OFFICE OF FINANCIAL RESEARCH

STAFF DISCUSSION PAPER

No. 2014-03 | September 18, 2014

A Flexible and Extensible Contract Aggregation Framework for Financial Data Stream Analytics

Bryan Ball
New York University
ball.bryan.b@gmail.com

Mark D. Flood
Office of Financial Research
mark.flood@treasury.gov

H.V. Jagadish
University of Michigan
jag@umich.edu

Joe Langsam
University of Maryland
jlangsam@umd.edu

Louiqa Raschid
University of Maryland
louiqa@umiacs.umd.edu

Peratham Wiriyathammabhum
University of Maryland
peratham@umiacs.umd.edu

A Flexible and Extensible Contract Aggregation Framework (CAF) for Financial Data Stream Analytics

Bryan Ball
New York University
ball.bryan.b@gmail.com

Mark D. Flood
Office of Financial Research
mark.flood@treasury.gov

H.V. Jagadish
University of Michigan
jag@umich.edu

Joe Langsam
University of Maryland
jlangsam@umd.edu

Louiqa Raschid
University of Maryland
louiqa@umiacs.umd.edu

Peratham Wiriyathammabhum
University of Maryland
peratham@umiacs.umd.edu

ABSTRACT

The paper presents the Contract Aggregation Framework (CAF) for the modeling and analysis of data streams representing arbitrary financial contracts, ranging from privately negotiated deals to exchange-traded securities. We discuss the need for a flexible and extensible data model and provide an exemplar representing trading in corporate equities and bonds. Using a measure of *Market volume*, we review several analytical methods to explore the data. Initial observations support the benefits of the framework to integrate and analyze disparate sources of data.

Keywords

Data stream, Market volume, Financial contract, Extensible model, Financial analytics, Tensor decomposition.

1. INTRODUCTION

Modern financial systems generate complex information flows ranging from loans to public trading to large-scale capital movements. Systemic financial analyses and systemic risk metrics must combine financial data streams. However, even a seemingly straightforward alignment task, *matching corporate bonds with the equity of the issuer of the bond* can result in unmatched trades. The resulting impact (selection biases) on statistical analyses is unknown. A more significant example is a "lamppost" problem that extrapolates results from some financial subsector where granular data are available to another for which only aggregated data exist.

Aggregating over individual financial contracts has important implications. For example, the diversification effect implies that portfolio return variance is less than the sum for the component securities. While the investments literature has produced a vast wealth of econometric models to study aggregate portfolio returns, there is paucity of research on other aggregations, such as systemic metrics.

DSMM'14, June 27, Snowbird, Utah, USA

In this paper, we present the Contract Aggregation Framework (CAF) for the modeling and analysis of financial data streams. We build upon fundamental data modeling principles that were used in developing the relational data model and the data cube model [14]. The framework is motivated on the twin facts that data representing financial activity are available at widely disparate levels of temporal and cross-sectional aggregation, while empirical analyses are greatly facilitated by a well chosen standardization of the data. CAF is designed to capture financial contract data at arbitrary levels of granularity. We propose a flexible and extensible data model based on a basic *container* and operators to measure financial activity; we demonstrate using a measure of *market volume* [12, 13].

We exhibit the usefulness of CAF with an initial analysis of bond and equity data. This proof of concept demonstrates the utility of the CAF in integrating financial activity data from different asset classes, over levels of granularity and temporal aggregation. In addition, preliminary results from tensor decomposition allow us to identify tensor factors that correspond to potential *latent* patterns of co-trading of individual equities or bonds.

2. CAF DATA MODEL AND OPERATORS

2.1 Motivation

Data are critical to the functioning of the financial system. Message standards exist as the primary vehicle for communicating decisions, obligations, payments and other facts and commitments within and between participants. These standards and the implementation architectures are typically crafted to meet specific localized requirements. They are not designed to support system-wide data integration. Data therefore appear with a wide range of observation frequencies and aggregation levels, and are represented in a dizzying array of idiosyncratic formats. This lack of structure presents a serious challenge for important tasks such as monitoring financial supply chains, systemic risk management, and macroprudential supervision.

