Dealer Specialization and Market Segmentation*

Chotibhak Jotikasthira
Southern Methodist University, USA

Christian T. Lundblad
University of North Carolina, Chapel Hill, USA

Jinming Xue
Southern Methodist University, USA

July 22, 2023

*We have benefited from helpful comments from Manuel Adelino, Mike Cooper, Kumar Venkataraman, and Kairong Xiao. We are grateful to the MSRB for providing the data. Jotikasthira (Email: cjotikasthira@smu.edu) is at Southern Methodist University, Lundblad (Email: christian_lundblad@kenan-flagler.unc.edu) is at University of North Carolina, Chapel Hill, and Xue (Email: Jinmingx@smu.edu) is at Southern Methodist University.

Dealer Specialization and Market Segmentation

Abstract

Using machine learning techniques, we uncover an important number of dealers in the U.S. municipal bond market who focus on geographically adjacent states, a characteristic distinct from dealer centrality. These "specialized" dealers enjoy larger market shares in states with greater local ownership and in local bonds with more complex features. We also find that trades intermediated by these specialized dealers have significantly larger markups than those intermediated by national dealers. For the average retail trade, about two-thirds of the differential markup are attributed to rent, with the remaining third to the unique benefits of specialization. Only the latter matters for institution-sized trades. Together, these results suggest that specialized dealers possess some monopoly power yet also provide important differentiated services. Specialized dealers provide immediacy, reward customers with an allocation of new bond offerings, help customers overcome information frictions, and facilitate access to local investor clienteles. The latter two account for the bulk of the specialization benefits. Over time, as transparency improves and local ownership declines, the average market share of specialized dealers decreases along with differential markups.

JEL classification: G12, G14, G24
Key words: Municipal bonds, Over-the-counter market, Segmentation, Clientele effects, Dealer network, Geographical specialization, Liquidity

1. Introduction

Most assets trade in over-the-counter markets, in which dealers operate within networks to provide liquidity and facilitate price discovery. The empirical literature documents that dealer networks for such assets as corporate bonds, municipal bonds, and asset-backed securities are well described by a core-periphery structure. The distinction between dealers who are at a network's core ("central dealers") and those who are not ("peripheral dealers") appears important in determining customer transaction costs and the speed of price discovery (Di Maggio et al. (2017), Hollifield et al. (2017), and Li and Schürhoff (2019)).[1] Several theoretical models, most of which are search-based, have been proposed to confront the empirical facts and endogenize the core-periphery structure. In equilibria with search frictions, central dealers emerge as those with high meeting rates (Üslü (2019)), those serving clients who trade frequently (Sambalaibat (2022)), or those with greater skills and risk appetite (Munyan and Watugala (2018)).[2]

The above models, and many others, assume a single market with one asset.[3] However, many asset markets around the world are segmented, at least partially, in the sense that different traders focus on distinct subsets of assets and potentially trade them in separate venues (though there may be arbitragers who trade across all subsets). International equity, bond, and foreign exchange markets are prime examples.[4] In these markets, dealer networks are likely to be more complex. Take the international bond market, for instance. If enough investors operate in a local market

[1] However, the determinants and the effects of an investor's choice of dealer depend on the nature of the assets being traded as well as general market circumstances. To appreciate the nuances, see, for example, the findings of Hollifield et al. (2017) versus those presented in Li and Schürhoff (2019) and Di Maggio et al. (2017). Hollifield et al. (2017) find that central dealers charge lower price markups in the ABS/MBS market, a phenomenon typically referred to as the centrality discount. Li and Schürhoff (2019) and Di Maggio et al. (2017), on the other hand, show evidence of a centrality premium in the municipal bond and corporate bond markets, respectively.

[2] Chang and Zhang (2021) assume homogeneous traders who form opinions, over multiple rounds of bilateral trading, about others' trading needs. Traders who provide immediacy to others eventually become central.

[3] In a different framework, Malamud and Rostek (2017) examine multiple assets trading in different decentralized exchanges and by different traders, and show that such decentralized exchanges may improve welfare if they induce more trades and allow high risk-tolerant traders to bear more risk.

[4] The international finance literature highlights the distinction between local and foreign investors, and the roles of the latter in transmitting both liquidity shocks (e.g., Jotikasthira et al. (2012)) information (e.g., Wongswan (2006)) local markets. Relatedly, French and Poterba (1991) and many subsequent papers study the geography preferences in investments across countries while Coval and Moskowitz (1999), Coval and Moskowitz (2001), Bernile et al. (2015), and Alok et al. (2020) focus on such preferences within the U.S. A strand of the microstructure literature studies the information advantages and disadvantages of local, relative to foreign, investors (e.g., Domowitz et al. (1997), Bailey and Jagtiani (1994), Brennan and Cao (1997), Brennan et al. (2005), Chan et al. (2005), Dvořák (2005), and Van Nieuwerburgh and Veldkamp (2009)).

and local dealers offer differentiated services (e.g., the opportunity to purchase new issues),[5] then the equilibrium is likely to feature both local and global dealers connected through a network with several local clusters. The same theoretical arguments based on clienteles that explain the coexistence of central and peripheral dealers also justify the coexistence of separate local and global dealers. However, other than the particular clients whom they serve, several questions remain about the prevalence of local dealers and their roles within a dealer network. From these perspectives, we build on Li and Schürhoff (2019) by exploring the empirical implications of market segmentation for market qualities associated with the presence of geographically specialized dealers.

In the U.S., the municipal bond market is generally characterized by segmentation across states and, plausibly, geographical regions. While segmentation is linked to tax barriers (Pirinsky and Wang (2011) and Babina et al. (2021)), it could also be associated with other political or behavioral effects. However, there are no real barriers to capital flows across states; investors and dealers can trade bonds issued by any state and with any other traders of their choices. The municipal bond market thus provides an interesting setup for us to study whether investors focusing on bonds issued in a particular state or region drive the organization of the dealer network. In addition, as tax barriers differ from state to state and over time, the municipal bond market also offers unique cross sectional and time-series variations to pin down the underlying economics by exploring how the market organization evolves.[6]

We start with an observation that dealers' market shares vary significantly across states. Some dealers may be in the top five in Georgia but not even in the top twenty in New York. Others may be in the top twenty in almost every state. A similar picture presents itself on the customer side. Apendix Table A.1 shows that the Nuveen fund family, for example, employs different sets of dealers for their different single state funds (funds that invest, for tax purposes, in municipal bonds issued by only one state).[7] In their N-CEN filing ending 2021-05-31, Bank of America serves as

[5] See a general discussion about the benefits of trading relationships by Hendershott et al. (2020).

