Relationship Trading in OTC Markets

Terrence Hendershott, Dan Li, Dmitry Livdan and Norman Schürhoff

October 3, 2016

RELATIONSHIP TRADING IN OTC MARKETS

Terrence Hendershott , Dan Li , Dmitry Livdan and
Norman Schürhoff

We examine the network of bilateral trading relations between insurers and dealers in the over-the-counter corporate bond market. Using comprehensive regulatory data we find that many insurers use only one dealer while the largest insurers have a network of up to eighty dealers. To understand the heterogeneity in network size we build a model of decentralized trade in which insurers trade off the benefits of repeat business against more intense dealer competition. Empirically, large insurers form more relations and receive better prices than small insurers. The model matches both the distribution of insurers' network sizes and how prices depend on insurers' size and the size of their dealer network.

Keywords: Over-the-counter market, corporate bond, trading cost, liquidity, decentralization, financial network.

We study investors' choice of trading networks in an over-the-counter (OTC) market. As an investor adds dealers to her network the dealers anticipate less future business, leading to inferior dealer prices. On the other hand, additional dealer competition affects bargaining and improves execution quality. In equilibrium investors' optimal network size trades off the benefits of repeat relations with individual dealers against the benefits of competition among dealers. A theoretical model with these tradeoffs generates predictions for optimal networks and execution prices. We empirically test these predictions qualitatively and quantitatively using insurance companies' secondary market trading of corporate bonds.

Corporate bonds are the corporations' primary source for raising capital

Haas School of Business, University of California, Berkeley, Berkeley, CA 94720
Board of Governors of the Federal Reserve System, Washington, DC 20551
Haas School of Business, University of California, Berkeley, Berkeley, CA 94720
Faculty of Business and Economics and Swiss Finance Institute, Université de Lausanne, CH-1015 Lausanne, Switzerland

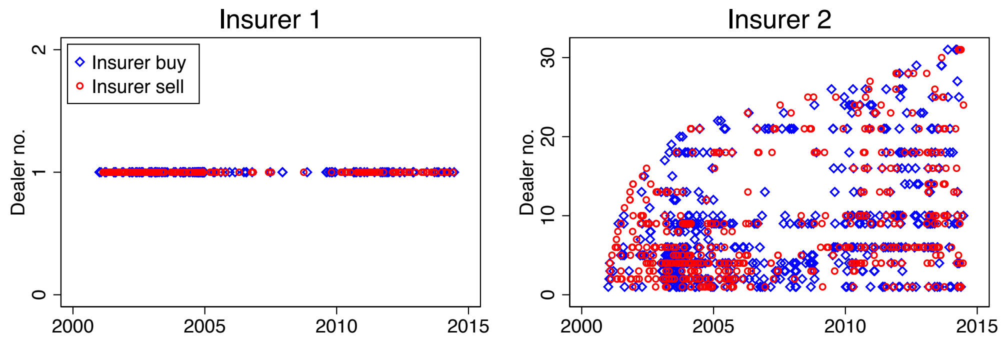

FIGURE 1.— Example of dealer-client trading relations

The figure shows the buy (blue squares) and sell (red circles) trades of two insurance companies with different dealers. We sort the dealers on the vertical axis by the first time they trade with the corresponding insurance company.

and trade in a classic OTC market with more than 400 active broker-dealers. Insurers are one of the largest investors in the $7.8 trillion market with more than 20 thousand bond issues. Insurers own roughly 30% of all corporate bonds. Insurers are long-term buy and hold investors. Insurers exhibit quite different trading needs depending on their balance sheet size and type of business, enabling the study of how investor heterogeneity impacts market outcomes. For more than 4,300 insurers regulatory data reveal the identities of the insurers and dealers on both sides of all transactions between them from 2001 to 2014.

Our empirical investigation starts by inspecting matching patterns between insurers and dealers. Insurers form few but long-lasting dealer relationships. Roughly 30% of insurers trade with a single dealer annually. Figure 1 provides two examples of client-dealer trading relations over time. Panel A shows buys and sells for an insurer trading with a single dealer and Panel B for an insurer trading with multiple dealers. Both insurers buy and sell bonds repeatedly from the same set of dealers.

Such trading behavior is unlikely under the undirected random search paradigm (Duffie, Garleanu, and Pedersen, 2005, 2007; Lagos and Rocheteau, 2007, 2009) where clients repeatedly search for best execution without forming long-term relationships with dealers. The empirical facts that insurers form finite-size dealer networks where many insurers use few dealers motivates an alternative model of OTC markets. The simple strategic model of finite network formation, which we call a Rolodex model, allows clients and dealer to share the benefits of repeated interactions. The model incorporates features from a variety of existing models, including search costs, exclusivity, loyalty,[1] and costs of maintaining links in the network.

In our model a single console bond is traded on an inter-dealer market which clients can only access through dealers. Dealers have search intensity λ and upon trading with a client then transact at the competitive inter-dealer bid/ask prices. Clients initially start without a bond but stochastically receive trading shocks with intensity η which cause them to simultaneously contact N dealers. The client's effective search intensity is $\Lambda = N\lambda$. Maintaining relations with a single dealer requires a sunk cost of K per search, which can be interpreted as wages for in-house traders and which precludes infinite network size. The first dealer to find the bond captures all benefits from the transaction. Thus, our trading mechanism is identical to repeated winner-takes-all races (Harris and Vickers (1985)) with participation by costly invitation with costs born by the client. The bond's purchase price is set by Nash bargaining. Once an owner, the client stochastically receives liquidity shock forcing her to sell the bond. The mechanics of the sell transactions are the same as the buy transaction. Both dealers and clients derive value from repeat transaction which help determine transaction prices, leading to price improvement for better clients in Nash bargaining. Both the Rolodex size and transaction prices are endogenously determined in equilibrium by maximizing buyers' utility.

[1]For a comprehensive model of loyalty see Board (2011).

The model is solved in steady state. While few fully analytical results are readily attainable, the model delineates the trade-offs leading to the optimal finite Rolodex. Clients adding additional dealers increase their speed of execution and enhances their bargaining position at the cost of less repeat business. On the bid side, the client sells the asset faster but may get less for it. When repeat business is important, the client chooses a single dealer. When repeat business is unimportant, the client uses all possible dealers.

The model provides an interesting set of comparative statics results. Clients' trading intensity η affects the optimal Rolodex size. Higher η increases repeated business with each dealer. Similar to Duffie et al. (2005) dealers then offer better prices. Higher trading intensity reduces the rate at which repeat business by each dealer falls when another dealer is added. In contrast, the benefits of speed are independent of trading intensity. Therefore, clients with higher trading intensity choose larger Rolodexes and receive better prices.

Both the size of the Rolodex and transaction prices are endogenous in our model. Existing OTC models provide predictions about network size or prices, but not both. Random search models assume investors may contact every other counterparty. Other models allow investors to choose specific networks or markets, but exogenously fix the the structure of those networks. For instance, in Vayanos and Wang (2007) investors chose to search for a counterparty between two markets for the same asset: a large market with faster execution but higher transaction costs, and a small market with slower execution but lower transaction costs. Neklyudov and Sambalaibat (2016) use a similar setup as in Vayanos and Wang (2007) but with investors choosing between dealers with either large or small dealer networks instead of asset markets.

The key qualitative pricing implications from our model are supported by the data. First, larger insurers trade more and have larger networks than smaller insurers. Second, larger insurers receive better prices due to more valuable repeat business. Third, these two effects combine to predict that clients with larger networks receive better prices. All these relations are present in the data.

We go beyond the qualitative predictions to see if the model can quantitatively match the insurers' observed Rolodex sizes and transaction prices. Doing so requires structural estimation of the model parameters not directly observable in the data, Θ. The clients' trading intensity, η, is the one parameter for which we observe the cross-sectional distribution, $p(\eta)$, in the data. Insurer n's trading shock intensity η_n can be estimated by the average number of bond purchases per year over the sample period. Utilizing multiple years of trade data, we perform the estimation separately for each insurer in the sample, which enables us to construct $p(\eta)$. The model in-turn provides the optimal Rolodex size, $N^*(\Theta, \eta_n)$, for client n, thus allowing us estimate the unobservable model parameters Θ by matching the model-implied N^* to its empirical counterpart. Using the structurally estimated parameters along with the distribution of trading intensities quantitatively reproduces the distribution of Rolodex sizes observed in the data and the dependence of trading costs on Rolodex size found in the data. The model estimates reasonable unobserved parameters, e.g, dealers' search period is between one and a half and three days and insurers' average holding period ranges from a few weeks to a month.

The paper complements the empirical literature on the microstructure of OTC markets and its implications for trading, price formation, and liquidity. Edwards, et al. (2005), Bessembinder et al. (2006), Harris and Piwowar (2007), Green et al. (2007) document the magnitude and determinants of transaction costs for investors in OTC markets. Our paper deepens our understanding of OTC trading costs by using the identities of all insurers along with their trading networks and execution costs. These help explain the substantial heterogeneity in execution costs observed in these studies. O'Hara et al. (2015) and Harris (2015) examine best execution in OTC markets without formally studying investors' optimal network choice.

There exists an empirical literature on the value of relationships in financial markets. Similar to our findings, Bernhardt et al. (2004) show that on the London Stock Exchange broker-dealers offer greater price improvements to more regular customers. Bernhardt et al. (2004) do not examine the client-dealer networks and in the centralized exchange quoted prices are observable. Afonso et al. (2013) study the overnight interbank lending OTC market and find that a majority of banks in the interbank market form long-term, stable and concentrated lending relationships. These have a significant impact on how liquidity shocks are transmitted across the market. Afonso et al. (2013) do not formally model the network and do not observe transaction prices. DiMaggio et al. (2015) study inter-dealer relationships on the OTC market for corporate bonds while our paper focuses on the client-dealer relations.

The role of the inter-dealer market in price formation and liquidity provision are the focus in Hollifield et al. (2015) and Li and Schürhoff (2015). These studies explore the heterogeneity across dealers in their network centrality and how they provide liquidity and what prices they charge. By contrast, we focus on the heterogeneity across clients and how trading needs affect their relationships and the execution terms they receive.

The search-and-matching literature is vast. Duffie et al. (2005, 2007) provide a prominent treatment of search frictions in OTC financial markets, while Weill (2007), Lagos and Rocheteau (2007, 2009), Feldhütter (2011), Neklyudov (2014), Hugonnier et al. (2015), Üslü (2015) generalize the economic setting. These papers do not focus on repeat relations and do not provide incentives to investors to have a finite size Rolodex.

Directed search models allow for heterogenous dealers and investors, as well as arbitrary trade quantities. These rely on a concept of competitive search equilibrium proposed by Moen (1997) for labor market. Prominent examples include Guerrieri, Shimer, and Wright (2010), Guerrieri and Shimer (2014), and Lester et al. (2015). These papers explain assortative matching between clients and dealers and show how heterogeneity affects prices and liquidity. However the matching technology employed by these papers is one-to-one, thus limiting the Rolodex size to a single dealer.

Our paper also relates to a growing literature studying trading in a network, e.g., Gale and Kariv (2007), Gofman (2011), Condorelli and Galeotti (2012), Colliard and Demange (2014), Glode and Opp (2014), Chang and Zhang (2015), Atkeson et al. (2015), Babus (2016), Babus and Hu (2016), and Babus and Kondor (2016). These papers allow persistent one-to-one dealer-client relationships, while the main focus of our model is on client networks.

1. DATA

Insurance companies are required quarterly to report trades of long-term bonds and stocks to the National Association of Insurance Commissioners (NAIC). For each trade the NAIC data include the dollar amount of transactions, par value of the transaction, insurer code, date of the transaction, the counterparty dealer name, and the direction of the trade for both parties, e.g., whether the trade was an insurance company buying from a dealer or an insurance company selling to a dealer. The NAIC trades preclude intraday analysis as the trades do not include time stamps of the trades. To focus on secondary trading for we only include trades more than 60 days after issuance and trades more than 90 days to maturity.

Our final sample covers all corporate bond transactions between insurance companies and dealers reported in NAIC from January 2001 to June 2014. We supplement the NAIC data with a number of additional sources.

