European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

Friend or Foe of Efficiency? The Impact of Algorithmic Trading on Price Discovery in Indian Base Metal Futures

Rashmi Jain*

*Research Scholar, Jain Deemed to be University, Bengaluru, India.

Dr. Chaya Bagrecha**

**Professor of Finance, CMS B - School Jain Deemed to be University, Bengaluru, India.

Abstract:

In this study, we introduce innovative methodologies to assess market-specific contributions to the process of price discovery within non-overlapping sequential markets, utilizing a structural Vector Autoregressive (VAR) model. The application of these methods is demonstrated through an empirical analysis focused on the trading dynamics of the futures and spot markets within the Indian base metal commodity market, spanning an extensive eight-year period.

Our findings reveal a prevailing dominance of price discovery by the futures market in the base metal sector, emphasizing its significant role in shaping market dynamics. Moreover, our investigation underscores the positive influence of algorithmic trading on the price discovery process, highlighting its role in enhancing market efficiency and contributing to a more robust discovery mechanism.

This study not only provides valuable insights into the intricate dynamics of base metal markets but also introduces a methodological framework that can be applied more broadly to assess the interplay of sequential markets in the context of price discovery. The implications of our research extend beyond the specific market examined, contributing to the broader understanding of the impact of algorithmic trading on market dynamics and efficiency.

Keywords- Algorithmic Trading, Market Efficiency, Price Discovery, Sequential Markets, Structural VAR Mode.

1. INTRODUCTION:

Price discovery, a crucial function in financial markets where the actions of buyers and sellers determine the asset's value, faces new dynamics in the era of computerized trading. Algorithmic Trading (AT) has introduced notable changes, given its higher speed and cost-effective monitoring capabilities, allowing for the swift incorporation of information into prices and an acceleration of the price discovery process. However, the rapid assimilation of short-term information may not always align with the fundamental value of assets. In the current high-frequency trading landscape, security prices become more susceptible to short-term fluctuations, such as liquidity shocks and inventory constraints.

This study explores the impact of Algorithmic Trading (AT) on the price discovery process, with a specific focus on the base metal market perspective. We delve into whether AT incorporates permanent information or transient noises into security prices, aiming to understand its contribution to both the permanent and noise components of price discovery, AT contributes more to the permanent price discovery process and less to transient pricing errors compared to non-AT counterparts.

Importantly, these findings extend to the base metal market, highlighting the role of AT in shaping the dynamics of price discovery within this specific commodity sector. The study demonstrates that the effects of AT on price discovery remain significant across all days, including high market stress days, both at the individual stock and market-wide levels. By incorporating the base metal market perspective, this investigation provides nuanced insights into the interaction between algorithmic trading and the persistence of price movements, contributing to a more comprehensive understanding of the dynamics within financial markets, particularly in the context of base metal commodities.

We leverage a novel dataset that quantifies algorithmic trading (Algotrading) intensity through the percentage of total turnover. Employing a state space framework, our study investigates the intricate relationship between Algotrading and the price discovery processes within the context of the base metal commodity market.

In our analysis, state space models play a pivotal role, enabling the decomposition of observed stock price series into unobserved permanent price components and transient pricing errors. The permanent component, associated with efficiency, is modeled as a martingale, capturing information arrivals influencing the lasting value of a stock. On the other hand, the transient pricing error component reflects short-lived price deviations not driven by fundamental value-related information. This component is modeled with an assumption of stationarity and an autoregressive component to capture its dynamic nature.

Our focus then shifts to examining the relationship between Algotrading order flows and the increments of the efficient and transitory components of daily prices. This analysis is conducted using a sample dataset from the Multi Commodity Exchange (MCX), specifically in the base metal commodity market, spanning from April 2015 to April 2023.