One goal of the CAF is consistent methodology for integration of legacy data across multiple levels of aggregation and data frequency. This includes relatively broad levels of aggregation, e.g., consolidated bank holding company financial statements [7], infrequent observation, e.g., annual

Start Date	July 1st, 2013
End Date	September 30th, 2013
Unique Equities	6795
Unique Bonds	20411
Industry Groups	24
Count of Equity Trades	448163063
Count of Bond Trades	2698330

Table 1: Statistics for Bond and Equity Data

Form PF reports to the Securities and Exchange Commission [15]), and time series transactional data sources that capture individual trades or order flow. While data aggregation has always been a problem for market participants, the increasing pace of financial activity combines with a new-found emphasis on system-wide stability to make large-scale information integration a special challenge.

2.2 CAF Principles

- The basic building block, a *container*, is a multidimensional representation in which each cell corresponds to transactions in one (or more) financial contracts.
- CAF must be extensible with respect to additional dimensions. This supports the later integration of data from multiple streams.
- Mapping rules must be defined to populate the container. The mapping must be disjoint and provide complete coverage (where possible) of all elements from the event or data stream(s).
- CAF must accommodate a range of granularity, including along the temporal dimensions. It must support aggregation over multiple elements, as well as the ability to reconfigure or partition relevant segments of the basic container. Examples are used to illustrate these principles.

2.3 Exemplar and Dataset

Our initial exemplar represents data on trading in equities from the Center for Research in Security Prices (CRSP) and data on trading in corporate bonds from the Trade Reporting and Compliance Engine (TRACE) service of the The Financial Industry Regulation Authority (FINRA). This simple example illustrates how CAF can integrate multiple data streams, across levels of granularity and temporal aggregation, and create interesting subsets.

```
- Companies.
    - SUBSET listed at NYSE, Amex, NASDAQ, or
      NYSE Arca (CRSP universe).
        - SUBSET with equities trading activity.
           - SUBSET with outstanding bonds
             (TRACE universe).
             - SUBSET with bond trading activity.
  GROUP BY NAICS code.
```

FINRA, an SEC-registered self-regulatory organization to regulate the securities industry, provides real-time information on over-the-counter (OTC) trading in U.S. corporate bonds through TRACE. TRACE covers all OTC activity for investment grade, high yield, and convertible debt corporate bonds—over 99% of the total U.S. corporate bond market activity. CRSP, founded in 1960 by the University of Chicago Booth School of Business, provides a comprehensive database of historical stock market data, dating to 1926. This includes end-of-day data on all common stocks listed on the NYSE, Amex, and NASDAQ and features a number of different fields including NAICS industry code, exchange, price, closing bid and ask, share volume, and number of trades. TRACE and CRSP is available through the Wharton Research Data Services (WRDS).

The summary statistics of the data used to populate the prototype CAF container is in Table 1.[1] We integrate these two data streams by first aggregating bond activity to daily totals. We then match bond and equity data according to the NAICS code for each company.

2.4 Market Size Metrics

A fundamental task for systemic analysis is to summarize the "size" of financial activity and risk exposures on a comparable scale across markets and institutions. A key intent (among several) of the Dodd-Frank Act, for example, is "to end too big to fail" [16]. Researchers have approached the concept of financial size in many ways. Lo and Wang [13] list a dozen metrics for trading volume in the secondary market for corporate equities alone.

We define an individual financial contract as an atomic unit. Size measures are then functions (simple or weighted aggregations) over some relevant attributes of a singular contract or set of contracts contained in a *ContractSet*.[2] Size may be computed with respect to a point-in-time snapshot of a portfolio inventory, or over an interval of time, e.g., portfolio turnover.

Size measures will typically be user-defined, and CAF must support a primitive set of operators that can be used to define customized size measures. For our initial analysis, we consider two basic measures of size, namely *notional value*, recorded in the legal terms of the contract and *market value*, the price at which a contract changes hands. Financial contracts typically have an explicitly defined measure, e.g., *principal value* for debt instruments, *par value* for equities, and *notional value* for derivatives. We group these measures under the label *notional value*. In trading contexts, a price is negotiated whenever a contract changes hands, reflecting its current economic value to the participants. In cases of infrequent trading, participants typically use a formula or algorithm to estimate this unobserved price. We refer to both a transaction price or an estimated price as the *market value* of a contract.[3]

Size measures may be parameterized as follows:

[1]The daily number of trades is not reported for about 61% of equities in CRSP, so the actual number is higher.
[2]Aggregations of risk exposures will typically be non-linear, for example due to diversification effects [3, 4]. In certain cases, this may constrain the applicability of linear tensor techniques described below.
[3]Accounting standards typically use the term *fair value* [11].