[6] For example, Eichengreen et al. (2016) show that improved cable connections between markets have net agglomerative effects on the foreign exchange markets, potentially boosting the foreign exchange turnover in London by up to one-third. In theory, the effects could go in either direction, as fiber-optic connections also reduce communication costs and may encourage transactions at the major financial centers to move to local markets (through a local sales desk).

[7] Item C.17 of SEC Form N-CEN, titled "Principal transactions," provides the identity and total trading amount for each of the top ten trading counterparties, in terms of dollar amount, with which the fund did principal transactions during the filing period. See the SEC filing: `https://www.sec.gov/Archives/edgar/data/1018972/000114554921045496/xslFormN-CEN_X01/primary_doc.xml`.

a top five counterparty for all the six single state funds, whereas Hilltop is a top five counterparty for their Virginia and Maryland funds, but not even in the top ten for their other single state funds. Using detailed transaction-level data with millions of observations, we comprehensively study dealers' business across all states. We then answer such questions as which states tend to share a common set of dealers and why some dealers might be specialized in some states while others spread their business across the nation. Further, we examine how institutions, such as Nuveen, choose dealers for bonds issued by different states and how such choices affect market outcomes, such as transaction costs and immediacy.

To fully characterize the municipal bond dealer landscape amidst market segmentation, we employ machine learning techniques to group states that tend to be intermediated by the same dealers. Specifically, using an *unsupervised topic analysis*, we find that the dealer market is indeed organized to match investor clienteles, with some dealers focusing on certain states ("specialized dealers"), while others spread their business proportionally across the nation ("national dealers"). Specialized dealers tend to focus on a few geographically adjacent states, and their relative shares of trading volume vary both cross-sectionally and over time.

From the topic results, we develop a specialization measure, the *Hellinger* distance, that captures how different, for a particular dealer, the shares of business across states are from the aggregate market shares. By aggregating, we then analyze how the (volume-weighted) average Hellinger distance varies across states. We find that the average Hellinger distance is greater, i.e., specialized dealers tend to have greater market shares, in states where municipal bonds are held more locally and are held less by institutions (more by retail investors). Consistent with the idea that specialized dealers are better at serving local clienteles, we also find that, as the relative importance of state tax privilege decreases and the markets become less segmented over time, more dealers become national, and the national dealers as a group increase their overall market share. This is also consistent with the increasing trend in the institutional holdings of municipal bonds as well as the share of holdings by national relative to single state bond funds (municipal bond funds that invest across the country versus funds that focus municipal bonds issued by a single state).

Specialized dealers serve both retail and institutional investors, although their shares of the market are larger for retail investors. We examine the transaction costs, when trades are intermediated by dealers with varying degrees of specialization, for retail-sized (smaller than \$100,000) and institution-sized ($100,000 or greater) trades. We find that specialized dealers are, on average, associated with higher trading costs, as measured by round-trip markups (see Green et al. (2007), for example). For small retail-sized trades, specialized dealers' markups are higher for bonds issued both by the states in which they are specialized and also those issued by other states, suggesting that specialized dealers capture clients with high search costs and can extract rents. For institution-sized trades, specialized dealers' markups are higher only for bonds issued by the states in which they are specialized. While specialized dealer market power appears to be at play (for retail trades at least), these results suggest that specialized dealers *also* differentiate themselves by offering unique benefits to their clients. Our estimates suggest that for an average retail-sized trade, about two-thirds of the differential markups (between specialized and national dealers) are attributed to rents and the other third to the unique benefits provided by local dealers.

Accordingly, we next explore the potential benefits of trading with dealers specialized in a particular state. We find that such benefits may include immediacy (willingness to take bonds into inventory), allocation of new bond offerings, expertise to help understand complex bonds, and access to a local pool of investors. First, like central dealers (Li and Schürhoff (2019)) but only for bonds issued by states in which they are specialized, local dealers seem more willing to provide immediacy to customers who sell the bonds. This is evidenced by a higher likelihood of holding inventories overnight, a higher likelihood of selling the bonds directly to customers (rather than another dealer), and as a result, a shorter intermediation chain. While compelling, the immediacy benefits alone are not sufficient to explain the specialization premium, as the premium remains highly significant even in situations in which dealers unload the bonds within the same day.

Second, using the holding and transaction data of insurance companies and mutual funds, we find that institutions that trade more with specialized dealers are more likely to receive an allocation of new bond offerings, which the literature has shown tend to be underpriced. Given that (i) states in which specialized dealers have high market shares tend to also be associated with high market shares for local underwriters, and (ii) specialization premia tend to be higher in local dealer and local underwriter dominant states, our result is consistent with dealers using their own or partnered underwriting business to reward their best customers (see Nikolova et al. (2020) for corporate bonds).

Third, more complex bonds (such as refundable bonds or sinkable bonds) and bonds that are more sensitive to issuer-specific information (such as revenue bonds) are associated with higher specialization premia. Market shares of specialized dealers are also higher among these bonds. These suggest that specialized dealers may be more knowledgeable about local bonds and may help investors overcome information frictions. Fourth, in states in which in-state bond ownership is high or state tax privilege is high, specialization premia also tend to be higher. As discussed, market shares of specialized dealers are also higher in these states. These suggest that specialized dealers may also help facilitate access to fragmented local investors, thereby commanding a higher markup.

Finally, we assess the relative importance of the above benefits of specialization by performing a Blinder-Oaxaca decomposition. We first sort transactions into quintiles by specialization of the dealers who intermediate them, and measure the total compensation for specialization benefits by comparing the average markups of the transactions in the top and bottom groups. Focusing on institutional-sized trades, we find that the specialization benefits are worth about 10 basis points in markups. We then decompose this markup differential into the endowment, coefficient, and interaction effects associated with the following four benefits provided by specialized dealers: immediacy, allocation of new bond offerings, help with complex or opaque bonds, and access to local investor pools. Our decomposition suggests that about two-thirds of the specialization premium are for access to local investor pools and the remaining third for help with complex or opaque bonds. Both benefits work through the coefficient rather than the endowment effects: for example, the markups are higher in the top decile of specialization not because the top decile contains more bonds issued by states with high local ownership, but rather because in the top decile of specialization, dealer markups are more strongly associated with local ownership.

