

Composite+[™]

Algorithmic Pricing Intelligence In Emerging Market Bonds



David Krein

Global Head of Research and Data

Julien Alexandre

Co-Head of Research

Introduction

Composite+ (CP+) is our AI-powered algorithmic pricing engine for corporate and Emerging Market bonds, combining both public (U.S. TRACE) and proprietary MarketAxess data. It is designed to support a variety of trading functions including pre-trade price discovery, liquidity provision, transaction cost analysis, auto-execution and crossing. It outputs an unbiased, two-sided market for more than 24,000 instruments globally. Updated every 15 to 60 seconds (depending on the liquidity of the instrument), the engine generates nearly 25 million levels per day covering 90-95% of trading activity in its markets.

In this paper, we will review CP+'s methodology, characteristics, and applications with a focus on Emerging Market (EM) corporate bonds. This transparency provides more specific insight into the CP+ design and potential roles in traders' workflows.

KEY DIFFERENTIATORS

MarketAxess global trading platform

- Data sourced from the MarketAxess trading system, which connects over 1,000 institutions across 70 countries to trade Emerging Markets debt.
- The only EM trading platform offering trading in both hard currency and 26 global local currencies with comprehensive coverage of all EM regions globally.

European trade data captured by Trax®

- Data sourced from over 1 billion cross-asset class transactions annually from 600+ entities, including approximately 12 million fixed income transactions.
- Covers nearly two-thirds (est.) of all EM transactions in Europe.

A.I.-powered pricing

- A robust machine learning algorithm, calibrated to the traded market.

Data and Construction

PROPRIETARY DATA SOURCES

MarketAxess Trading Platform

Inquiry data from the world's largest electronic marketplace

Trax

European Regulatory Reporting and Trade Matching engine, agnostic of execution source (voice and electronic)

Dealer Runs

Streaming dealer inventory pricing

PUBLIC DATA SOURCES

TRACE

FINRA-developed engine that facilitates the mandatory reporting of over-the-counter secondary market transactions in eligible fixed income securities

OVERNIGHT

Data Processing

Calibration

Algorithmically selected trades



Feature Engineering

Signal generation from aggregated content



Risk Mitigation

Error reduction through triangulation of bond features



Machine Learning

Training

Train GBM (Gradient Boosting Machine model) to predict the trade level on minimized Median Absolute Error (MAE) and Bias

Output

Prediction engine for next trading day



TRADING DAY

Output

Real-time bid and offer pricing predictions for over 24,000 bonds, updated every 15-60 seconds

CONSTRUCTION

CP+ works in two steps: (a) the prediction engine is trained daily using historical data and (b) the engine generates its predictions in real-time while incorporating all available intra-day information. All predictions are unique combinations of time, bond, and direction.

The prediction engine rests on a supervised machine learning algorithm. The supervisory variable is the price of all qualifying reported client-to-dealer TRACE/Trax trades. Qualifications help narrow the subset of TRACE/Trax trades to those that best represent dealer risk pricing for institutional clients in a competitive market. For example, we limit consideration to institutionally-sized trades \$150,000 or larger, and exclude seemingly non-principal risk trades; e.g. paired transactions for which there is a corresponding client-to-dealer trade of the same size at roughly the same time in the opposite direction. These conditions are essential because they define the calibration of the algorithm.

The prediction engine employs unique sources of bond trading data: earlier Trax and TRACE trades, indicative bond price data streamed by dealers and RFQ responses sent by liquidity providers via the MarketAxess trading platform. RFQ responses include trade levels, cover levels and all levels on inquiries that Did Not Trade (DNT).

We begin by creating a set of features, best defined as explanatory signals that aggregate the data from the past and the present to help predict the future. Each feature has a different predictive power that will depend on the vicinity to the prediction.

The most proximate features include TRACE/Trax traded prices, indicative levels and RFQ responses from the same day on the same side for the bond of interest. This information is valuable in generating the prediction, but the narrow constraints – a specific bond on a specific side on a specific day – may limit the coverage.

