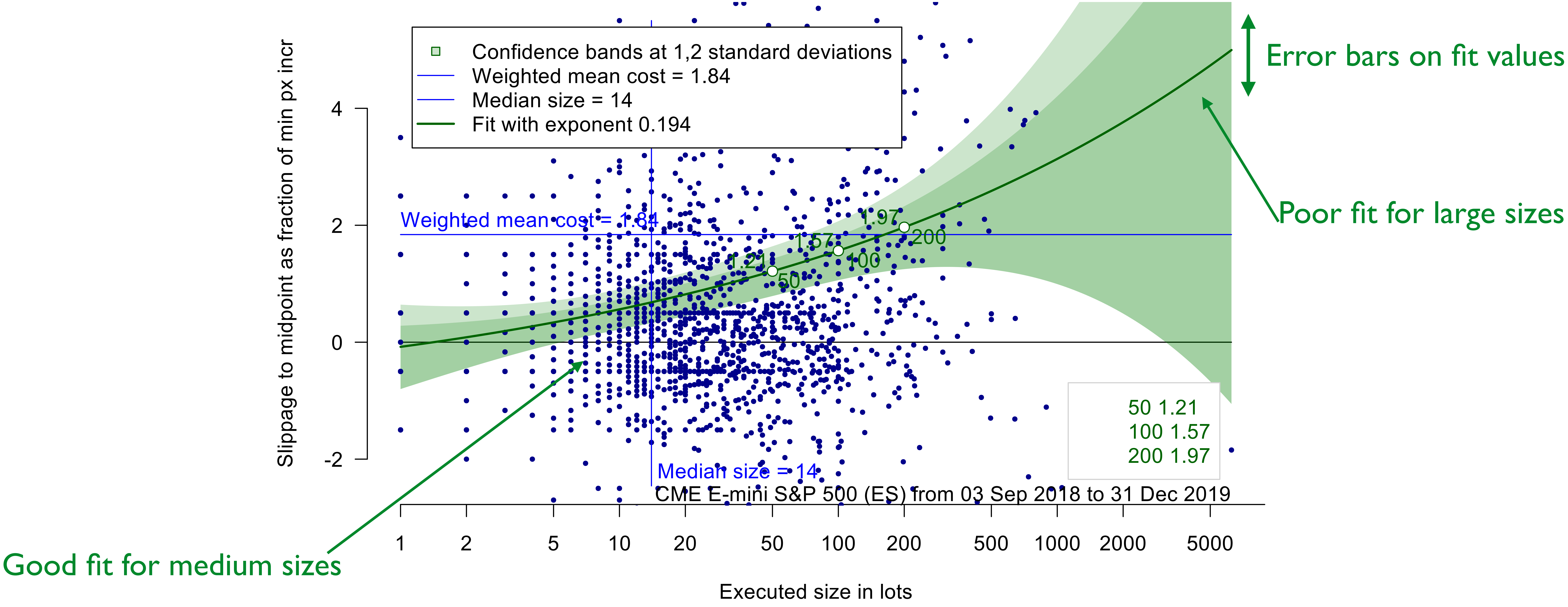
γ determined by one-dimensional minimization (easy!)

One-dimensional search



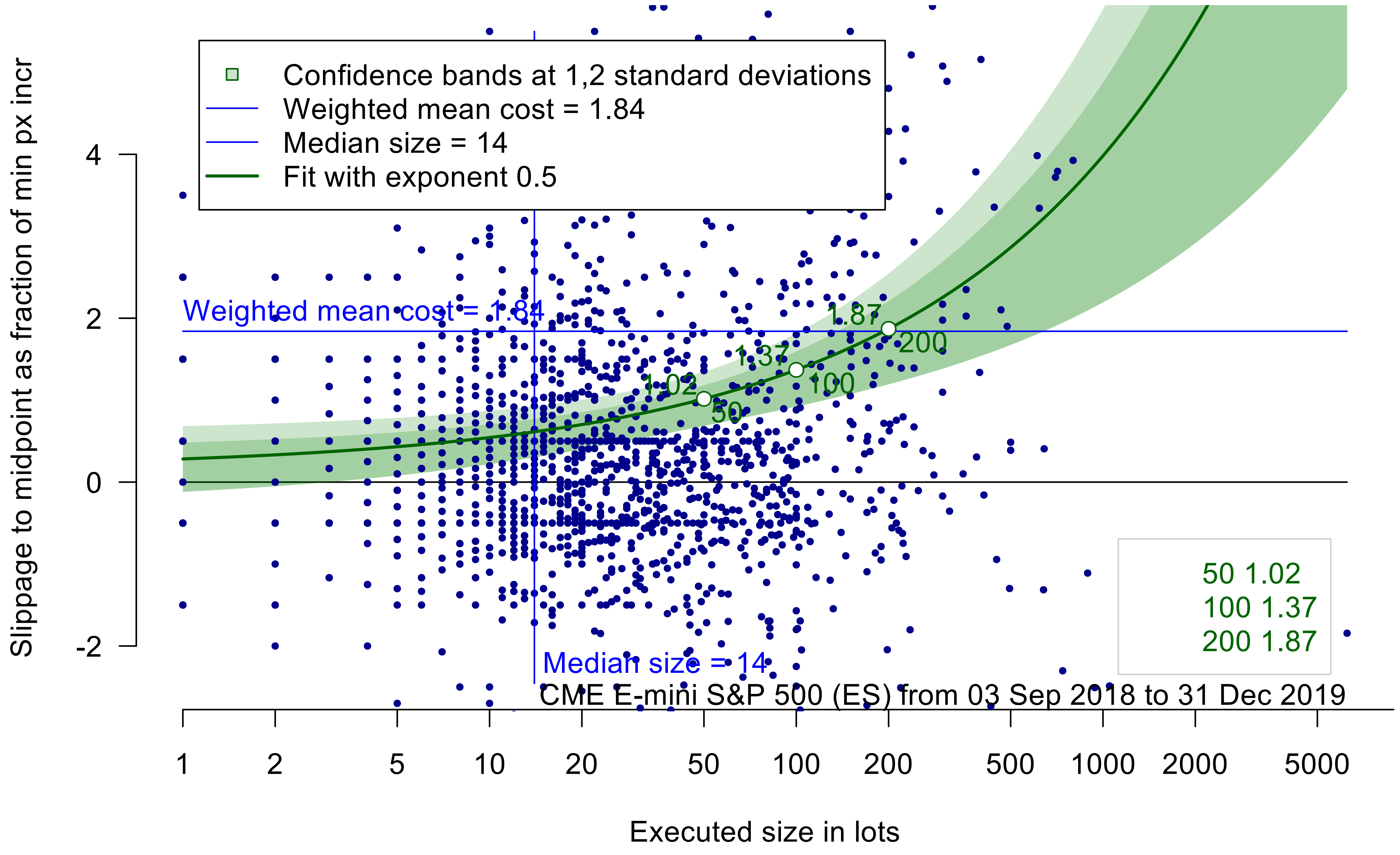
0.3% reduction
from worst
residual
to best

Fit curve: exponent = 0.812

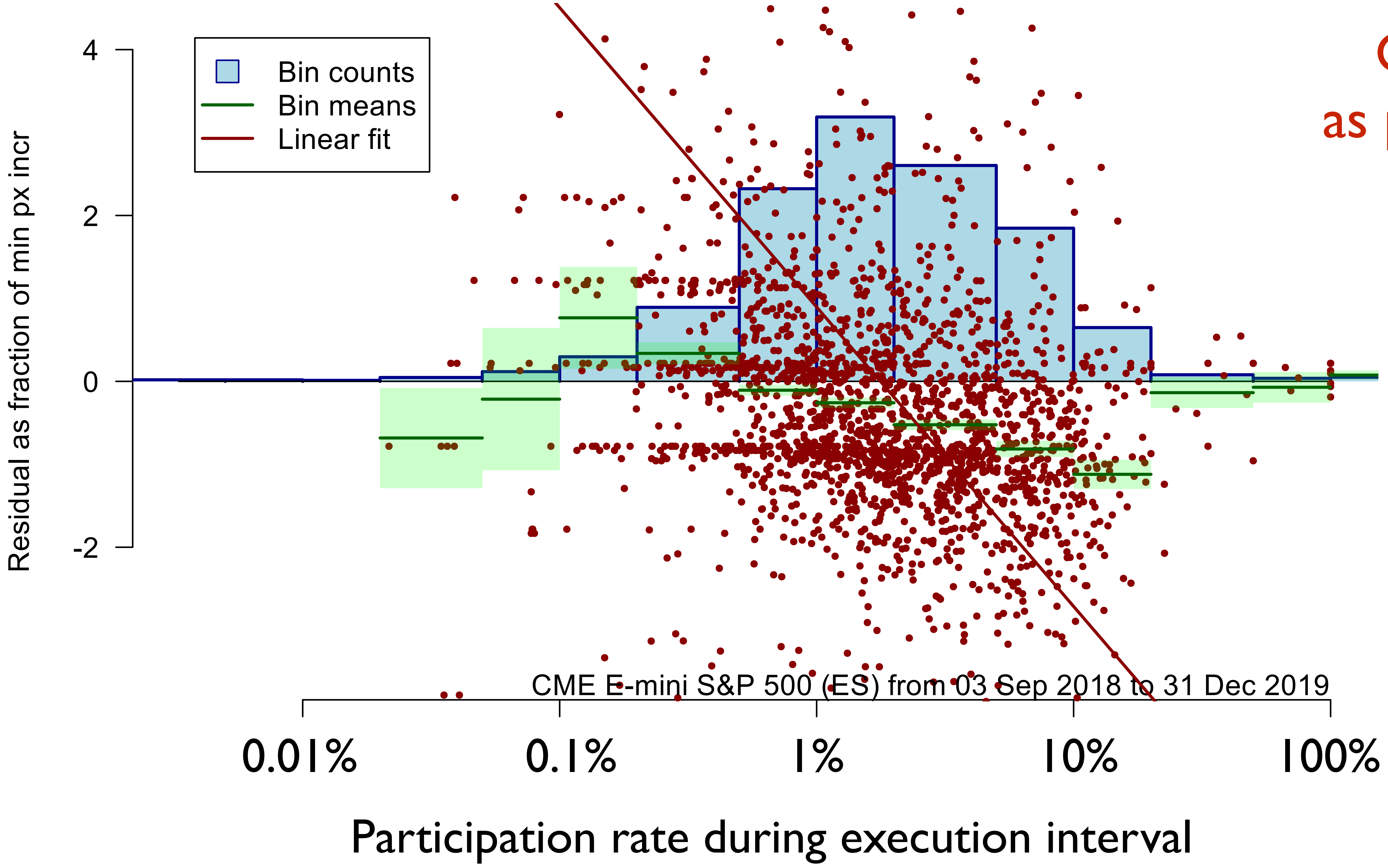


What should the exponent be?

