**Multi-Level Order-Flow Imbalance in a Limit Order Book (2019)**

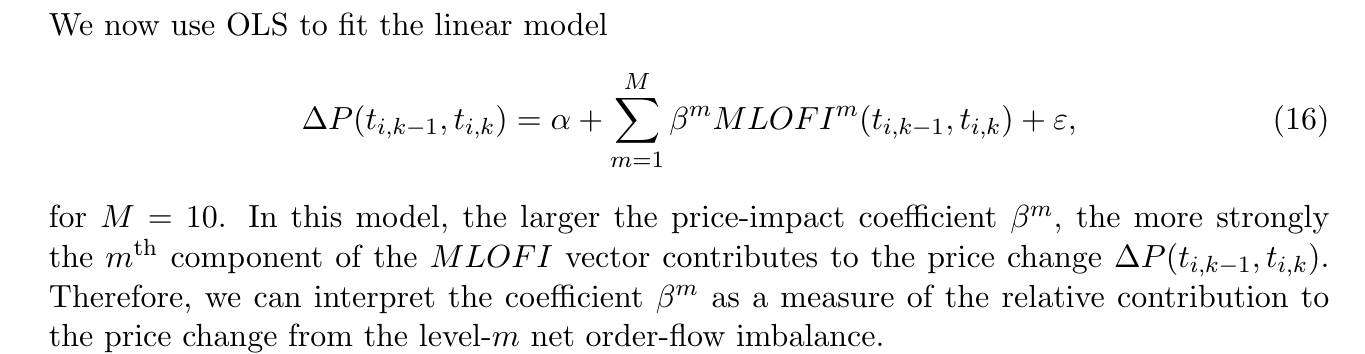
**Focused Question:**

Some previous models provide have conducted researches on the interplay between order flow, liquidity and price formation, but they typically encounter the drawback of being extreme complexity.

This paper is to address such problem by uncovering and fitting the simple statistical relationship which explain price formation in an LOB in terms of Net Order Flow (an easily understood and easily measurable input).

**Conclusions:**

1. MLOFI equation (16) reveals a simple linear relationship between net order-flow at the first M populated price levels in an LOB and the contemporaneous change in mid-price.



1. Perform OLS regressions and Ridge regressions to fit the MLOFI equation (16), and use both the adjusted R2 and the RMSE to assess the goodness-of-fit of the fitted relationships and compare the results quantitatively
2. Analyze both the sample correlations in the MLOFI vector and the out-of-sample RMSE, and find that Ridge regression performs better than OLS regression in fitting the MLOFI equation (16)
3. With either R2 or out-of-sample RMSE, the goodness-of-fit of the fitted MLOFI equation (16) was considerably stronger for large-tick stocks than small-tick stocks: -> the overall goodness-of-fit increased as M increased, but its rate declines

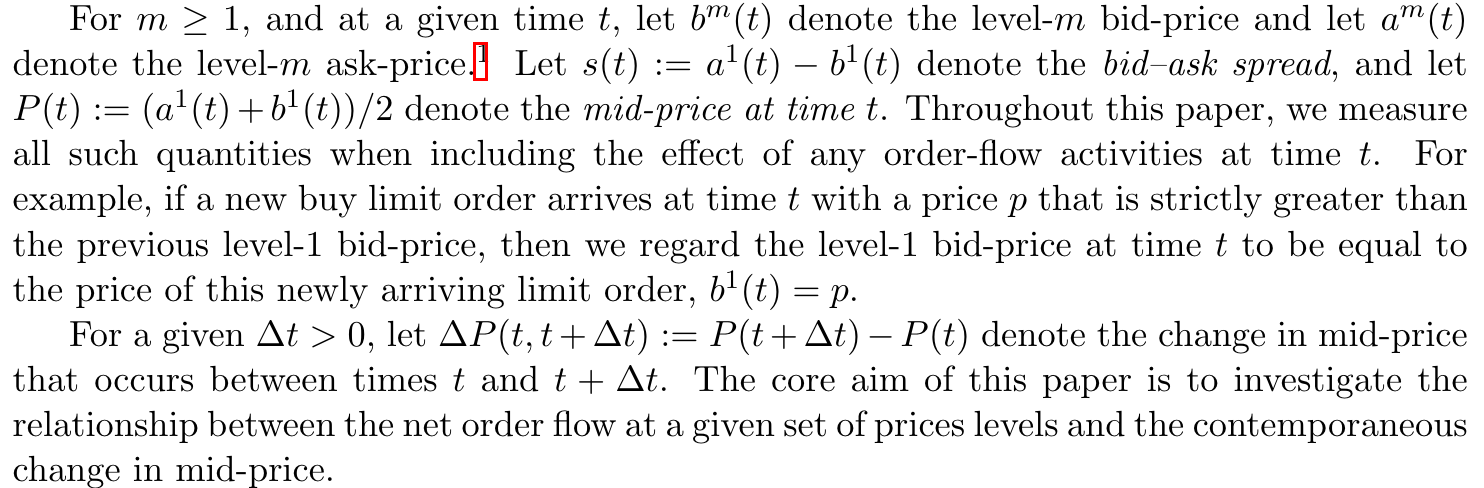
**Possible approaches to improve the goodness-of-fit of the MLOFI equation:**

Method 1: refine the definition of the MLOFI vector to address some of the weaknesses in Section 6.

