Interbank Lending Networks and the Urgency to Borrow

Celso Brunetti◊, Jeffrey H. Harris○ and Shawn Mankad□

Abstract

PRELIMINARY: PLEASE DO NOT QUOTE.

Key words: Banking Networks, Interconnectedness, Credit Default Swaps

JEL:

◊Federal Reserve Board (celso.brunetti@frb.gov); ○American University (jharris@american.edu); Cornell University (spm263@cornell.edu) corresponding author, this material is based upon work supported by the National Science Foundation under Grant No. 1633158 (Mankad).

The views in this paper should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. All errors and omissions, if any, are the authors’ sole responsibility.
1 Introduction

Network analysis has been proven as an effective tool for understanding financial markets. Correlation networks among stock returns are shown to reflect financial interconnectedness and crises (Billio, Getmansky, Lo and Pelizzon (2012)) and physical, interbank trading (borrowing/lending) networks are shown to forecast market liquidity problems (see Brunetti, Harris, Mankad and Michailidis (2018)).¹ In this paper we examine an alternative dimension of the physical overnight lending market--the network of liquidity demand/supply. In our setting, liquidity networks capture the urgency to trade while trading networks simply reflect borrowing and lending. We first demonstrate that liquidity networks constructed from executed trades reveal information that differs from information gleaned from simple trading networks. We then show that combining liquidity networks (the urgency to trade) with trading networks (borrowing/lending) reveals important information about regional and country-specific market conditions.

The fact that the liquidity network contains useful information beyond that embedded within the trading network is important. For example, Brunetti et al. (2018) show that trading networks directly reflect balance sheet flows in the interbank market. Therefore, one might expect information from liquidity networks (which contains no direct information about bank balance sheets) to be largely subsumed by the corresponding trading network.² While we generate both networks from the same set of overnight interbank transactions, we find that the

¹ Adamic, Brunetti, Harris and Kirilenko (2017) explore how physical trading networks can be used to forecast short-term market conditions as well.
two types of networks exhibit marked differences across time, primarily in the
time-series of the Largest Strongly Connected Component (LSCC) and
reciprocity.\(^3\) The LSCC, which measures aspects of liquidity and information
revelation, is systematically higher in the liquidity network than in the trading
network. Likewise, reciprocity, which measures the likelihood of banks to be
mutually linked, is systematically higher in the liquidity network as well.\(^4\) To the
extent that networks are used to understand and diagnose market activity, our
findings show that liquidity networks help provide a more complete
characterization of markets and their participating institutions.

Moreover, we find that each network contains unique information that
complements the other by identifying the urgency (liquidity network) and
direction (trading network) of trading among banks in the overnight lending
market. Leveraging these different dimensions, we combine liquidity and trading
networks into the log odds ratio as a new metric that simultaneously combines
information from both networks. The log odds ratio reveals the urgency to
borrow/lend aggressively/passively for each bank in the interbank market.\(^5\)

Importantly, the log odds ratio can be examined for each bank in the
market, and can also easily be aggregated across banks to characterize urgency to
borrow aggressively by country, by region (e.g., Core Europe versus Peripheral
countries), or for the entire market (here, all of Europe).\(^6\) In fact, grouping banks

---

\(^3\) LSCC is defined as the maximum number of traders that can be reached from any other trader by following
directed edges (see Adamic et al. (2017) and Brunetti et al. (2018)) with further details in Section 2.
\(^4\) Consequently, our results indicate a potential negative bias in studies that characterize
interconnectedness using the trading network LSCC and/or reciprocity.
\(^5\) The details behind the log odds calculation are explained further in Section 3 below.
\(^6\) Log odds ratios can be constructed for other markets as long as trade counterparties and standing quotes
can be identified.
by country, we find that the urgency to borrow Granger-causes corresponding CDS spreads for sovereign debt in countries where banks are domiciled. The Granger causal relationship is most often one-way—sovereign CDS spreads do not lead the log odds ratio for banks in the same country.

From a policy perspective, our findings imply that institutions and scholars should analyze both networks together when possible to better understand interconnectedness and liquidity flows. For example, our analysis reveals that banks from Ireland, Greece, Portugal, and the United Kingdom borrowed very aggressively at different points from 2008 onwards.

We validate our new methodology and statistic using a series of econometric models and note that our methodology departs from conventional financial network analysis paradigms by allowing edges to be heterogeneous. Specifically, each directed edge in our proposed network representation distinguishes both the buyer/seller and executing/quoting relationships between banks. Therefore, the degree of each bank becomes a vector that counts the total amount of aggressive buys, aggressive sells, passive buys, and passive sells for each bank.

For illustration, consider the hypothetical network below in Figure 1, where the banks are labeled A through E. In our trading network, Bank A would be a dominant seller. In our liquidity network, Bank A passively trades with all other firms by providing quotes that other firms execute against through market orders.

**Figure 1:** A hypothetical banking network, where the banks (nodes) are labeled A through E.
Our examination of liquidity networks expands the view of physical networks beyond trading networks (of buyers and sellers) that have been studied elsewhere in the financial and economics literature. Importantly, we integrate liquidity characteristics into network analysis in finance, an important contribution, given the importance of market liquidity and liquidity risk during the recent financial crisis and beyond. Our work is among the first in financial economics to systematically analyze the time-series of multiple network types that share the same nodes (here, banks). In addition to defining a new physical network, we introduce a compact form to represent data that yields a network statistic (the log odds ratio) specifically designed for financial markets and thus has a practical and meaningful interpretation.