Strong reasons to prefer k = 0.5



Residual as function of participation rate

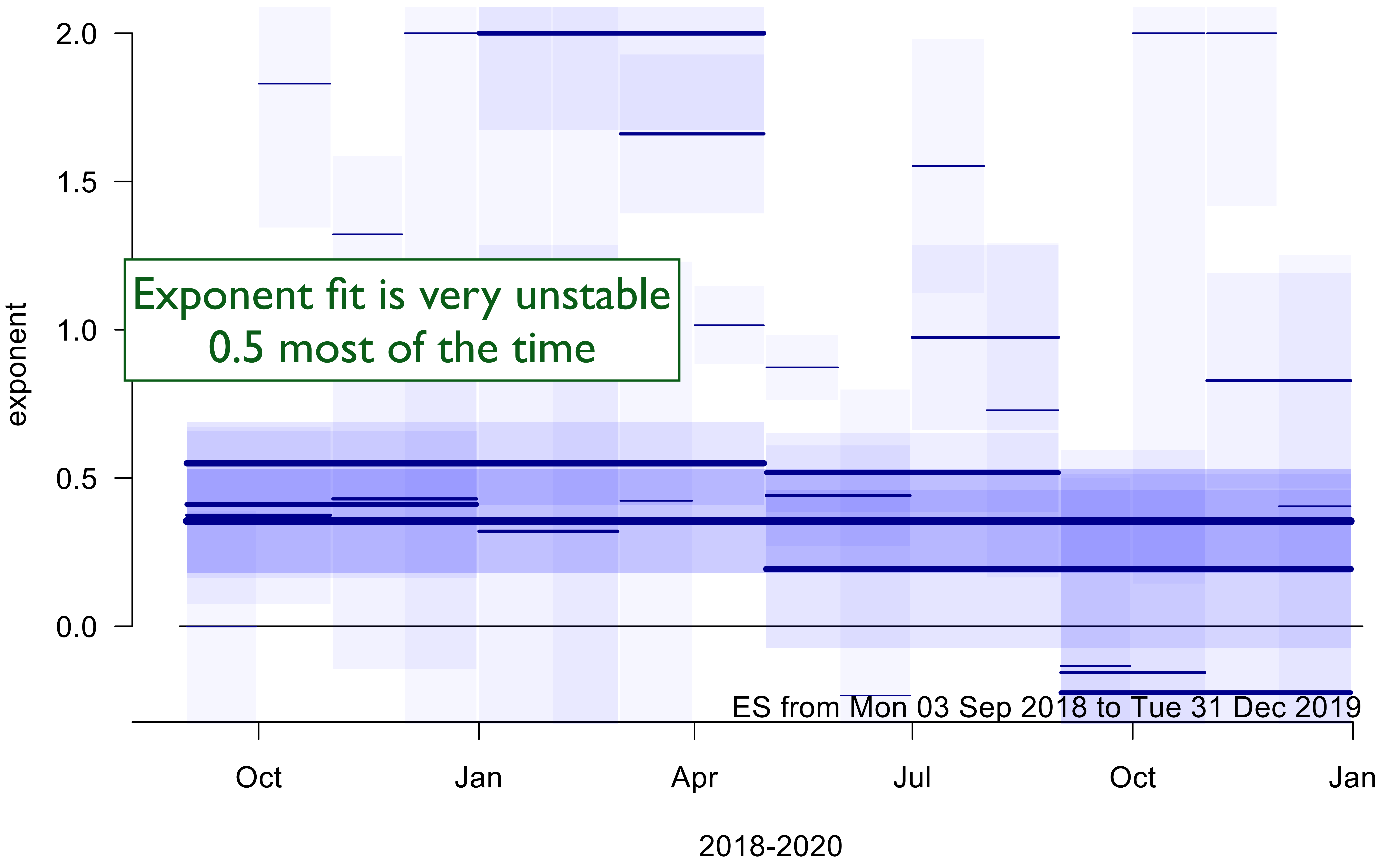


Cost decreases
as participation rate
increases

Participation rate depends on how quickly the order fills.

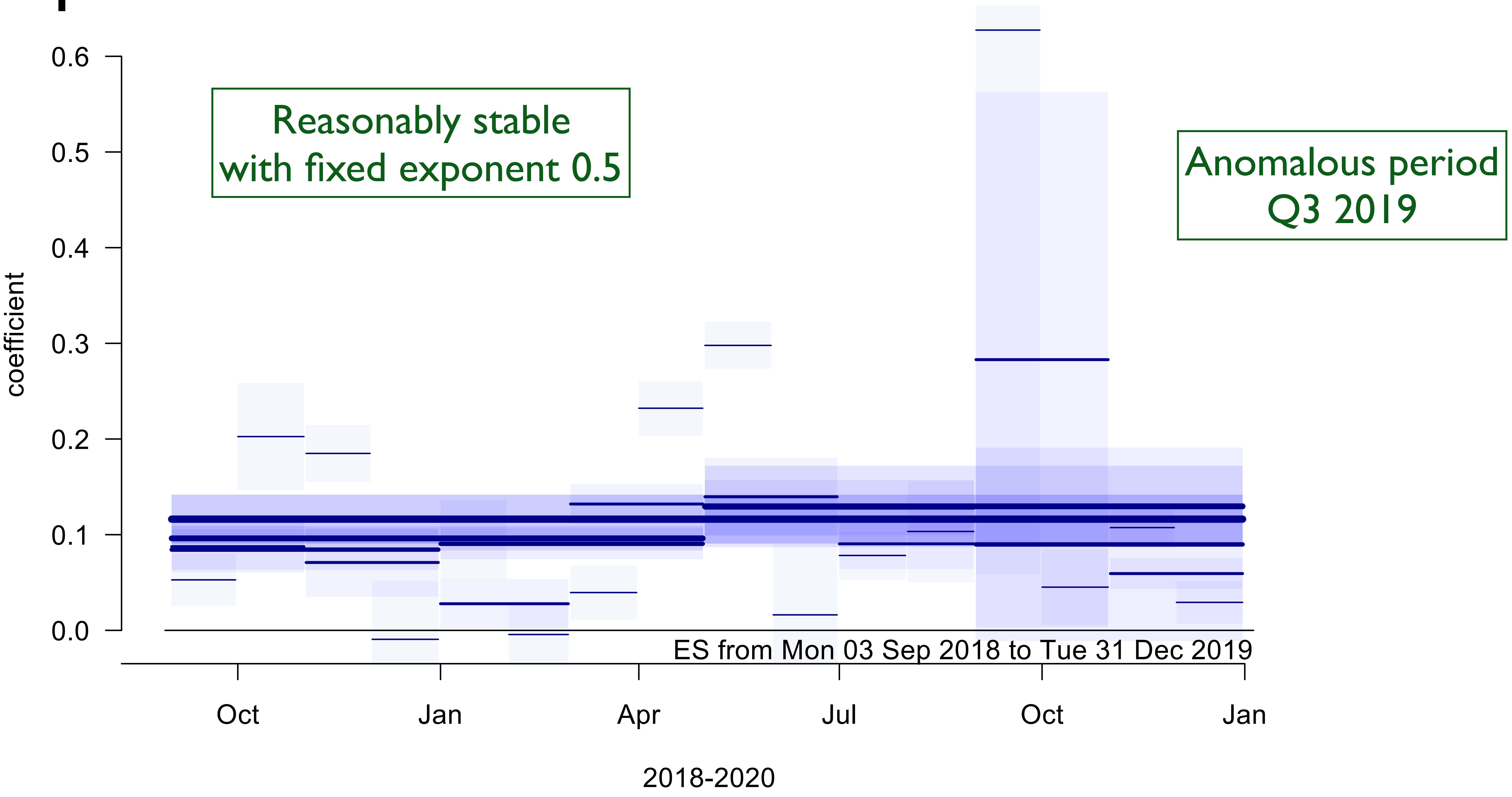
Participation rate is a *dependent* variable, not *independent.*

Dependence of exponent on time



Fit values on
- 16 months
- 8 months
- 4 months
- 2 months
- 1 month

Dependence of coefficient on time



Conclusions of single-asset fitting

Fractional-power model gives reasonable agreement

- Settle on exponent k = 0.5

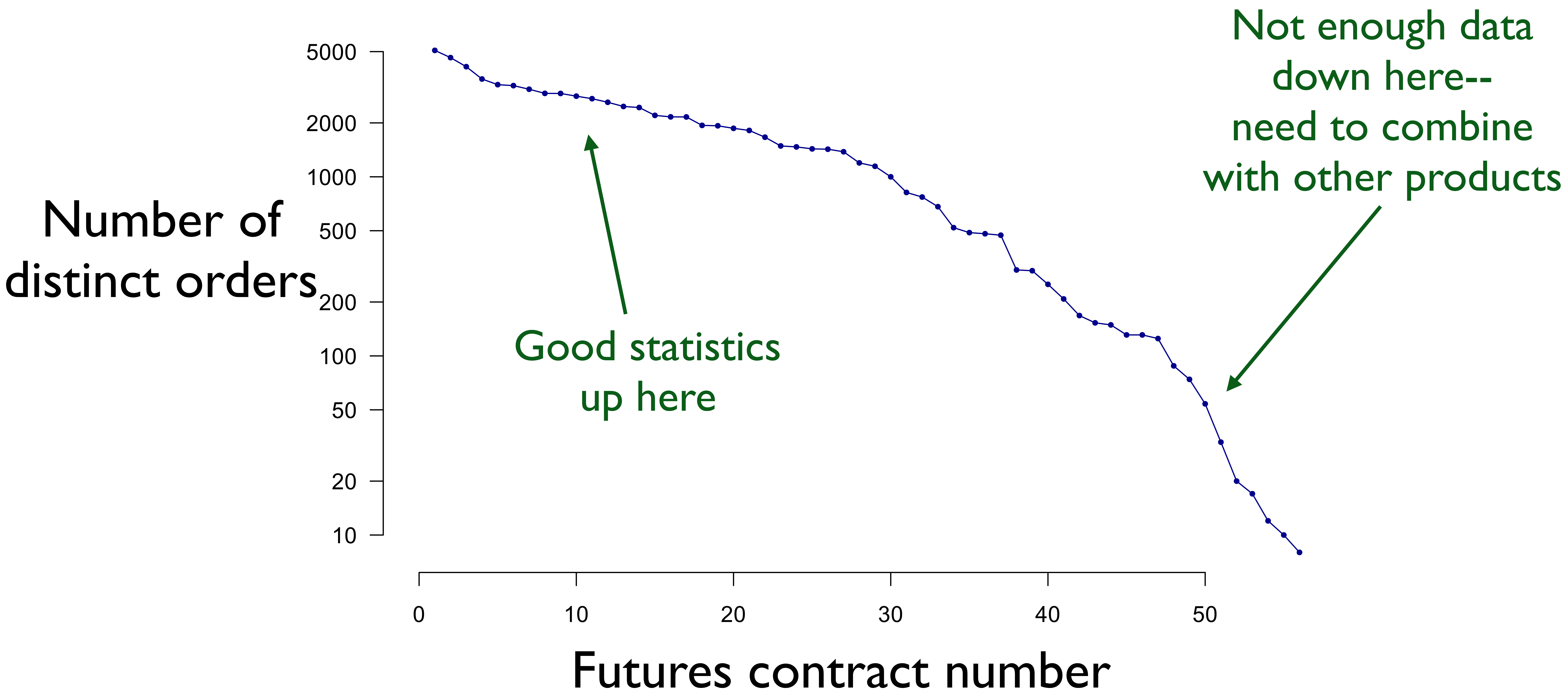
Want to fit important part of parameter range (50-100)

Neglect participation rate

Multi-asset fitting

Challenge: wide range of products
Not enough data for each one to fit individually
How to group together?