- A singleton contract or a set, *ContractSet*.
- A temporal *Interval*. If the start and stop times of an interval are identical, then we have a point-in-time snapshot.
- A *SizingRule* to determine the size (volume or value) of the contracts occurring in the ContractSet.

Typical examples of size measures are as follows:

- *TradeCount*: Cardinality of the trades executed against the contracts in the ContractSet.
- *ValueNotional*: Sum of notional values extracted from contract-level attributes in the ContractSet.
- *ValueMarket*: Sum of market values extracted from contract-level attributes in the ContractSet.

2.5 Extensibility of the Framework

A key motivation of CAF is extensibility to additional sources and dimensions; we briefly discuss some extensions:

- Ownership and other relationships among the institutions that are counterparties to contracts.
- Contracts that reference other contracts, e.g., an equity call option is a contract that references an equity.
- Contracts that are composed of other contracts, e.g., mortgage-backed securities.

3. PRELIMINARY OBSERVATIONS ON EQUITY AND BOND VOLUMES

As a proof of concept, we consider (i) daily dollar-weighted market volume, defined as price per share (or bond) times number of units traded, aggregated to daily totals, and further aggregated by industry category according to the first two digits of the 6 digit NAICS code. We use the NAICS code associated with the corporate equity. Forty four groupings (22 each for bonds and equities) remained after filtering out sparsely populated categories; due to space limitations we do not provide the 22 NAICS code labels.[4] We also use (ii) the proportion of market volume for that NAICS code and (iii) market volume adjusted with the NYSE daily volume.

3.1 Key Observations from the Analyses

We note that due to the very limited sample period of data collection of 3 months, our analysis merely demonstrates the feasibility and utility of CAF; we do not attempt to produce conclusive econometric results.

- There is a strong common component to trading of all financial instruments; this leads to positive correlations in trading volume between all industries for bonds and equities.
- Correlations within the same asset class of different industries were positive. Typically, bond trading across a pair of industries better, compared to bond trading and equity trading in the same industry.
- Adjustments to compute the proportion of market volume, or normalize with respect to the NYSE daily volume, reduced the correlation across industry pairs within the same asset class. However, pairs with high (or low) correlation retained their high (low) correlation. These adjustments did not affect correlation across differing asset classes. Normalizing the market volume using the previous day's S&P return had no impact. Normalizing on the current day's S&P return had noticeable effects on the finance and insurance industry (FIN) and the construction industry (CON).
- Autoregressive models showed positive autoregressive coefficients for market trading volume for each industry and asset class using AR(1) models. They were not always statistically significant and they showed a large range of values. Equity volumes were less volatile from day to day and were easier to predict. The AR coefficient of an industry's bond trading volume was not indicative of how high its AR coefficient for equity trading would be and vice versa.

3.2 Methods and Results

An industry-by-industry correlation matrix (44x44, bonds and equities), shows a positive correlation in almost all cells. There is naturally a strong common component to market-wide trading volume underlying these mostly positive correlations. Average industry correlations within the bond market were slightly higher than in the equity market. While still positive, average bond-equity correlations were significantly lower. Correlation of bond trading volume with industry-matched equity trading volume were higher than the overall average bond-equity cell, but lower than the average bond-bond or equity-equity cell. In other words, bond trading in a given industry correlates better with bond trading in another industry than equity trading in the same industry. This suggests that trading volume is driven less by firm-specific news, which should affect bonds and equities of a firm simultaneously, and more by factors, such as portfolio rebalancing, specific to the asset class (bond or equity).