Additional subsets of features include progressively relaxed constraints. For example, features variously consider activity on the opposite side and in the inter-dealer market; and activity in other bonds with matching attributes. This last category allows features to incorporate activity in bonds with the same issuer, sector, rating and maturity. By aggregating a broad and diverse set of data sources, including MarketAxess' proprietary trading platform activity and the TRACE/Trax dataset, it predicts subsequent trade levels with minimal error.

Even when the most proximate signals are missing, the algorithm will learn to triangulate the less predictive features successfully. This approach leverages all the available information to increase coverage without impacting quality.

Overall, the algorithm consumes more than 200 features since it can make good use of even low information content factors to improve its prediction. The process also calls an outlier detection mechanism to identify and exclude off-market levels.

Each night, the features are fed into a tree-based machine learning algorithm called Gradient Boosting Method (GBM). GBM provides better accuracy than other algorithms like deep learning and neural networks since it is better at identifying the signal within noisy bond market data.

“Overall, the algorithm consumes more than 200 features since it can make good use of even low information content factors to improve its prediction. The process also calls an outlier detection mechanism to identify and exclude off-market levels.”

– DAVID KREIN, GLOBAL HEAD OF RESEARCH

COVERAGE

MARKET SEGMENT	AVG # OF BONDS	% OF TRACE/TRAX PRINTS
Asia Corporates (USD)	1,859	88.9%
Asia Sovereigns (USD)	158	92.6%
EMEA Corporates (USD)	450	94.3%
EMEA Sovereigns (USD)	272	98.9%
LATAM Corporates (USD)	802	95.1%
LATAM Sovereigns (USD)	213	97.0%
MENA Corps and Sovs (USD)	426	95.0%
EUEM Corps and Sovs (EUR)	435	96.5%

February 2019 Data
Source: MarketAxess Research

BENCHMARKING

The CP+ algorithm – designed to predict, with minimal error, the price of qualified client-to-dealer TRACE/Trax trades that are \$150,000 or larger – can easily be tested out-of-sample with historical TRACE/Trax data.

Specifically, we benchmark TRACE/Trax traded prices above \$150,000 to the CP+ prediction 15 minutes prior. We use a 15-minute buffer because TRACE prints have to be reported within 15 minutes; further, when an RFQ is underway on MarketAxess, dealers can modify their inventory due to the new information they received from the existence of the inquiry. (This information shouldn't be taken into account when evaluating the accuracy.)

Further, we can benchmark those same TRACE/Trax traded prices to the TRACE/Trax dataset itself, using the most recent, or last, TRACE/Trax traded price up to 60 days prior with the same constraints.

First we define the delta (Δ).

$$\Delta = \text{Price}_{\text{trade}} - \text{Price}_{\text{CP+}/\text{last trade}} \text{ if Dealer Sells}$$

$$\Delta = \text{Price}_{\text{CP+}/\text{last trade}} - \text{Price}_{\text{trade}} \text{ if Dealer Buys}$$

For example, if a dealer sells a bond at \$99.70, while CP+ Offer is at \$99.80, then the Δ for this trade is negative 10 cents; if a dealer buys a bond at \$99.70 and CP+ Bid is at \$99.80, then the Δ for this trade is positive 10 cents.

The error (Median Absolute Error or MAE) is defined as the median of all absolute distances (Δ).