Distribution of number of orders



Comparison of coefficients: normalize by volume and volatility

$$\frac{C}{\sigma} = f\left(\frac{X}{V}\right) + \text{error} \qquad\qquad f(x) = a + bx^{\gamma}$$

$$\begin{aligned} C &= \text{slippage} \\ X &= \text{trade size} \\ \sigma &= \text{forecast daily volatility} \\ V &= \text{forecast daily volume} \end{aligned}$$

Consistent with literature

PHYSICAL REVIEW X **1**, 021006 (2011)

Anomalous Price Impact and the Critical Nature of Liquidity in Financial Markets

B. Tóth, Y. Lempérière, C. Deremble, J. de Lataillade, J. Kockelkoren, and J.-P. Bouchaud
Capital Fund Management, 6, blvd Haussmann 75009 Paris, France
(Received 9 May 2011; published 31 October 2011)

We propose a dynamical theory of market liquidity that predicts that the average supply/demand profile is V shaped and *vanishes* around the current price. This result is generic, and only relies on mild assumptions about the order flow and on the fact that prices are, to a first approximation, diffusive. This naturally accounts for two striking stylized facts: First, large metaorders have to be fragmented in order to be digested by the liquidity funnel, which leads to a long memory in the sign of the order flow. Second, the anomalously small local liquidity induces a breakdown of the linear response and a diverging impact of small orders, explaining the "square-root" impact law, for which we provide additional empirical support. Finally, we test our arguments quantitatively using a numerical model of order flow based on the same minimal ingredients.

sqrt(volume) and volatility
both scale linearly with time
Incorporates changing
market conditions

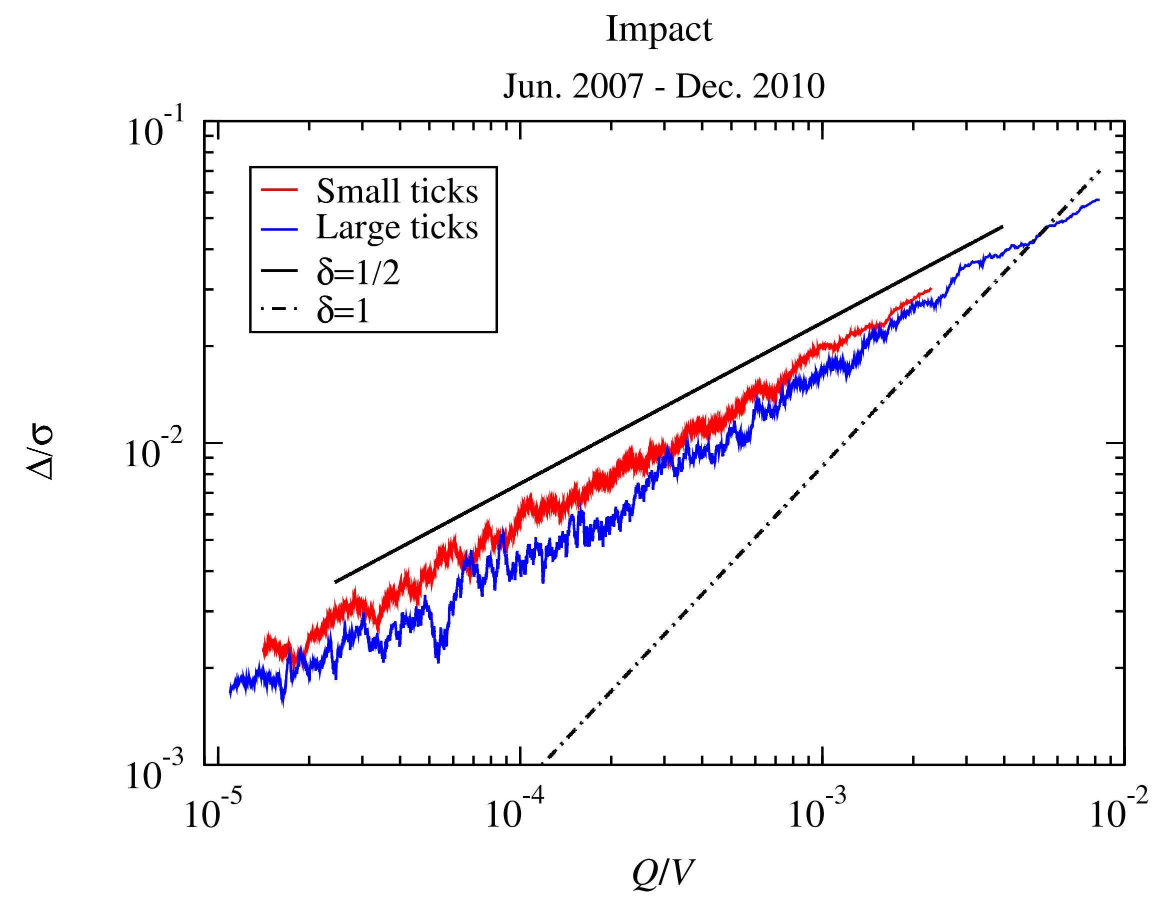


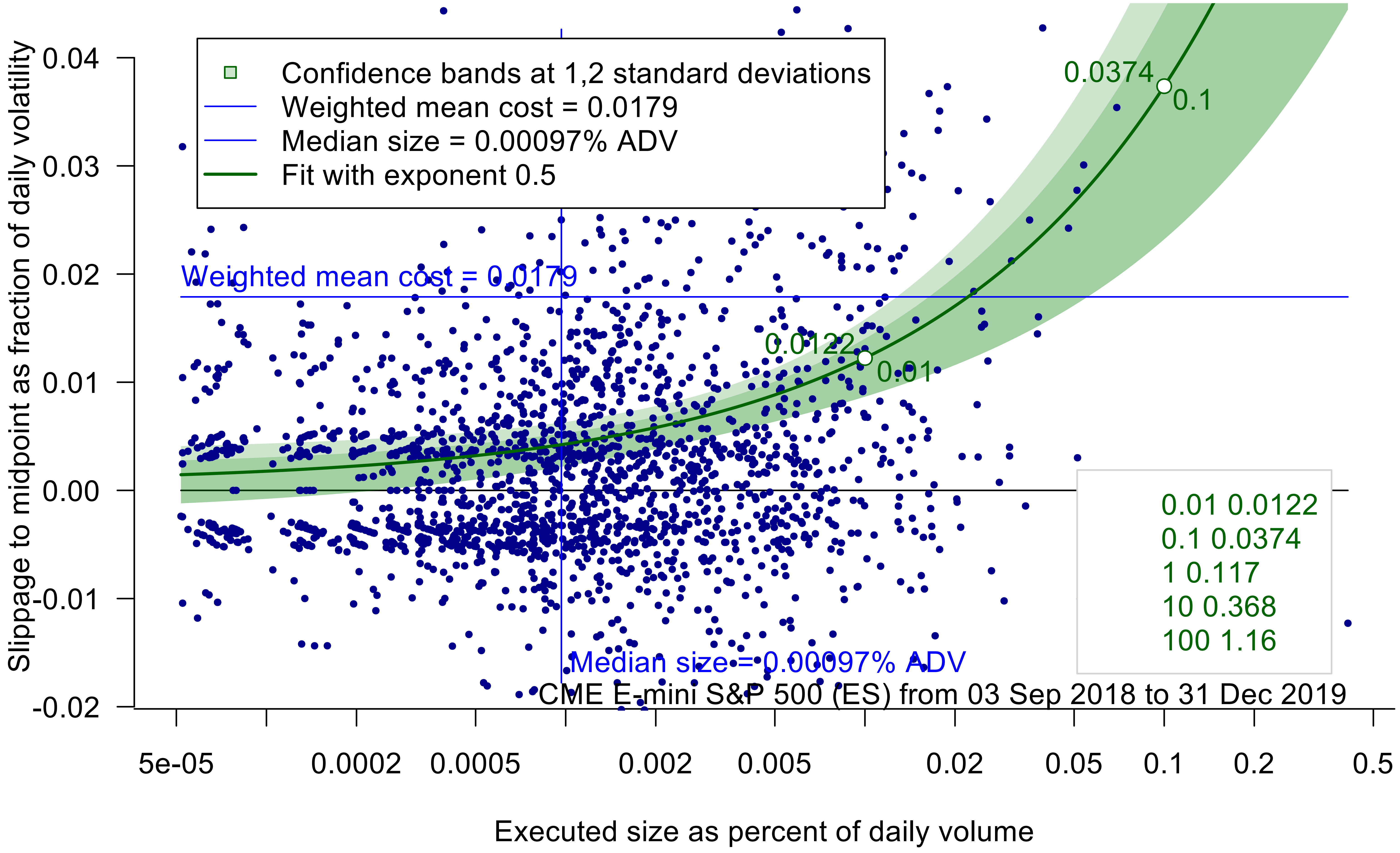
FIG. 1. The impact of metaorders for Capital Fund Management proprietary trades on futures markets, in the period from June 2007 to December 2010. Impact is measured here as the average execution shortfall of a metaorder of size Q.

the average relative price change Δ between the first and the last trade of a metaorder of size Q is well described by the so-called "square-root" law:

$$\Delta(Q) = Y\sigma\sqrt{\frac{Q}{V}}, \qquad (1)$$

where σ is the daily volatility of the asset and V is the daily traded volume, and both quantities are measured contemporaneously to the trade. The numerical constant Y is of order unity. Published and unpublished data suggest slightly different versions of this law; in particular, the $\sqrt{Q}$ dependence is more generally described as a power-law relation $\Delta(Q) \propto Q^{\delta}$, with δ in the range 0.4 to 0.7

Scaled fit


Confidence bands at 1,2 standard deviations
Weighted mean cost = 0.0179
Median size = 0.00097% ADV
Fit with exponent 0.5
Weighted mean cost = 0.0179
Median size = 0.00097% ADV
0.0374
0.1
0.0122
0.01
0.01 0.0122
0.1 0.0374
1 0.117
10 0.368
100 1.16
CME E-mini S&P 500 (ES) from 03 Sep 2018 to 31 Dec 2019
Slippage to midpoint as fraction of daily volatility
0.04
0.03
0.02
0.01
0.00
-0.01
-0.02
5e-05
0.0002
0.0005
0.002
0.005
0.02
0.05
0.1
0.2
0.5
Executed size as percent of daily volume

How to group?