It is instructive to consider the correlation outliers. For bond volumes, industries 21:OGM, 33:MCH, and 52:FIN are the most correlated with the other bond sectors. These are all capital-intensive industries, so we may be seeing the effect of the business cycle playing out in volumes; this warrants deeper investigation. Bond volumes for industries 11:AGF and 49:WAR show low correlations with other industries' bond trading, and their equity trading showed low correlations with other industries' equity trading. Bond trading in industry 56:WST was the least correlated with other sectors' bond volume. On the equities side, industry 45:REC correlated poorly with other industries' equity trading.

Figure 1 plots the eight most highly correlated bond industries by aggregated daily trading volume using a log scale on the y-axis, i.e., the eight industries involved in the seven highest correlation pairs. Note that many industries are in-

[4] A table of these three-letter codes is available upon request.

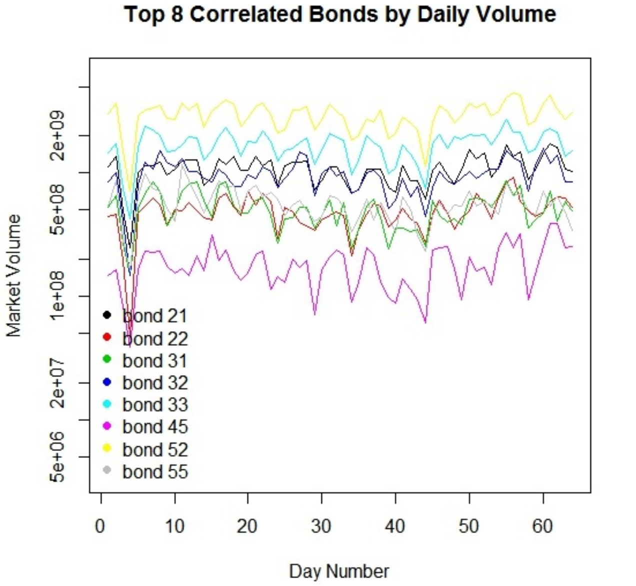

Figure 1: Top Correlated Bonds by Daily Volume (log scale) – 21:OGM; 22:UTL; 31:EDB; 32:CHM; 33:MCH; 45:REC; 52:FIN; 55:MGT.

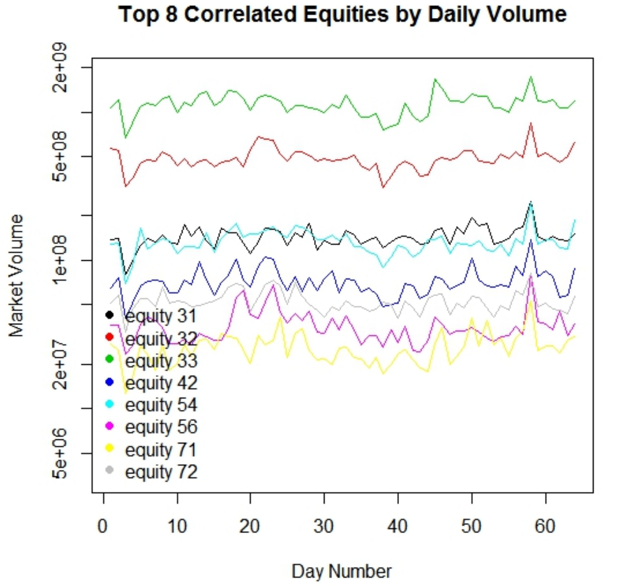

Figure 2: Top Correlated Equities by Daily Volume (log scale) – 31:EDB; 32:CHM; 33:MCH; 42:WHL; 54:PRO; 56:WST; 71:ENT; 72:TRV.

cluded in multiple pairs. The pairs are (with a "B" prefix to indicate bonds):

```
(B52:FIN-B21:OGM)
(B33:MCH-B22:UTL)
(B52:FIN-B31:EDB)
(B33:MCH-B32:CHM)
(B52:FIN-B33:MCH)
(B33:MCH-B45:REC)
(B52:FIN-B55:MGT)
```

Again here we see the dominant presence of sectors 21:OGM, 33:MCH, and 52:FIN as highly correlated.