Beyond these methodological contributions, our log odds ratio (which captures the urgency to borrow) provides a quantitative measure of the motivations behind traders who initiate trades. In this light, our work relates closely to Sarkar and Schwarz’s (2009) market sidedness measure (the correlation between the standardized number of initiated buys and standardized number of initiated sells) which can identify periods when traders demand immediate executions due to information revelation or trader impatience. However, our log

---

8 Since edges in networks from previous studies tend to be homogenous (i.e., defined by buy-sell relationships only or correlations only), the field has largely and successfully repurposed methods from social network analysis for the financial context.
odds ratio differs from market sidedness in important ways. First, market sidedness has been used exclusively to explore the motives for trade initiation in the stock market where the urgency to trade is driven primarily by trader impatience or information revelation. The urgency to trade in the interbank market more likely results from liquidity demands generated by bank business practices rather than exogenous information events. Second, the methodological underpinnings are also very different. Our approach builds on financial network analysis, while market sidedness is correlation-based. We compare our log odds ratio with market sidedness on the e-MID data and find that the log odds ratio is preferred for forecasting movements in CDS spreads—the log odds ratio more often Granger-causes corresponding CDS spreads for sovereign debt more often.

2 Trading and Liquidity networks

Any network can be represented canonically using a set of nodes and corresponding edges. We denote nodes as banks and edges as connections generated by executed trades between banks. This paper is the first to distinguish between trading networks and liquidity networks, which differ on how the edges are defined. Trading networks are defined by mapping lenders to borrowers. Specifically, if Bank B borrows from Bank A within the time interval of interest, then an edge is drawn from A to B. Liquidity networks are defined by trade initiation, so that the initiating bank is mapped to its quoting counter-party. If
Bank B initiates a trade with Bank A within the time interval of interest, then an edge is drawn from A to B.\textsuperscript{10}

After constructing the networks, there are several methodological approaches one could take to analyze both network structures. The simplest approach is to analyze each network separately with network statistics and, in a second stage, identify correlations between the two sets of results. One potential downside of this approach is that it can be difficult to measure interactions between the networks (e.g., how connected were Spanish banks that borrowed when initiating the transaction?). An alternative approach that addresses this challenge is to jointly model the networks. Note that each network is generally constructed from the same sequence of transactions so that both trading and liquidity networks capture behavior from the same set of banks with heterogeneous edges, allowing for a more nuanced analysis. We explore both approaches in this paper.

\textbf{2.1 Network Statistics}

In addition to sharing the same set of nodes, several commonly used network statistics will be identical between trading and liquidity networks. For example, aggregate statistics like the \textit{overall degree} of the network (total number of connections) remains fixed between the two networks. In general, any network statistic that ignores the directionality of the edges will be invariant between the two types of networks. Note that this includes measures, such as degree and \textit{clustering coefficient} (the percentage of closed triangles), which have been used to

\textsuperscript{10} Edges can also be optionally weighted by the size of the transaction. It is convention for multiple transactions between the same two banks in the time interval of interest to be combined into a single directed edge weighted by the total amount of all transactions.

Network statistics that account for directionality will tend to be different between the networks. For example, a bank’s in and out degree at the node-level will in general be different between the trading and liquidity networks. More importantly, the interpretation of these statistics change. For the trading network, in-degree represents the total amount borrowed on the interbank market, while for the liquidity network this statistic corresponds to the amount transacted (borrowed or lent) using market orders. Likewise, out-degree corresponds to limit orders on the liquidity network and lending on the trading network.

Another important network variable that has been shown to characterize interconnectivity is the Largest Strongly Connected Component (LSCC), which is defined as the maximum number of banks that can be reached from any other bank on the network by following the network’s directed edges. The last measure we highlight here is Reciprocity, which measures the likelihood that pairs of nodes are linked in both directions. Note that on the trading network both LSCC and Reciprocity will be closer to their maximum value of one when banks are both buying and selling and with a larger number of counter-parties. On the liquidity network, larger values of LSCC and Reciprocity correspond to multiple trades between the same pair of banks, where the initiator alternates. Thus, higher values of LSCC and Reciprocity on both networks indicate higher levels of intermediation. Values close to zero for these statistics indicate a highly fragmented market and a breakdown of interconnectivity.
2.2 Results

In this section we present time-series of network statistics for trading and liquidity networks constructed from the e-MID platform, the only electronic market for interbank deposits in the Euro region, offering interbank loans ranging from overnight (one day) to two years in duration, with overnight contracts representing 90% of total volume during our sample period (see Brunetti, diFillippo and Harris (2011)). The e-MID market is open to all banks admitted to operate in the European interbank market and non-European banks can access the market through their European branches. As of August 2011, the e-MID market had 192 members from European Union countries and the U.S., including 29 central banks acting as market observers (Finger, Fricke and Lux, 2013).

Our detailed trading data spans from January 2006 through December 2012 and includes 212 unique banks and 464,772 trades. Each e-MID transaction includes the time (to the second), lender, borrower, interest rate, quantity, and an indication of which party is executing the trade.