1. Based on intrinsic properties of product:
   - tick size
   - liquidity
   - etc

This does not work

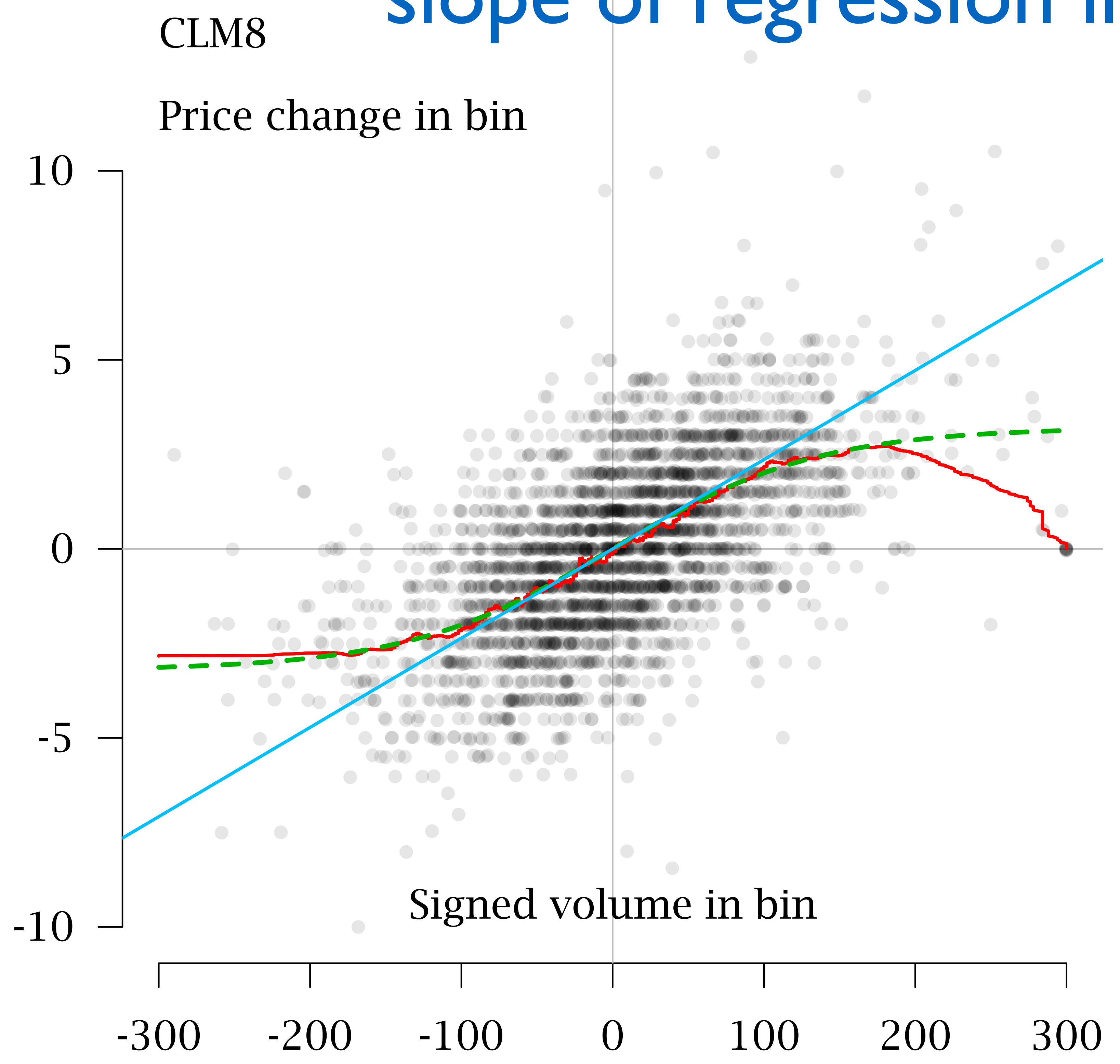
2. Based on regression fit itself:
   - mean and variance of coefficients

This works

Grouping by intrinsic properties

Illiquidity =
slope of regression line

Liquidity
= price change per volume traded



Tick size
= average spread in terms of minimum price increment
= reversion ratio (Robert/Rosenbaum)
= average quote size / average trade size

Large tick vs small tick

Journal of Financial Econometrics, 2011, Vol. 9, No. 2, 344–366

A New Approach for the Dynamics of Ultra-High-Frequency Data: The Model with Uncertainty Zones

Christian Y. Robert
CREST and ENSAE Paris Tech

Mathieu Rosenbaum
CMAP-École Polytechnique Paris UMR CNRS 7641

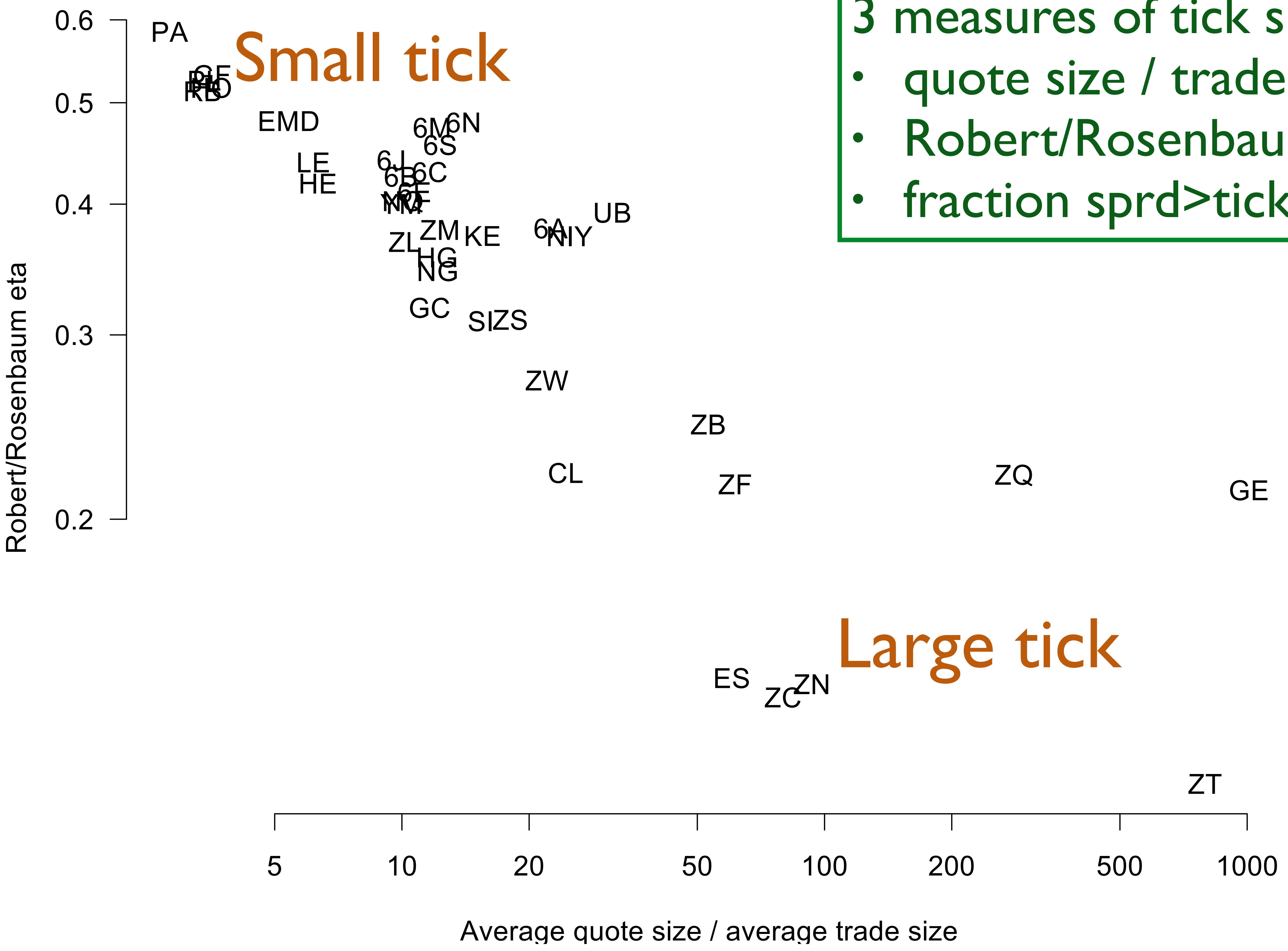
ABSTRACT
In this paper, we provide a model which accommodates the assumption of a continuous efficient price with the inherent properties of ultra-high-frequency transaction data (price discreteness, irregular temporal spacing, diurnal patterns...). Our approach consists in designing a stochastic mechanism for deriving the transaction prices from the latent efficient price. The main idea behind the model is that, if a transaction occurs at some value on the tick grid and leads to a price change, then the efficient price has been close enough to this value shortly before the transaction. We call uncertainty zones the bands around the mid-tick grid where the efficient price is too far from the tick grid to trigger a price change. In our setting, the width of these uncertainty zones quantifies the aversion to price changes of the market participants. Furthermore, this model enables us to derive approximated values of the efficient price at some random times, which is particularly useful for building statistical procedures. Convincing results are obtained through a simulation study and the use of the model over 10 representative stocks.

One can also see the parameter η as a measure of the relevance of the tick size on the market. Indeed, if $\eta < 1/2$, market participants are convinced they have to trade at a new price before the efficient price crosses this new price on the tick grid. So, it means that the tick size appears too large to them. Conversely, a large η ($\eta > 1/2$) means that the tick size appears too small. From the tick size perspective, an ideal market is consequently a market where η is equal to $1/2$.