Similarly, Figure 2 plots the eight most highly correlated equity industries by aggregated daily trading volume using a log scale on the y-axis. The pairs are ("E" prefix to indicate equities):

```
(E71:ENT-E31:EDB)
(E42:WHL-E32:CHM)
(E42:WHL-E33:MCH)
(E32:CHM-E54:PRO)
(E54:PRO-E56:WST)
(E56:WST-E72:TRV)
```

Notably, we see a quite different subset of the industry sectors dominating the correlations. Sector 56:WST, which is the least correlated among the bond sectors, even appears twice among the high-correlation equity pairs.

As a simple control for general market activity, we regressed the daily trading volume in each industry for both bonds and equities against the NYSE trading volume, percent change in the 2-year U.S. Treasury yield, and percent change in the 10-year Treasury yield. We then recalculated the correlation matrix for the regression residuals, which we call "adjusted volume." Since NYSE trading volume is strongly correlated to aggregate equity trading volume, this filtering removed most of the general market-related trading activity for equity categories. The equity-equity correlations averaged across all industry sectors dropped when considering the adjusted volume. 51 of the (22*21/2=231) equity-equity correlations showed a fairly large change in absolute value and 58 pairs change sign. The impact of adjusted volume was much weaker for bonds.

Interestingly, when considering bond-equity adjusted volume correlation, the average correlation shows an increase. Volume tends to cluster in generally high-volume days for bonds, and separate high-volume days for equities. This fact tends to reduce the average correlations in raw bond-equity volumes, because volume variance is tied less to industry specific news and more to overall activity in bond or equity markets. Filtering out these general factors partially removes this effect, exposing the industry-level patterns in the residuals. While the impact of filtering was sharply different for average equity-equity versus bond-bond correlations, the same industry pairs with the lowest raw-volume correlations also had the lowest adjusted-volume correlations.

Contemporaneous industry-matched bond-equity correlations were highest (exceeding 0.400) for the following industries: 23:CON, 32:CHM, 33:MCH, 48:TRN, 56:WST, 62:HLT, and 71:ENT. At the other extreme, industry-matched correlations were lowest (below 0.060) for the inventory-intensive industries: 11:AGF, 21:OGM, and 49:WAR.

Based on the assumption that large changes in market volume in one day may lead to increased trading activity in the following day, we tested the hypotheses that the previous day's S&P return might be a good feature in the regression to create adjusted volume This hypotheses did not hold; the effect was negligible for all industries. A regression based on the current day's S&P return showed that there was a larger effect for the finance and insurance (FIN) and construction (CON) industries.

Finally, we rerun the correlations using the proportion of daily bond volume within each industry, and similarly for equities. This pushes the average correlation coefficients close to zero and restricts any particular industry from correlating strongly with all the others. On the other hand, it emphasizes those industry pairs that tend to trade jointly, after conditioning over all other factors. Bonds from industries B49:WAR and B53:REL showed a high correlation. Similarly, equities from industries E32:CHM and E62:HTL showed high correlation in proportionate volume.

We also examined correlations versus lagged volume. There were strong correlations along the diagonals, indicating activity persistence, meaning that volume in a particular industry on one day is predictive of its trading volume the following day. To capture this more formally, we modelled log volume using three different autoregressive moving-average (ARMA) time series models.[5]. We tried three models: AR(1), MA(1), and ARMA(1,1). The AR(1) models appeared to perform well overall.

Equity volumes were more consistent day to day and also easier to predict. Equity AR(1) coefficients ranged from 0.160 for E51:INF to 0.646 for industry E48:TRN. All were positive and most were statistically significant. For bonds, they ranged from 0.084 for industry E54:PRO (Professional, scientific, and technical services) to 0.679 for industry E51:INF. Again, all were positive and most were statistically significant. It is interesting that the information sector had the highest autoregressive coefficient for bonds, but the lowest for equities.

These results are preliminary, and there are many obvious ways to strengthen the statistical analysis. Nonetheless, even this simple and coarse correlation study reveals some of the possibilities for systematic integration of consistent financial activity measures to surface intriguing patterns and promising questions for further research. We expect the power of this method to improve as we continue to build out both the data and tooling in the CAF.