Bond and issuer characteristics come from the Mergent Fixed Income Security Database (FISD). Insurer holdings and bond ratings come from Lipper eMAXX data. Insurer financial characteristics come from A.M. Best and SNL Financial. The final sample contains 506 thousand insurer buys and 497 thousand insurer sells.

Table I reports descriptive statistics for the corporate bond trades (Panel A) and insurers (Panel B) in our sample that ranges from January 2001 to June 2014. There are 4,324 insurance companies in our sample. Insurance companies fall into three groups based on their products: (i) Health, 617 companies (14% of the sample); (ii) Life, 1,023 companies (24% of the sample); (iii) P&C 2,684 companies (62% of the sample). Health insurance companies account for 16.6% of trades and 4.7% of trading volume. They trade on average with 17 dealers. Life insurance companies account for the majority 46% of trades and 69.5% of the total trading volume. They trade on average with 19 dealers. P&C insurance companies comprise 37.4% of trades and 25.8% of trading volume. They trade on average with 14 dealers.

The distribution of trading activity is skewed with the top ten insurance companies accounting for 6.9% of trades and 15.7% of trading volume. They use on average almost 30 dealers which is much higher than the sample average of 5.83 dealers per insurer. The top 100 insurers account for almost half of trades (46.3%) as well as for 32.8% of trading volume. The 3,000 smallest insurers use on average 3.76 dealers.

Insurers trade in a variety of corporate bonds. The average issue size is quite large at $917 million and it is similar across insurer's buys and sells. The average maturity is nine years for insurer buys and eight years for insurer sells. Bonds are on average 2.88 years old with sold bonds being a little older at 3.09 years. Finally, 75% of all bonds are investment grade while only 1% of bonds are unrated with the remainder being high yield. Privately placed bonds form a small minority of our sample at 8%.

TABLE I

DESCRIPTIVE STATISTICS

The table reports descriptive statistics for trades (Panel A) and insurers (Panel B) in our sample from 2001 to 2014. Panel A reports the average across all trades over the sample period. Panel B reports the yearly average across insurers.

Panel A: Trades			
	All trades	Insurer buys	Insurer sells
No. of trades (k)	1,003	506	497
Trade par size ($mn)	1.80	1.73	1.87
Bond issue size ($mn)	916.66	921.37	911.87
Bond age (years)	2.88	2.67	3.09
Bond remaining life (years)	8.54	8.94	8.13
Private placement (%/100)	0.08	0.08	0.07
Rating (%/100)			
IG	0.74	0.76	0.72
HY	0.25	0.23	0.28
Unrated	0.01	0.01	0.01

Panel B: Insurers (N= 4,324)			
	Volume ($mn)	No. of trades	No. of dealers
All insurers	17.32	9.52	5.83
Insurer type			
Health (617, 14%)	10.66	21.74	6.59
Life (1,023, 24%)	103.00	37.71	8.06
P&C (2,684, 62%)	14.08	11.29	4.81
Insurer activity			
Top 10	2,111.88	517.92	29.89
11-100	509.49	233.04	22.07
101-1000	75.66	46.22	11.56
1001+	3.80	4.24	3.76
Insurer characteristics:	Mean (SD)		
Insurer size	4.97 (0.90)		
Insurer RBC ratio	3.36 (0.35)		
Insurer cash-to-assets	3.49 (10.79)		
Life insurer	0.24 (0.42)		
P&C insurer	0.62 (0.48)		
Insurer rated A-B	0.37 (0.38)		
Insurer rated C-F	0.01 (0.07)		
Insurer unrated	0.53 (0.39)		

The RBC ratio measures an insurer's capital relative to the riskiness of its business. The higher the RBC ratio, the better capitalized the firm. Insurer size is reported assets. The cash-to-asset ratio is cash flow from the insurance business operations divided by assets.

Overall, there exists a large degree of heterogeneity on the client side in our sample. Insurance companies buy and sell large quantities of different corporate bonds and execute these transactions with the number of dealers ranging on average from one to as many as 80.

2. EMPIRICAL RESULTS ON INSURER TRADING NETWORKS

This section empirically characterizes insurers' trading needs and the size of their trading networks.

2.1. *Insurer trading activity*

We investigate the determinants of both the extensive margin, i.e., the number of trades, and intensive margin, i.e., the total dollar volume traded, of the insurer trading in a given month and, where applicable, in a given year. Both margins reveal that insurers have heterogenous trading needs.

We start with univariate analysis. The majority of insurers do not trade often at the annual frequency about 30% of insurers trade just once per year while 1% of the insurers make 25 trades per year. This is consistent with the evidence from Table I that while the top 100 insurers constitute just 2.31% of the total sample, they account for as much as 33% of all trades in our sample. The mean number of trades per year is 16, with a median of 14, with several insurers making more than 1,000 trades in some years and up to the maximum of 2,200 trades in a year.

Figure 2 shows the distribution in the average number trades per year across insurers. A large fraction of insurers do not trade in a given month and we therefore report an annual figure. The annual distributions follow a power

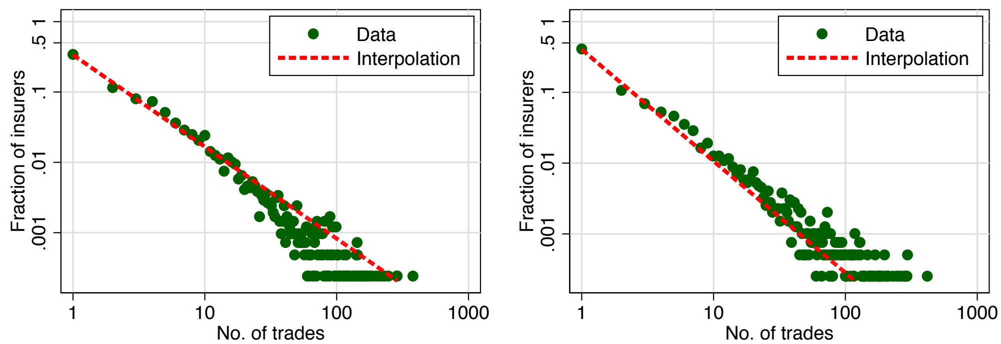

FIGURE 2.— Insurer trading activity

The figure shows the distribution in the number of insurer buys per year (left) and insurer sales (right). We use a log-log scale.

law with $p(X) \propto .27 \times X^{-1.21}$ for all insurer trades combined. The power law is $.34 \times X^{-1.31}$ for insurer buys (depicted in Panel A) and $.40 \times X^{-1.58}$ insurer sales (Panel B). Visually the two power law distributions for insurer buys and sales in Figure 2 look remarkably similar. This suggests insurers buy and sell at similar rates, even though these rates vary significantly across insurers.

We next examine what characteristics explain the heterogeneity in trading needs. Table II documents the determinants of the intensive margins (trading volume in \$bn, column (1)) and extensive margins (number of trades in a month, column (2)) of the annual trading by insurance companies using pooled regressions with time fixed effects. The specification consists of the trade par size, insurer and bond characteristics, as well as the variation in the trade size and bond characteristics across all trades of the insurer during a given month. Insurer's characteristics include its size, cash-to-assets ratio, type, risk-based capital (RBC) ratio, and rating. Bond characteristics include most of the usual suspects such as size, age, maturity, rating, a private placement dummy, and the trade size. All variables are log-transformed and all regressors are averaged across all trades of the insurer during the period and lagged by one time period.

Logarithms of both measures of trade intensity are persistent; the coefficient on the lagged log-volume is 0.67 and the coefficient on the lagged log-number-of-trades is 0.76. Both coefficients are statistically significant at 1% levels. This evidence is consistent with insurance companies rebalancing their holdings over several months.

Insurer trading strongly correlates with insurer size, type, and quality, with bond types and bond varieties and these variables explain 79% of the variation in annual trading volume and 65% variation in annual number of trades. A ten-fold increase in insurer's size increases trading volume by $2.2 billion. Larger insurance companies and insurers with higher cash-to-assets ratio also trade more often and submit larger orders. Insurers with higher RBC ratios trade less often insurers with low RBC ratios. Both margins of trading increase with the insurer's rating, i.e., insurers with the lowest rating (C-F) trade less than higher-rated insurers. Life insurers tend to submit larger orders.

Both margins of bond turnover increase as bond ratings decline; lower rated bonds are traded more often and in larger quantities. Insurers tend to trade privately placed bonds less since potentially they just own fewer of them than publicly placed bonds. Both margins of bond turnover decline with par size and bond age indicating that the majority of insurers are long-term investors demanding specifically structured cash flows from their bond holdings. Both measures of trade intensity do not depend on bond issue size and remaining life as their coefficients are not statistically significant.

Finally, both trading volume and number of trades decline if overall more bond varieties are traded, i.e., trading is concentrated around few bonds with similar characteristics. However, a specific variety can have an opposite effect on the trading intensity. For instance, both measures of trading intensity increase with variation in bond rating and bond life. Once again this finding is consistent with insurers increasing trading intensity when rebalancing their portfolios, i.e. shifting from high-yield to lower yield bond or from younger to older bonds.

Overall, these analysis highlight large heterogeneity in trading needs across different insurers. Trading needs depend on the variety of bonds in the order, bond specific characteristics, and on the insurer type and quality. We now turn to how these characteristics affect the insurers choice of network.

2.2. *Properties of insurer networks*

The previous section's results demonstrate that insurer characteristics explain the intensity of their trading. There is large degree of heterogeneity in the insurers' trading intensity with some insurers trading on average twice per day while others trade just once per year. This suggests that insurers may have similarly heterogeneous demands for their dealer network. This section studies how many dealers they trade with over time and how persistent are these networks. These results describe the basic network formation mechanism in the OTC markets.

We start with the examples of the insurer-dealer relationship depicted in Figure 1. These show that insurers do not trade with a dealer randomly picked from a large pool of corporate bonds dealers. Instead, these insurers buy from the same dealers that they sell bonds to and they engage in long-term repeat, but non-exclusive, relations. We analyze how representative are the examples in Figure 1 and how insurer characteristics determine their network size.

Figure 3 plots the degree distribution across insurers by year, i.e., the fraction of insurers trading on average across all years with the given number of dealers, using a log-log scale. The figure shows insurers trade with up to 40 dealers every year; a few observation are higher than 40, going up to 80, but they represent less than 1/10,000 of the sample. Exclusive rela-

TABLE II

Insurers' trading activity

The determinants of insurance company trading activity are reported. We measure trading activity by the total dollar volume traded in a given year and, alternatively, by the number of trades over the same time horizon. All dependent variables are log-transformed by 100*log(1+x). All regressors are averaged across all trades of the insurer during the period and lagged by one time period. Estimates are from pooled regressions with time fixed effects. Standard errors are adjusted for heteroskedasticity and clustering at the insurer and time level. Significance levels are indicated by * (10%), ** (5%), *** (1%).