A natural estimation procedure for the parameter η is given in Robert and Rosenbaum (2010a). We define an alternation (resp. continuation) of one tick as a price jump of one tick whose direction is opposite to (resp. the same as) the one of the preceding price jump. Let $N^{(a)}_{\alpha,t}$ and $N^{(c)}_{\alpha,t}$ be respectively the number of alternations and continuations of one tick over the period $[0,t]$. An estimator of η over $[0,t]$ is given by

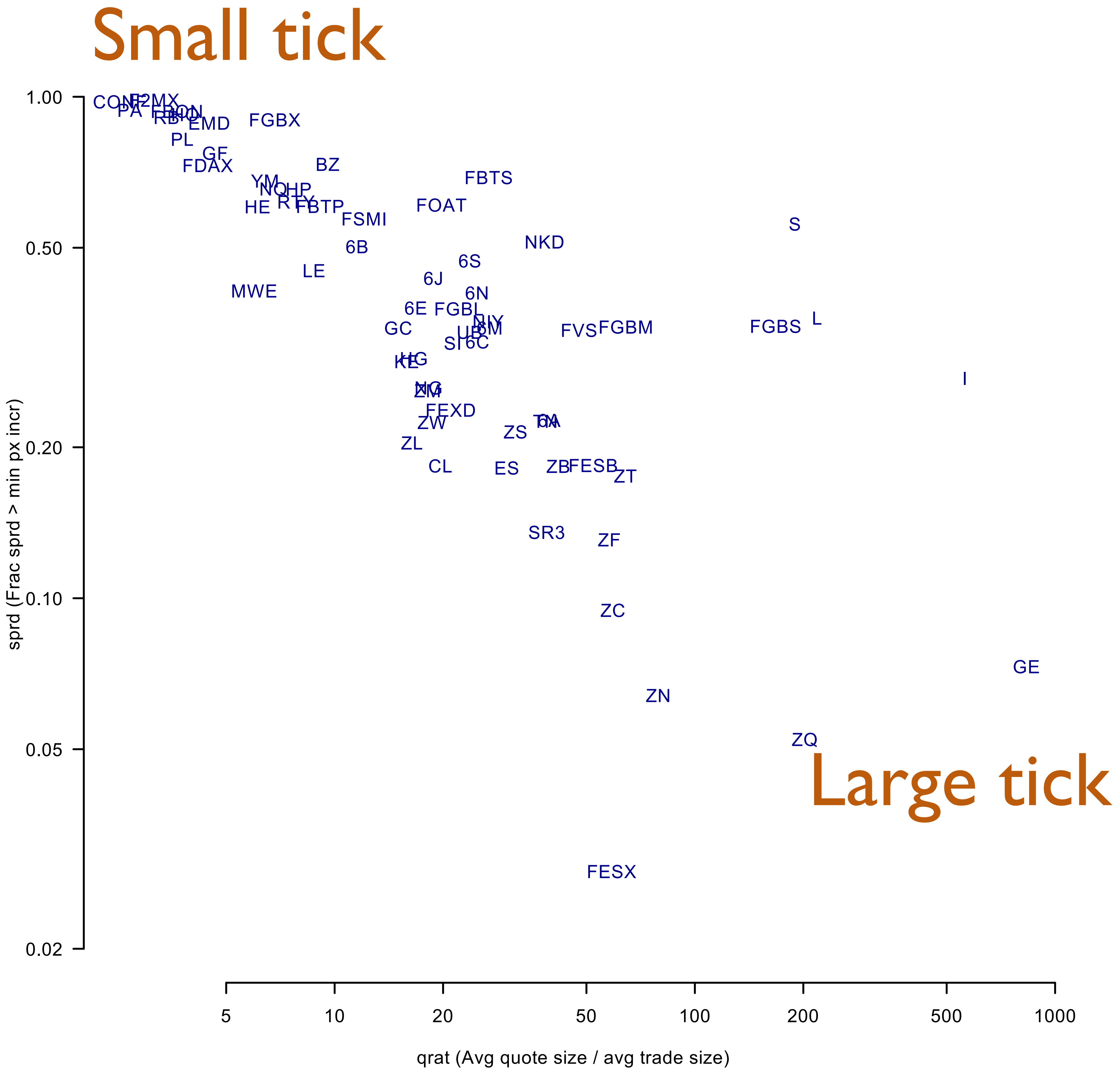
$$\hat{\eta}_{\alpha,t} = \frac{N^{(c)}_{\alpha,t}}{2N^{(a)}_{\alpha,t}}.$$

Tick size spectrum

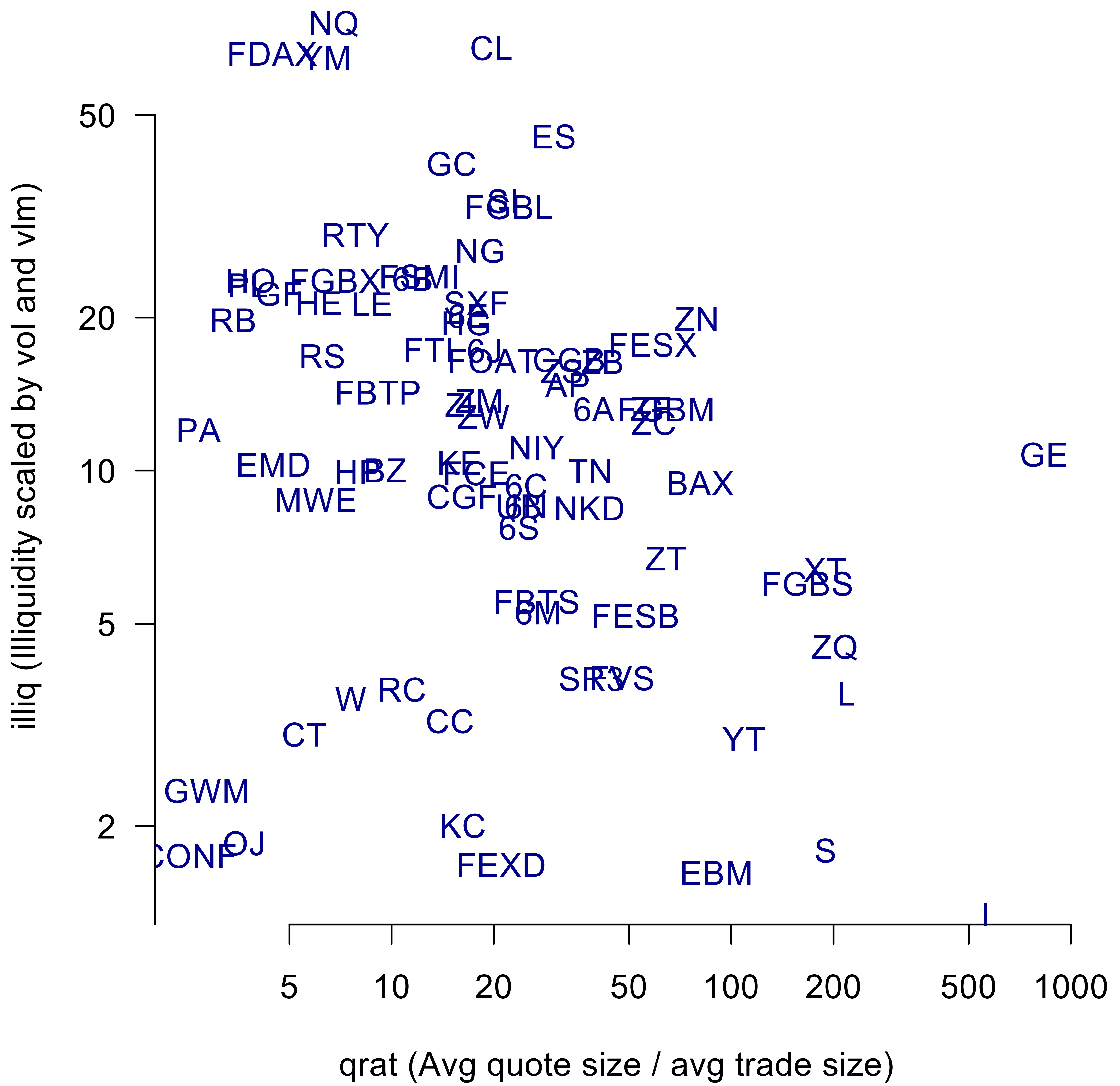


3 measures of tick size:

- quote size / trade size
- Robert/Rosenbaum eta
- fraction sprd>tick



Tick size vs nondimensional liquidity

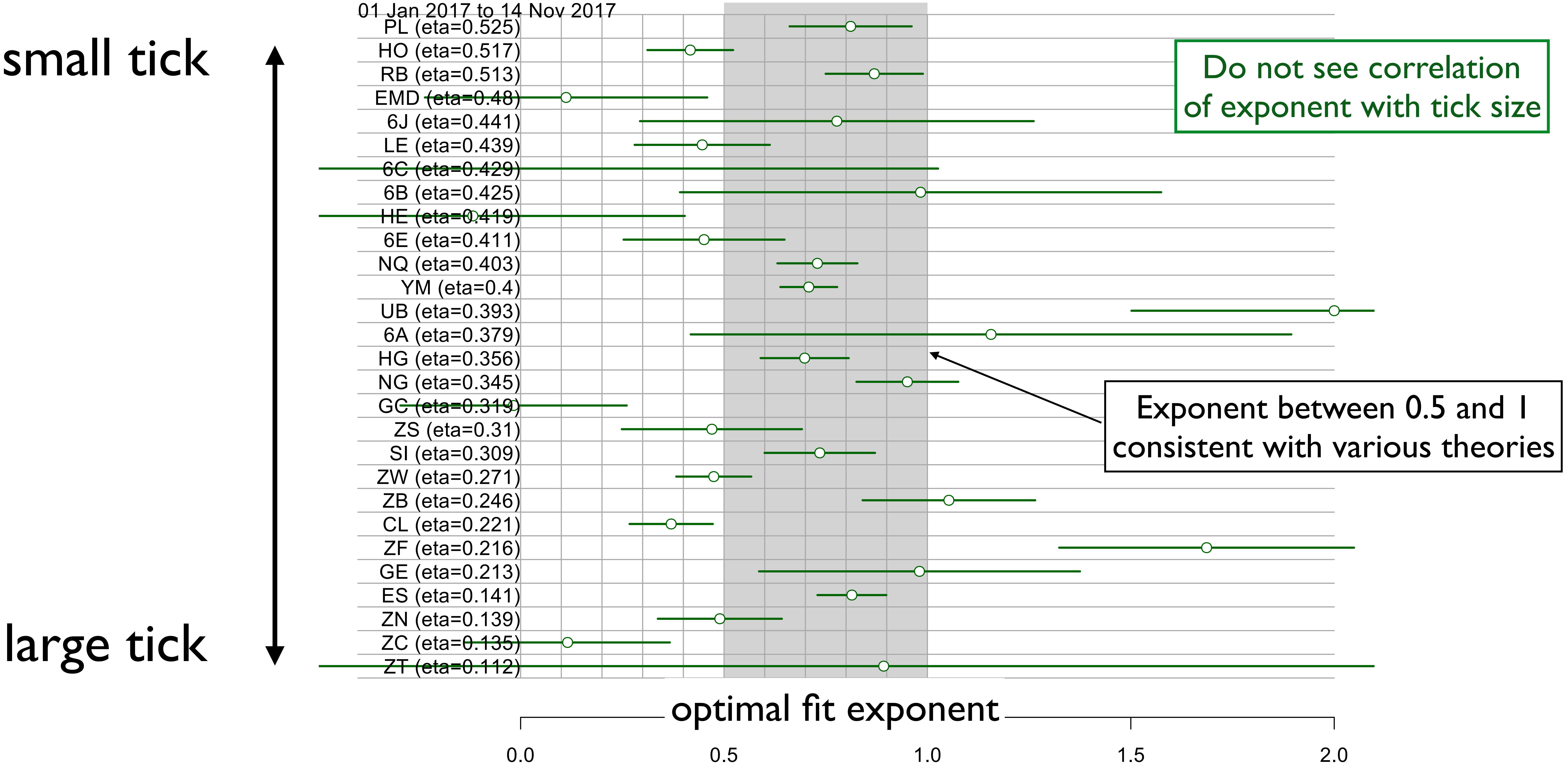


Subdivide based on these parameters

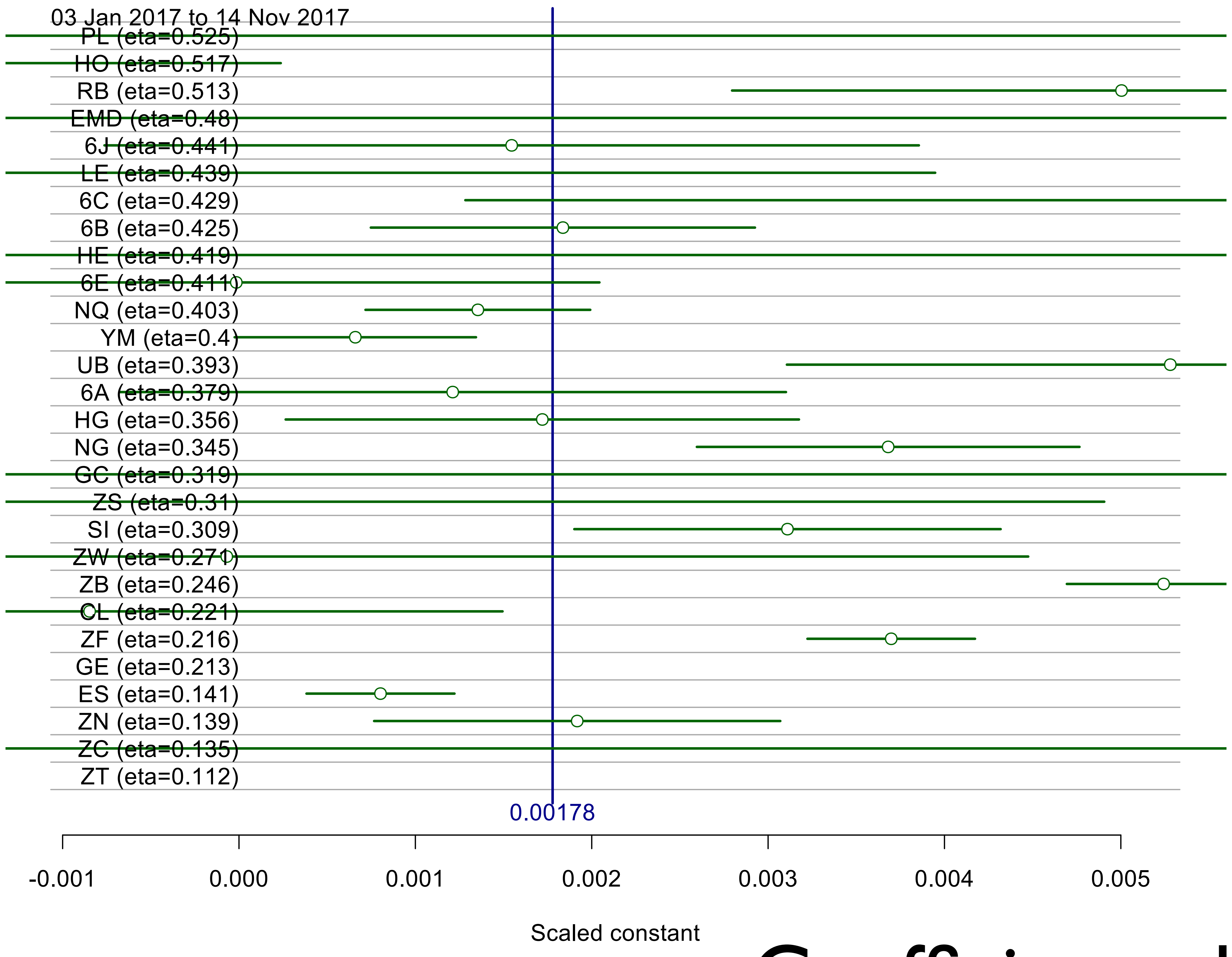
Does not work (does not give meaningful results)
because points that are close in parameters
are not close in cost models

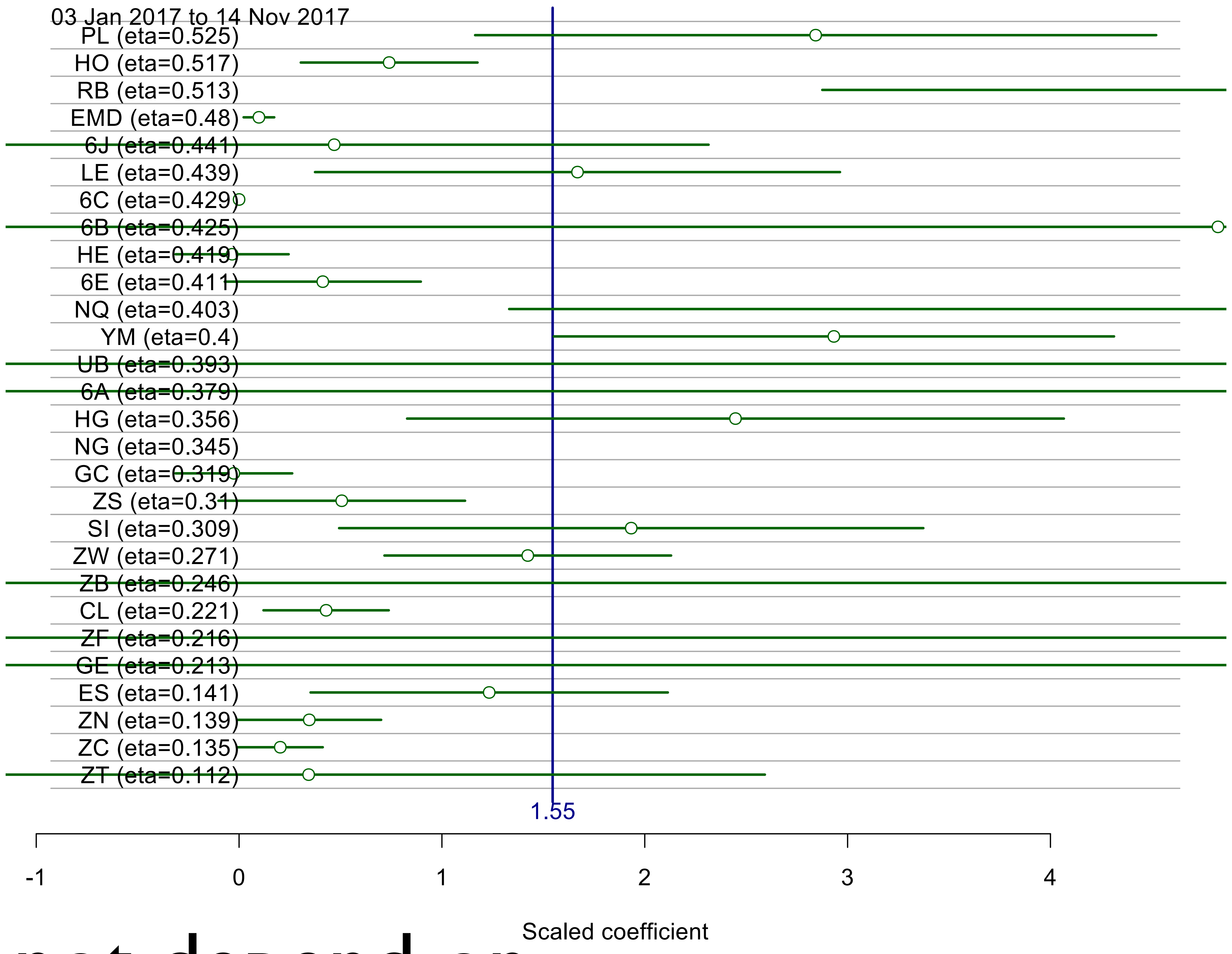
Problem: market impact model depends on properties
that are not part of market data
for example, size of underlying asset.

Variation of exponent across products



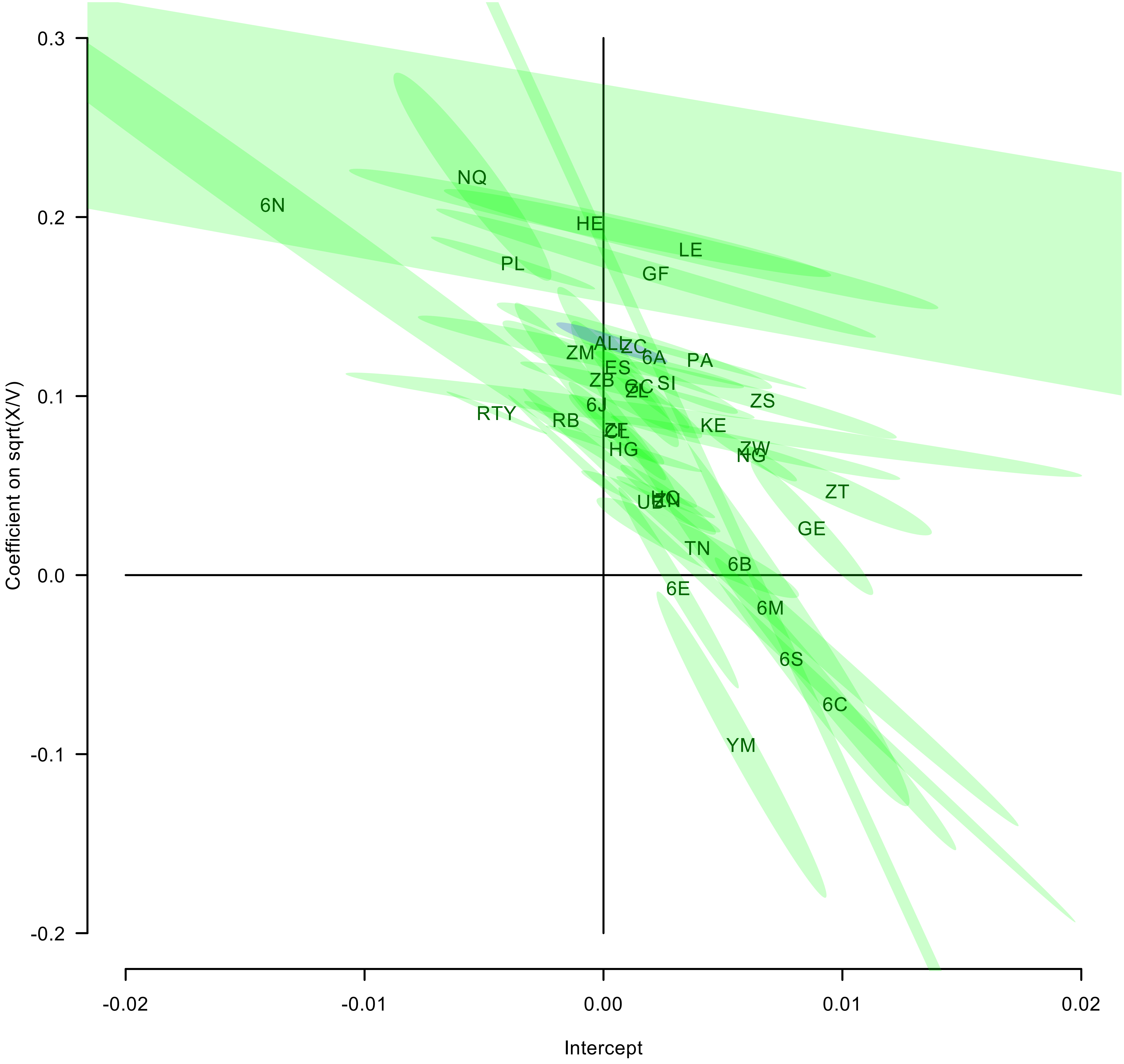
Variation of scaled coefficients across products