![Figure 2](image-url)

**Figure 2**: Financial statistics at the daily resolution from the e-MID interbank market.

Figure 2 shows several market statistics at the daily level for the e-MID. We see that interest rates fell starting with the onset of the 2007-09 financial crisis. Rates started to recover as the crisis abated, but in 2012 fell again to crisis levels as Europe experienced a weak recovery. Effective spreads remain relatively
stable across our sample period, suggesting that interbank market liquidity did not suffer appreciably during the crisis. On the other hand, there is a clear negative trend in both the number of active banks trading and daily volume. Signed volume is also negative throughout our sample period with a clear increasing trends towards zero. These patterns indicate that banks actively used the e-MID platform for selling funds, though by the end of our sample period liquidity levels are poor. Trade imbalance (scaled by volume) shows a greater proportion of aggressive lending during the 2007-09 crisis. During the weak recovery in Europe, trade imbalance even became positive for a handful of days indicating that banks were aggressively borrowing. Lastly, likely driven by the reduction in active banks, the Herfindahl index rises consistently over our sample period, reflecting greater concentration among banks using e-MID.

![Network statistics](image)

**Figure 3:** Network statistics corresponding to 30-day rolling liquidity (solid red) and trading (dashed blue) networks from the e-MID interbank market.
Next we present results from the network analysis to gain further insights into how the e-MID market evolved in our data. As shown in Figure 3, the degree and clustering coefficient resemble the trend in volume – dropping precipitously as the 2007-09 crisis unfolded. As the crisis abated, the clustering coefficient nearly recovered pre-crisis levels though, as Europe suffered a weaker recovery from the crisis, it fell again indicating lower levels of interconnectivity. Note also that the time-series for degree and clustering coefficient are identical for both networks, as discussed above, since the statistics are network-wide and undirected.

LSCC and Reciprocity show divergent values for each network, where the liquidity network has higher average values. While LSCC drops for both networks following the failure of Lehman Brothers, the liquidity network LSCC rises slowly to end near pre-crisis levels. In contrast, the LSCC for the trading network continues to fall. Similarly, we see that Reciprocity decreased following the failure of Lehman Brothers for both networks. Reciprocity for the trading network continued a steady decline, showing a nearly 50% drop during the span of our data. This decrease indicates that banks became less willing to both borrow and lend funds on the e-MID, instead preferring to trade in one direction as the crises unfolded. For the liquidity network, after the Lehman Brother failure, reciprocity recovers to a level above its pre-crisis starting point. Thus, banks that continued to utilize e-MID were willing to both initiate trades with market orders and reveal their preferences by providing quotes through limit orders.

Altogether we see consistent evidence that liquidity decreased significantly in the e-MID market. As the crisis unfolded, banks tended to initiate lending less and/or banks were less willing to reveal their desire to borrow with limit orders.
This led to a drop in overall activity and the decline of network interconnectivity in the e-MID. The trading networks became less dense and more fragmented by the end of our sample than they were at the start. On the other hand, in spite of the overall drop in activity, the liquidity networks show evidence that following the crisis, trust levels recovered and were high between banks that continued to operate in the e-MID.

Decomposing trading activity further and by type of bank would provide further granularity into the dynamics of the e-MID. For example, by participating in the interbank market, banks may (inadvertently) signal their relative health status through their trading and execution patterns. Thus, it is of interest to study whether a bank’s trades were more likely to be initiated borrowing, initiated lending, passive borrowing, or passive lending. Especially given the eventual European sovereign debt crisis, it is also desirable to segment the results by country or region of origin for the bank. In the next section, we present a novel network framework that combines both physical networks together and satisfies these desiderata.

3 A Joint Network Analysis Framework

Recall that the banks represented in the execution and trading networks are the same because each network is constructed from the same sequence of transactions. Building on this fact, we present next a joint representation of both networks, where the edges are allowed to be heterogeneous. Specifically, an undirected edge between two banks is created if they participate in a trade, where the edge has several variables associated with it that identify the buyer, seller,
initiator, quoter, and volume of the transaction. For illustration, in Figure 4 the network representation is an aggregation of the raw transactional data over the period of interest stratified by the trading and execution relationships.

![Network statistics corresponding to 30-day rolling execution (solid red) and trading (dashed blue) networks from the e-MID interbank market.](image)

The main advantage of this abstraction is that the degree, arguably the most widely used statistic in all of network science, can be organized into a two-way (2x2) table as shown below, where $v_{ij}$ represents the volume for trades of type $i,j$.

<table>
<thead>
<tr>
<th>Execution</th>
<th>Trading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiated</td>
<td>Bought</td>
</tr>
<tr>
<td>Quoted</td>
<td>Sold</td>
</tr>
<tr>
<td></td>
<td>$v_{11}$</td>
</tr>
<tr>
<td></td>
<td>$v_{12}$</td>
</tr>
<tr>
<td></td>
<td>$v_{21}$</td>
</tr>
<tr>
<td></td>
<td>$v_{22}$</td>
</tr>
</tbody>
</table>