We contribute to the microstructure literature that examines the contours of dealer networks in relatively illiquid over-the-counter markets. Much of the literature focuses on a core-periphery network structure (for example, Di Maggio et al. (2017), Hollifield et al. (2017), Li and Schürhoff (2019), Friewald and Nagler (2019), Eisfeldt et al. (2023) and Chernenko and Doan (2020)), where dealer centrality is primary. In the face of market segmentation, we show that dealer networks are more nuanced. Specifically, we detect an important role for specialized dealers, quite distinct from central dealers, who disproportionately focus on specific market segments.

By examining the specific trades intermediated by specialized dealers, we augment the understanding of the drivers of trading costs. Specialized dealers are associated with elevated markups that suggest some monopoly power. However, we show that specialized dealers also earn part of their markups by providing important differentiated services to their retail and institutional clients. Specifically, they provide immediacy, reward customers with new offerings, help customers overcome information frictions, and facilitate access to local customers. We also show that the latter two are most important in explaining our documented specialization premia, consistent with market segmentation and opacity explaining the presence of specialized or local dealers.

The municipal bond market, where search frictions are high and trading is infrequent (Harris and Piwowar (2006), Green et al. (2007), Schultz (2012), Schultz (2013), Chalmers et al. (2021), and Griffin et al. (2022)), serves as a natural empirical setting for our examination. Accordingly, we highlight the role of geographically-focused specialized dealers as an important determinant of trading costs in of the municipal bond market. Further, as transparency improves and the market becomes more geographically integrated over time, the average market share of specialized dealers decreases along with differential markups.

Last, given that we highlight the advantages of geographically specialized dealers, we connect to a long literature in international finance on the relative informational disadvantages of foreign versus local investors. While this is hardly a settled issue (see Lundblad et al. (2023), for example), there are a number of papers that show that foreign investors are at an informational disadvantage relative to locals in an emerging market setting (i.e., Kang and Stulz (1997), Choe et al. (2005), Dvořák (2005), Bena et al. (2017) and Ferreira et al. (2017)). We draw a parallel to the municipal finance dealer network, where specialized, largely local agents, help overcome information frictions in local markets, and in doing so, play an important role in facilitating price discovery across geographically segmented markets.

2. Data and Specialization Measures

We obtain data on municipal bond transactions from the MSRB's Academic Historical Transaction Database, data on municipal bond characteristics from Mergent Municipal Bond Securities Database, data on municipal bond issuance from SDC Platinum, data on mutual funds' holdings from Morningstar, and data on insurance companies' transactions and holdings from NAIC Schedule D. The state-level economic and market data are from standard sources. Given the limitations of the MSRB data, our sample period is from January 2006 to June 2017. Below, we describe how we construct several key variables and provide summary statistics. Detailed variable definitions are in Appendix A.

2.1 Geographical Organization of Municipal Bond Market

In each transaction reported by the MSRB, we observe the bond identifier (CUSIP), the direction of the trade (dealer purchasing from customer, dealer selling to customer, or inter-dealer), trade size, trade date/time, trade price (or, in some cases, also yield), anonymized identifiers of dealers involved in the trade, and various other bonds/trade/reporting characteristics.[8] As the anecdotes (Nuveen mentioned above) from the SEC's Form N-CEN data suggest, state municipal bond funds tend to employ distinct dealers, even if they are run by the same family. That is, a particular dealer may be good at trading bonds issued by some states but not others. To see which states tend to be accommodated by the same dealers, we perform an unsupervised topic analysis. Specifically, we apply the Latent Dirichlet Allocation (LDA), a well-developed approach in machine learning (e.g., Blei et al. (2003), Blei and Jordan (2006)), to the dealer-state dimension of municipal bond transactions. LDA is a generative probability model for collections of discrete data. One common application of LDA is topic modeling, which provides methods for automatically organizing, understanding, and summarizing large electronic archives. Provided with a large set of documents, or "dealers" in our case, LDA helps group the words, or "states", based on their co-occurrence among the documents, thus generating topics to which we refer as "regions" with distinct contents.

For our application, the outcomes of the LDA algorithm are summarized by two sets of probabilities:

1. $p_k(state = s)$: the probability that a bond in region k is issued by state s, and
2. $p_d(region = k)$: the probability that dealer d trade bonds in region k,

which together characterize the dealer-state dimension of each municipal bond transaction in the data. Our estimation approach follows that of Hoffman et al. (2010). The total number of regions ($k = 0, 1, \cdots, K$) in the model is a parameter pre-specified by researchers. We have tried various choices of K from 1 to 5 and settled on $K = 3$ (i.e., four regions) since it explains the data well

[8] See https://www.msrb.org/Academic-Historical-Transaction-Data-Product-Format for the full list of variables.

and generates the most stable classification. Further details on our topic modeling approach are in Appendix B.

Appendix Figure A.1 presents the estimated probabilities $p_k(state = s)$ for $k = 0, 1, \cdots, 3$ as four separate heat maps. In each region k, the sum of the probabilities across states must equal 1. The highest probability state in Region 0 is Texas (15.7%), while those in Regions 1, 2, and 3 are California (82%), New York (29%), and Illinois (12%), respectively. These states are popular, even outside of the region in which they are dominant, because they disproportionately represent large shares of trading volume. Small states, on the other hand, show up infrequently in all four regions. To neutralize the influence of state size so that we can see more clearly the main region with which each state is associated, we normalize $p_k(state = s)$ by the unconditional probability that a randomly drawn transaction is on a bond issue by state s, $p(state = s)$. The normalized probability measures the over- or under-representation of state s in region k. If the state shows up in the region at the same probability as it does nationally, then the normalized probability will be 1. If the value of the normalized probability is greater (smaller) than 1, then the state shows up disproportionately more (less) frequently than normal.

[Insert Figure 1]

Figure 1 Panel A plots the normalized probabilities of the states in each region as a heat map. The clusters of dark colors show that the four regions resemble the Southeast, California, the Northeast, and the Midwest/West. In addition, if we assign each state to only one region in which it has the highest representation, we have a clear definition of the four regions as in Figure 1 Panel B. Interestingly, the four regions consist of geographically contiguous states despite the fact that our model is completely blind to the states' geographical locations. The revealed patterns of geographical business clustering reflect the trading activities of municipal bond broker-dealers. Bonds issued by contiguous states are likely to be intermediated by the same dealers, consistent with anecdotes about mutual funds' dealer choices from the SEC's N-CEN data. For example, Hilltop is a top five counterparty for Nuveen's Virginia and Maryland funds and number eight for its New Mexico fund. However, Hilltop is not in the top ten for Nuveen's Colorado, Arizona, and Pennsylvania funds.