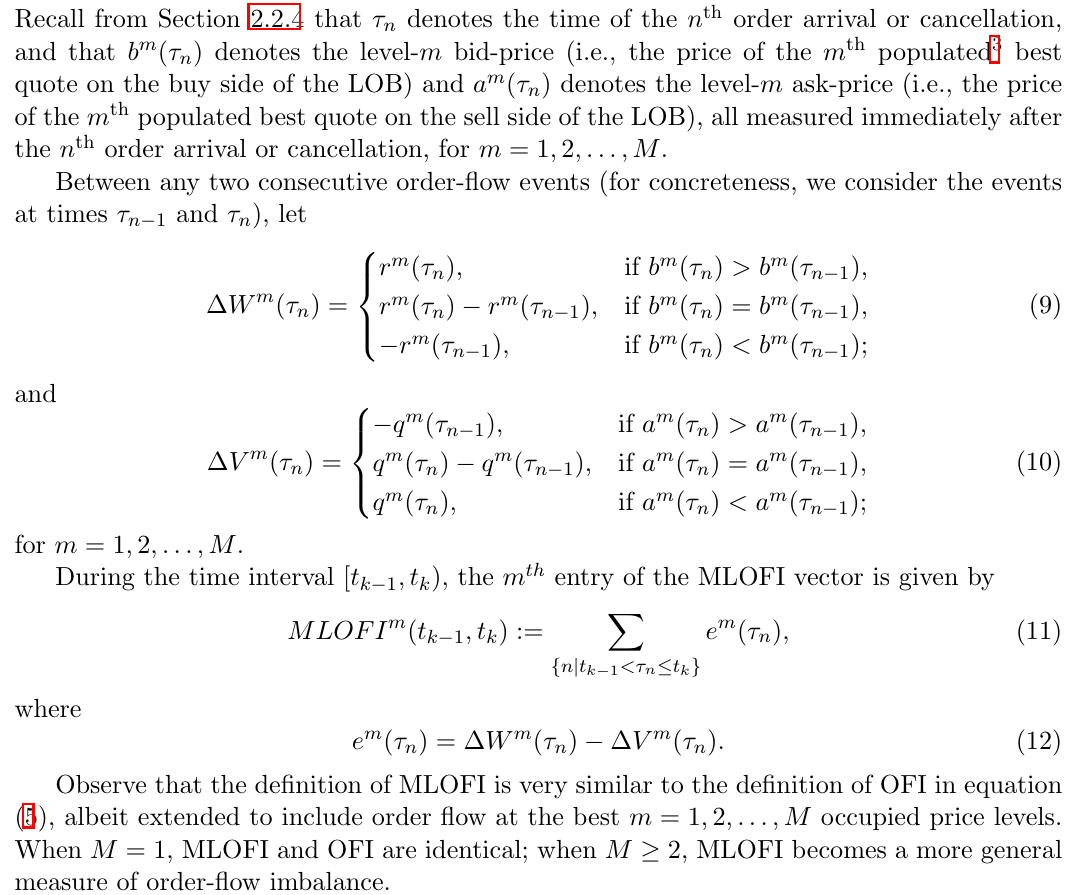
Method 2: use the MLOFI vector to model price impact as a dynamical and latent variable, using the framework recently introduced by Mertens et al. [2019].

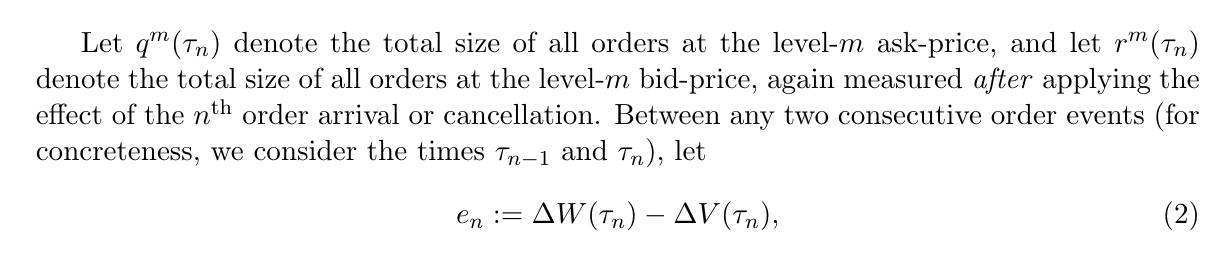
Method 3: include other input variables into the regression.

**Prerequisite of Order Flow, Liquidity and Price Formation**



**Multi-Level Order-Flow Imbalance: Calculating MLOFI**





**Results**