By employing this rigorous methodology, our research aims to shed light on how Algotrading influences both the enduring and short-lived aspects of price changes in the dynamic landscape of the base metal commodity market. This study contributes valuable insights into the nuanced dynamics of market processes, particularly in the context of algorithmic trading and its impact on price discovery.

Our research significantly contributes to the existing body of literature in multiple key dimensions. Primarily, we advance the scholarly discourse by conducting a thorough examination of the Algorithmic Trading (AT) price discovery process. This study serves as a pioneering effort, uniquely analyzing the specific impact of AT on the efficient price discovery process within commodity markets. Our findings unveil a noteworthy influence of futures market activity from AT, signifying a more substantial role in the permanent price discovery process. Importantly, our study establishes a robust and heightened relationship between AT and the overall price discovery process.

Furthermore, our contribution extends to documenting the effects of AT, allowing for a nuanced and direct comparison with those attributed to High-Frequency Trading (HFT). While prior research has highlighted similarities in the effects of AT and HFT, such as enhanced liquidity and increased price discovery, our study offers a distinctive perspective. Through a deliberate contrast of the findings on AT with those on HFT, notably referencing the work of Brogaard et al. (2014), our research enriches the literature on the heterogeneity of computerized trading strategies. This comparative analysis proves particularly insightful, shedding light on the unique characteristics inherent in these trading approaches.

High-frequency traders, characterized by their deployment of expensive low-latency technologies and engagement in a competitive race for milliseconds of advantage (Moinas et al., 2013), are typically adopted by a select group of sophisticated proprietary traders and electronic market makers. In contrast, AT technology finds wider adoption among buy-side funds and brokerages due to its cost-effectiveness and reduced reliance on intricate infrastructure. Consequently, AT tends to adopt longer trading horizons, emphasizing the efficient execution of liquidity-demanding trades. Our findings align with this perspective, illustrating that AT liquidity-demanding order flows exhibit a more positive correlation with future returns compared to non-AT, diverging from the observed relationship in HFT studies (Brogaard et al., n.d.).

In summary, our study contributes valuable insights to the nuanced understanding of the interplay between algorithmic trading strategies, market efficiency, and price discovery, particularly within the context of commodity markets.

The subsequent sections of this paper are structured as follows. Section 2 delves into an exploration of the relevant literature, providing context and insights. In Section 3, we expound upon the data utilized in our study present descriptive statistics, and elucidate the intertemporal relationships between Algorithmic Trading (AT) and the Base Metal market. Following this, Section 4 conducts an in-depth analysis of the effects of AT on price discovery within the context of Base Metal. Lastly, Section 5 encapsulates the paper with concluding remarks.

2. LITERATURE REVIEW

Our study contributes to the expanding literature on Algorithmic Trading (AT) price discovery. The theoretical framework established by Foucault et al., 2016 posits that fast traders, such as those engaged in AT, can integrate information by responding to short-term news and long-term price forecasts, distinguishing them from slower traders who are limited to long-term price movements. Hendershott et al., 2011demonstrate that AT enhances the informativeness of quotes,establish its positive impact on informational efficiency.

However, existing literature predominantly focuses on quote-driven markets, with a noticeable gap in evidence concerning algorithmic trading in emerging order-driven economies. Given the distinct market structures of developed and emerging markets, our study aims to fill this void by examining the impact of algorithmic trading on liquidity, volatility, and price discovery in an emerging market, specifically India.

Brogaard et al., delved into the High-Frequency Trading (HFT) realm, addressing its impact on market quality in the US equity market. While focusing on HFT, their study sheds light on aspects common to Algorithmic Trading. The research addresses various questions related to HFT, including market activity, determinants influencing buying and selling decisions, industry profitability, anticipatory trading, strategy correlation, herding behavior, and positive feedback loops.

Chaboud et al., (2009) to the literature by studying the impact of algorithmic trading on price discovery and volatility in the foreign exchange market. Their investigation delves into order persistence, discerning between human and computer-generated trades to identify their respective impacts on prices, volatility, and liquidity. The study extends the understanding of algorithmic trading to the forex market, addressing issues akin to those explored in the equity market.