Benchmark Performance

	TRADE COUNT	CP+ MEDIAN ABSOLUTE ERROR (MAE)	LAST PRICE MEDIAN ABSOLUTE ERROR (MAE)
--	-------------	---------------------------------------	--

MARKET SEGMENT

Asia Corps (USD)	17,170	0.14	0.17
Asia Sovs (USD)	1,824	0.14	0.20
EMEA Corps (USD)	5,481	0.13	0.20
EMEA Sovs (USD)	5,580	0.12	0.22
LATAM Corps (USD)	22,535	0.17	0.23
LATAM Sovs (USD)	4,932	0.16	0.26
MENA Corps and Sovs (USD)	5,643	0.14	0.21
EUEM Corps and Sovs (EUR)	3,651	0.10	0.15

LIQUIDITY SCORE (USD)

No Score	9,156	0.14	0.21
0 to 3	7,621	0.19	0.25
4 to 5	11,127	0.16	0.24
6 to 7	13,754	0.15	0.20
8 to 10	22,434	0.13	0.19

SIDE (USD)

B	29,678	0.15	0.21
S	34,414	0.15	0.20

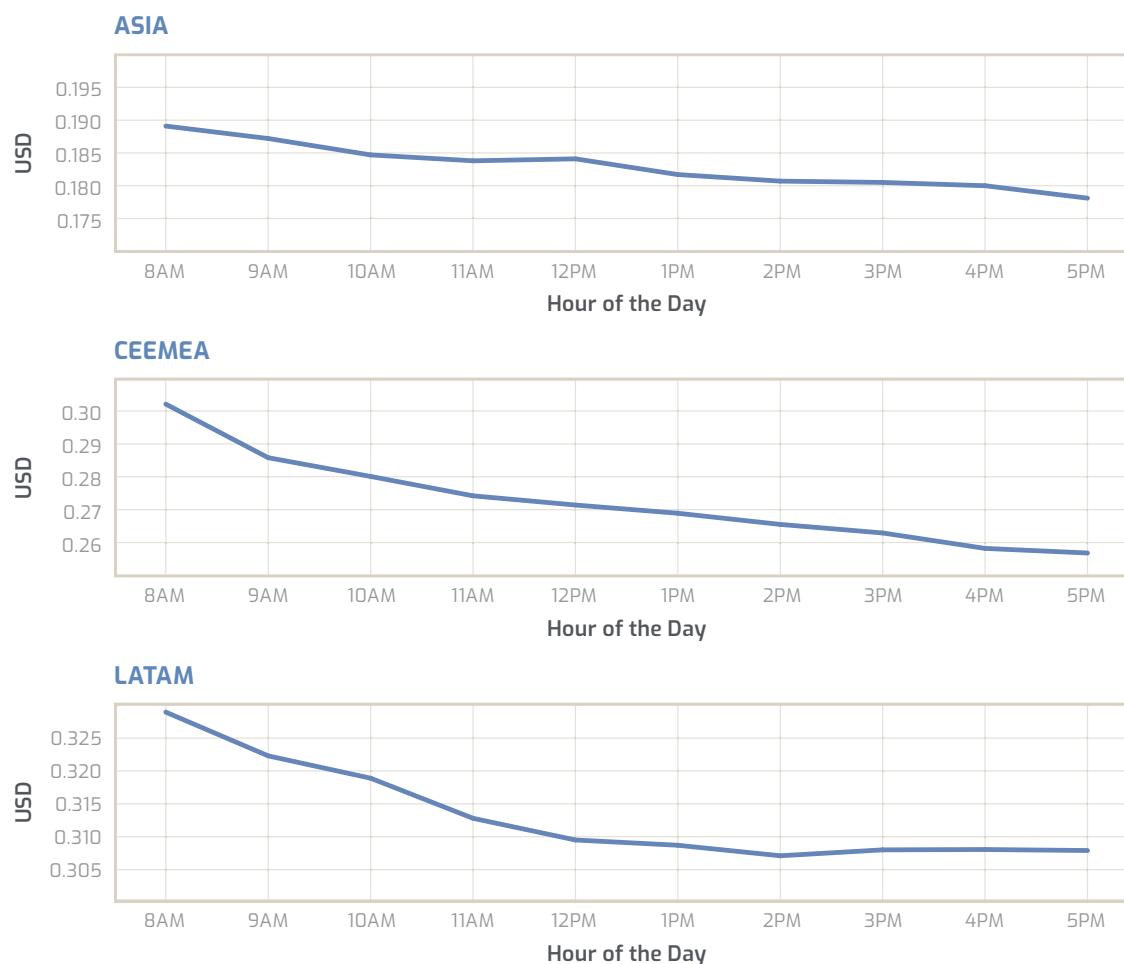
SIZE BUCKET (USD)