	(1)	(2)
Determinant	Volume ($mn)	No. of trades
Insurer size	21.95***	14.51***
Insurer RBC ratio	-1.53	-4.68***
Insurer cash-to-assets	0.27***	0.26***
Life insurer	4.96***	0.27
P&C insurer	-1.06	-3.89**
Insurer rated A-B	5.13**	5.80***
Insurer rated C-F	1.97	-0.43
Insurer unrated	6.39***	5.79**
Trade par size	-3.62***	-3.22***
Bond issue size	-0.00	0.00
Bond age	-0.79***	-1.25***
Bond remaining life	-0.04	-0.05
Bond high-yield rated	4.62***	4.65***
Bond unrated	-6.67	-12.81*
Bond privately placed	-5.37	-3.56
Variation in trade size	4.50***	1.52***
Variation in issue size	0.00	0.00
Variation in bond age	0.37	0.39
Variation in bond life	0.52***	0.65***
Variation in bond rating	-0.35	0.04
Variation in rated-unrated	39.60*	6.16
Variation in private-public	1.62	-1.60
No varieties traded	9.23***	13.64***
Lagged volume	0.67***	
Lagged no. of trades		0.76***
Year fixed effects	Yes	Yes
r2	0.789	0.646
N	30,029	30,029

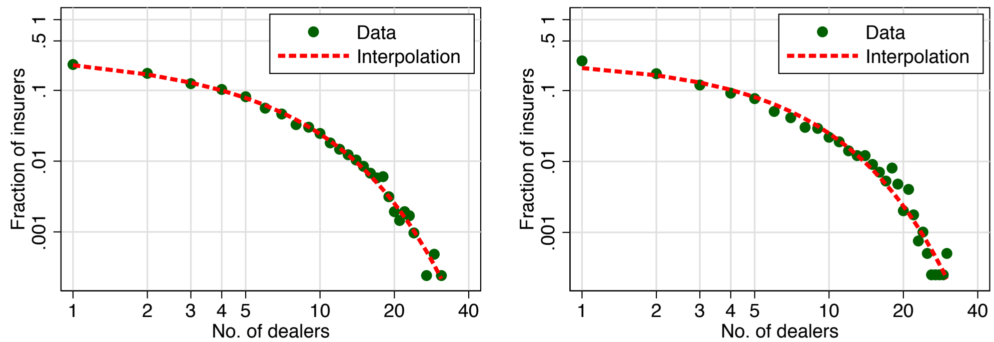

FIGURE 3.— Size of insurer-dealer trading networks

The figure shows the degree distribution for insurer-dealer relations by year for insurer buys (left) and insurer sales (right). We use a log-log scale.

tions are dominant with almost 30% of insurers trading with a single dealer in a given year. The degree distributions in Figure 3 follows a power law with exponential tail starting at about 10 dealers. This is consistent with insurers building Rolodexes that they searching randomly within. Fitting the degree distribution to a Gamma distribution by regressing the log of the probabilities of each N on a constant, the logarithm of N, and N yields the following coefficients:

$$\begin{aligned} \text{For all insurer trades combined:} \quad & p(N) \propto N^{.15} e^{-.20N}, \\ \text{For insurer buys:} \quad & p(N) \propto N^{-.12} e^{-.22N}, \qquad (1) \\ \text{For insurer sales:} \quad & p(N) \propto N^{.01} e^{-.24N}. \end{aligned}$$

Table III reports the determinants of insurer dealer networks sizes using pooled regressions with time fixed effects. We measure the size of the trading network by the number of dealers that an insurance company trades with in a given month. We log-transform all dependent variables by $100 * log(1+x)$ and average all regressors across all trades by the same insurer during the sample period and lag them by one time period. We perform the estimation on the whole sample (Column (1)) and, in order to control for the insurer's

TABLE III

Size of insurers' trading network

The table reports the determinants of the size of insurers' trading network. We measure the size of the trading network by the number of different dealers that an insurance company trades with in a given month. See caption of Table II for additional details. Standard errors are adjusted for heteroskedasticity and clustering at the insurer and time level.

	(1)	(2)	(3)
Determinant	All insurers	Small insurers	Large insurers
Insurer size	9.95***	7.93***	5.04***
Insurer RBC ratio	-3.58***	-3.64***	0.11
Insurer cash-to-assets	0.15***	0.15***	0.11**
Life insurer	-0.13	0.75	-2.77**
P&C insurer	-2.60***	-1.20	-3.59***
Insurer rated A-B	4.12***	6.00***	0.13
Insurer rated C-F	-1.54	-0.18	-5.42***
Insurer unrated	3.05*	3.98*	-2.86**
Trade par size	-1.92***	-1.23***	-1.24***
Bond issue size	-0.00	-0.00	0.00
Bond age	-0.91***	-0.77***	-1.10***
Bond remaining life	-0.08	-0.11*	0.10
Bond high-yield rated	1.23	-2.93*	2.56
Bond unrated	-13.74***	-12.50***	-7.73
Bond privately placed	-3.85	1.10	-10.22
Variation in trade size	0.51	-0.10	0.81*
Variation in issue size	0.00	0.00	0.00
Variation in bond age	0.50**	0.56**	0.16
Variation in bond life	0.50***	0.41***	0.11
Variation in bond rating	0.05	0.19	0.36*
Variation in rated-unrated	3.60	-11.38	-24.50
Variation in private-public	1.20	-4.33	3.17
No varieties traded	5.88***	2.71***	12.50**
Lagged no. of dealers	0.75***	0.62***	0.75***
Year fixed effects	Yes	Yes	Yes
r2	0.614	0.350	0.593
N	30,029	18,033	11,996

TABLE IV

Persistence in insurers' trading network

Switching probabilities p(No. of dealers in $t+1$|No. of dealers in t) for using a network size conditional on the insurer's past behavior for each year.

	No. of dealers next year			
No. of dealers this year	1	2-5	6-10	>10
1	0.61	0.30	0.06	0.03
2-5	0.20	0.54	0.20	0.06
6-10	0.06	0.31	0.40	0.24
>10	0.01	0.07	0.17	0.75

size, on sub-samples of small and large insurers based on asset size. We classify an insurer as small if it falls in the bottom three size quartiles and, respectively, as large if it falls in the top quartile of the size distribution.

Column (1) indicates that insurer size and type, bond characteristics, and bond varieties matter for the size of the dealer network. Large insurers with more trading needs have more dealers. Insurers with demand for larger bond variety have larger networks even controlling for their size (column (2) and (3)). Higher quality insurers, i.e., insurers with higher cash-to-assets ratio or higher ratings, have larger networks but it matters exclusively for smaller insurers as column (2) indicates. It is potentially due to the fact that it is easier for higher quality insurers to find a dealer through referrals and it is cheaper for a dealer to set up a credit account for higher quality insurers. These factors matter more for small insurers since they face larger adverse selection problems in forming permanent links with dealers. Insurer with greater variety in the bond they trade have larger netwokrs. Overall these findings suggest insurers' network choice is endogenous and dependent on multiple factors. Competition and specialization jointly determine investors' trade choices.

Table III show persistence in the size of the network with the coefficient on the lagged network size is 0.75 (column (1)). This result is mostly due to large insurers, since this coefficient is equal only to 0.62 for small insurers. Table IV examines this in more detail by reporting statistics for the frequency with which insurers adjust their network size. We compute the likelihood that an insurer uses a certain number of dealers in a given and year and compare it the corresponding number in the next year. The transition probabilities are reported in Table IV. Trading relations are persistent from year to year. This is especially true for exclusive relations as the probability of staying with a single dealer each month and year is equal to 0.61. Insurers with more than one dealer are very unlikely to switch to a single dealer as monthly switching probabilities are equal to 0.20 for insurers with 2 to 5 dealers and 0.06 for insurers with 6 to 10 dealers. Insurers with the largest networks (> 10 dealers) tend to maintain large networks over time, with the probability of staying with a large network being 0.75. The distribution of insurers shown in Figure 3 together with the stable network sizes are difficult to reconcile with a "pure" random search model à la Duffie et al. (2005, 2007).

The next section uses this evidence to motivate a Rolodex model of the OTC markets. The model is stylized as our goal is a parsimonious model capable of producing the empirical facts thus far as well as delivering further predictions.

3. MODEL

Existing OTC models generate results about network size or prices, but not both. In random search models investors do not limit the number of counterparties they contact. Directed search models in labor and other areas contain richer pricing rules and endogenous quantities, but have one-to-one matching. An intermediate approach is to allow investors to choose specific networks or markets, but exogenously fix the the structure of those networks. Our model retains many of the features of financial OTC market models while incorporating a simple mechanism for endogenous one-to-many matching. The simplicity of matching is facilitated by assuming dealers are symmetric and a reduced-form source of prices and supply and demand in an interdealer network.

We consider a setting where clients transact with dealers in an OTC market. Dealers transact with each other in an inter-dealer market with exogenous bid and ask prices. Due to search and bargaining clients receive inferior prices relative to dealers' acquisition/liquidation cost. Clients optimally form a finite-size dealer network, or Rolodex, by trading off increased search intensity against bargaining costs and variable costs of maintaining a larger active dealer network. Dealers participating in such networks anticipate repeated business from the same clients. As in Duffie et al. (2005) we abstract from in-depth modeling of traded asset's fundamentals to focus on the value generated by the client-dealer relations.

3.1. *Setup and Solution*

The economy has a single risk-free perpetual bond paying a coupon flow C. The risk-free discount rate is constant and equal to r, so that the present value of the bond is $\frac{C}{r}$. To model client-dealer interactions, we keep several attractive features of Duffie et al. (2005) type models such as liquidity supply/demand shocks on the client side and random search with constant intensity. Following Lester et al. (2015), the bond trades on a competitive market accessible only to dealers. Unlike the frictionless inter-dealer market in Lester et. al. (2015), in our model dealers face search frictions as in Duffie et al. (2005). Dealers buy bonds at an exogenously given price M^{ask} from other dealers and sell it to other dealers at an exogenously given price M^{bid}. A bid-ask spread $M^{ask} - M^{bid} \geq 0$ reflects trading costs or cost of carry.

Each client chooses a network of dealers, N, without knowledge of other clients' decisions. When a client wants to buy/sell a bond, she simultaneously contacts all N dealers in her network. Upon being contacted each dealer starts searching the competitive dealer market for a seller/buyer with a search intensity λ.[2] All dealers in the client's network search independently of each other. Therefore, the effective rate at which a client with N dealers in her Rolodex finds a counterparty equals $\Lambda = \lambda N$. When the client receives a subsequent trading shock all dealers in the Rolodex are contacted to reverse the initial transaction.

Each client pays a cost K per dealer per transaction, treated as a sunk cost.[3] Because client-dealer contacts are simultaneous across all dealers in the Rolodex, the cost K are any cost of a clients contacting more dealers. These consist of both the time required to make each phone call and any fixed costs required to hire more in-house traders. Clients' search mechanism can be viewed as a winner-takes-all race with a cost per invitation K paid by the client with the dealer first to find the bond winning the race. The prize is the spread $P^b - M^{ask}$ when the clients buys and $M^{bid} - P^s$ when the clients sells, where P^b (P^s) is the price at which the client buys (sells) the bond from (to) the dealer.

Clients transition through ownership and non-ownership based on liquidity shocks. At these transitions clients act as buyers and sellers. The discounted transition probabilities and transaction prices link their valuations across the owner, non-owner, buyer, and seller states.

A client starting as a non-owner with valuation V^{no} is hit by stochastic trading shocks to buy with intensity η. The client contacts her network of

---

[2]Bessembinder et al. (2016) show that corporate bond dealers increasingly hold less inventory and facilitate trade via effectively acting as brokers by simultaneously buying and selling the same quantity of the same bond.

[3]The costs of additional dealers can be modeled as per dealer or per dealer per transaction. Per dealer costs consist of costs of forming a credit relationship and any other costs of maintaining the relationship independent of the number of trades. Such per dealer costs will immediately lead to clients with larger trading needs using more dealers.

N dealers leading to her transitioning to a buyer state with valuation V^b. In steady state valuations in these two states are related by

$$V^{no} = V^b \frac{\eta}{r+\eta}. \tag{2}$$

The buyer purchases the bond from her network at the expected price $\mathrm{E}[P^b]$ and transitions into being an owner with valuation V^o. In steady state valuations in these two states are related by

$$V^b = (V^o - \mathrm{E}[P^b] - K)\frac{\Lambda}{r+\Lambda}. \tag{3}$$

While clients are owners they they receive a coupon flow C and have valuation V^o. Non-owners do not receive the coupon flow. With intensity κ an owner receives liquidity shock forcing her to become a seller with valuation V^s. In steady state valuations in these two states are related by

$$V^o = \frac{C}{r+\kappa} + \underbrace{V^s \frac{\kappa}{r+\kappa}}_{\text{Value From Future Sale}}, \tag{4}$$

where the second term captures the value from future sales. The liquidity shock received by the owner reduces the value of the coupon to $C(1-L)$ until she sells the bond. After receiving the liquidity shock she contacts her dealer network expecting to sell the bond for $\mathrm{E}[P^s]$. Upon selling she becomes a non-owner, completing the valuation cycle. Valuations V^s and V^{no} are related by

$$V^s = \frac{C(1-L)}{r+\Lambda} + (\mathrm{E}[P^s] + V^{no} - K)\frac{\Lambda}{r+\Lambda}, \tag{5}$$

This sequence of events continues in perpetuity and, therefore, we focus on the steady state of the model.