4. TENSOR DECOMPOSITION

Given the multi-dimensional representation associated with the basic financial container, tensor decomposition appears to be a relevant analytical method that can be applied to identify potential *latent factors*, e.g., the co-trading of groups of equities or groups of bonds. Tensor factors will cluster individual elements, e.g., corporate equities or corporate bonds, across one or more industry sectors, in some specfic time intervals. We note that this is preliminary analysis to show the potential of this method.

Consider a trade against a bond *CorpBond* or an equity

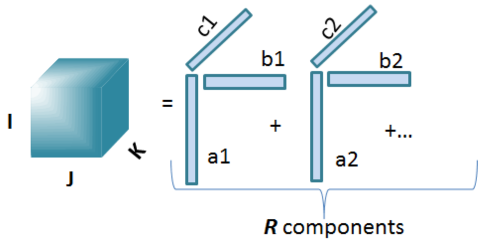

Figure 3: Example of CP Decomposition

$$\mathbb{X} \approx \sum_{r=1}^{R} \lambda_r \times a_r \circ b_r \circ c_r$$

$$= [\![A, B, C]\!] \qquad \begin{array}{l} A \in \mathbb{R}^{I*R} = [\![a_1, \ldots, a_R]\!] \\ B \in \mathbb{R}^{J*R} = [\![b_1, \ldots, b_R]\!] \\ C \in \mathbb{R}^{k*R} = [\![c_1, \ldots, c_R]\!] \end{array}$$

Figure 4: R Components from the CP Decomposition

CorpEquity associated with a company *CorpEntity* in an industry sector *IndustrySector*. Let *DataTime* represent the time of the trade, *CorpEqTrade* or *CorpBondTrade*. We can construct the following tensors to capture this information where each element of the tensor will represent a market size metric. We note that the third tensor combines the data from the first two tensors and that this is only one among many alternate representations:

- (DateTime, CorpEntity, IndustrySector, CorpEquity, CorpEqTrade)
- (DateTime, CorpEntity, IndustrySector, CorpBond, CorpBondTrade)
- (DateTime, CorpEntity, IndustrySector, (CorpBond | CorpEquity), (CorpEqTrade | CorpBondTrade)).

Among the many methods for tensor based analytics [6], we consider a class of closely related constructions, known collectively as CP decomposition, named after the two most popular and general variants, CANDECOMP and PARAFAC [10]. Such decompositions represent a tensor as the sum of the N-fold outer products of Rank-1 tensors, where N is the dimension of the original tensor indices. Figures 3 and 4 illustrate an example of the CP decomposition to produce R tensor factors, where each factor is represented by 3 *Rank-1* tensors along the I, J and K indices of the original 3 dimensional tensor.

Figure 5 illustrates the Top 3 elements of 3 *Rank-1* tensors, *DateTime*, *CorpEquity* and *IndustrySector*. There are 3 equities, `SPY`, `BAC` and `EEM` that occur in this factor and all are in the NAICS section 52 (Finance, Insurance). Figure 6 and 7 plot the equity prices and trading volume reported by the site *Yahoo! Finance* for `SPY` and `EEM` in a corresponding time interval. It is notable that there appears to be a strong correlation of activity around these two equities on these dates.

Score	Date
0.025852	9/18/2013
0.021639	8/27/2013
0.021223	9/20/2013

Score	CUSIP	Company
0.071315	78462F10	SPY-SPDR S&P 500 ETF Trust
0.063712	6050510	BAC
0.043936	46428723	EEM-iShares MSCI Emerging Markets Indx (ETF)

Score	NAICS	Industry
0.97295	52	Finance and Insurance
0.018698	33	Manufacturing
0.008349	51	Information

Figure 5: Tensor Factor Corresponding to Equity Trades

Figure 6: Equity Trading Price and Volume for SPY

Similarly, we identify a tensor factor comprising 4 corporate bonds. The tensor factor is summarized in Table 4. Again, using data from *Yahoo! Finance*, we observe a high correlation of activity for these four bonds, in the corresponding time interval.