This comprehensive review sets the stage for our study, positioned at the intersection of algorithmic trading and price discovery within the unique context of the emerging base metal commodity market scenario in India

3. DATA AND DESCRIPTIVE STATISTICS

This study leverages a unique dataset of daily base metal orders and trades from the MCX exchange in India, spanning eight years from 2015 to 2023. What sets this dataset apart is that each trade is tagged with an "AT intensity" label, indicating whether it originated from an automated trading (AT) system or not. This direct measure of algorithmic trading activity surpasses the limitations of prior research that relied on indirect proxies like message traffic or order clustering.

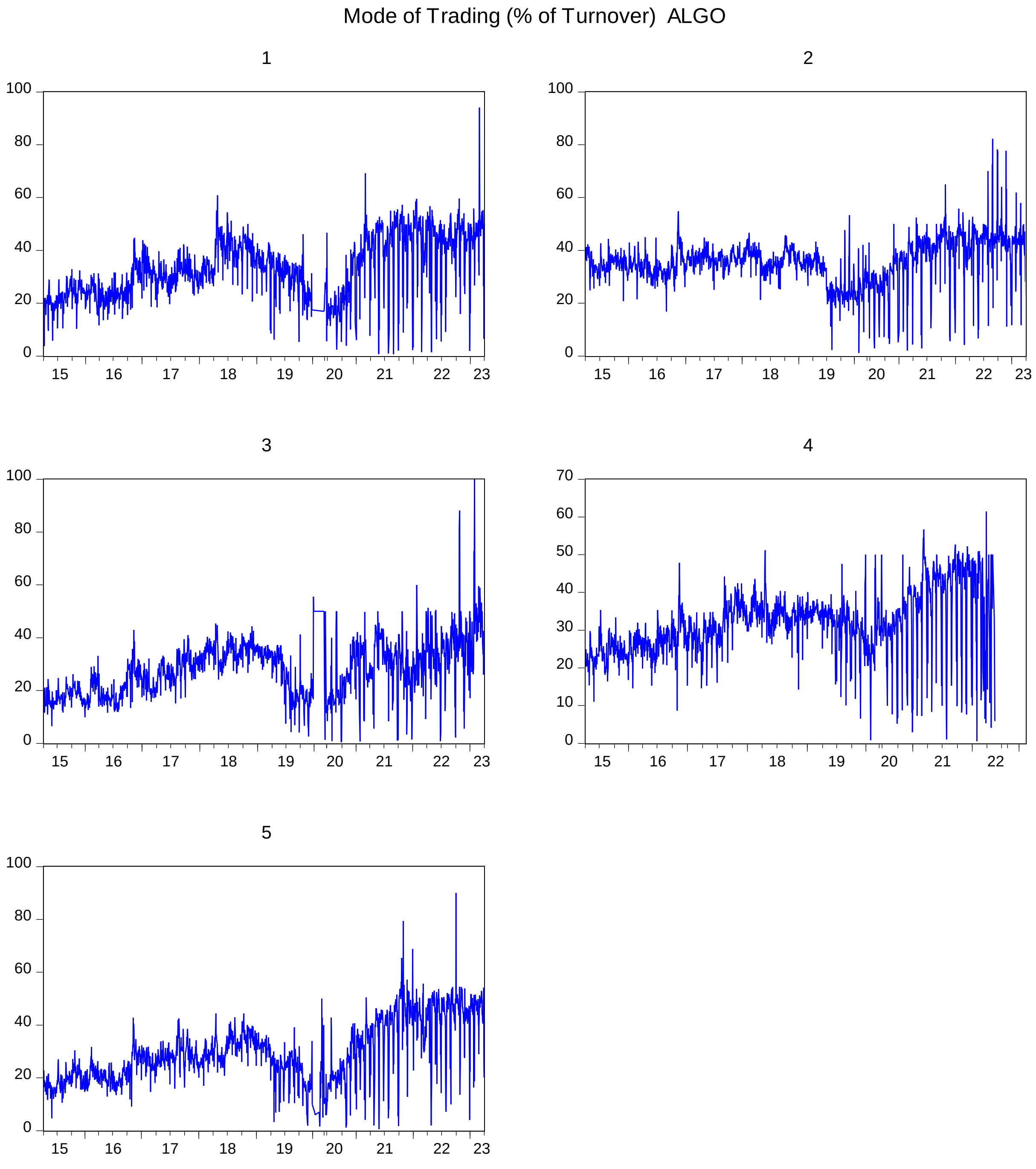
Traditionally, researchers have gauged the impact of AT through proxies like electronic message traffic (Hendershott et al., 2011; Bohemer et al., 2012) or by analyzing message clusters within 10-minute intervals (Hasbrouck and Saar, 2013). Another approach involved identifying exchange-labeled high-frequency trading firms (Brogaard, 2010; Brogaard et al., 2012; Carrion, 2013). However, this method only captured trades executed by a specific set of firms across 120 chosen securities, excluding a vast portion of the overall AT activity. Similarly, Hendershott and Riordan (2013) examined all AT orders on the German DAX exchange, but their analysis was limited to just 30 securities over 13 trading days.