2.2 Dealer's Specialization and Other Characteristics

Dealers are masked in the academic MSRB, so we can only measure their characteristics with respect to the transactions that they intermediate. While there are over 600 unique masked dealers in the data at most points in time, a few dealers capture most of the volume. On average, the top ten and top fifty dealers by volume capture 65% and 89% of the volume, respectively. Dealer size as measured by trading volume is highly positively correlated with dealer centrality, as measured in Li and Schürhoff (2019). Table I shows that the cross sectional correlation between the two variables is about 0.83, and is highly stable from year to year.

The set of probabilities $p_d(region = k)$, for $k = 0, 1, \cdots, 3$ characterize the distribution of each dealer's volume across the four regions. On average, the distribution is fairly even with about 28% in Region 0 (the Southeast), 16% in Region 1 (California), 33% in Region 2 (the Northeast), and 23% in Region 3 (the Midwest/West). However, the degree of geographical concentration varies widely across dealers, with larger dealers distributing business more evenly across the U.S., while small dealers often focus on certain regions. For instance, in 2010, the largest masked dealer (with annual volume of \$388 billion) distributes 22%, 17%, 52%, and 8% of its volume across Regions 0 to 3, respectively. By contrast, another dealer with an annual volume of \$7 billion specializes in Region 3 with 86% of its trading volume coming from that region alone.

To measure the general degree of geographical specialization, we calculate the $Hellinger$ distance (see, for example, Chapter 3 of Pollard (2002)) between dealer d's distribution of volume and the national average in year y:

$$Hellinger_{d,y} = \sqrt{1 - \sum_{k=0}^{3} \sqrt{p_{d,y}(region = k) p_{\bar{d},y}(region = k)}}, \quad (1)$$

where $\bar{d}$ denotes the average dealer in the sample. From equation (1), we can see that the $Hellinger$ measure reaches a minimum value of 0 when a dealer's regional distribution is exactly the same as the market average. When the dealer instead dedicates 100% of the volume to the smallest region, the maximum value is slightly below 1. In general, a high value of the $Hellinger$ measure indicates that the dealer's volume distribution across the four regions is very different from the national average, suggesting a high degree of regional specialization. For ease of exposition and economic inference, we will refer to $Hellinger$ as dealer "specialization", where dealers with low (high) $Hellinger$ can be thought of as national (specialized) dealers.

[Insert Table I]

Table I Panel A reports cross-sectional summary statistics for the $Hellinger$ measure over time. A few observations are worth noting. First, there is significant variation in the overall degree of specialization across dealers. For example, in 2017, the 25th and 75th percentiles of the $Hellinger$ measure are about 0.177 and 0.508, respectively. Given the natural minimum and maximum values of the $Hellinger$ measure, the statistics suggest that at least a quarter of the 555 dealers in 2017 appear to trade municipal bonds almost proportionally to the aggregate shares of the four regions, while another quarter or more largely focus on one region.

Second, dealer specialization is negatively correlated with both dealer size and centrality (with the latter measured as in Li and Schürhoff (2019)) Over time, the correlation of the $Hellinger$ measure and logged dealer volume varies between -0.418 and -0.271 while the correlation of the $Hellinger$ measure and dealer centrality lies between -0.383 and -0.221. Intuitively, the more the dealer trades, the more states he is likely to cover and the more he begins to look like a representative national dealer. In addition, as the dealer covers more states, his network expands and hence his centrality increases. Despite these natural relationships, there remains a significant degree of variation in specialization within large and small dealers as well as across central and peripheral dealers.

[Insert Figure 2]

Figure 2 illustrates dealer networks in three sample states: CA, MD, and NY. The size of the dot reflects dealer centrality while the color of the dot captures $Hellinger$, with dark blue being the lowest and light yellow being the highest. The graphs visually shows that in each of the three states, national and (nationally) central dealers tend to also be at the core of the local network. On the other hand, specialized dealers (lighter colored dots) are both at the core and in the periphery. This suggests that geographical specialization is a distinct characteristic, consistent with the statistics in Table I. The correlations of $Hellinger$ with both logged dealer volume and dealer centrality, while robustly negative, are far smaller in magnitude than the correlation between the latter two.

Finally, Table I Panel A also shows that as dealers become less specialized (or, more national) over time, their overall specialization becomes more strongly correlated with both their volume and centrality. The average $Hellinger$ decreases by almost 20% during our sample period, from 0.431 in 2006 to 0.356 in 2017. Over the same period, the correlation between $Hellinger$ and logged volume increases in magnitude by over 50% (from 0.271 to 0.418), as does the correlation between $Hellinger$ and dealer centrality. That is, in recent years, the distinction between large, central, and national dealers and small, peripheral, and specialized dealers becomes much clearer. To examine the reasons, we separate dealers by volume into three groups: top ten, next forty (11-50), and the rest, and investigate their volume share and specialization trends. Figure 3 Panel A shows that the top ten dealers, who account for about 70% of market volume and whose average $Hellinger$ is the lowest, lose market share over time. Their average $Hellinger$ does not decrease either. Thus, the decreasing average specialization is not driven by the largest dealers amassing greater market shares. We next examine the middle group of dealers, those ranked 11 to 50. Figure 3 Panel B shows that they gain the most market share yet at the same time become less specialized. As shown in Figure 3 Panel C, the remaining small dealers, those ranked above 50, have also gained market share and exhibit a gradual reduction in $Hellinger$. Overall, the decreasing trend in dealers' specialization appears to be due to some consolidation of mid-sized and small dealers into larger ones, as well as the transformation of their business to become more national.[9]

[Insert Figure 3]

While $Hellinger$ reflects the degree of dealer specialization, it does not pinpoint in which state or region a particular dealer is specialized. This level of detail is important to understand the costs and benefits of specialization. The same dealer may not be able to offer the same expertise nor any differentiated services for bonds issued outside of their main states of specialization. To measure the specialization of a dealer in a particular state, we start with the LDA-implied probability that a particular dealer d trades a bond issued by state s (which can occur to varying degrees through all four regions): $p_d(state = s) = \sum_{k=0}^{3} p_d(region = k)p_k(state = s)$. We then normalize the size effect (large states have naturally higher probabilities) by scaling $p_d(state = s)$ by the probability

[9]This is also consistent with the findings that the municipal bond market has become more nationally integrated over time (for example, see Figure IA.1 of Babina et al. (2021)). Since the global financial crisis, the share of national municipal bond funds, as a group, has increased significantly relative to that of single state municipal bond funds.