Such tables have a long history in the statistical analysis of categorical data (Agresti 2003) under the topic of contingency tables, where one uses the observed trading frequencies in the 2x2 contingency table to infer how the execution and trading patterns are related, if at all. For example, imagine a bank decided to borrow funds on the interbank market, which resulted in multiple transactions. We would then expect $v_{11}$ and $v_{21}$ to be the dominant frequencies in the table, as these values correspond to amount borrowed through market and limit orders. Moreover, if the bank preferred market orders over limit orders, a signal of
urgency to borrow, then \( v_{11} \) relative to \( v_{21} \) should be the most dominant frequency in the table. A rigorous statistical measure that captures this intuition is the odds ratio, which is defined as

\[
\Theta = \frac{odds_1}{odds_2} = \frac{\frac{v_{11}}{v_{12}}}{\frac{v_{21}}{v_{22}}} = \frac{v_{11}v_{22}}{v_{12}v_{21}},
\]

(1)

where \( odds_1 = \frac{v_{11}}{v_{12}} \) is the sample probability of an initiated buy relative to the probability of an initiated sell, and \( odds_2 = \frac{v_{21}}{v_{22}} \) is the sample probability of a limit order buy relative to the probability of a limit order sell. The odds ratio takes values in \([0, \infty)\].

The value \( \theta = 1 \) serves as a baseline for comparison and corresponds to statistical independence between the trading and execution variables. Thus, an odds ratio on either side of one reveals associations between the execution and trading behavior. As shown in Table 1 on three illustrative scenarios, values between zero and one mean that the odds of borrowing through limit orders is greater than the odds of borrowing through market orders. Values greater than one indicate “borrower urgency”, i.e., that the odds of borrowing through market orders were greater for the time interval of interest than the odds of borrowing through passive orders.

<table>
<thead>
<tr>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Execution</strong></td>
<td><strong>Trading</strong></td>
<td><strong>Execution</strong></td>
</tr>
<tr>
<td>Initiated</td>
<td>Bought</td>
<td>Sold</td>
</tr>
<tr>
<td>Quoted</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>( \text{Odds: } \hat{\theta} = \frac{4/1}{1/3} = 12 )</td>
<td>( \text{Odds: } \hat{\theta} = \frac{1/10}{1/1} = 0.1 )</td>
<td>( \text{Odds: } \hat{\theta} = \frac{9/2}{9/2} = 1 )</td>
</tr>
<tr>
<td>( \text{Log Odds: } \log(\hat{\theta}) = 2.48 )</td>
<td>( \text{Log Odds: } \log(\hat{\theta}) = -2.30 )</td>
<td>( \text{Log Odds: } \log(\hat{\theta}) = 0 )</td>
</tr>
</tbody>
</table>
Table 1: Contingency tables and odds ratio for three illustrative cases. Case 1 shows a bank that borrowed by initiating trades than through limit orders. Case 2 shows a bank that tended to lend money through initiated trades. Case 3 shows a bank that borrows equally through both execution options.

Note also that the odds ratio is not impacted by arbitrary decisions about the orientation of the table. For example, the odds ratio does not change when the table is transposed (switching the rows and column definitions). When the order of the rows is reversed or the order of the columns is reversed, the new value of $\theta$ is the inverse of its original value.

There is one property, however, that can be statistically troublesome. The sampling distribution of the odds ratio is highly skewed. When $\theta = 1$, since the estimated odds ratio $\hat{\theta}$ is bounded below by zero, it cannot be much smaller than $\theta$, but it could be much larger with non-negligible probability. To solve this skewness issue, we utilize an alternative but equivalent measure, the log odds ratio, which is computed by taking the natural logarithm of $\hat{\theta}$. Consequently, independence corresponds to $\log(\theta) = 0$ and the sign of the log odds now indicates the association between trading and execution patterns. Larger positive values are a signal of borrower urgency, whereas larger negative values indicate a lack of urgency. With the skewness issue resolved, the log odds ratio has a sampling distribution that can be approximated at large-samples with a Gaussian distribution (Agresti 2003)

$$\log(\hat{\theta}) \sim \text{Normal}\left(\log(\theta), \frac{1}{v_{11}} + \frac{1}{v_{12}} + \frac{1}{v_{21}} + \frac{1}{v_{22}}\right), \quad (2)$$

which allows us to rigorously test the relationship between execution and trading patterns and compare between banks.
To summarize, we construct a new network representation that aggregates the raw transactional data stratified by the trading and execution relationships over a period of interest. The network is summarized by degree, which is now a multidimensional vector of total volume transacted (market order buy/sells and limit order buy/sells). The degree vector is organized into a 2x2 contingency table for each node (bank). Then for each contingency table, the log odds ratio is computed to summarize each bank’s trading activity. Larger positive values indicate urgent borrowing and larger negative values indicate lack of urgency to borrow. Finally, the log odds ratios can be compared rigorously through hypothesis testing using asymptotic Normality and/or aggregated to characterize the trading activity by country or region of origin for the bank.