150k – 1 MM	45,984	0.15	0.21
> 1MM	18,108	0.15	0.20

February 2019 Data
Source: MarketAxess Research

APPLICATIONS

Intraday Bid-Ask Spread



February 2019 Data

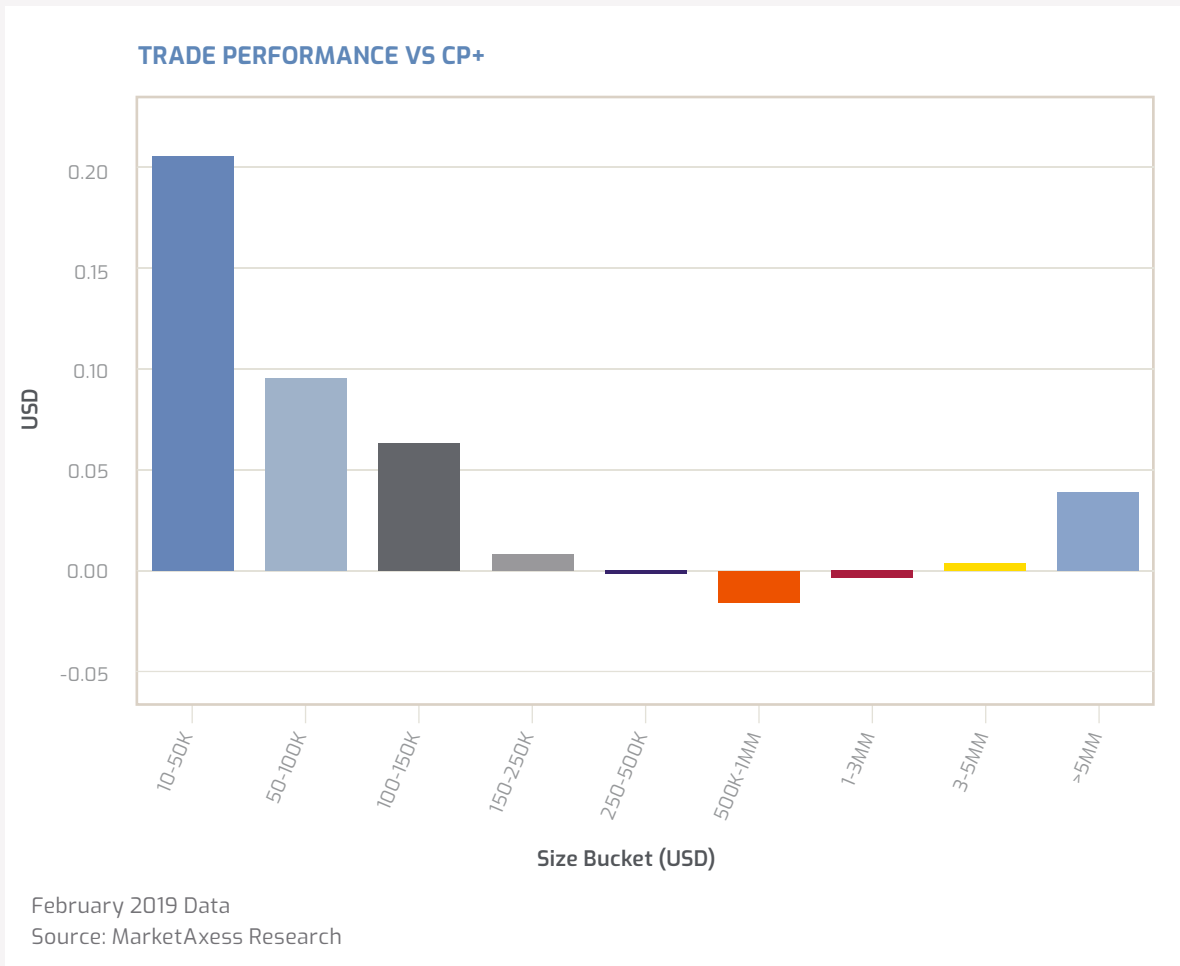
Source: MarketAxess Research

The time for each region is local time. We use Singapore for Asia, London for CEEMEA, New York for LATAM.