The above valuations equations depend upon the expected transaction prices. The realized transaction prices are determined by bilateral Nash bargaining. The clients reservation values are determined by the differences in values between being an owner and non-owner and a buyer and seller. Similarly, the dealers reservation values arise from their transaction cycle. Each dealer acts competitively, i.e., without taking into account the effect of her actions on the actions of other dealers. When a client contacts her dealer network each dealer simultaneously starts looking for the bond at rate λ and expects to pay the inter-dealer ask price M^{ask} for the bond. The value to the dealer searching for the bond is

$$\begin{aligned}(6)\qquad U^b &= \frac{1}{1+rdt}[\lambda dt(P^b - M^{ask}) + \Lambda dt U^o + (1-\Lambda dt)U^b] \\ &= \underbrace{(P^b - M^{ask})\frac{\lambda}{r+\Lambda}}_{\text{Transaction Profit/Loss}} + \underbrace{U^o\frac{\Lambda}{r+\Lambda}}_{\text{Value of Future Business}}.\end{aligned}$$

The last term in the expression for U^b captures the expected value of the future business with the same client which happens with frequency Λ. This client, who is now the owner of the bond, becomes a seller with intensity κ and contacts dealers in her Rolodex to sell the bond. This generates a value $U^o = U^s\frac{\kappa}{r+\kappa}$ per dealer, where U^s represents the valuation of the dealer searching to sell the bond. The dealer expects to resell the bond at rate λ for the inter-dealer bid price M^{bid} and earn the realized spread, $M^{bid} - P^s$. She also anticipates that with frequency Λ the same client will approach her in the future to buy back the bond. Future business from the same client generates $U^{no} = U^b\frac{\eta}{r+\eta}$ in value to the dealer, thus leading to the following expression for U^s:

$$(7)\qquad U^s = \underbrace{(M^{bid} - P^s)\frac{\lambda}{r+\Lambda}}_{\text{Transaction Profit/Loss}} + \underbrace{U^{no}\frac{\Lambda}{r+\Lambda}}_{\text{Value of Future Business}}.$$

Valuations U^{no} and U^o lead to price improvement for repeat business.

As in most OTC models, prices are set by Nash bargaining resulting in:

$$\begin{aligned}(8)\qquad P^b &= (V^o - V^b)w + (M^{ask} - U^o)(1-w), \\ (9)\qquad P^s &= (V^s - V^{no})w + (M^{bid} + U^{no})(1-w).\end{aligned}$$

Prices are the bargaining-power (w) weighted average of the reservation values of the client and dealer. The above equations assume that the dealer loses all future business from the client if the bilateral negotiations fail. Upon dropping a dealer the client maintains her optimal network size by forming a new link with another randomly picked identical dealer. Thus, by agreeing rather than not, the dealer receives U^o. As a consequence, dealers face intertemporal competition for future clients. This is a novel assumption missing from the existing models of OTC markets.

Bargaining power could differ for buys and sells. For ease of exposition, we equate them here. In the subsequent structural estimation we allow for different bargaining powers when the insurer is looking to buy a bond, w^b, and when the insurer is selling a bond, w^s. The valuations and prices provide ten equations and ten unknowns. Proposition 1 provides the closed form solutions for prices.

PROPOSITION 1: *Bid and ask prices are*

$$\begin{aligned} P^b &= \left[\frac{C(1-L)}{\Lambda}\Psi_1 + \frac{C}{r+\kappa}\Psi_2 + K\Pi_1(\kappa,\eta)\right] w \\ &\quad + \left[M^{ask}\Pi_2 + M^{bid}\Pi_3\frac{\kappa}{r+\kappa}\right](1-w), \end{aligned} \tag{10}$$

$$\begin{aligned} P^s &= \left[\frac{C(1-L)}{\Lambda}\Psi_3 - \frac{C}{r+\kappa}\Psi_4 - K\Pi_1(\eta,\kappa)\right] w \\ &\quad + \left[M^{bid}\Pi_2 + M^{ask}\Pi_3\frac{\eta}{r+\eta}\right](1-w). \end{aligned} \tag{11}$$

where the coefficients Ψ_1*,* Ψ_2*,* Ψ_3*,* Ψ_4*,* $\Pi_1(x,y)$*,* Π_2*, and* Π_3 *are given in the Appendix.*

Proof: See Appendix.

Expressions (10) and (11) are nonlinear functions of the model primitives and N, which makes the analytical analysis impossible. However, we can verify that prices are well-behaved functions of the Rolodex size, N, in the large network limit, $N \to \infty$. $N \to \infty$ implies $\frac{\Lambda}{r+\Lambda} \to 1$ and clients' search friction in terms of time is zero. Dealers' valuations, in this case are denoted with subscript $N \to \infty$, satisfy the system of equations

$U^b_{N\to\infty} = U^s_{N\to\infty}\frac{\kappa}{r+\kappa}$ and $U^s_{N\to\infty} = U^b_{N\to\infty}\frac{\eta}{r+\eta}$. These only have a trivial solution $U^s_{N\to\infty} = U^b_{N\to\infty} = 0$, implying that dealers compete away all rents from future relations with clients. Clients have valuations $V^o_{N\to\infty} - V^b_{N\to\infty} = P^b_{N\to\infty} + K$ and $V^s_{N\to\infty} - V^{no}_{N\to\infty} = P^s_{N\to\infty} - K$ yielding the following expressions for transaction prices:

$$P^b_{N\to\infty} = M^{ask} + \frac{w}{1-w}K, \tag{12}$$

$$P^s_{N\to\infty} = M^{bid} - \frac{w}{1-w}K. \tag{13}$$

Dealers receive no relationship-based rents, but they charge clients a spread $\frac{w}{1-w}2K$ per roundtrip transaction over the inter-dealer spread. The coefficient $\frac{w}{1-w}$ indicates that the non-zero bargaining power enables dealers to extract some value from a client. This is because every dealer charges the same price thus making clients' threat of ending the relationship and forming a new link not credible (Diamond's (1971) paradox). Equations (12) and (13) illustrate the importance of making $K > 0$ in the model. If K is equal to zero a client can choose an infinitely large network. This increases dealer competition such that the client trades at the inter-dealer prices, effectively becoming a dealer herself. Therefore, the finite network is never optimal if K is equal to zero.

Overall, dealers' surplus comes from both the immediate value of trade (the spread) and the value of future transactions. As the probability of transacting with the same client in the future declines with the size of the client's network, dealers charge higher spreads. Therefore, we postulate that there exists a set of model parameters for which the transaction prices are monotonic functions of N and $\frac{dP^b}{dN} > 0$ and $\frac{dP^s}{dN} < 0$. This result is similar to findings by Vayanos and Wang (2007) where an asset with more buyers and sellers has lower search times and trades at a worse prices relative to its identical-payoff counterpart with fewer buyers and sellers. However, Vayanos and Wang (2007) do not model client-dealer repeated interactions.

We verify that the model can yield a finite Rolodex by considering a limiting case of a very large search intensity, $\lambda \to \infty$. In this case a single dealer can instantaneously find the bond and the optimal size of the Rolodex is one. Taking this limit in equations (3)–(7) yields $(V^o - V^b)_{\lambda\to\infty} = P^b_{\lambda\to\infty} + K$, $(V^s - V^{no})_{\lambda\to\infty} = P^s_{\lambda\to\infty} - K$ and, hence,

$$(14) \qquad U^{no}_{\lambda\to\infty} = \frac{1}{N}\frac{\frac{\eta}{r+\eta}}{1-\frac{\kappa}{r+\kappa}\frac{\eta}{r+\eta}}\left[P^b_{\lambda\to\infty} - M^{ask} + \frac{\eta}{r+\eta}(M^{bid} - P^s_{\lambda\to\infty})\right],$$

$$U^{o}_{\lambda\to\infty} = \frac{1}{N}\frac{\frac{\kappa}{r+\kappa}}{1-\frac{\kappa}{r+\kappa}\frac{\eta}{r+\eta}}\left[M^{bid} - P^s_{\lambda\to\infty} + \frac{\kappa}{r+\kappa}(P^b_{\lambda\to\infty} - M^{ask})\right].$$

Solving for transaction prices from (8-9) we immediately obtain

$$(15) \qquad P^b_{\lambda\to\infty} = M^{ask} + \frac{w}{1-w}K - U^o_{\lambda\to\infty},$$

$$P^s_{\lambda\to\infty} = M^{bid} - \frac{w}{1-w}K + U^{no}_{\lambda\to\infty}.$$

Expressions (15) further confirm that $U^o_{\lambda\to\infty}$ and $U^{no}_{\lambda\to\infty}$ represent the repeat relation buy discount and sell premium, respectively. Equation (14) shows that both $U^o_{\lambda\to\infty}$ and $U^{no}_{\lambda\to\infty}$ are strictly decreasing with the size of the Rolodex, N. Therefore, the client optimally chooses a single dealer.

In order to find the optimal Rolodex size N^* in general case, we maximize the valuation of the "first-time" owner, i.e., the client buying the asset, V^b. This is because the client has to take possession of the asset in the first place. Proposition 2 shows that, for a given client type described by the model primitives $\{L, K, w, \kappa, \lambda, \eta\}$, there may exist an optimal network size $N^* = N(L, K, w, \kappa, \lambda, \eta)$.

PROPOSITION 2: *The optimal size of the insurer's dealer network N^* is given by the following condition*

$$(16) \qquad \frac{(\lambda N^*)^2}{1+\frac{\kappa}{r+\kappa}\frac{\eta}{r+\eta}(\frac{\lambda N^*}{r+\lambda N^*})^2}\left[\left(\frac{C(1-L)}{r} - P^s + K\right)\frac{\kappa}{r+\kappa}\frac{r}{(r+\lambda N^*)^2} + \frac{dP^b}{d\Lambda} - \frac{dP^s}{d\Lambda}\frac{\kappa}{r+\kappa}\frac{\lambda N^*}{r+\lambda N^*}\right] = rV^b.$$

Proof: See Appendix.

Next we provide intuition for why the model yields finite size networks greater than one. Because equation (16) is nonlinear in N^* the existence of the solution is not guaranteed for an arbitrary parameter set $\{L, K, w, \kappa, \lambda, \eta\}$. Since V^b is nonnegative for all N, the term $\frac{V^b}{N^2}$ in the derivative $\frac{dV^b}{dN} = \frac{\Lambda}{r+\Lambda}(\frac{V^b}{N^2} + \frac{dV^o}{dN} - \frac{dP^b}{dN})$ is strictly decreasing with N, and using our earlier argument that there exists a set of parameter values under which P^b is a monotonically increasing function of N, all we need for the optimum to exists, i.e., $\frac{dV^b}{dN} = 0$ for some $N = N^* > 0$, is to have V^o as a non-increasing function of N. $\frac{dV^s}{dN}|_{N=N^*}$ at the optimum,[4] i.e., where $\frac{dV^b}{dN}|_{N=N^*} = 0$, is

$$(17) \qquad \frac{dV^s}{dN}|_{N=N^*} = -\frac{C(1-L)}{N^*(r+\lambda N^*)} + \frac{\lambda N^*}{r+\lambda N^*}\frac{dP^s}{dN}|_{N=N^*} + \frac{rV^s}{N^*(r+\lambda N^*)}.$$

Therefore, as long as the following inequality holds at the optimum

$$(18) \qquad V^s \leq \frac{C(1-L)}{r} - \frac{\lambda\,(N^*)^2}{r}\frac{dP^s}{dN}|_{N=N^*},$$

the derivative $\frac{dV^o}{dN}|_{N=N^*}$ is nonpositive and the solution to (16) exists. Inequality (18) implies that at the optimal Rolodex size the value of sales is no greater than the value of holding the discounted asset and the present value of all marginal sale price improvements from not increasing the size of the network.

A finite optimal size of the Rolodex follows from several important tradeoffs. When a client adds another dealer to her Rolodex the speed of execution, measured by $\frac{\Lambda}{r+\Lambda}$, improves. The additional dealer worsens prices as the value of the relationship is shared by more dealers. In addition, the transaction prices include indirect effects additional dealers have on both client and dealer reservation values. When deciding to add another dealer the client trades off the effects on speed and prices. More dealers lead to faster transactions. For there to be a finite network size the transaction price in aggregate must worsen. However, it is not necessarily the case that both buy and sell prices worsen. The optimal size of the Rolodex is set when the benefits from transacting with the larger number of dealers equal the costs.

[4]We have that $\frac{dV^o}{dN}|_{N=N^*} = \frac{\kappa}{r+\kappa}\frac{dV^s}{dN}|_{N=N^*}$.