In contrast, the MCX data employed in this study offers unparalleled comprehensiveness. It encompasses all base metal trades across the entire eight-year period, providing a detailed and nuanced view of AT activity within the Indian commodity market. This unique dataset empowers researchers to conduct more precise and insightful analyses of algorithmic trading's influence on market dynamics.

Figure 1 to 5 AT intensity between 2015 and 2023

These graphs depict the Algorithm-based Trading (AT) intensity within the overall base metal futures market at the Multi Commodity Exchange (MCX) from 2015 to 2023. AT intensity is measured as the percentage of daily trading value attributable to algorithmic trading activities compared to the total trading value for that day. Essentially, it shows how much of the base metal futures market activity at MCX is driven by algorithmic trading over these eight years

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk



In this investigation, we delineated two distinct periods for all Base Metal commodities utilizing five distinct figures (Figure 1, Figure 2, Figure 3, Figure 4, and Figure 5). These graphical representations illustrate the fluctuations in Algorithmic Trading (Algo) intensity spanning from April 2015 to February 2023.

In Figure 1, our observations indicate a period of low Algo trading intensity for Aluminium from April 2015 to March 2020. Conversely, from April 2020 to February 2023, Algo trading intensity exhibited a steady increase. Figure 2 highlights the Algo trading intensity for copper, with a low-intensity phase observed from April 2015 to August 2019, followed by a high-intensity period from September 2019 to February 2023. A similar pattern is observed in Figure 3 for Lead, where the low-intensity period spans from April 2015 to March 2020, and the high-intensity phase occurs from April 2020 to February 2023. For Nickel, as depicted in Figure 4, the low-intensity period extends from April 2015 to March 2020, followed by a high-intensity period from April 2020 to February 2023. Lastly, in Figure 5, focusing on Zinc, the quiet intensity period is noted from April 2015 to February 2020, and the high-intensity period emerges from March 2020 to February 2023.

This comprehensive analysis provides a lucid depiction of the dynamic shifts in Algo trading intensity across all Base Metal commodities. These insights will serve as a foundation for the subsequent examination of the impact of Algo trading on price discovery.