3.1 Estimation

While the expression of the odds ratio in Equation (1) is straightforward to calculate, in practice this estimator can be of poor quality due to zeros in the observed trading data. Specifically, if the trading volume of passive borrowing or aggressive lending (the denominator cells \( v_{12} \) or \( v_{21} \)) is zero, which could occur simply because a bank decides to trade in a different manner within the time interval of interest, then the estimator in Equation (1) is undefined due to division by zero. Moreover, zero volume observations can also be caused by banks systematically choosing not to participate in the e-MID market. For example, those banks who borrowed in the e-MID received interest rates below their reservation rate and thus chose not to pursue funding from other venues. If urgency to borrow has a positive effect on rates (i.e., these banks are charged more to borrow potentially because they were viewed as riskier counterparties during
the crisis (Heider, Hoerova, and Holthausen 2009), then banks who have greater need for funds will on average be offered a higher rate and therefore be more likely to trade elsewhere. Consequently, we are less likely to observe bank activity as their urgency to borrow increases. Since the observed sample is no longer fully representative of the population of interest, our estimation will become biased. Thus, zero inflation in the data can cause numerical instabilities and be an indicator of sampling bias – significant issues that we address through (i) a regression framework that provides an alternative and computationally stable estimator of the log odds ratio; and (ii) a Heckman (1976, 1977) two-step correction that mitigates potential sample selection bias.

3.1.1 Regression Framework

To derive the regression framework, we model each cell of the 2x2 contingency table as sampled from a population with mean $\mu_{ij}$. Letting $n = \sum_{ij} v_{ij}$ be the total number of contracts traded by the bank (or unit of observation) of interest, we decompose the mean as

$$\mu_{ij} = np_{ij},$$

where $p_{ij}$ is the probability that a trade is of type $i,j$ ($\sum_{i,j} p_{ij} = 1$). If buying and selling are statistically independent of initiated and quoted behavior, then

$$\mu_{ij} = np_{ij} = np_{i+}p_{+j},$$

where $p_{i+} = \sum_{j} p_{ij}$ and $p_{+j} = \sum_{i} p_{ij}$. Then taking the natural logarithm of both sides yields an additive model

$$\log(\mu_{ij}) = \log(n) + \log(p_{ij}) = \log(n) + \log(p_{i+}) + \log(p_{+j}).$$

---

11 A similar argument can also be made for sampling bias on lenders.
\[ \log(\mu_{ij}) = \log(n) + \log(p_{i+}) + \log(p_{+j}). \tag{3} \]

Note that \( \log(n) \) does not depend on \( i \) or \( j \), \( \log(p_{i+}) \) captures the effect of trades with urgency type \( i \) (initiating or quoting trades), and \( \log(p_{+j}) \) captures the effect of trading type \( j \) (buying or selling). Thus, viewed as a two-way analysis of variance regression model, \( \log(n) \) is an intercept term (global mean), and \( \log(p_{i+}) \) and \( \log(p_{+j}) \) can be viewed as row and column effects, and thus modeled using dummy variables. In other words, an equivalent form of the model in Equation (3) is

\[ \log(\mu_{ij}) = \beta_0 + \beta_1 \text{Passive} + \beta_2 \text{Sold} + u_1, \]

where \( u_1 \) is random Gaussian noise and Passive and Sold are dummy variables capturing trade characteristics.

If the rows and columns of the contingency table are statistically dependent (i.e., if the log odds ratio is non-zero), then \( p_{ij} \) cannot be decomposed as a product of \( p_{i+} \) and \( p_{+j} \), meaning that the row and column effects interact. This insight motivates the inclusion of an interaction term to the regression model to capture the level of association or dependence between buys/sells and active/passive.

\[ \log(\mu_{ij}) = \beta_0 + \beta_1 \text{Passive} + \beta_2 \text{Sold} + \beta_3 \text{Passive} \ast \text{Sold} + u_1. \tag{4} \]

Using the contingency table representation below,

<table>
<thead>
<tr>
<th>Execution</th>
<th>Trading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bought</td>
<td>Sold</td>
</tr>
</tbody>
</table>

17
Initiated & $\beta_0$ & $\beta_0 + \beta_2$
\hline
Quoted & $\beta_0 + \beta_1$ & $\beta_0 + \beta_1 + \beta_2 + \beta_3$
\hline

it is straightforward to show that the interaction coefficient $\beta_3$ is a valid estimator of the log odds ratio

$$E(\log(\theta)) = \log\left(\frac{\mu_{11}\mu_{22}}{\mu_{12}\mu_{21}}\right) = \log(\mu_{11}) + \log(\mu_{22}) - \log(\mu_{12}) - \log(\mu_{21}) = \beta_3.$$ 

Note when $\beta_3 = 0$ the log odds is zero indicating, as discussed in Section 2, statistical independence between the trading and liquidity variables.

Lastly, we note that in practice one can construct multiple 2x2 contingency tables – one for each unit of observation. For example, in our empirical analysis presented in Section 4 below, we analyze the data at the country level (due to confidentiality restrictions) by constructing a single contingency table for all banks originating from the same country. We use the following regression equation to jointly model all ten countries within a time interval of interest

$$\log(\mu_{ijct}) = \beta_0 + \beta_c + (\beta_1 + \beta_{2c})\text{Passive}_t + (\beta_3 + \beta_{4c})\text{Sold}_t + \\
(\beta_5 + \beta_{6c})\text{Passive}_t \ast \text{Sold}_t + u_{1t}, \quad (5)$$

where $c$ indexes the country of origin in time interval $t$ with e-MID volume of type $i,j$. Analogous to Equation (4), the log odds ratio for each country is equal to $\beta_5 + \beta_{6c}$. To estimate the regression model, one uses the trading volume $1 + v_{ijct}$ as sample observations for $\mu_{ijct}$ and constructs the appropriate dummy variables as
the independent variables. Then standard linear regression techniques can be utilized to obtain estimates.