that the average dealer $\bar{d}$ trades a bond in that state, $p_{\bar{d}}(state = s)$. Together, our dealer state concentration, or DSC, measure for each dealer d in state s in year y can be written as

$$DSC_{d,s,y} = \frac{p_{d,y}(state = s)}{p_{\bar{d},y}(state = s)} = \frac{\sum_{k=0}^{3} p_{d,y}(region = k) p_{k,y}(state = s)}{\sum_{k=0}^{3} p_{\bar{d},y}(region = k) p_{k,y}(state = s)}. \tag{2}$$

DSC is a dealer-state pair measure, for which the value is high when the dealer is specialized in the region that exhibits a high exposure to the state. Because specialized dealers, or dealers with a high $Hellinger$ measure, by definition have disproportionately high exposures to particular regions and low exposures to others, their DSC's are particularly high in some states and low in others. In this sense, to use the international finance analogy, we think of specialized dealers as being "local" in some states and "foreign" in others. By contrast, national dealers, or dealers with low $Hellinger$, spread their business roughly evenly across regions, and as a result, should have DSC close to 1 in most states. That is, national dealers are more or less the average dealer we use as the benchmark in calculating DSC in equation (2). Table I Panel B confirms these observations. We separate dealers by $Hellinger$ at the median. Both the high and low $Hellinger$ groups have the mean DSC very close to 1 every year, but the interquartile range is much larger for the high group than the low group.[10] In 2017, for example, the 25th percentile of DSC for the high group, or more specialized dealers, is only 0.23, while the 75th and 95th percentiles are as high as 1.57 and 3.66, respectively. That is, an above-median $Hellinger$ dealer may trade as little as a fifth of the average dealer in some states and as much as 2-4 times the average dealer in a few select other states.

Table I Panel B also shows that DSC exhibits the same time trend as the $Hellinger$ measure. While the mean remains close to 1 over time, the interquartile range becomes narrower in more recent years, gradually decreasing from 0.66 to 0.41 for the low $Hellinger$ group and from 1.44 to 1.34 for the high $Hellinger$ group. The especially pronounced decrease in DSC dispersion among the low $Hellinger$ group is consistent with what we observe for the $Hellinger$ measure itself; dealers ranked between 11 to 50, or dealers with low but not the lowest $Hellinger$ measure, have become increasingly national over time.

[10]By construction, if all dealers distributed business in exactly the same way, DSC would be 1 for all dealer-state pairs.

2.3 Transaction and Bond Characteristics

We study transaction costs associated with specialized versus national dealers. Using the academic version of the MSRB data, we follow Li and Schürhoff (2019) and build chains of transactions that flow from a customer selling a bond, through the dealer network, to another customer buying the same bond. Using the same notations as Li and Schürhoff (2019), we denote a customer as C and a dealer as D, and take each well-defined chain of related transactions, written as $C(N)DC$ (where $N = 1, 2, \cdots$) as one observation. In total, we have slightly over 7.3 million chains. For each chain, we calculate the round-trip markup as the customer purchase price less the customer sale price divided by the customer sale price, and use this as our measure of transaction costs. Consistent with the findings in the literature, Table II Panel A shows that round-trip markups for municipal bonds vary widely, with a 25th percentile of just 42 basis points and a 75th percentile of 2.36%. The mean is 1.53%, which is relatively large compared to other types of bonds possibly due to the prevalence of retail investors. For reference, the 75th percentile transaction size, as measured by par value (often used to identify retail trading) is only $50,000. Institutional-sized transactions, defined as those with par value of $100,000 or larger, represent less than 20% of all observations.

[Insert Table II]

While the transaction chain length can be very long, most are not. The majority of transaction chains involve just one dealer (i.e., CDC). Even at the 75th percentile, the chain length is just 3 (i.e., $CDDDC$). In most of our analyses, we associate the transaction chain with the first dealer in the chain. Across the transaction chains in our sample, dealer size is heavily skewed; the average trading volume per dealer per year is $62 billion while the median is only $12 billion. Dealer centrality is also very high; even the 25th percentile is 0.94 (i.e., the 94th percentile of raw centrality). This suggests that a small number of large and central dealers intermediate most of the transactions. The first dealer, which in most cases is the only one in the chain, captures about 1% in markup on average, with the 25th and 75th percentiles being 23 basis points and 1.58%, respectively. This suggests that the large variation in round-trip markup is not due to the variation in transaction chain length. The markup captured by the first dealer alone varies almost as much. Li and Schürhoff (2019) attribute part of the variation in markup to dealer centrality. Given the important role for municipal bond market segmentation, we seek to understand the complementary role of dealer specialization.

On average, the traded bonds have about 10 years of remaining maturity, and have aged for 5-6 years since issuance. The issue size is highly skewed with the median of just $6.35 million and the mean of over $28 million. About 63% of the transactions involve revenue bonds, 70% of the traded bonds are refundable, 48% are sinkable, and 6% are bank-qualified. In terms of credit quality, 47% are insured, 0.4% are below-investment grade, and 20% are not rated. The average credit rating is between AA and AA+.

2.4 State Characteristics

States differ significantly in size, as measured for example by GDP, municipal bonds outstanding, fiscal and other policies, and many other characteristics. Some of these aspects have been shown to be important in explaining bond parameter choices, e.g., whether or not a given bond is insured, its investor clientele, and its pricing. For example, Gore et al. (2004) find that when states require more financial disclosure by municipal bond issuers, the issuers use less bond insurance. Gao et al. (2020) find that state policies on distressed municipalities affect municipal bond yields. Specifically, yields are lower in states that have proactive policies to help distressed municipalities. Pirinsky and Wang (2011) and Babina et al. (2021) find that tax privilege for in-state holders of local municipal bonds increases local ownership of the bonds, making the bonds' yields more sensitive to changes in supply and local political risk. Table II Panel B presents summary statistics for state characteristics. At this point, we simply note that our state variables are standard and have typical distributional properties.[11]

3. Where Do We See More Specialized Dealers?

By construction, dealers with high $Hellinger$ scores conduct disproportionately more business in some regions than others. Do they cater to local investors or do they serve equally both local and non-local customers? Do they provide unique benefits, and if so, are these benefits more valuable

[11] In most of our analyses, we absorb time-invariant state characteristics using state fixed effects, and where possible, time-varying state characteristics using state *x* year fixed effects.

for trading some bonds than others? In this section, we address these questions by examining the relative shares of specialized dealers in different states and bond types.