3.1 The Impact of AT on Price discovery: Fixed-effects panel regression

In our analysis, we examine the influence of Algorithmic Trading (AT) on price discovery by employing a fixed effect panel regression. The sample encompasses dates characterized by both low and high algo trading intensity during specific periods. Through this regression model, we aim to quantify the impact of AT on the process of price discovery across the given set of dates and periods under consideration.

$$PD_{it} = \alpha + \beta 1 AT_{it} + \beta 2\ TD_{it} + \epsilon_{it}$$

The variable PD_{it} represents the price discovery for commodity 'i' at time 't'. The term AT_{it} signifies algo trading intensity, while the time dummy variable takes the value 1 if time 't' corresponds to high algo trading intensity and 0 otherwise. The goal is to explore the relationship between algo trading intensity and price discovery through the specified model.

Fixed effects panel regression is advantageous for controlling unobserved individual-specific effects over time, addressing issues of omitted variable bias, endogeneity, and autocorrelation. By capturing time-invariant heterogeneity and providing more accurate estimates, fixed effects enhance the identification of causal relationships in panel data analysis.

The coefficient of interest, denoted as β1, represents the estimate of the treatment effect, specifically the impact of high Algorithmic Trading (AT) on price discovery. A significant β1 suggests that AT enhances price discovery, while a zero value implies no impact of AT intensity. The hypothesis test can be articulated as follows:

Null Hypothesis (H_0):

$H0: \beta 1=0$

Alternative Hypothesis (H_1):

$H1: \beta 1 \neq 0$

This formulation implies that the test aims to assess whether the coefficient β1, associated with the treatment effect of high AT, is significantly positive or negative, indicating a potential impact on price discovery.

3.2 Descriptive statistics

Table 1- Summary statistics of AT intensity in the low-at and high-at periods

All values in %

	Low Algo Trading	High Algo Trading
Mean	*28.59*	*34.22*
Median	*29.59*	*35.82*
Maximum	*60.83*	*100*
Minimum	*0.24*	*0.39*
Std. Dev.	*8.51*	*13.19*
Observations	*6072*	*3338*

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

Key observations from the summary statistics of AT intensity in the low-AT and high-AT periods:

- Increased AT activity in the high-AT period: Both the mean and median AT intensity is significantly higher during the high-AT period, suggesting a general upsurge in algo trading activity compared to the low-AT phase. This difference of 5.63% in means is statistically significant, implying a substantial shift in market dynamics.
- Potential bimodality in high-AT distribution: The higher standard deviation in the high-AT period hints at a potentially bimodal distribution, with both low and high-intensity strategies coexisting. Further analysis using histograms or kernel density plots would be valuable to confirm this and explore the relative prominence of each mode.
- Wider range and presence of diverse strategies in high-AT: The maximum AT intensity reaching 100% in the high-AT period indicates the presence of entities utilizing comprehensive algo trading strategies, while the lower maximum in the low-AT period (60.83%) suggests a prevalence of less aggressive or limited approaches. This highlights the wider range and diversity of algo trading strategies employed during the high-AT phase.

  While these observations provide valuable insights, a more in-depth understanding can be achieved by exploring potential relationships. Additional analysis could investigate the correlation between heightened Algo Trading (AT) intensity during the high-AT period and factors such as market price discovery.

3.3Measurement

In this research setting, we embark on innovative approaches to measurement and research design, aiming to achieve robust causal inference. Our methodological journey begins with the quantification of Algorithmic Trading (AT) intensity within the market. Subsequently, we proceed to employ measures derived from trades and orders data to assess market price discovery dynamics

3.3.1 AT intensity

The MCX provides daily data on algorithmic trading activity, expressed as a percentage, for all securities traded on the exchange. This data uniquely identifies both buy and sell orders executed by algorithmic traders, offering a comprehensive picture of their participation in the market.

3.3.2 Modelling Price Discovery

Consider a s base metal commodity traded in two adjoining and non-overlapping markets during a trading day. Market 1 represents the futures market, and Market 2 represents the spot market.