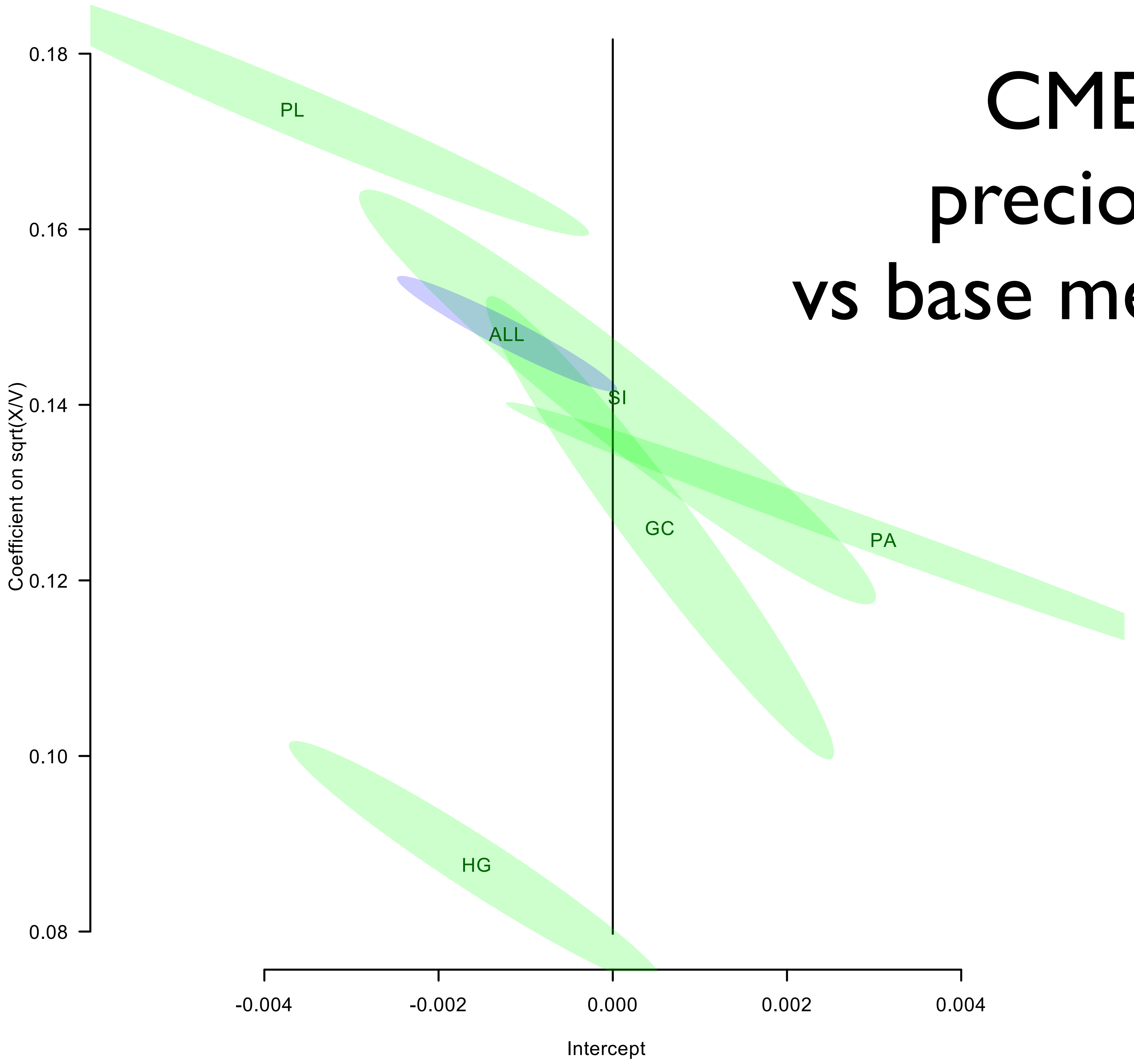


Coefficients do not depend on tick size in consistent way

Subdivide based on fit itself

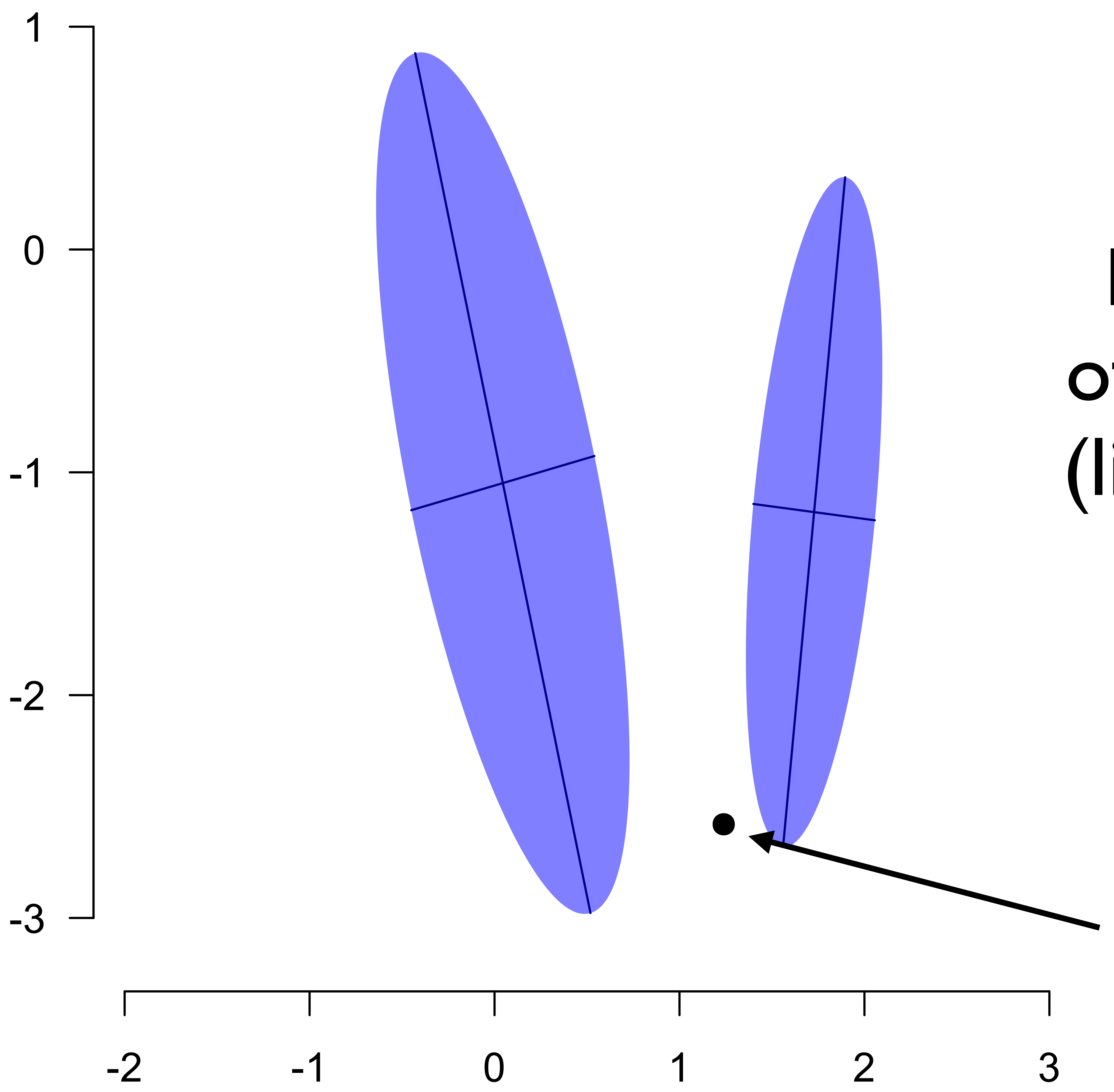


Commonality in products



CME metals:
precious metals
vs base metals (copper)