3.1 Customer Characteristics

Reflecting some degree of segmentation across states, it is well known that municipal bonds are disproportionately locally held. We conjecture that geographically specialized dealers match with geographically focused investors in the same partially segmented states. Specifically, we test whether specialized dealers, as a group, enjoy relatively higher market shares in states in which local ownership is more pronounced by running the following regression:

$$\overline{Hellinger}_{s,y} = \beta_0 LO_{s,y} + \sum_{i=1}^{N} \beta_i X^i_{s,y} + \alpha_y + \epsilon_{s,y}, \tag{3}$$

where $\overline{Hellinger}_{s,y}$ is the volume-weighted average $Hellinger$ measure of all dealers trading the bonds of state s in year y and $LO_{s,y}$ is a candidate measure of local ownership. At the state level, $\overline{Hellinger}_{s,y}$ captures the relative market shares of specialized dealers because it increases in the shares of volume intermediated by dealers with high $Hellinger$. Following Pirinsky and Wang (2011) and Babina et al. (2021), we measure local ownership using either state fund holdings or state tax privilege.

X is a vector representing other time-varying state-level variables, which include the volume-weighted average dealer centrality, institutional holdings, the debt-to-income ratio, a proactive policy dummy (from Gao et al. (2020)), and other standard controls. As discussed, dealer centrality is negatively correlated with dealer $Hellinger$, and therefore, to isolate the specialization component, we employ centrality as a control variable. Because specialized dealers also tend to be smaller, they may disproportionately serve retail investors. To the extent that local investors are more likely to be retail investors, we include a measure of institutional holdings to help separate the effects of specialized dealers catering to retail versus local investor clienteles. The debt-to-income ratio captures credit risk, and the proactive policy dummy captures whether states are proactive in helping distressed municipalities. Other standard controls include proxies for trading volume, the size of the market, the size of the economy, and economic strength. We also include year fixed effects, α_y, and cluster standard errors by year. The variation in which we are interested is largely cross-sectional in nature.

[Insert Table III]

Table III Panel A reports the results. We find that the state-year average $Hellinger$ measure is higher in states with higher single state fund ownership (relative to national funds) in column 1 and higher tax privilege in column 2. The coefficient estimates for both state fund holdings and state tax privilege are positive and statistically significant at the 1% level. In economic terms, a one standard deviation increase in state fund holdings is associated with an increase of 4.44×10^{-3} in the state-year average $Hellinger$ measure (11% of its standard deviation). Of the many coefficients for the control variables, two are worth highlighting. First, the coefficient estimates for state-year average dealer centrality are negative and statistically significant, confirming our earlier findings that specialized dealers tend to be less central in the national network. Second, the coefficient estimates for institutional holdings are negative and statistically significant at the 1% level, suggesting that states whose bonds are disproportionately held by institutions tend to be served less by specialized dealers. A one-standard deviation increase in institutional holdings is associated with a decrease of 6.23×10^{-3} in the state-year average $Hellinger$ measure (16% of its standard deviation). Together, the results suggest that specialized dealers tend to cater to and may potentially provide access to local and retail clients.

3.2 Bond Characteristics

Some municipal bonds are more difficult to value and more expensive to trade than others (see, for example, Harris and Piwowar (2006) in the case of municipal bonds, and Edwards et al. (2007), in the case of corporate bonds). To the extent that specialized dealers may have unique expertise in understanding and locating customers for harder-to-value bonds, they should enjoy greater market shares in these bonds. The literature suggests that hard-to-value bonds are bonds with complex features, such as being callable, sinkable, or paying interest at irregular intervals.

Revenue bonds, whose payments are tied to revenues of specific projects, require more information to assess than general obligations bonds. Similarly, bonds with higher credit risk also have prices that are more sensitive to information. If specialized dealers are better informed about local bond issuers and can help investors overcome information frictions, they should also have greater market shares in revenue bonds and bonds characterized by relatively elevated credit risk.

We study the relative market shares of specialized dealers across bonds with different features discussed above using the following regression:

$$\overline{Hellinger}_{b,y} = \sum_{i=1}^{N} \beta_i X^i_{b,y} + \alpha_{s(b),y} + \epsilon_{b,y}, \tag{4}$$

where $X_{b,y}$ is a vector of bond b's time-invariant as well as time-varying characteristics in year y. To ensure that the relationships we uncover are not driven by omitted factors driving both the $Hellinger$ measure and bond characteristics, we also include state x year fixed effects, $\alpha_{s(b),y}$. We double-cluster standard errors by bond and year.

Table III Panel B reports the results. The coefficient estimates are largely consistent with our conjectures. In column (1), we find that specialized dealers are over-represented in the trading of sinkable bonds, bank-qualified bonds, revenue bonds, unrated bonds, and bonds with lower credit ratings in general. In contrast, specialized dealers have lower market shares in the trading of insured bonds. The coefficients on the callable dummy and the high-yield dummy are not significant. Most municipal bonds are refundable so that feature may not be viewed as particularly complex. High-yield dummy is a function of numerical credit rating, and the latter, included in a discrete linear form, may have absorbed much of the credit quality effect. In column (2), we include the count of complexity features such as callable, sinkable, having extraordinary redemption or call provisions, having non-standard interest payment frequencies, etc., following Harris and Piwowar (2006). The coefficient on complexity count is positive and statistically significant, confirming that specialized dealers have greater representation in the trading of more complex bonds.

4. Are Specialized Dealers More or Less Expensive with which to Trade?

If markets are competitive and dealer services are all the same, then dealers with the lowest marginal costs should win the most trades. This implies that all dealers that coexist, regardless of whether they are national or specialized, should earn about the same on average. On the other hand, if dealers' services are differentiated or different dealers capture separate clienteles, then it is possible that transaction costs systematically vary across transactions intermediated by national dealers and those intermediated by specialized dealers. Dealers who provide greater benefits, such as allocations of new bond offerings, and those that serve clienteles facing greater frictions, such as local retail investors, are likely to be associated with elevated transaction costs.