The closing prices of the futures and spot markets are denoted as p_{1t} andp_{2t} respectively. All prices are in logarithm. The open-to-close returns of the markets are defined as:

These returns are subject to market-specific price shocks $\eta 1t$ and $\eta 2t$. The first subscript i (where i=1,2) indicates the market (futures or spot), and the second subscript t (where t=1,2, 3…) indicates day. This study employs a Structural Vector Autoregressive (SVAR) model to analyse the dynamic relationships between base metal futures and the spot market. The returns for the two markets, denoted $rt=[r_{1t}, r_{2t}]'$, are modeled using the following SVAR process:

$$\sum \boldsymbol{B_0}\, \boldsymbol{r_t} = \sum_{k=1}^{k} \boldsymbol{B_k} \boldsymbol{r_{t-k}} + \boldsymbol{\eta_t}$$

Here, B_0 is a lower triangular matrix capturing the lagged effects, and η_t represents market-specific price innovations. The model assumes uncorrelated structural shocks $\eta 1,t$ and $\eta 2,t$ with normalized variances.

Reduced Form Representation: The SVAR model is transformed into a reduced form to facilitate analysis-

$$\boldsymbol{r_t} = \sum_{k=1}^{k} \boldsymbol{A_k}\, \boldsymbol{r_{t-k}} + \epsilon_t$$

This reduced form allows for a clearer understanding of the dynamic interactions between base metal futures and the spot market.

European Economic Letters
ISSN 2323-5233
Vol 14, Issue 2 (2024)
http://eelet.org.uk

3.3. Efficient Price and Information Shares

The study explores efficient price changes and information transmission using the SVAR framework. The efficient price (*mt*) is defined as the limit of the expected future prices:

$$mt = \lim_{q\to\infty} E(pt+q|Ft) = p_0+\iota' A(1)^{-1} B^{-1} \eta_t$$

Additionally, the daily change in efficient price (Δmt) is examined as a combination of structural shocks:

$$\Delta mt = \iota' A(1)^{-1} B^{-1} \eta_t$$

The cross-market distribution of information shares (*IS*(*SVAR*)) is computed based on the impact coefficients of structural shocks, providing insights into the relative contributions of each market to overall price dynamics.

Where m_{it} is the efficient price reflecting new information on economic fundamentals, and *uit* is a noise term resulting from transitory factors. Changes in the efficient price in markets 1 and 2 are denoted as:

$$\Delta m_{it} = m_{2t} - m_{1t}$$

These changes are independent of each other and of historical events, capturing the information components in price innovations $\eta 1t$ and $\eta 2t$.

The information flow in each market is measured by the variance of Δm_{it} where i=1,2. The information share (IS_{it}) of market i is defined as:

$$IS_{it} = var(\Delta m_{it}) / (var(\Delta m_{1t}) + var(\Delta m_{2t}))$$

This measure quantifies the proportion of total information variance attributable to each market. The higher the IS_{it}, the more influential the information flow from market i in determining the overall integrated variance (var(Δmt)).

In the context of a base metal commodity market, this framework allows for understanding how information is shared between the futures and spot markets, and how each contributes to the overall variance of price changes.

4. Analysis and interpretation

In the fixed-effect regression outlined in Section 3, we utilize the sample as input. The estimation is conducted for all the price discovery variables calculated at a daily frequency

4.1 Impact of Algo trading on price discovery

$$PD_{it} = \alpha + \beta 1AT_{it} + \beta 2\ TD_{it} + \epsilon_{it}$$

H1: Algorithmic trading impacts of price discovery of base metal futures market

variable	IS
α	4735.808*
β1	-115.041*
β2	-1157.395*
Model Stats	
R-squared	0.501

Hypothesis	**Variable $\beta 1$ (Coefficient)**	**Significance**	**Reject/Accept H0**
H1: Algo Trading will impact price discovery of base metal futures market	-115.041	0.00	Reject

4.2 Hypothesis Testing (H1):

- **H1.1: Algo Trading Impact on Price Discovery:**
  - Coefficient ($\beta 1$): -115.041, highly significant (p-value = 0.00).
  - **Conclusion:** Reject the null hypothesis (H0). There is evidence that algorithmic trading impacts the price discovery of the base metal futures market.