We briefly highlight the main advantages of using Equation (5) to estimate the log odds ratio. First, the regression model can be estimated using ordinary least squares, which has advantageous statistical properties and is computationally stable with respect to having zero observations in the contingency table. In fact, a country can technically be included in the estimation regardless of whether banks from the country trade over the period of interest. Second, unlike the original expression in Equation (1), the regression model allows us to easily combine daily observations over a rolling window. For example, in our analysis we use daily observations from the current and previous 30 days for each country when estimating Equation (5). Utilizing variation in time helps improve the accuracy of the estimates, as otherwise with a single day the model would have as many parameters as data points. There still remains a major weakness, as argued above, of sampling bias. We address this issue next.

### 3.1.2 Heckman Correction

Suppose we only observe a bank’s activity in the e-MID when its available rates for either borrowing or lending exceed their reservation price in the appropriate direction. In other words, \( \log(\mu_{ijct}) \) in Equation (5) is censored according to the following probit-type selection equation

\[
\text{WillingnessToTrade}_{ijct} = \gamma_0 + \gamma_1 \left( \text{Lend Rate}_{ct-1} - \text{Lend Rate}_{t-1} \right) + \\
\gamma_2 \left( \text{Borrow Rate}_{ct-1} - \text{Borrow Rate}_{t-1} \right) + u_{2t}, \quad (6)
\]
where non-zero values for Equation (5) are observed when the \( \text{WillingnessToTrade}_{ijct} > \lambda \) for some unknown \( \lambda \). Additionally, driven by the economic argument of how bank urgency to trade can influence their ability to participate in the e-MID, we assume that the Gaussian error terms are correlated

\[
\begin{bmatrix}
    u_{1t} \\
    u_{2t}
\end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}\right).
\]

The variables \( \text{Lend Rate}_{ct-1} \) and \( \text{Borrow Rate}_{ct-1} \) denote the observed rate for lending and borrowing averaged over all trades in the previous time period for banks originating from country \( c \), and \( \bar{\cdot} \) denotes the average for the entire market.

Then under the sampling mechanism described above, the usual estimate of Equation (5) is unbiased if \( \mathbb{E}(u_{1t} \mid \text{WillingnessToTrade}_{ijct} > \lambda) = 0 \). Expanding the expectation yields

\[
\mathbb{E}(u_{1t} \mid \text{WillingnessToTrade}_{ijct} > \lambda) = E(u_{1t} | u_{2t} > \lambda + f(\text{Lend Rate}_{ct-1}, \text{Borrow Rate}_{ct-1})).
\]

However, due to the correlation between \( u_1 \) and \( u_2 \), the expectation is non-zero and can in fact be written in closed form as a function of the Inverse Mills Ratio (Heckman 1979)

\[
\mathbb{E}(u_{1t} \mid \text{WillingnessToTrade}_{ijct} > \lambda) = \frac{\sigma_{12}}{\sigma_{22}} \text{IMR}(\text{Lend Rate}_{ct-1}, \text{Borrow Rate}_{ct-1}, \sigma_{22}).
\]

This is a key result from Heckman (1976, 1977), since it allows the sampling bias to be represented as an omitted variable. Consequently, a two-stage correction procedure can be used to obtain accurate and consistent estimates of Equation (5). First the Inverse Mills Ratio is estimated with a probit model, where the response variable is a dummy variable indicating whether banks from a given country
participated in the e-MID in the time interval of interest. The estimated Inverse
Mills Ratio can in turn be included in the second stage as an additional variable
in the estimation of Equation (5) to obtain consistent estimates for countries that
did have banks participating in the interbank market.

4 Log Odds Ratio Results on the e-MID

In this section we utilize the trading data from the e-MID platform to
calculate the daily log odds ratio over a 30-day rolling window. On each day
we average (volume-weighted) the log odds ratios for each bank at three
different levels of spatial aggregation: (i) the entire market; (ii) by region
(e.g., Core Europe versus Peripheral countries, and so on); (iii) by country.
Results at the highest level of granularity (bank-level) are not presented due
to data confidentiality restrictions.

Figure 5: The time-series of log odds ratio (30 day rolling window) for the entire
e-MID market.

Figure 5 shows the log odds ratio averaged for the entire e-MID
market. The log odds ratio is consistently negative from 2006 to mid-2010,
indicating that, on average, the odds of borrowing through limit orders was
higher than the odds of borrowing aggressively through market orders. However, as the 2007-09 crisis unfolded, the log odds ratio increased towards zero signaling more initiated borrowing. Persistently high levels of the log odds ratio from 2010 onwards indicate that trading and execution patterns never returned to their pre-crisis dynamics and that effects of the crisis continued to affect the banking system in post-recession Europe. In fact, in 2010 and again in 2012, there are episodes where the log odds ratio became positive, signaling major increases in initiated borrowing. Since these intervals fall within a period when Euro-area bank CDS premia rose significantly and sovereign bond spreads widened appreciably for Greece, Ireland, Italy, Portugal, and Spain (relative to Germany), we present next a country-level breakdown of the log odds ratio.

Note that the network framework and the subsequent log odds ratio calculations allow us to readily decompose trading activity and provided new insights into how banks from different countries traded over time. Results in Figure 6 show there is evidence that the increase in the log odds ratio for Greece, Ireland, Italy, Portugal, Spain, and the United Kingdom coincided

![Figure 6: The time-series of log odds ratio (30 day rolling window) aggregated by country.](image)
with the rise in their CDS premia and sovereign bond spreads. We test this relation more formally in the next section.