[Insert Figure 4 and Table IV]

To examine whether specialized dealers are more or less expensive with which to trade, we start by sorting markups in each month of the sample into deciles by the dealer-level $Hellinger$ measure. Figure 4 plots the average markup in each decile, which shows a clear increasing trend from deciles 1 (low) to 10 (high). The average markups are 1.48% for the lowest $Hellinger$ decile and 1.73% for the highest $Hellinger$ decile. The difference of 0.25% is economically meaningful, given that the average markup for the full sample is 1.53%. In Table IV, we isolate the effects of specialization from those of centrality by double sorting. We first sort markups in each month into terciles by dealer centrality, and then in each centrality tercile, we further sort markups into terciles by $Hellinger$. The results show that specialized dealers are more expensive with which to trade, regardless of their degree of centrality. The differences between the average markup in the top and bottom $Hellinger$ terciles are 0.23%, 0.19%, and 0.42%, respectively, for the low, medium, and high centrality groups. These differences are all statistically significant at the 1% level. Taken together, these sorted results suggest that there is a premium associated with trading with specialized dealers.

It is important to keep in mind that markups are well known to be correlated with various bond and transaction characteristics. For example, markups tend to be higher for smaller transactions and for more complex bonds. To the extent that specialized dealers disproportionately serve retail investors and intermediate relatively complex bonds, our evidence may be driven by customer size or bond complexity. To isolate the effects of specialization, we add dealer $Hellinger$ to the regression specification of Li and Schürhoff (2019):

$$Markup_c = \beta_0 Hellinger^N_{d(c),y(c)} + \sum_{i=1}^{N} \beta_i X^i_{b(c),d(c),t(c)} + \alpha_{s(c),m(c)} + \epsilon_c, \tag{5}$$

where $Markup_c$ is the round-trip markup (or, in some cases, part of the round-trip markup) associated with transaction chain c, which is on bond b at time t (in month m and year y) and intermediated by dealer d. For ease of interpretation, we normalize $Hellinger$ to obtain $Hellinger^N$ by subtracting its monthly mean and dividing by its monthly standard deviation. To the extent that the chain involves more than one dealer, we associate the chain with the first dealer who buys the bond from the customer to start the chain. The coefficient of interest is β_0, which captures the markup differential for trades intermediated by more or less specialized dealers, holding constant other transaction and bond characteristics.

Accordingly, X is a vector of transaction and bond controls, which include all variables used in the original specification of Li and Schürhoff (2019). However, we use dealer annual trading volume, as opposed to dealer asset size as in Li and Schürhoff (2019), because dealer identities are masked in our version of MSRB data. Dealer centrality is not our focus, but to highlight the distinct effects of specialization, we also include it and its interaction with chain length (i.e., number of dealers in the chain) in the regression. We include state x month fixed effects, $\alpha_{s,m}$, to absorb the general variation in transaction costs across states and over time, and double-cluster the standard errors by dealer and month.

[Insert Table V]

Table V reports the results. In column (1), we simply replicate the results of Li and Schürhoff (2019) for all transaction chains, confirming their evidence of a centrality premium that decreases in transaction chain length. In column (2), to capture the average incremental markup associated with dealer specialization, we add dealer $Hellinger^N$. The coefficient on the $Hellinger^N$ measure is both positive and statistically significant at the 1% level. In economic terms, a standard deviation increase in $Hellinger$ is associated with an increase in full markups of about 0.13 percentage points. Assuming a normal distribution, this estimate would imply a slightly larger difference between the markups in deciles 1 and 10 than that presented in Figure 4 (about 0.43% vs. 0.25%). In column (4), we focus only on CDC chains, which Li and Schürhoff (2019) argue are more cleanly associated with a particular dealer. The coefficient on the $Hellinger$ measure is practically unchanged and remains highly statistically significant. In column (6), we include all chains but measure only the markup associated with the first dealer, whose $Hellinger$ measure we employ in the regression. The result is again practically unchanged, confirming that the differential markup is indeed attributed to the dealer assigned to the chain. Taken together, our results provide robust evidence of a "specialization premium".

5. Does the Specialization Premium Reflect Rents or Unique Benefits?

A natural next question is then what explains the "specialization premium"? Does it reflect any unique benefits that specialized dealers provide or simply the rents that they extract due to the relatively high search-cost customer clientele that they typically serve? To differentiate the two explanations, we make a reasonable assumption that specialized dealers can only provide transaction-related benefits (e.g., immediacy or an allocation of new bond issues) in the specific states in which they are specialized. When specialized dealers intermediate bonds issued by states outside their specialization, we assume that they do not offer any special benefits above and beyond what can be offered by national dealers. In that setting, any incremental markups they charge must be due to rent.

We gauge a dealer's state-level specialization using our dealer-state concentration (DSC) measure. As described above in Section 2, DSC captures the share of a dealer's business in a particular state relative to a national benchmark. Specialized dealers, or dealers with high $Hellinger$, have high DSC in specific states in which they are specialized and low DSC in states in which they are not. Thus, high markups associated with dealers with a high $Hellinger$ measure in states in which they have low DSC can be viewed as monopoly rent, while high markups associated with dealers with a high $Hellinger$ measure in states in which they have high DSC may instead reflect compensation for the unique benefits that they provide. We isolate the incremental effects of DSC on markups by adding a normalized version of it to equation (5):

$$Markup_c = \beta_0 DSC^N_{d(c),y(c)} + \beta_1 Hellinger^N_{d(c),y(c)} + \sum_{i=2}^{N} \beta_i X^i_{b(c),d(c),t(c)} + \alpha_{s(c),m(c)} + \epsilon_c. \quad (6)$$

Specifically, we normalize DSC by subtracting its state-month-level mean and dividing by its state-month-level standard deviation (i.e., focusing on the cross-sectional variation). The reason we need to normalize DSC this way is that in addition to the ease of interpretation, the benchmark share of a small state (e.g., Idaho) is so small and so its DSC, as measured by (2), has an extremely high standard deviation. This renders the effects of DSC per unit of the raw measure much lower in small states than in others. The equalization scales the unit to one standard deviation and thus helps to level the variation across states.

Column (3) of Table V reports the full sample estimates. The coefficient on DSC^N is positive and statistically significant at the 5% level. In column (5), we repeat the analysis for the cleaner CDC-only sub-sample. The coefficient estimate for DSC^N has the same sign but is now significant at the 1% level and, compared to the full sample estimate, twice as large in magnitude. The full round-trip markup may not move one-to-one with the compensation captured by the first dealer and hence may not accurately reflect the effects of DSC. Hence, we employ the full sample but use, as the dependent variable, the markup attributed to the first dealer whose DSC is used in the regression in column (7). The coefficient of DSC is again positive and statistically significant at the 1% level. Using the estimate in column (7), a standard deviation increase in DSC is associated with a markup increase of about 0.06 percentage points. This may appear small, but the reality is that it is not. As discussed, most specialized dealers have extremely high DSC in a few states and close to zero DSC in others. The distance between states within and out of their specialization, especially among dealers with high $Hellinger$ measures, is several times the standard deviation of DSC in the full sample. For example, assuming a normal distribution, the estimate would imply a difference in markups of about 0.20% between dealers in DSC deciles 1 and 10.