4.3 Implication:

Statistical Significance: The coefficient associated with algorithmic trading is statistically significant, indicating a significant adverse impact on the price discovery of the base metal futures market.

Practical Significance: Assess the economic significance of the coefficient to understand the magnitude of the impact of algorithmic trading on price discovery.

5. CONCLUSIONS

Over the past three decades, technological advancements have reshaped financial markets, with algorithmic trading (AT) becoming a prominent tool for trade execution on electronic exchanges. While initially praised for its efficiency and investor benefits, AT, especially in the realm of base metal commodities, is now subject to regulatory scrutiny. Despite a growing body of research, establishing a definitive causal relationship remains challenging.

The research design strategically identifies pairs of base metal commodities with similar fundamentals but varying levels of algorithmic trading. This approach, treating such pairs as a natural experiment, aims to discern differences in price discovery due to algorithmic trading. It acknowledges the complexity of the base metal commodity market and seeks a rigorous analysis considering unique characteristics and algorithmic trading intricacies.

Analyzing the regression results, the significant constant term implies an unexplained baseline effect on price discovery. The negative coefficient for the mode of trading turnover suggests a potential adverse impact on price discovery as turnover in this mode increases. The insignificant dummy variable for time indicates that temporal changes have no significant impact in this model.

References

1. Biais, B., Foucault, T., & Moinas, S. (2015). Equilibrium fast trading. Journal of Financial Economics, 116(2), 292-313.
2. Boehmer, E., Fong, K. Y. L., & Wu, J. (Julie) (2020, May 12). Algorithmic Trading and Market Quality: International Evidence. Journal of Financial and Quantitative Analysis. Retrieved from https://ssrn.com/abstract=2022034 or http://dx.doi.org/10.2139/ssrn.2022034
3. Chaboud, A., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2009). Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. International Finance Discussion Paper No. 980. Retrieved from https://ssrn.com/abstract=3884249
4. Foucault, T., Hombert, J., & Rosu, I. (2016, January 19). News Trading and Speed. Journal of Finance, 71, 335-382. HEC Paris Research Paper No. 975/2013. Retrieved from https://ssrn.com/abstract=2188822 or http://dx.doi.org/10.2139/ssrn.2188822
5. Hendershott, T. J., Jones, C. M., & Menkveld, A. J. (2010, August 30). Does Algorithmic Trading Improve Liquidity? Journal of Finance, 66, 1-33. WFA 2008 Paper. Retrieved from https://ssrn.com/abstract=1100635
6. https://www.mcxindia.com/market-data/reports-on-historical-data

Appendix 1

Summary of Returns: Returns in the Futures (rt1) and spot (rt2) markets are defined as 100*[ln(P_{close})–ln($P_{close-1}$)].

	Rt1	Rt2
Mean	8.97E-05	9.01E-05
Median	0.000222	8.55E-05
Maximum	1.427190	1.399343
Minimum	-2.931321	-2.885812
Std. Dev.	0.042624	0.042589
Skewness	-39.35523	-38.33429
Kurtosis	2923.030	2793.152
Observations	9475	9475

Appendix 2

Structural VAR Model. The number of lags K is determined by the Akaike Information Criterion.

Sample (adjusted): 4/06/2015 3/31/2023
Included observations: 9465 after adjustments
Convergence achieved after 11 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I

A =

1	0
C(1)	1

B =

C(2)	0
0	C(3)

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	-0.520783	0.007390	-70.47488	0.0000
C(2)	0.016596	0.000121	137.5863	0.0000
C(3)	0.011931	8.67E-05	137.5863	0.0000

Log likelihood 53849.48

Estimated A matrix:

1.000000	0.000000
-0.520783	1.000000

Estimated B matrix:

0.016596	0.000000
0.000000	0.011931