5 Validation and Policy Implications

5.1 Country Level: Log Odds Ratio and CDS Spreads

In this section we study how the log odds ratios at the country level is related to changes in the corresponding sovereign CDS spreads using bivariate Granger causality. Shown in Figure 7, we use data for 5-year CDS spreads for 11 countries operating in the e-MID. While every country experienced a sudden increase to CDS spreads in 2009, Greece, Ireland, Italy, Portugal, and Spain experienced a larger shock and persistent rise.

Since the levels are non-stationary, when testing the null hypothesis of Granger-non-causality, both the CDS spreads and log odds ratios were first differenced to satisfy the Augmented Dicker-Fuller and Philips-Perron unit root tests. The number of lags were chosen using the AIC.

Figure 7 presents results for Granger-causality tests among the full set of countries we examine. We focus on the interactions between the log odds ratio and

Figure 7: Time-series of 5-year sovereign CDS spreads.
CDS spreads. As Figure 7 shows, the urgency to trade by German banks play an important role in Granger-causing and being Granger-caused by CDS spreads in a wide variety of European countries. Conversely, the log odds ratio of Italian banks largely leads CDS spread changes in Portugal, Ireland, and Greece. The urgency to trade in other countries is largely unrelated to sovereign CDS spread changes.

Figure 7: Granger causality between the log odds ratio and sovereign CDS spreads, by country. A→B indicates that variable A Granger-causes variable B at the 10% significance level. The number of lags is chosen by AIC and all variables are represented in the first difference (levels are non-stationary).

Figure 7 also sheds light on the lead-lag relations between banks in different countries and the changes in sovereign CDS spreads. British sovereign CDS spread changes are largely Granger-caused by the log odds ratios in Belgium, Greece and Spain. Among other European countries, Greek, Irish, and Portuguese
CDS spreads are each significantly related to Italian and German bank odds ratios while other sovereign CDS spreads are linked primarily to German banks.

We further explore the relations between German banks and other European sovereign CDS spreads by estimating bivariate vector-autoregressions (VAR) among the German log odds ratios and CDS spreads. In Figure 8 below we present the orthogonal impulse-response coefficients from the German log odds ratio. We see that an increase in the German log odds ratio is associated with a persistent decline in sovereign CDS premia, not only for Germany itself, but for several other countries including the Netherlands, Portugal, and the United Kingdom. Note that Portuguese CDS spreads do have an initial period of increase prior to the lasting decline. Sovereign CDS spreads for France, Ireland, and Belgium decrease temporarily after a 10-20 day lag. For Italian and Spanish sovereign CDS spreads, we see sequential periods of increase and decrease, each lasting approximately 10 days, before ultimately the effect of the increase in German log odds ratio fades away. The response of Greek CDS spreads to an increase in the German log odds ratio is fairly volatile with several sharp periods of movement in either direction.

Figure 8: Cumulative orthogonal impulse-responses from German log odds ratio to European sovereign CDS spreads with 95% bootstrapped confidence intervals (shaded). The number of lags is chosen by AIC and all variables are represented in the first difference (levels are non-stationary).
The results above provide evidence that the German log odds ratios has a meaningful impact on sovereign CDS spreads for several European countries. Yet, since the analysis above is done using bivariate models, it is possible that the relationship between the log odds ratio and CDS spreads is fully moderated by the German CDS spread. To address this concern, we estimate a tri-variate VAR model, where the German CDS spread and German log odds ratio are always included in the model. We rotate the third series between the other sovereign CDS spreads and test whether the log odds ratio is statistically significant. If the German log odds ratio does not have predictive power beyond German CDS series, then the Granger causality relationships should vanish after controlling for the German CDS series.

Results in Table 2 show that the relationship between the German log odds ratio and CDS spreads is robust. Specifically, even after controlling for the German CDS series, the German log odds ratio Granger causes every country except the Netherlands and the United Kingdom at the 20% significance level. Furthermore, after controlling for the German CDS series, Greek and Italian CDS spreads are the only series that Granger cause the German log odds ratio. Thus, we find that the Granger causal relationship is most often one-way—the German log odds ratio tends to lead sovereign CDS spreads.
<table>
<thead>
<tr>
<th>Response variable</th>
<th>Belgium CDS</th>
<th>France CDS</th>
<th>German CDS</th>
<th>Greece CDS</th>
<th>Ireland CDS</th>
<th>Italy CDS</th>
<th>Netherlands CDS</th>
<th>Portugal CDS</th>
<th>Spain CDS</th>
<th>United Kingdom CDS</th>
<th>German Log Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.008</td>
</tr>
<tr>
<td>France CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.153</td>
</tr>
<tr>
<td>Greece CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
</tr>
<tr>
<td>German CDS</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.022</td>
<td>Always &lt; 0.16</td>
</tr>
<tr>
<td>Ireland CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.085</td>
</tr>
<tr>
<td>Italy CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.074</td>
</tr>
<tr>
<td>Netherlands CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.391</td>
</tr>
<tr>
<td>Portugal CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.108</td>
</tr>
<tr>
<td>Spain CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.116</td>
</tr>
<tr>
<td>United Kingdom CDS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.428</td>
</tr>
<tr>
<td>German Log Odds Ratio</td>
<td>0.651</td>
<td>0.660</td>
<td>Always &lt; 0.10</td>
<td>0.000</td>
<td>0.593</td>
<td>0.046</td>
<td>0.894</td>
<td>0.274</td>
<td>0.220</td>
<td>0.639</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2:** Tri-variate VAR partial F-test results. P-values are shown corresponding to tests of whether the German log odds ratio Granger causes the response variable conditional on German CDS spreads being in the model. The number of lags is chosen by AIC and all variables are represented in the first difference (levels are non-stationary). Smaller p-values indicate statistically significant Granger causal relations controlling for the 3rd set of variables.
5.2 Market Level: Log Odds Ratio, Sidedness, and Network Statistics