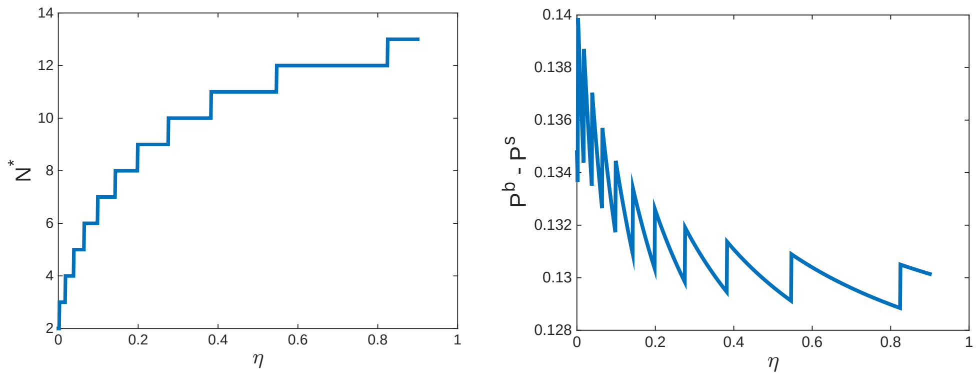

FIGURE 4.— Model predictions

The figure illustrates the comparative statics with respect to the trading intensity η. The left plot shows that the optimal "Rolodex" size N^* increases with client's trading needs proxied by η. The right plot shows that clients with larger trading needs proxied by η receive better execution as measured by the realized spread $P^b - P^s$, calculated at the optimal N^*. The model is solved numerically using the parameters provided in the Online Appendix.

3.2. *Model predictions for execution costs*

The model provides a number of comparative statics results that we can examine in the data. While there is not client heterogeneity in the model, clients who trade more frequently with having higher trade intensity η. Thus, we can examine how larger trading needs impact clients' choice of network size and the transaction prices they receive.

We first discuss the effect of η on transaction prices while keeping the size of the Rolodex fixed. When $\eta = 0$ the client who is currently an owner will never be an owner again after she sells the bond. Therefore, all of the dealers' surplus comes from the price at which the sale occurs and not from the repeated business. They charge high realized spread to compensate for it. When η is large, dealers derive significant value from the repeated business, leading to smaller realized spreads.

HYPOTHESIS 1: *Larger clients and clients with more trading needs (high η) should get better execution than smaller clients and clients with less trading needs (low η).*

The optimal size of the Rolodex depends on the trade propensity η through two channels. First, increasing η improves transaction prices thus allowing the client to increase N, which, at the optimal network size, has an opposite effect on transaction prices. The second channel is through $\frac{dV^s}{dN}|_{N=N^*}$. It follows from expression (17) that $\frac{dV^s}{dN}|_{N=N^*}$ increases with η for a given N, as both $\frac{dP^s}{dN}|_{N=N^*}$ and V^s are increasing with η.[5] Intuitively, dealers' repeated business improves with the increasing trading intensity and dealers offer better execution. Clients who need to trade more due to increasing trading intensity benefit from a larger Rolodex which improves the speed of execution. There exist a set of parameters for which these benefits dominate losses from the degrading execution quality as N increases. Figure 4 shows that N^* is increasing with η. The model, therefore, predicts that larger insurers with more frequent trading needs should have larger dealer networks, which is consistent with our empirical findings from the previous section.

Together with Hypothesis 1 this result implies that clients with a larger Rolodex should get better execution thus leading to our second hypothesis. This prediction is not present in other OTC models because they lack an endogenous network size.

HYPOTHESIS 2: *Execution quality should be increasing with the size of the client-dealer network.*

Both plots in Figure 4 have a "stair-like" appearance. When we numeri-

[5] V^s is increasing in the sales price P^s which improves with larger η. V^s is also proportional to $V^{no} = V^b \frac{\eta}{r+\eta}$, which increases with η.

cally solve for the optimal Rolodex size, N^*, we use a grid of integer values for N, while the grid for η is very refined. As a result, a small change in η leads to the same optimal value of N^*. When N^* stays the same, the realized spread monotonically decreases with η in accordance with our intuition. However, when N^* increases the realized spread jumps up as the negative effect from the increase in N overwhelms the positive effect coming from the increase in η.[6] These "stairs" are smoothed out if we allow for non-integer values for N.

To summarize, the model predicts that clients should establish permanent links with a finite number of dealers and clients who have more trading needs, e.g., larger clients, select larger dealer networks. Therefore, larger clients get better prices despite choosing larger networks which result in the loss of repeat business. In the next section we test our hypotheses for execution costs in the NAIC data.

4. INSURER EXECUTION COSTS AND NETWORKS

Tables II and III suggest that bond characteristics impact insurers' trading needs. If different types of insurers hold different bonds, or trade them at different times, it is difficult to test the model's prediction regarding trading costs. To control for bond, time, and bond-time variation, we compare transaction prices to daily bond-specific Bank of America-Merrill Lynch (BAML) bid (sell) quotes. BAML is the largest corporate bond dealer, transacting with more than half of all insurers for almost 10% of both the trades and volume. BAML (bid) quotes can be viewed as representative quotes for insurer sales and enable us to measure prices relative to a transparent benchmark price. The BAML quotes essentially provide bond-time fixed effects, which

[6] Increasing η by a small amount $d\eta$ should result by a small increase in N, dN. However, since we use a grid of integers for N, the increase in N, when it happens, is disproportionately large.

would be too numerous to estimate in our sample. Our relative execution cost measure in basis points is defined as

$$\text{(19)}\quad \text{Execution cost (bp)} = \frac{\text{BAML Quote} - \text{Trade Price}}{\text{BAML Quote}} * (1 - 2 * \mathbf{1}_{Buy}) * 10^4,$$

where $\mathbf{1}_{Buy}$ is an indicator for whether the insurer is buying or selling. Because some quotes may be stale and lead to extreme costs estimates, we windsorize the distribution at 1% and 99%.

Execution costs depend on the bond being traded, time, whether the insurers buys or sells, the insurer's characteristics, dealer identity and characteristics, and the insurer network size. To examine the relationship specific effects on execution costs we control for bond and time fixed effects. In principle if the BAML perfectly controls for bond-time effects, the additional bond and time fixed effects would be unnecessary.

The relationship component of transaction costs depends on the properties of the insurers' networks. Figure 3 and equation (1) indicate that the degree distribution for insurer-dealer relations follows a Gamma distribution. Therefore, we include both the size of the network, N, and its natural logarithm, $\log(N)$, as explanatory variables. We control for seasonality using time fixed effects, α_t, and for unobserved heterogeneity using either bond characteristics or bond fixed effects, α_i. The other explanatory variables consist of either insurers' or dealers' characteristics, or both. We estimate the following panel regression for execution costs in bond i at time t:

$$\text{(20)}\qquad \text{Execution cost}_{it} = \alpha_i + \alpha_t + \beta N + \gamma lnN + \theta X_{it} + \epsilon_{it}.$$

The set of explanatory variables X includes characteristics of the bond, as well as features of the insurer and dealer.

Table V provides trading costs estimates from panel regressions. We adjust standard errors for heteroskedasticity and cluster them at the insurer, dealer, bond, and day level. The coefficient on insurer buy captures the

TABLE V

EXECUTION COSTS AND INVESTOR-DEALER RELATIONS

The table reports the determinants of execution costs. Execution costs are expressed in basis points relative to the Merrill Lynch quote at the time of the trade. Standard errors are adjusted for heteroskedasticity and clustered at the insurer, dealer, bond, and day level. See caption of Table II for additional details.

Determinant	(1)	(2)	(3)	(4)	(5)	(6)
Insurer size	-4.89***		-3.72***	-3.59***	-3.52***	-3.95***
Insurer no. of dealers		-0.37***	-0.22***	0.32***	0.32***	0.32***
ln(Insurer no. of dealers)				-6.29***	-6.51***	-6.55***
Insurer RBC ratio	-3.57***	-0.67	-3.51***	-4.19***	-4.68***	-5.40***
Insurer cash-to-assets	-0.04**	-0.02	-0.04**	-0.04**	-0.04**	-0.03*
Life insurer	4.66***	3.32***	4.43***	4.47***	5.73***	7.21***
P&C insurer	2.26***	2.10***	1.72**	1.73**	1.99**	2.82***
Insurer rated A-B	-0.51	1.71***	-0.47	0.01	-0.18	-0.50
Insurer rated C-F	11.67*	14.97**	11.29*	11.13*	10.90*	12.18*
Insurer unrated	0.93	3.88***	0.74	0.62	0.55	0.14
Insurer buy	39.77***	39.67***	39.56***	39.27***	39.71***	40.17***
Trade size × Buy	-0.19**	-0.64***	-0.26***	-0.25***	-0.20**	-0.18*
Trade size × Sell	0.59***	0.19*	0.53***	0.50***	0.48***	0.52***
Bond issue size					-0.00***	-0.00***
Bond age					0.58***	0.63***
Bond remaining life					0.81***	0.82***
Bond HY rated					4.54***	4.16***
Bond unrated					-5.96***	-6.53***
Bond privately placed					3.38***	3.26***
Dealer size						-5.37***
NYC dealer						-6.66***
Primary dealer						2.43
Dealer leverage						-5.68**
Dealer diversity						0.41***
Dealer dispersion						0.07
Local dealer						0.23
Dealer distance						-0.20
Dealer leverage missing						-6.41**
Dealer dispersion missing						6.09
Bond fixed effects (16,823)	Yes	Yes	Yes	Yes	No	No
Dealer fixed effects (401)	Yes	Yes	Yes	Yes	Yes	No
Day fixed effects (3,375)	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.154	0.154	0.154	0.155	0.103	0.098
N	918,279	918,279	918,279	918,279	918,987	891,875

average bid-ask of roughly 40 basis points. Consistent with hypothesis 1, column (1) of Table V shows that large insurers pay on average lower execution costs. An unsurer with 10 times as many assets has trading cost 4.89 basis points lower. Better capitalized insurers (higher RBC ratio) get better prices. Column (2) of Tables V replaces insurer size with the number of dealers in the insurer's network. Consistent with hypothesis 2 execution costs decline with insurer network size. An insurer with an additional dealer has trading cost 0.37 basis point lower. Column (3) includes both insurer size and network size and shows that conditional on network size the insurer size still matters.

Columns (5) and (6) replace bond and dealer fixed effects with bond and dealer characteristics. NYC-located dealers offer better prices to all insurers and more diversified dealers charge, on average, higher prices. Bond characteristics matter for execution costs as insurers receive worse prices for special bonds and better prices for bonds with larger issue size. Insurers get better prices on unrated bonds.

In summary, Table V is consistent with our model's qualitative prediction that insurers get better execution if they have larger network. We next examine whether the model can quantitative match the distribution of insurer's network sizes and execution costs.

4.1. *Model estimation*

Settings such as labor markets and marriages markets involve one-to-one matching. Papers in these literatures structurally estimate their models to examine how quantitatively well they can fit the data (for example, Poste-Vinay and Robin (2002) for labor search and Hitsch et al. (2010) and Choo (2015) for marriage search. We follow this approach for our one-to-many network formation model. The model is solved assuming the client maximizes her value as a buyer. It is possible to solve the model assuming the client maximizes her value as a sell, but to simplify exposition we structurally estimate the model using the buying network sizes. We can then use these estimated parameters to examine how well the model fits the distribution of selling network sizes.

Figures 2 and 3 shows insurers heterogeneous trading needs and network sizes. To estimate the model's distribution of network sizes we must infer the distribution of trading needs, η_n, $n = 1, ..., \mathcal{N}$, across insurance companies n. Section 2.1 provides the compound distribution of trading activity. If trading shocks occur at Poisson times, the intensity η_n of the shock can be estimated by the expected number of buy trades per year. This yields the maximum likelihood estimator $\widehat{\eta_n} = \frac{1}{T}\sum_{t=1}^{T} X_{nt}$, where X_{nt} is the number of bond purchases by insurer n in year t. To utilize the multiple years of trade data, we perform the estimation separately for each insurer in the sample. This yields a cross-sectional distribution of trading intensities, which we index by $p(\eta)$. Section 2.1 and Figure 2 show the distribution in the insurer trades per year follows approximately a power law. For insurer buys, this distribution is best described by $p(\eta) = .34 \times \eta^{-1.31}$. Thus, wide variation in trading needs exists across insurers.

