Simple Cost Model

Charles Pehlivanian

September 9, 2016

Abstract

A simple trading cost model is specified and fit based on realized trading trajectories of parent orders on exchanges within CMEGroup for clients of Quantitative Brokers. The fit incorporates Spread, Temporary, and Permanent Cost assumptions, although those costs are not explicitly estimated, only aggregate cost. As a next step we would estimate these costs separately, and fit a true intra-order model. The current model is meant to be a first step or proof-of-concept.

Dataset

The dataset consists of summary statistics for realized child trade confirms by parent order. All trades occur on CME Globex and are classified by the Intrument Codes as given in Table 1.

The trade universe consists of the Firm's *BOLT* orders, which are arrival-price-benchmarked and attempt to minimize implementation shortfall. See any standard textbook. The remainder of the trades are *STROB* trades which attempt to replicate vwap (or twap, without the volume curve) over the trading window.

Define

$$\begin{aligned} Cost = & \quad 1e4 * Sidesign * (vwap - arrivalprice)/arrivalprice && (1) \\ POV = & \quad (parentorderparticipation) * (duration) * (durationwtdvolatility) && (2) \\ & && (3) \end{aligned}$$

where

$$\begin{aligned} vwap &= \frac{1}{\sum_{i=0}^{n} v_i} \sum_{i=0}^{n} p_i v_i && (4) \\ arrivalprice &= midpriceatorderconception && (5) \\ & && (6) \end{aligned}$$

Inst Code	Description	#Parent Orders
AG	CBOT Ags	5724
EN	NYMEX Energy	6957
EQ	Equity Index	8155
FX	FX Pairs	2925
IR	Interest Rate	17760
MT	NYMEX Metals	9169
TOTAL		45690

Table 1: CME Instrument Code Summary.

Inst Code	Beta	Gamma
AG	9.5706	0.4309
EN	2.7326	0.1087
EQ	4.0216	0.4277
FX	0.9242	0.2169
IR	0.7273	0.2258
MT	3.6563	0.3815
ALL	4.266	0.3561

Table 2: CME Instrument Code Summary.

etc.

Model Specification

The model is a power law in the POV factor, with constant, and spread-based additional summands. After consideration of several models, this was found to give the best fits (see following for basic summary metrics) without sacrificing simplicity. In summary

$$Cost = C + \theta S + \beta X^{\gamma} \tag{7}$$

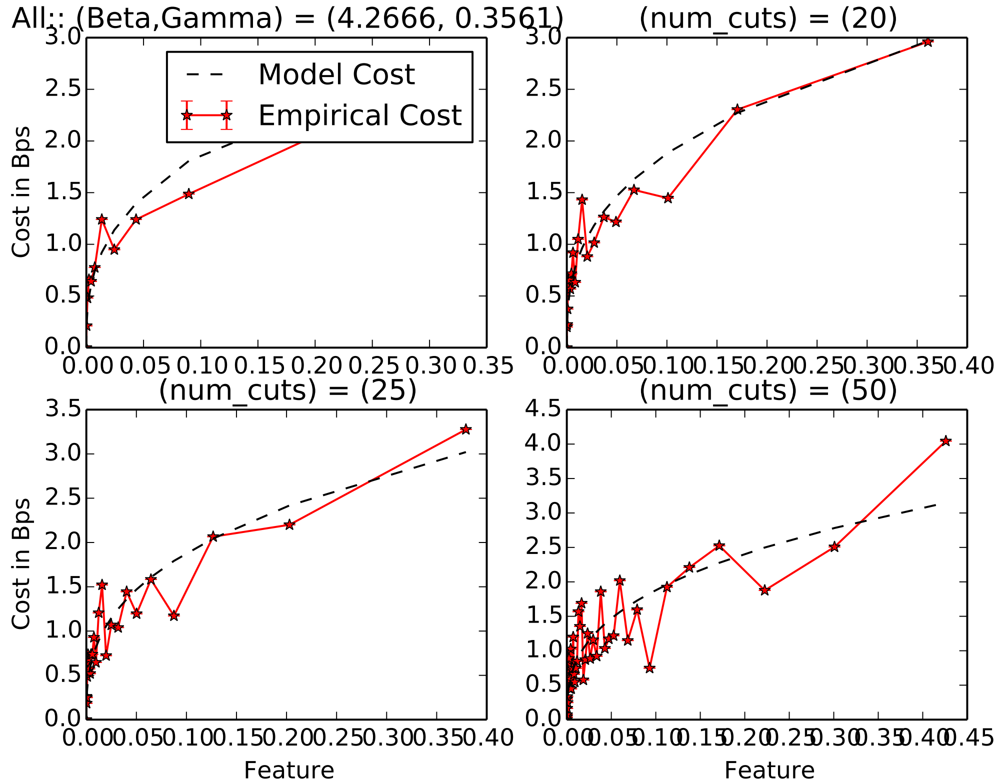
The model fits were carried out in Python using the *scipy.ODR* Orthogonal Distance Regression module.

Fits

The model fits by Instrument Class follow. The important (β, γ) are shown in Table 2., as well as quality of fit graphs - theoretical v empirical. The empirical fit is averaged by bin, where the number of bins (num_cuts) varies from 10 to 50. The convexity of the empirical cure is evident in these graphs especially at the low-cost end. Residual R-squareds are given for a theoretical linear fit as well as for the power-law fit. These further justify the choice of a sublinear cost function, which can also be justified on theoretical grounds by a no trading-cost arbitrage argument.

	Group	beta (est)	gamma (est)	lin r^2	pow_law r^2
0	AG_All	9.5706	0.4309	0.0266	0.0306
0	EN_All	2.7326	0.1087	0.0178	0.036
0	EQ_All	4.0216	0.4277	0.0352	0.0459
0	FX_All	0.9242	0.2169	0.0305	0.0493
0	IR_All	0.7273	0.2258	0.0228	0.0376
0	MT_All	3.6563	0.3815	0.0377	0.0521

Coefficients by Instrument Class.



Empirical v Theoretical Plots,, Empirical plots are binned averages.