The graph shows the hourly bid/ask spread for Emerging Market securities in February 2019 measured in USD cent terms. The universe of bonds is held constant on a daily basis. We see that the hourly bid/ask decreases throughout the day as the trading information increases.

This is consistent with expectations. In early trading, there is little information available to the market, and by extension the CP+ model reflects that risk with wider bid-ask spread. As the day goes by, more traded prices and RFQ responses come into play and we observe the bid-ask spread tighten accordingly. The decline from high to low is about 1 cent for Asian bonds, 4 cents for CEEMEA and 2 cents for LATAM.

Transaction Cost Analysis

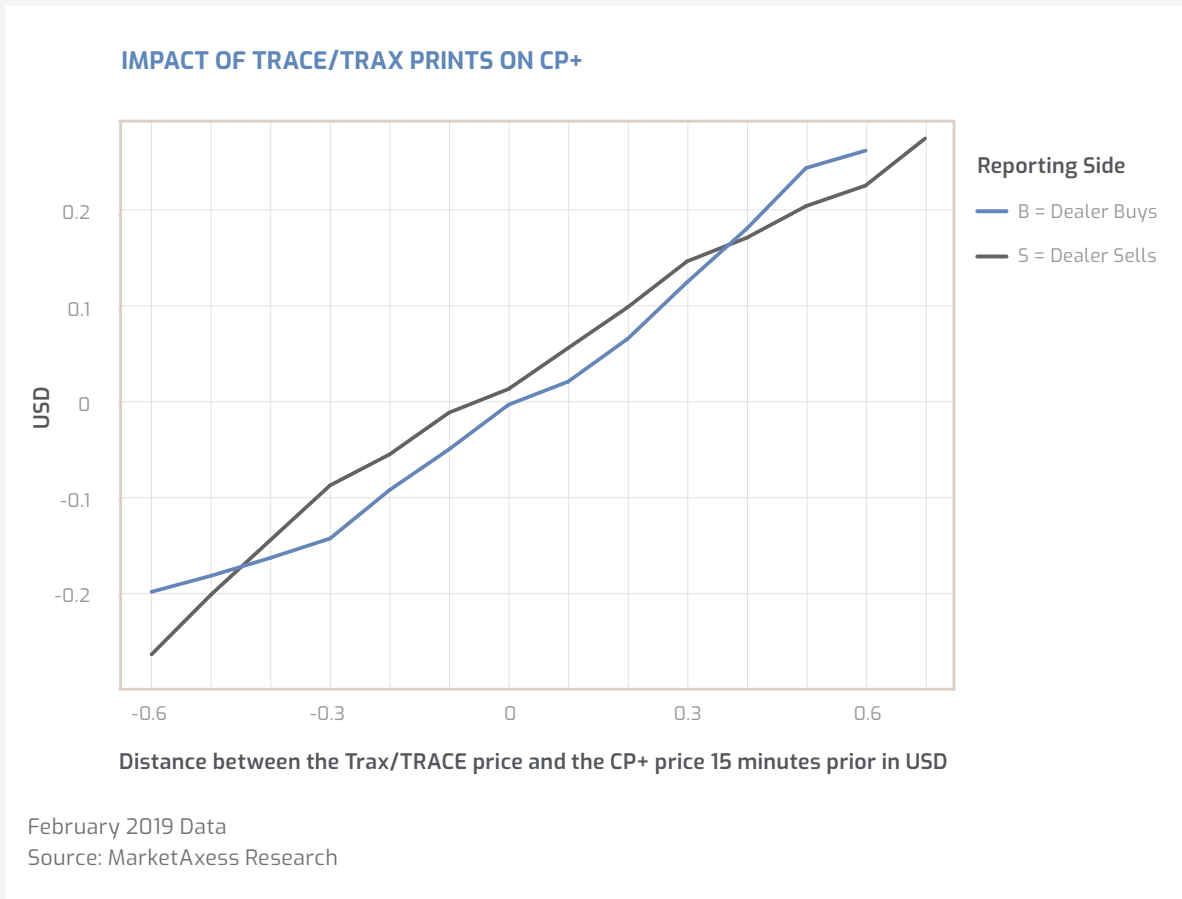


The graph shows the transaction cost for Emerging Market (USD) securities by size bucket. Transaction cost is the distance between the trade price and the CP+ price on the corresponding side. A negative value on the y-axis means that the trade was inside (more competitive than) CP+. A positive value means that it was outside (less competitive than) CP+. Each size bucket's value in the chart is the median of all observations. For example, if a dealer or liquidity maker sells a bond at \$99.7, while CP+ Offer is at \$99.8, then the dealer's transaction cost is negative 10 cents;

if a dealer buys a bond at \$99.7 and CP+ Bid is at \$99.80, then the dealer's transaction cost is positive 10 cents.

The graph demonstrates that trades under 150k are priced wider than CP+ while trades above 150k are generally in-line with CP+. This confirms the assumption that institutional-sized trades start at 150k and that calibrating the algorithm on these trades only, reduces the bias while including the maximum number of realized samples.

Price Impact



Characterized as the movement of CP+ after a trade hits Trax or TRACE, the price impact graph above reveals the Bayesian nature of CP+. On the x-axis, we have the distance between the TRACE/Trax price and the CP+ price 15 minutes prior. On the y-axis lays the price impact, defined as the median difference between the CP+ 15 minutes after the trade and the CP+ 15 minutes prior.