The remaining model parameters $\Theta = (L, K, \kappa, \lambda, w^s, w^b)$ are not directly observable in the data, so they need to be estimated structurally. Section 2.2 and Figure 3 show the degree distribution of client-dealer relations follows a mixed power-exponential law in the data. The model-generated distance between the empirical distribution and the network sizes N^* depends on the parameters Θ. We estimate this by minimizing the probability-weighted distance between the data and model:

$$\sum_{N=1}^{max(N)} p(N)[N - N^*(\Theta, \eta(p(N)))]^2, \tag{21}$$

where $p(N)$ is the network-size distribution in the NAIC data and $N^*(\Theta, \eta)$ is the model-implied Rolodex size for the set of parameters Θ given η.

Inverting the power law distribution for trading intensities from Panel A of Figure 2 yields $\eta(p) = (\frac{p}{.34})^{-\frac{1}{1.31}}$. It then follows from equation (1) that for insurer buys $p(N) = .28e^{-.22N}N^{-0.12}$. Therefore, by substituting $p(N)$ into the expression for $\eta(p)$ we obtain the mapping of trading intensities into the network sizes, $\eta(p(N)) = (\frac{.28}{.34})^{-\frac{1}{1.31}}e^{\frac{.22}{1.31}N}N^{\frac{0.12}{1.31}}$. This is the relation we use in the expression (21).

The model also predicts how percentage trading costs, $(P^b - P^s)/.5(P^b + P^s)$, depend on the parameters Θ and η. The model's optimal solution incorporates prices at which clients both buy and sell. Empirically, the estimated relation between trading costs and network size is in column (4) of Table V as $c(N) = 51 + 0.32N - 6.28\, lnN$. The probability-weighted distance between the empirical and the model-generated trading costs c^* depends on the parameters Θ and is:

$$\sum_{N=1}^{max(N)} p(N)[c(N) - c^*(\Theta, \eta(p(N)))]^2. \tag{22}$$

A minimum-distance estimator sets the model parameters to minimizes the two distances (21) and (22); we choose equal weights. Because the optimal N^* in the model is a step function in Figure 4, the objective functions (21) and (22) are non-smooth. To accommodate these complications we optimize using simulated annealing with fast temperature decay and fast reannealing.

To ensure that the estimated parameters remain in their natural domains, i.e., positive or between zero and one, in the estimation we transform model parameters as follows:

$$\begin{aligned} L &= \Phi(\theta_L) \in (0,1), \quad K = e^{\theta_K} \geq 0, \quad \kappa = e^{\theta_\kappa} \geq 0, \quad \lambda = e^{\theta_\lambda} \geq 0, \\ w^s &= \Phi(\theta^0_{w^s} + \theta^1_{w^s} ln\, \eta) \in (0,1), \quad w^b = \Phi(\theta^0_{w^b} + \theta^1_{w^b} ln\, \eta) \in (0,1), \end{aligned} \tag{23}$$

where $\Phi \in (0,1)$ is the normal cdf. The propensity of insurers receiving trading shocks overall is heterogeneous so we allow overall trading intensity η to vary across insurers. In the model κ measures clients' propensity to sell after having already bought. Because empirically the number of buys and sells are equal and buys and sells follow a similar power law there is not evidence of heterogeneity in κ. Therefore, we assume κ is constant across insurers.

The evidence in Table V shows trading costs depend on insurer size even after controlling for network size. The most straightforward way to accommodate this is to allow dealer bargaining power to vary across insurers. We report the model fit under both constant and variable bargaining power by holding w^b and w^s constant and allowing w^b and w^s to be functions of η. Because we fit the model to the two observable outcomes of network size and trading costs we limit the number of model parameters that vary across insurer by assuming L and K are constant. Because insurers can potentially access all dealers λ is constant across insurers. Our results are robust to alternative transformations in place of Φ and e .

We set the competitive interdealer spread to match the average execution costs of insurers by choosing $C = 1$, $r = .1$, $M^{ask} = (1 + 0.002)(\frac{C(1-L)}{r} + \frac{C}{r})/2$, and $M^{bid} = (1 - 0.002)((\frac{C(1-L)}{r} + \frac{C}{r})/2$.

Table VI reports the estimates for the model parameters $\Theta = (L, K, \kappa, \lambda, w^s, w^b)$.[7] All parameters are significantly different from zero and appear well identified as the standard deviations of the residuals are small. Specification 2 allows bargaining power to vary across insurers and generates a substantially lower minimum distance between model and empirical distribution than specification 1.

The liquidity shock parameter L is estimated to be 89% (84%) of the flow income from the bond in specification 1 (2), suggesting a high willingness to pay for immediacy. The flow cost K for obtaining quotes from an additional dealer is small, 0.4 and 2.6 basis points per bond and unit of time, respec-

[7]Estimates for transformed parameters $\theta = (\theta_L, \theta_K, \theta_\kappa, \theta_\lambda, \theta_w)$ are available in Online Appendix.

tively. The intensity κ of the selling shock is estimated between 15 and 20, a holding period of few weeks. Longer holding periods, smaller κ, reduces the value of repeat relations and increases network size. Without heterogeneity in κ the model cannot reproduce the large fraction of insurers using a single dealer with holding periods longer than a few weeks. The dealers' search efficiency λ is estimated between 80 and 140. This suggests dealers take between one and a half and three days $(250/\lambda)$ to locate a bond. This number seems realistic if not a bit large given the search frictions present in the market.

The bargaining power of the dealers is estimated to be different between the buy and the sell side, suggesting in both specifications that the buyer has little bargaining power while the seller captures most of the trade surplus. In specification 2, when we allow bargaining power to depend on the insurer's type, we find that larger insurers with higher trading frequency η have higher bargaining power than small insurer with more infrequent trading. In both specifications dealers earn significant profits on insurer buys. This could arise from short selling constraints whereby a dealer can more easily locate a buyer than a seller. The heterogeneity in bargaining across insurers and across buy and sell transactions suggests richer modelling of the price setting process may enable deeper understanding of OTC trading.

Panel B of Table VI and Figure 5 examine the models' fit to the data in detail. Given the network sizes follow a mixed power-exponential law, we first visually examine how well the models' network distribution fits its empirical counterpart. Figure 5 shows the degree distribution for insurer-dealer relations in the data (circles) and the interpolated power law with exponential tail (dashed line) compared to the model-implied distribution under the estimated parameters (solid line). The distance between the two lines captures the model fit. The two lines are virtually on top of each other for all values of N less than 20. The Rolodex model thus fits the distribution of client-dealer relation well for all insurers with small to medium-sized networks. The model has some difficulty in matching the exponential tail of large dealer networks perfectly as it slightly overpredicts the propensity to see very large networks in the data. The right plot for the distribution of selling network sizes is produced using model estimates based only on the buying networks. The fit for the selling network is similar to the buying networks except for the largest selling networks.

To verify more formally that the network sizes from the model correlate strongly with the data, Panel B of Table VI runs the regression $N\,(\text{data}) = \alpha + \beta N^*\,(\text{model}) + \varepsilon$ to see if α is 0 and β equals 1. Specification 1 generates a 91% correlation with the empirical distribution, while specification 2 matches the data very closely, as depicted in Figure 5.

Another approach to examining the models' fit with the data is to test how closely the model-implied probability distribution of networks sizes fits the Gamma distribution in (1). Replicating the fitting regression, $ln\,p(N^*) = \alpha + \beta N^* + \gamma lnN^* + \varepsilon$, on the model-generated distribution of network sizes is given in Panel B of Table VI. Specification 1 provides a 95% fit while specification 2 has a R^2 of 99%.

The next set of regression coefficients in Panel B of Table VI examines the models fit for trading costs in a similar way to network sizes. First, we examine if the model can generate the distribution of trading costs in the data If the model exactly fits the data, the coefficients in the regression $\text{Trading cost}\,(\text{data}) = \alpha + \beta\,\text{Trading cost}\,(\text{model}) + \varepsilon$ should be $\alpha = 0$ and $\beta = 1$. In specification 2 has a $R^2 = 0.91$ and β is larger than one, 1.13. The suggests that, given the approximate power law for η in the data, a single set of parameters (L, K, κ, λ) are unlikely to be the same across all insurers. Second, we examine if the model can reproduce the dependence of trading costs on network size in Table V. In both specifications, the coefficients α, β, and γ in the regression $\text{Trading cost} = \alpha + \beta N^* + \gamma lnN^* + \varepsilon$ have the

TABLE VI

Estimated model parameters

Panel A reports the estimated model parameters $\theta = (L, K, w^i, \kappa, \lambda)$ with $w^i = \Phi(\theta^0_{w^i} + \theta^1_{w^i} ln\, \eta)$ for $i = s, b$, where $\Phi \in (0, 1)$ is the normal cdf. Estimates are from the minimum-distance estimator (21) and (22). Standard errors are reported in parenthesis and scaled by 100. Panel B reports model-implied distribution of network sizes and execution costs at the estimated parameters. Corresponding empirical values are reported in parenthesis.

	Specification 1	Specification 2
Panel A: Implied model parameters		
L	0.89	0.84
K*100	0.04	0.26
κ	15.04	19.96
λ	141.46	79.05
w^s (S.D.)	0.05 (0.00)	0.16 (0.31)
w^b (S.D.)	0.94 (0.00)	0.45 (0.21)
Minimum distance	588.31	38.19
S.D. residuals	2.73	0.70
S.D. residuals network	2.55	0.09
S.D. residuals prices	0.96	0.69
Panel B: Model fit		
Data-model correlation in network sizes: $N\,(\text{data}) = \alpha + \beta N^*\,(\text{model}) + \varepsilon$		
α (0), β (1)	-12.08, 2.94	0.58, 0.97
R^2	0.91	0.99
Network distribution: $ln\, p(N^*) = \alpha + \beta N^* + \gamma lnN^* + \varepsilon$		
α (-1.27), β (-0.22), γ (-0.12)	-4.34, -1.28, 5.34	-1.06, -0.20, -0.36
R^2	0.95	0.99
Data-model correlation in prices: $\text{Trading cost}\,(\text{data}) = \alpha + \beta\, \text{Trading cost}\,(\text{model}) + \varepsilon$		
α (0), β (1)	-40.08, 1.92	-6.11, 1.13
R^2	0.98	0.91
Price distribution: $\text{Trading cost} = \alpha + \beta N^* + \gamma lnN^* + \varepsilon$		
α (51), β (0.32), γ (-6.29)	47.59, 0.21, -3.56	49.65, 0.19, -4.50
R^2	0.99	0.95

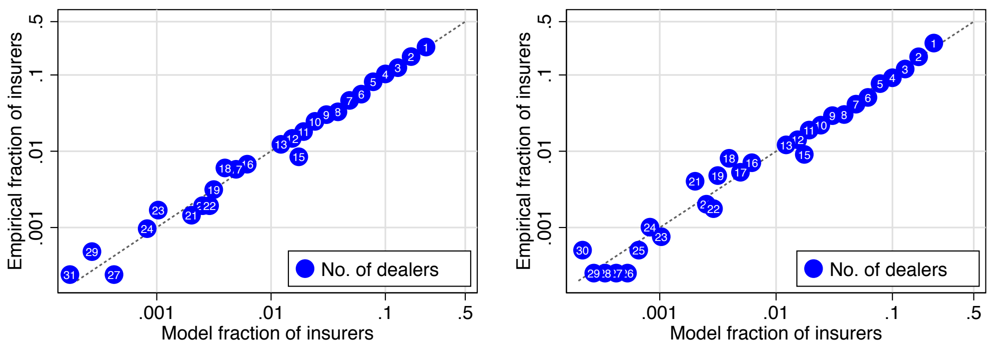

FIGURE 5.— Model fit

The figure shows on a log-log scale the degree distribution for insurer-dealer relations in the data compared to the model-implied distribution under the estimated parameters. The left (right) plot is for insurers' buys (sells). Both plots are generated under the same set of parameters. The deviation from the 45-degree line measures model fit.

same sign as in the data in Table V. Their magnitudes are not exact. Prices decline less steeply with N in the model than in the data.