5. RELATED WORK

Data cubes [14] represent data as measures or dimensions; they are similar to a tensor representation of an array that projects data from the database. The typical multidimensional analysis is to compute aggregated statistics, e.g., *How many equities have trading activity that exceeds X during a time interval, grouped by industry sector?*

The scale of financial activity is important in many contexts. Quantity transacted, for example, is a fundamental input to the most basic market phenomenon, the equilibration of supply and demand; [12, 9]. Aggregated volume and market turnover (volume divided by shares outstanding) are common heuristic measures of market liquidity; [2]. However, despite the significance of volume, it has not received

Date	Normalized weight for Date
20130912	0.08975
20130911	0.06533
20130913	0.04847
20130918	0.04743

Bond name	Normalized weight for Bond name
VZ 43	0.21082
VZ 23	0.19129
VZ 18C	0.10806
VZ 33	0.10754

Table 2: A tensor factor comprising 4 Verizon (VZ) corporate bonds with different maturity dates.

Figure 7: Equity Trading Price and Volume for EEM

nearly the level of research attention as price, particularly in the asset pricing literature; [13]. This is due in part to the fact that trading volumes are not always reported with the same high-frequency timeliness as market prices.

The scale of financial activity also matters at the systemic level. For example, it factors in discussions of aggregate imbalances [1], the *too-big-to-fail* problem [16], and the growth of shadow banking [8]. The CAF helps address the general issue of scale by providing a unifying framework for measuring and managing a very general class of size metrics across a very broad range of financial activity.

6. CONCLUSIONS AND FUTURE WORK

The Contract Aggregation Framework can model a wide range of financial activity measured at disparate levels of aggregation. We provide a proof of concept, applying a preliminary version of the CAF to bond and equity volume data. Our preliminary results demonstrate both the feasibility of data integration through the CAF, and its ability to identify interesting patterns and research questions. We further apply tensor decompositions to the data, to identify potential latent patterns of co-trading of individual equities or bonds. Future work includes the following: (a) Developing the CAF proof of concept into a more robust implementation. (b) Refining both the tensor and volume analytics. (c) Adding visualizations. (c) Populating the container with a much broader range of financial data streams.

7. ACKNOWLEDGMENTS

This research was partially supported by NSF grant CNS1305368 and the Smith School of Business.

8. REFERENCES

[1] L. Alessi and C. Detken. ‘real time’ early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. ECB Working Paper 1039, European Central Bank, March 2009.

[2] Y. Amihud and H. Mendelson. Liquidity and stock returns. *Financial Analyst Journal*, 42(3):43–48, 1986.

[3] B. Basel Committee on Banking Supervision. Developments in modelling risk aggregation. Technical report, BIS, 2010.

[4] B. Basel Committee on Banking Supervision. Principles for effective risk data aggregation and risk reporting. Technical report, BIS, 2013.

[5] G. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control (Third ed.)*. Prentice-Hall, 1994.

[6] C. Faloutsos, T. Kolda, and J. Sun. Mining large graphs and streams using matrix and tensor tools. In *Proceedings of the ACM SIGMOD Conference*, page 1174, 2007.

[7] Federal Reserve. Fr y-9c: Consolidated financial statements for holding companies. Technical report, 2014.

[8] G. Gorton and A. Metrick. Regulating the shadow banking system. *Brookings Papers on Economic Activity*, 2010:261–312, Fall 2010.

[9] J. Karpoff. The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis*, 22:109–126, 1987.

[10] T. Kolda and B. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[11] C. Laux and C. Leuz. The crisis of fair-value accounting: Making sense of the recent debate. *Accounting, Organizations and Society*, 34(6-7):826–834, 2009.

[12] A. Lo and J. Wang. Trading volume: Definitions, data analysis, and implications of portfolio theory. *The Review of Financial Studies*, 13(2):257–300, 2000.

[13] A. Lo and J. Wang. Stock market trading volume. In *Handbook of Financial Econometrics, Volume 2: Applications*, pages 241–342. 2010.

[14] T. Pedersen. Multidimensional modeling. In *Encyclopedia of Database Systems*, pages 1777–1784. Springer, 2009.

[15] Securities and Exchange Commission. Form pf: Reporting form for investment advisers to private funds and certain commodity pool operators and commodity trading advisors. Technical report, 2011.

[16] U. S. Congress. Dodd-frank wall street reform and consumer protection act. Technical report, One Hundred Eleventh Congress of the United States of America, July 2010.