Distance between Gaussian distributions



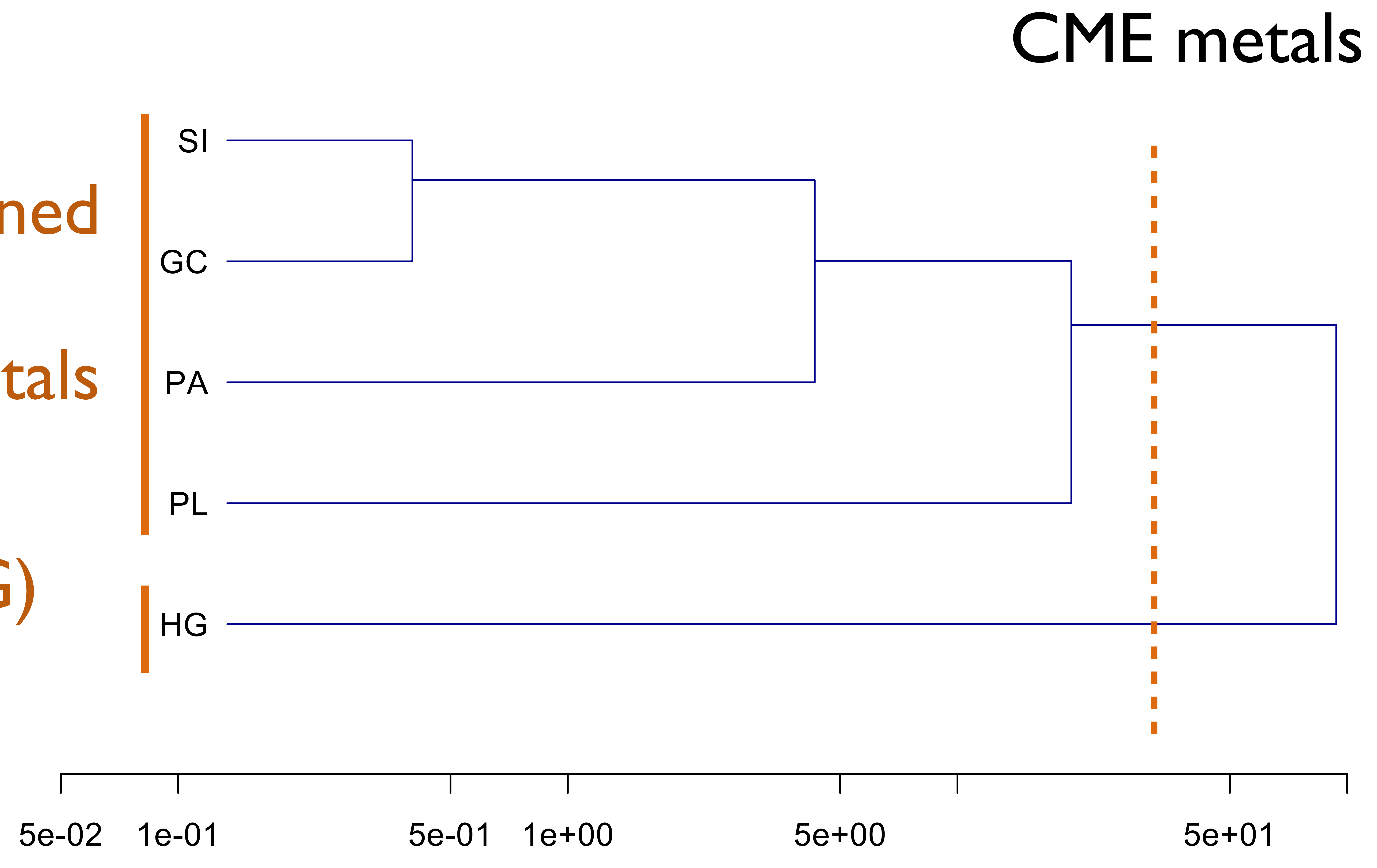
Distance = -log prob of most probable point (like a 2-variable t-test)

Most probable point

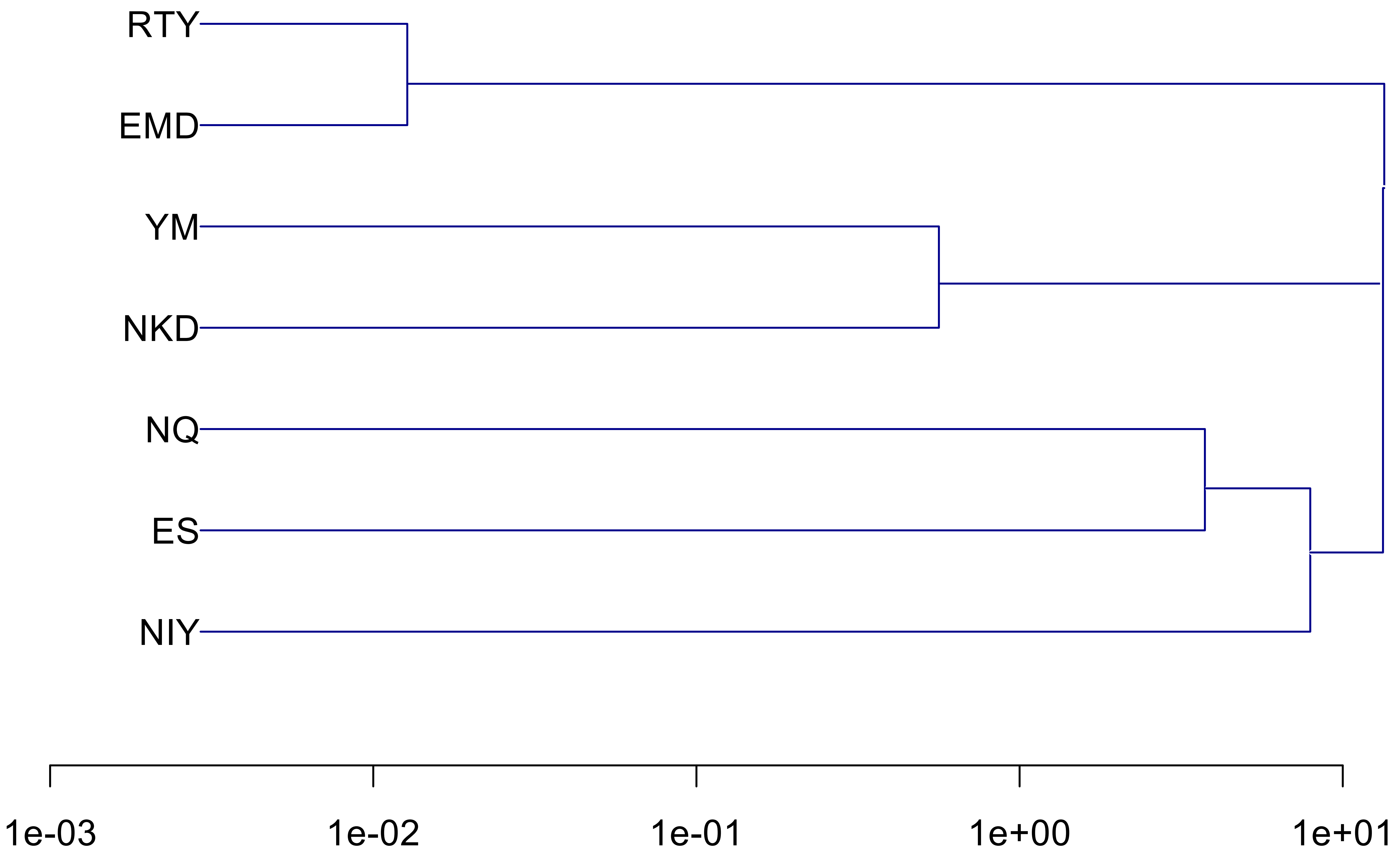
Clustering based on this distance

Make combined fit for precious metals

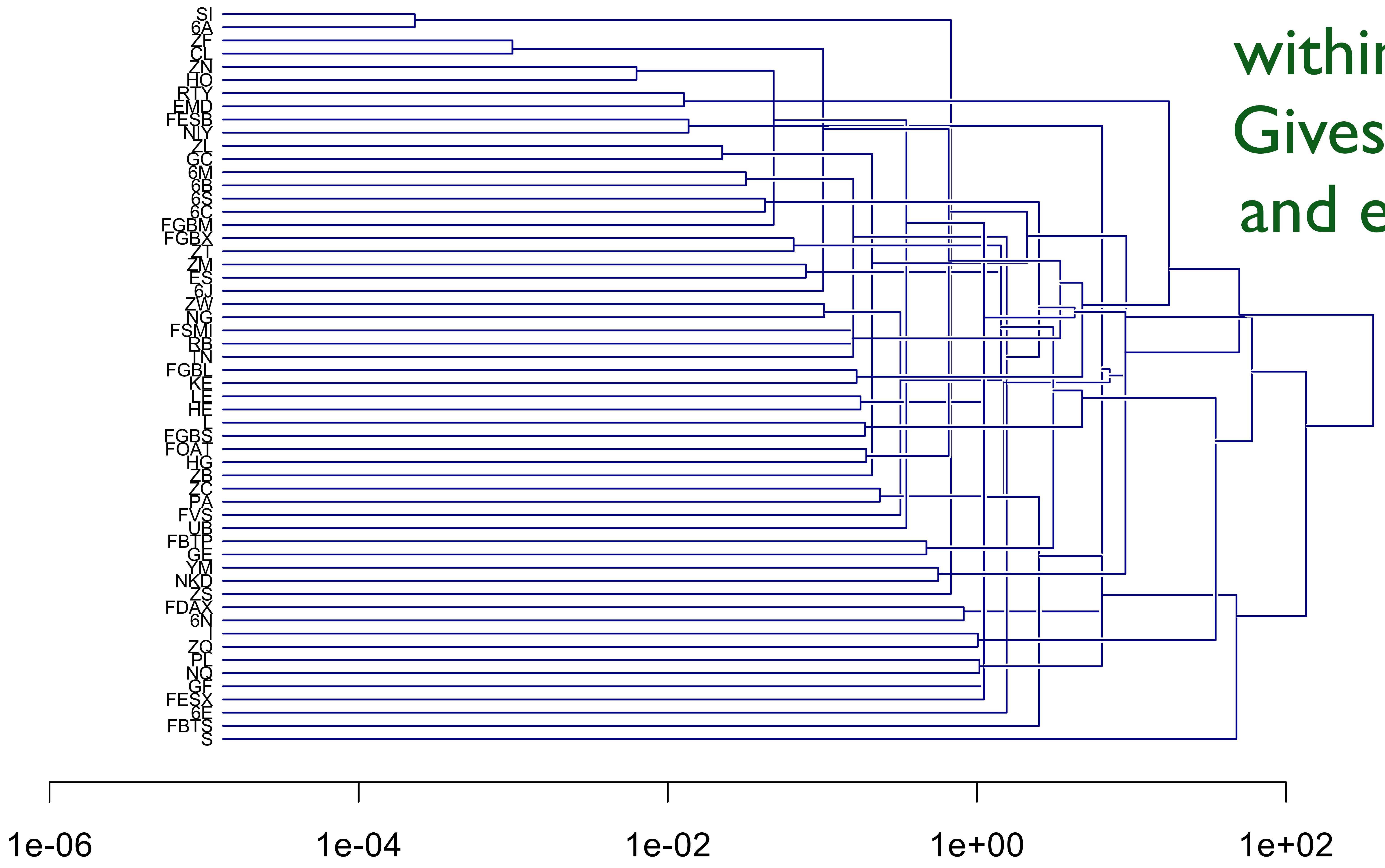
Copper (HG) on its own



CME equity index futures



Clustering is not stable on whole data set



Conclusion: cluster within exchange and class. Gives reasonable accuracy and economically sensible

Conclusions

Market impact modeling is noisy

- R^2 terrible, t-stats good
- ability to predict any particular trade is poor

Need to use physical reasoning and ad hoc decisions

- focus on parameter ranges that are economically important

Futures challenge is hetergeneous products

- need to cluster based on economic properties and fit