A single trade can't be the only source that sets the price of a bond. Every other meaningful data point, even those less proximate to the prediction, will have an influence on CP+. Consequently, when a trade hits TRACE/Trax and deviates from CP+ by Δ , the impact on CP+ will be less than Δ . The algorithm still weights the other features when adjusting its price.

Conclusion

- Thanks to the latest AI techniques, CP+ functions to transform proprietary and public data points into an unbiased 2-way reference price for more than 24,000 bonds including 4,500 Emerging Market bonds.
 - By aggregating a broad and diverse set of data sources, including the MarketAxess trading platform, Trax, and TRACE, it predicts subsequent trade levels with minimal error.
 - CP+'s calibration on risk-based Trax and TRACE trades of \$150,000+ gives traders a very robust, precise and consistent price point.
-

For more information, please contact research@marketaxess.com



MarketAxess®

MarketAxess.com | ©2019 MarketAxess Holdings Inc. (the "Company"). MarketAxess, the MarketAxess logo, Trax, Composite+ and "Now you're in the market" are trademarks of the Company. MarketAxess Corporation is a member of FINRA and SIPC. MarketAxess Europe Limited and MarketAxess Capital Limited are authorised and regulated by the UK Financial Conduct Authority. MarketAxess Singapore Pte. Limited is recognised by the Monetary Authority of Singapore. This information is not intended to represent an offer or solicitation of any financial instrument.

INFORMATIONAL PURPOSES ONLY. This material is for informational purposes only. It is not intended to form the basis of any investment decision, should not be considered a recommendation by MarketAxess or any other person and does not constitute an offer or solicitation to enter into a transaction with respect to the purchase or sale of any investment or security, nor is it a confirmation of terms. Sources of information are believed to be reliable, but MarketAxess makes no representation and gives no warranty that the information contained herein is accurate or complete. Any opinions or estimates given may change. MarketAxess undertakes no obligation to provide any **additional information or any update to or correction of the information contained herein.**