We also examine the Granger-causality among our variables at the market-level, including market sidedness, and present results in Figures 9 and 10 below. Since market sidedness is typically used with much higher frequency data and in much finer time intervals, we briefly describe the procedure used to compute market sidedness. First for each day we count the total number of aggressive buys and total number of aggressive sells. Using the previous 30 days, each count is standardized to have mean zero and unit variance. We then compute market sidedness as the absolute value of the correlation between the standardized buy and sell counts.

Figure 9 shows that the log odds ratio aggregated for the entire market using a volume-weighted average of country-level ratios significantly leads several network statistics in the interbank network, including the interbank market clustering coefficient, reciprocity for both the trading and liquidity network, and the LSCC for the trading network. Conversely, the LSCC in the liquidity network significantly leads the log odds ratio. Interestingly, these lead-lag tests show that the log odds ratio is largely unrelated to market sidedness, reinforcing the notion that these variables capture different aspects of the interbank lending market.
Figure 9: Granger causality tests between the log odds ratio and other variables. $A \rightarrow B$ indicates that variable $A$ Granger-causes variable $B$ with dark (and light) arrows representing Granger-non-causality at 10% (20%) significance levels. The number of lags is chosen by AIC and all variables are represented in the first difference (levels are non-stationary).

To further examine lead-lag relations in our data, Figure 10 presents Granger-causality test results between our various variables and market sidedness specifically. As shown, market sidedness is significantly related to the LSCC and the clustering coefficient. However, it significantly lags behind LSCC and degree for both networks.
Figure 10: Granger causality between market sidedness and other variables. A→B indicates that variable A Granger-causes variable B with dark (and light) arrows representing Granger-non-causality at 10% (20%) significance levels. The number of lags is chosen by AIC and all variables are represented in the first difference (levels are non-stationary).

To home in on the relationship between market sidedness and the log odds ratio, Table 3 provides detailed estimation results for Granger causality. We see weak evidence that the log odds ratio leads sidedness, with this hypothesis test approaching significance, whereas the opposite lead-lag relationship has a much larger p-value. Yet, the lack of overall significance in their pairwise Granger causal relationship along with the different connectivity patterns displayed in Figures 9 and 10 underscore that these variables capture different aspects of the interbank lending market.

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>F-Statistic</th>
<th>Lag</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sidedness</td>
<td>Log Odds Ratio</td>
<td>0.892</td>
<td>4</td>
<td>0.602</td>
</tr>
<tr>
<td>Log Odds Ratio</td>
<td>Sidedness</td>
<td>1.764</td>
<td>4</td>
<td>0.224</td>
</tr>
</tbody>
</table>

Table 3: Granger Causality between market sidedness and the log odds ratio. The number of lags is chosen by AIC and all variables are represented in the first difference (levels are non-stationary).

6 Conclusion

In the last decade within financial economics, network analysis has been utilized aiming to better understand how the interconnectedness between market participants results in spillovers, amplifies or absorbs shocks, and/or creates other non-linearities that affect key markers of market health, including the supply and demand of liquidity. In this paper, we propose a new network construct, the liquidity network that represents the urgency banks express for borrowing, and use it to examine the physical overnight lending market in Europe.
Combined with the standard trading network that connects sellers to buyers, we demonstrate that both types of networks complement each other to more comprehensively characterize interconnectivity in the e-MID overnight lending market. Conceptually this paper is, to our knowledge, the first to jointly model two physical financial networks simultaneously. Specifically, we develop a new network statistic, the log odds ratio, to simultaneously combine information from both liquidity and trading networks. The log odds ratio provides a quantitative measure of the urgency of participants to trade in the market—in our setting the urgency of banks to borrow in the e-MID interbank market. Note that the log odds ratio is widely applicable and can be constructed for any market so long as trade counterparties and standing quotes can be identified.

Beyond methodological contributions, our econometric validation reveals interesting insights with policy implications. For example, we show that interbank activity by German banks statistically leads (typically without feedback) movements in other European sovereign CDS spreads. At the market level, we find that the log odds ratio tends to lead other popular network metrics used in finance. Moreover, the log odds ratio complements market sidedness (Sarkar and Schwarz 2009), another measure of the motivations behind trader activity. Thus, from a policy perspective, our findings imply that institutions and researchers should analyze both types of networks together when possible to better understand interconnectedness and liquidity flows.
References


Heckman, J.J., 1976. The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In Annals of Economic and Social Measurement, Volume 5, number 4 (pp. 475-492). NBER.