The parameter estimates can help quantify the value of repeat client-dealer relations. One approach to capture the impact of repeat business on prices is to compute counterfactual bid-ask prices under the assumption that the dealers do not give price improvement for repeat business from the same insurer. This implies that the search and bargaining proceeds in the same way, but that the dealers in the network are essentially re-chosen with every trade. In this counterfactual scenario, bid and ask prices can be calculated in the model by setting $U^o = U^{no} = 0$. After determining price-improvement-free bid and ask prices for insurers with different values of η, we can compare them to their model-implied values with price improvement due to repeat business. Figure 6 plots the bid-ask prices in the model under the estimated parameters from specification 2 in Table VI for different values of N (solid lines). Figure 6 also graphs the counterfactual bid-ask prices (dashed lines) that would arise under the same parameters without dealer

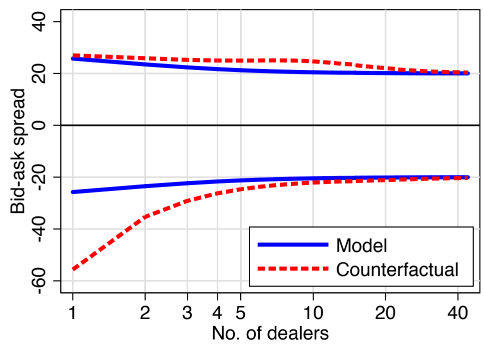

FIGURE 6.— Counterfactual analysis

The figure shows the bid-ask prices in the model under the estimated parameters (solid line) compared to the counterfactual bid-ask spread that would arise under the estimated parameters, if the dealers did not give price concessions for repeat business from the insurer. In the counterfactual scenario, we determine bid and ask prices by setting $U^o = U^{no} = 0$. The distance between the solid and the dashed lines can be interpreted as a measure of the value of repeat relations. We use a log-linear scale.

price-improvement for repeat business. The distance between the solid and dashed lines captures the value of repeat relations.

Figure 6 shows the impact of price improvement is significant and asymmetric. When dealers are buying, insurers with a single dealer capture price improvement of up to 30 basis points by committing to a singleton network. The price improvement they receive is significant compared to their average bid-ask spread of 50 basis points. Small insurers thus reduce their trading costs from 80 to 50 basis points through order flow concentration. By contrast, dealers give little price improvement for repeat relations when they are selling bonds to insurers. The maximum is 5 basis points for insurers with medium-sized dealer network (the rebate peaks at $N = 8$) but low portfolio turnover ($\eta < 10$). The key advantage of small client-dealer networks in the data thus seems to be the provision of liquidity at low cost for insurers selling. By contrast, prices for buying bonds are not so much driven by repeat business considerations.

5. CONCLUSION

Over-the-counter markets are pervasive across asset classes and often considered poorly functioning due to lack of transparency and fragmented trading imposing search frictions. Regulators have attempted to address these concerns through increases in transparency and the implementation of Dodd-Frank legislation. Competitors have entered with more centralized markets. Changes in regulations and market structure will impact heterogenous investors differently. However, there is limited theory closely linked to empirical work to guide these decisions.

We use comprehensive regulatory corporate bond trading data for all U.S. insurance companies to study how investors choose their size of trading networks and how this impacts prices they receive. To understand these empirical outcomes we develop a parsimonious model for OTC trading between clients and dealers. Clients trade off the value of repeat relations with dealers against the benefits of competition among dealers. The value of repeat relations diminishes more slowly with the addition of dealers for clients with larger trading needs. Therefore, these larger clients use more dealers. Dealers provide better prices to these larger clients because their repeat business is more valuable. Insurance company data shows that 30% of insurers only trade with a single dealer each year. Using the structurally estimated model parameters we find that small insurers benefit most from repeat relations. This suggest that care is needed in regulations affecting heterogenous investors.

The paper provides a first step in understanding how investor heterogeneity impacts observed patterns of trading and prices in OTC markets. A number of other important dimensions require further study. First, what is the impact of dealer heterogeneity on network formation and prices? Second, how do long-lasting client-dealer relations form and severe? Third, is Nash bargaining a realistic price setting mechanism in OTC markets?

APPENDIX

Data Filters

TABLE VII

DATA FILTERS

Filter	Full sample	Corp. bonds
1. All trades from original filings (includes all markets, all trades since 2001)	19.1	4.5
2. Remove all trades that do not involve a dealer (e.g., paydown, redemption, mature, correction)	6.6	3.1
3. Remove duplicates, aggregate all trades of the same insurance company in the same CUSIP on the same day with the same dealer	6.5	3.1
4. Map dealer names to SEC CRD number, drop trades without a name match, drop trades with a dealer that trades less than 10 times in total over the sample period	6.1	2.9
5. Drop if not fixed coupon (based on eMaxx data), drop if outstanding amount information is in neither eMaxx nor FISD	5.3	2.5
6. Drop if trade is on a holiday or weekend	5.2	2.5
7. Drop if counterparty is "various"	4.1	2.1
8. Drop trades less than 90 days to maturity or less than 60 days since issuance (i.e., primary market trade)	2.8	1.5
9. Merge with FISD data, keep only securities that are not exchangeable, preferred, convertible, issued by domestic issuer, taxable muni, missing the offering date, offering amount, or maturity, and offering amount is not less than 100K	1.0	1.0

Parts 3 and 4 of Schedule D filed with the NAIC contains purchases and sales made during the quarter, except for the last quarter. In the last quarter of each year, insurers file an annual report, in which all transactions during the year are reported. Part 3 of Schedule D reports all long-term bonds and stocks acquired during the year, but not disposed of, while Part 4 of Schedule D reports all long-term bonds and stocks disposed of. In addition, all long-term bonds and stocks acquired during the year and fully disposed of during the current year are reported in the special Part 5 of Schedule D. NAIC's counterparty field reports names in text, which can sometimes be mistakenly typed. The bank with the most variation in spelling is DEUTSCHE BANK. We manually clean the field to account for different spellings of broker-dealer names.

We compile the information in Parts 3, 4, and 5 of Schedule D to obtain a comprehensive set of corporate bond transactions by all insurance companies regulated by NAIC.

We apply various FISD-based data filters based on Ellul et al. (2011) to eliminate outliers and establish a corporate bond universe with complete data. The data filters are in Table VII which summarizes the number of observations that is affected by each step. We exclude a bond if it is exchangeable, preferred, convertible, MTN, foreign currency denominated, puttable or has a sinking fund. We also exclude CDEB (US Corporate Debentures) bonds, CZ (Corporate Zero) bonds, and all government bond (including municipal bonds) based on the reported industry group. Finally, we also drop a bond if any of the following fields is missing: offering date, offering amount, and maturity. We restrict our sample to bonds with the offering amount greater than $10 million, as issues smaller than this amount are very illiquid and hence are rarely traded. Ellul et al. (2011) have used $50,000 which we find restrictive for our purpose. We windsorize the cash-to-asset ratio at 80 percent to remove extreme values.

Proofs

Proof of Proposition 1: Define $\widetilde{\lambda} = \frac{\lambda}{r+\Lambda}$, $\widetilde{\Lambda} = \frac{\Lambda}{r+\Lambda}$, $\widetilde{\eta} = \frac{\eta}{r+\eta}$, and $\widetilde{\kappa} = \frac{\kappa}{r+\kappa}$. We can rewrite both the client's and the dealer's valuations as

$$V^b = \frac{C}{r+\kappa}\widetilde{\Lambda} - (P^b + k)\widetilde{\Lambda} + V^s\widetilde{\kappa}\widetilde{\Lambda} = \frac{\widetilde{\Lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[\frac{C(1-l)}{r+\Lambda}\widetilde{\kappa} + \frac{C}{r+\kappa} - (P^b + K) + (P^s - k)\widetilde{\kappa}\widetilde{\Lambda}\right],$$

$$V^s = \frac{C(1-l)}{r+\Lambda} + (P^s - k)\widetilde{\Lambda} + V^b\widetilde{\eta}\widetilde{\Lambda} = \frac{\widetilde{\Lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[\frac{C(1-l)}{\Lambda} + \frac{C}{r+\kappa}\widetilde{\eta}\widetilde{\Lambda} + (P^s - K) - (P^b + K)\widetilde{\eta}\widetilde{\Lambda}\right],$$

and

$$U^s = (M^{bid} - P^s)\widetilde{\lambda} + U^b\widetilde{\eta}\widetilde{\Lambda} = \frac{\widetilde{\lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[(M^{bid} - P^s) + (P^b - M^{ask})\widetilde{\eta}\widetilde{\Lambda}\right],$$

$$U^b = (P^b - M^{ask})\widetilde{\lambda} + U^s\widetilde{\kappa}\widetilde{\Lambda} = \frac{\widetilde{\lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[(P^b - M^{ask}) + (M^{bid} - P^s)\widetilde{\kappa}\widetilde{\Lambda}\right].$$

After some algebra one obtains

$$
\begin{aligned}
V^o - V^b &= \frac{C}{r+\kappa} + V^s\widetilde{\kappa} - V^b \\
&= \frac{C}{r+\kappa} + \frac{\widetilde{\Lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[\frac{C(1-l)}{\Lambda}\widetilde{\kappa} - \frac{C(1-l)}{r+\Lambda}\widetilde{\kappa} - \frac{C}{r+\kappa}(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) + \right. \\
&\quad \left. + (P^b+K)(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) + (P^s-K)\widetilde{\kappa}(1-\widetilde{\Lambda})\right], \\
V^s - V^{no} &= V^s - V^b\widetilde{\eta} \\
&= \frac{\widetilde{\Lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[\frac{C(1-l)}{\Lambda}(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) - \frac{C}{r+\kappa}\widetilde{\eta}(1-\widetilde{\Lambda}) + \right. \\
&\quad \left. + (P^b+K)\widetilde{\eta}(1-\widetilde{\Lambda}) + (P^s-K)(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda})\right].
\end{aligned}
$$

Substituting these expressions into (8) and (9) yields

$$
\begin{aligned}
P^b &= \frac{C}{r+\kappa}w \\
&\quad + \frac{\widetilde{\Lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[\frac{C(1-l)}{\Lambda}\widetilde{\kappa} - \frac{C(1-l)}{r+\Lambda}\widetilde{\kappa} - \frac{C}{r+\kappa}(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) + (P^b+K)(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) + (P^s-K)\widetilde{\kappa}(1-\widetilde{\Lambda})\right] w \\
&\quad + M^{ask}(1-w) - \frac{\widetilde{\lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[(M^{bid}-P^s) + (P^b-M^{ask})\widetilde{\eta}\widetilde{\Lambda}\right]\widetilde{\kappa}(1-w), \\
P^s &= \frac{\widetilde{\Lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[\frac{C(1-l)}{\Lambda}(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) - \frac{C}{r+\kappa}\widetilde{\eta}(1-\widetilde{\Lambda}) + (P^b+K)\widetilde{\eta}(1-\widetilde{\Lambda}) + (P^s-K)(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda})\right] w \\
&\quad + M^{bid}(1-w) + \frac{\widetilde{\lambda}}{1-\widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2}\left[(P^b-M^{ask}) + (M^{bid}-P^s)\widetilde{\kappa}\widetilde{\Lambda}\right]\widetilde{\eta}(1-w).
\end{aligned}
$$

These expressions can be rewritten as

$$
\begin{aligned}
P^b A_1(w) &= \left[\frac{C(1-l)}{\Lambda}\widetilde{\kappa}\widetilde{\Lambda} + \frac{C}{r+\kappa}\right](1-\widetilde{\Lambda})w + KA_2(\widetilde{\kappa}, w) \\
&\quad + M^{ask}A_3(w) - M^{bid}\widetilde{\kappa}\widetilde{\lambda}(1-w) + P^s\widetilde{\kappa}A_4(w) \\
P^s A_1(w) &= \left[\frac{C(1-l)}{\Lambda}(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}) - \frac{C}{r+\kappa}\widetilde{\eta}(1-\widetilde{\Lambda})\right]\widetilde{\Lambda}w - KA_2(\widetilde{\eta}, w) \\
&\quad - M^{ask}\widetilde{\eta}\widetilde{\lambda}(1-w) + M^{bid}A_3(w) + P^b\widetilde{\eta}A_4(w),
\end{aligned}
\tag{24}
$$

where we have defined

$$
\begin{aligned}
A_1(w) &\equiv 1 - \widetilde{\Lambda}w - \widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2(1-w) + \widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}\widetilde{\lambda}(1-w), \\
A_2(x, w) &\equiv [1 - \widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda} - x(1-\widetilde{\Lambda})]\widetilde{\Lambda}w, \\
A_3(w) &\equiv [1 - \widetilde{\kappa}\widetilde{\eta}(\widetilde{\Lambda})^2 + \widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}\widetilde{\lambda}](1-w), \\
A_4(w) &\equiv (1-\widetilde{\Lambda})\widetilde{\Lambda}w + \widetilde{\lambda}(1-w), \\
A_5(w) &\equiv A_1(w) - \widetilde{\kappa}\widetilde{\eta}\frac{A_4(w)^2}{A_1(w)}.
\end{aligned}
\tag{25}
$$

The system of equations (24) can be solved to yield expressions (10) and (11), where we have defined

$$
\begin{aligned}
(26) \qquad \Psi_1 &\equiv \frac{\widetilde{\kappa}\widetilde{\Lambda}}{A_5(w)}\left[1-\widetilde{\Lambda}+(1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda})\frac{A_4(w)}{A_1(w)}\right],\\
\Psi_2 &\equiv \frac{(1-\widetilde{\Lambda})}{A_5(w)}\left[1-\widetilde{\kappa}\widetilde{\eta}\frac{A_4(w)}{A_1(w)}\widetilde{\Lambda}\right],\\
\Psi_3 &\equiv \frac{\widetilde{\Lambda}}{A_5(w)}\left[1-\widetilde{\kappa}\widetilde{\eta}\widetilde{\Lambda}+\widetilde{\kappa}\widetilde{\eta}(1-\widetilde{\Lambda})\frac{A_4(w)}{A_1(w)}\right],\\
\Psi_4 &\equiv \frac{\widetilde{\eta}(1-\widetilde{\Lambda})}{A_5(w)}\left[\widetilde{\Lambda}-\frac{A_4(w)}{A_1(w)}\right],\\
\Pi_1(x,y) &\equiv \frac{1}{A_5(w)}\left[A_2(x)-xA_2(y)\frac{A_4(w)}{A_1(w)}\right],\\
\Pi_2 &\equiv \frac{1}{A_5(w)}\left[A_3-\widetilde{\kappa}\widetilde{\eta}\frac{A_4(w)}{A_1(w)}\widetilde{\lambda}\right],\\
\Pi_3 &\equiv \frac{1}{A_5(w)}\left[A_3\frac{A_4(w)}{A_1(w)}-\widetilde{\lambda}\right].
\end{aligned}
$$

Proof of Proposition 2: We need to calculate the derivative of V^b with respect to N. We start by rewriting expression (3) as

$$
(27) \qquad V^b=\left(\frac{C}{r+\kappa}+\frac{C(1-l)}{r+\Lambda}\frac{\kappa}{r+\kappa}+\left(E[P^s]+V^b\frac{\eta}{r+\eta}-K\right)\frac{\kappa}{r+\kappa}\frac{\Lambda}{r+\Lambda}-P^b-K\right)\frac{\Lambda}{r+\Lambda},
$$

and solve it for V^b to obtain

$$
(28) \qquad V^b=\frac{\frac{\Lambda}{r+\Lambda}}{1-\frac{\kappa}{r+\kappa}\frac{\eta}{r+\eta}(\frac{\Lambda}{r+\Lambda})^2}\left[\frac{C}{r+\kappa}+\frac{C(1-l)}{r+\Lambda}\frac{\kappa}{r+\kappa}+(P^s-K)\frac{\kappa}{r+\kappa}\frac{\Lambda}{r+\Lambda}-(P^b+K)\right].
$$

Taking into account that $d\frac{\Lambda}{r+\Lambda}/d\Lambda=\frac{r}{(r+\Lambda)^2}$, we obtain the following expression for the derivative in question:

$$
\begin{aligned}
(29) \qquad \frac{dV^b}{d\Lambda} &= \left(\frac{r+\Lambda}{\Lambda}-\frac{\eta}{r+\eta}\frac{\kappa}{r+\kappa}\frac{\Lambda}{r+\Lambda}\right)^{-1}\left[\frac{V^b}{\Lambda}\frac{r}{\Lambda}-\frac{C(1-l)}{(r+\Lambda)^2}\frac{\kappa}{r+\kappa}+\right.\\
(30) \qquad & \left.+\left(P^s+V^b\frac{\eta}{r+\eta}-k\right)\frac{\kappa}{r+\kappa}\frac{r}{(r+\Lambda)^2}+\frac{dP^s}{d\Lambda}\frac{\kappa}{r+\kappa}\frac{\Lambda}{r+\Lambda}-\frac{dP^b}{d\Lambda}\right],
\end{aligned}
$$

which after setting $\frac{dV^b}{d\Lambda}=0$ and some algebra yields expression (16).

ACKNOWLEDGEMENTS

We thank seminar audiences at California Polytechnic University, Carnegie Mellon University, Higher School of Economics-Moscow, IDC Herzliya, Imperial College, Michigan State University, Rice University, University of Houston, University of Illinois-Chicago, University of Oklahoma, University of Wisconsin at Madison, WU Vienna, and the 2015 Toulouse Trading and post-trading conference, 2016 AFA, 2016 SFI Research Day for valuable comments and feedback.

References

Afonso, Gara, Anna Kovner, and Antoinette Schoar, 2013, "Trading partners in the interbank lending market," Federal Reserve Bank of New York Staff Report.

Atkeson, Andrew, Andrea Eisfeld, and Pierre-Olivier Weill, 2015, "Entry and exit in OTC Derivatives Markets," *Econometrica* 83, 2231-2292.

Babus, Ana and Tai-Wei Hu, 2016, "Endogenous Intermediation in Over-the-Counter Markets," *Journal of Financial Economics*, Forthcoming.

Babus, Ana, 2016, "The Formation of Financial Networks," *RAND Journal of Economics* 47, 239-272.

Babus, Ana and Peter Kondor, 2013, "Trading and Information Diffusion in Over-The-Counter Markets," Working Paper, Federal Reserve Bank of Chicago.

Bernhardt, Dan, Vladimir Dvoracek, Eric Hughson, and Ingrid Werner, 2005, "Why do large orders receive discounts on the London Stock Exchange?" *Review of Financial Studies* 18, 1343-1368.

Bessembinder, Hendrik, Stacey Jacobsen, William Maxwell, and Kumar Venkataraman, 2016, "Capital Commitment and Illiquidity in Corporate Bonds," working paper.

Bessembinder, Hendrik, William Maxwell, and Kumar Venkataraman, 2006, "Market transparency, liquidity externalities, and institutional trading costs in corporate bonds," *Journal of Financial Economics* 82, 251-288.

Board, Simon, 2011, "Relational Contracts and the Value of Loyalty," *American Economic Review* 101, 3349-3367.

Chang, Briana and Shengxing Zhang, 2015, "Endogenous market making and network formation," Mimeo.

Choo, Eugene, 2015, "Dynamic marriage matching: An empirical framework,", *Econometrica* 83, 1373-1423.

Colliard, Jean-Edouard and Gabrielle Demange, 2014, "Cash Providers: Asset Dissemination Over Intermediation Chains," Paris School of Economics Working Paper No. 2014-8.

Condorelli, Daniele and Andrea Galeotti, 2012, "Endogenous Trading Networks," Working Paper, University of Essex.

Diamond, Peter, 1971, "A model of price adjustment," *Journal of Economic Theory* 3, 156-168.

DiMaggio, Marco, Amir Kermani, and Zhaogang Song, 2015, "The Value of Trading Relationships in Turbulent Times ," Mimeo.

Duffie, Darrell, Garleanu, Nicolae, and Lasse Pedersen, 2005, "Over the counter markets," *Econometrica* 73, 1815-1847.

Duffie, Darrell, Garleanu, Nicolae, and Lasse Pedersen, 2007, "Valuation in over the counter markets," *Review of Financial Studies* 20, 1865-1900.

Edwards, Amy, Lawrence Harris, and Michael Piwowar, 2005, "Corporate bond markets transparency and transaction costs," *Journal of Finance* 62, 1421-1451.

Andrew EllulRegulatory pressure and fire sales in the corporate bond market," *Journal of Financial Economics* 101, 695-620.

Feldhutter, Peter, 2012, "The same bond at different prices: Identifying search frictions and selling pressures," *Review of Financial Studies* 25, 1155-1206.

Gale, Douglas M., and Shachar Kariv, 2007, "Financial Networks," *American Economic Review* 97, 99-103.

Glode, Vincent and Christian C. Opp, 2014, "Adverse Selection and Intermediation Chains," Working Paper, University of Pennsylvania.

Gofman, Michael, 2011, "A Network-Based Analysis of Over-the-Counter Markets," Working Paper.

Green, Richard C., Burton Hollifield, and Norman Schürhoff, 2007, "Financial Intermediation and the Costs of Trading in an Opaque Market," *Review of Financial Studies* 20(2), 275-314.

Green, Richard C., Dan Li, and Norman Schürhoff, 2010, "Price Discovery in Illiquid Markets: Do Financial Asset Prices Rise Faster Than They Fall?," *The Journal of Finance* 65(5), 1669-1702.

Guerrieri, Veronica, Robert Shimer, and Randall Wright, 2010, "Adverse selection in competitive search equilibrium," *Econometrica* 78, 1823-1862.

Guerrieri, Veronica and Robert Shimer, 2014, "Dynamic adverse selection: A theory of illiquidity, fire sales, and flight to quality," *American Economic Review* 104, 1875-1908. .

Harris, Christopher and John Vickers, 1985, "Perfect Equilibrium in a Model of a Race," *Review of Economic Studies* 52, 193-209.

Harris, Larry, 2015, "Transaction Costs, Trade Throughs, and Riskless Principal Trading in Corporate Bond Markets," Mimeo.

Hitsch, Gunter J., Ali Hortaçsu, and Dan Ariely, 2010, "Matching and Sorting in Online Dating Markets," *American Economic Review* 100, 130-163.

Hollifield, Burton, Artem Neklyudov, and Chester Spatt, 2015, "Bid-Ask Spreads and the Pricing of Securitizations: 144a vs. Registered Securitizations," Working Paper, Carnegie Mellon University.

Hugonnier, Julien, Benjamin Lester, and Pierre-Olivier Weill, 2015, "Heterogeneity in Decentralized Asset Markets," Working Paper, Swiss Finance Institute.

Lagos, Ricardo and Guillaume Rocheteau, 2007, "Search in asset markets: Market structure, liquidity, and welfare," *American Economic Review* 97, 198-202.

Lagos, Ricardo and Guillaume Rocheteau, 2009, "Liquidity in asset markets with search frictions," *Econometrica* 77, 403426.

Lester, Benjamin, Guillaume Rocheteau, and Pierre-Olivier Weill, 2015, "Competing for Order Flow in OTC Markets," *Journal of Money, Credit and Banking* 47, 77-126.

Moen, Espen R., 1997, "Competitive Search Equilibrium," *Journal of Political Economy* 105, 385-411.

Neklyudov, Artem, 2014, "Bid-Ask Spreads and the Over-the-Counter Interdealer Markets: Core and Peripheral Dealers," Working Paper, Swiss Finance Institute.

Neklyudov, Artem and Batchimeg Sambalaibat, 2016, "Endogenous Specialization and Dealer Networks," Working Paper, Swiss Finance Institute.

O'Hara, Maureen, Yihui Wang, and Xing Zhou, 2015, "The Best Execution of Corporate Bonds," Working paper, Cornell University.

Postel-Vinay, Fabien, and Jean-Marc Robin, 2002, "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica* 70, 2295-2350.

Vayanos, Dimitri and Tan Wang, 2007, "Search and endogenous concentration of liquidity in asset markets," *Journal of Economic Theory* 136, 66-104.

Üslü, Semih, 2015, "Pricing and Liquidity in Decentralized Asset Markets," Working Paper, UCLA.

Weill, Pierre-Olivier, 2007, "Leaning against the wind," *Review of Economic Studies* 74, 1329-1354.