With the exception of the regression in column (5), the coefficients on the $Hellinger$ measure, after DSC is added to the regression, remain positive and statistically significant, albeit somewhat smaller in magnitude. Taken together, the results for $Hellinger$ and DSC show that specialized dealers charge higher markups both in the specific states in which they are specialized (high DSC and high $Hellinger$) and those in which they are not (low DSC but high $Hellinger$). This suggests that specialized dealers both provide unique benefits, plausibly made possible by their specialization, and earn rents, plausibly for serving some clienteles with few outside options.

[Insert Table VI and Figure 5]

To assess the relative significance of the rent versus benefit explanations and to confirm our associated economic interpretation, we next repeat the regression (6) separately for sub-samples of retail-sized ($< \$100K$) and institutional-sized ($\geq \$100K$) trades. We report the results in Table

VI. In columns (1) to (3), we focus on the retail-sized trades, which account for about 80% of the observations. We find that the coefficients on both the $Hellinger$ and DSC measures are positive in all specifications and statistically significant in most. Figure 5 Panels A and B show that the results are robust when we perform the regressions year by year (with the exception of 2008 where the coefficient on DSC is negative but insignificant). To measure the economic effects of $Hellinger$ and DSC, we perform the Blinder-Oaxaca decomposition for the specialization premium. Specifically, we (sequentially) double-sort transactions into three-by-three groups first by $Hellinger$ and then by DSC. We then take the transactions in the lowest $Hellinger$ tercile as those intermediated by a representative national dealer, and the transactions in the highest $Hellinger$ and highest DSC group as those intermediated by a representative specialized dealer in its specialized state. The difference in full markups between the two groups captures the specialization premium. In the final step, we apply the Blinder-Oaxaca decomposition to split the specialization premium into two components: one explained by $Hellinger$ and the other by DSC.

[Insert Figure 6]

Figure 6 Panel A presents the decomposition results. After stripping out the effects of control variables using the model in equation (6) but omitting $Hellinger$ and DSC, we find that the difference in full markups between transactions intermediated by representative national and specialized dealers, i.e., the specialization premium, is on average 27.5 basis points. $Hellinger$ explain 19 basis points, and DSC explains 8.5 basis points. The former reflects the markup differential between specialized and national dealers holding DSC at the sample average. Since the average DSC is about the same as DSC of a typical dealer in a typical neutral state in which it has no special expertise, we interpret the 19-bp $Hellinger$ effects as monopoly rents, plausibly due to specialized dealers capturing customers with low bargaining power. The 8.5-bp DSC effect, on the other hand, reflects the markup differential for a typical dealer in a state in which it has special expertise versus another state in which it does not. To the extent that dealers can render unique benefits in their states of specialization, the 8.5-bp DSC effect should capture these benefits. Taken together, for a retail-sized trade, customers pay about 27.5 basis points more to trade with a dealer who is specialized in that bond than a national dealer, and of that difference, 69% is due to rents and 31% due to unique benefits of specialization.

The last three columns of Table VI report the estimates of regression (6) for the sub-sample of institutional-sized trades, which account for about 20% of the observations. A noticeable change from the first three columns is that the coefficients on $Hellinger^N$ are negligible and no longer statistically significant. This is in line with our interpretation that when DSC and $Hellinger$ are included in the same markup regression, the latter measure captures dealer rents. This is also consistent with the finding in the literature that dealer rents are significantly larger for smaller than larger trades (e.g., Green et al. (2007) and Green et al. (2010)). On the other hand, the coefficients on DSC^N, which we view as capturing the unique benefits of specialization, are positive and statistically significant across columns (4) - (6), with the magnitudes of 0.04 - 0.05, not far from the estimates in the sub-sample of retail trades. Figure 5 Panels C and D show that the coefficient estimates for both $Hellinger^N$ and DSC^N are stable over time, with the coefficients on $Hellinger^N$ being insignificant in all but one years while those on DSC being positive and significant in all years. Consistent with these results, the Blinder-Oaxaca decomposition in Figure 6 Panel B shows that the specialization premium is practically entirely driven by DSC. Taken together, the results show that although institutions, like retail investors, pay more to trade with specialized dealers, they do so only in the subset of bonds in which the dealers have special expertise. That is, institutions do not appear to pay dealer rents; they only compensate specialized dealers for the unique benefits that the dealers provide.

6. What Unique Benefits do Specialized Dealers Provide?

In this section, we explore a few aspects of execution quality and other benefits that the literature has shown may be provided by dealers to their customers. We focus only on institutional-sized trades since the markup differentials for these trades mostly reflect differential benefits of trading with specialized versus national dealers, with limited contamination from dealer rents.

6.1 Immediacy

Li and Schürhoff (2019) argue that a key benefit of trading with central dealers is that they are willing to provide immediacy. Investors pay a price for immediacy because they can sell the bond to the dealers immediately without having to wait and bear the risk that the bond price may change.

Central dealers' provision of immediacy is likely facilitated by their location in the dealer network. Given their centrality, these dealers have more connections and observe more of the order flow, together enabling them to directly match sellers and buyers without having to prearrange trades. Given the segmented nature of the municipal bond market, we conjecture that specialized dealers may be able to provide immediacy better than national dealers, not because of their location in the dealer network but because of their superior understanding of the local bond market. By focusing on local bonds, specialized dealers may be able to locate likely buyers more efficiently, perceive the inventory as being less risky, and therefore are more willing to immediately take the bond into their inventory.

To test our conjecture, we follow Li and Schürhoff (2019) and run the following regression:

$$Y_c = \beta_0 DSC^N_{d(c),y(c)} + \beta_1 Hellinger^N_{d(c),y(c)} + \sum_{i=2}^{N} \beta_i X^i_{b(c),d(c),t(c)} + \alpha_{s(c),m(c)} + \epsilon_c, \tag{7}$$

where the outcome variable Y_c may be one of the following variables: a dummy for overnight inventory by any of the dealers in the chain, a dummy for overnight inventory by the first dealer, a dummy for chain length equal to one (one dealer in the chain), and the natural log of the chain length (number of dealers in the chain).

[Insert Table VII]

Table VII reports the results. Consistent with our conjecture, dealers with higher DSC measures are more likely to hold inventory overnight, more likely to sell directly to customers without using the dealer network, and hence are associated with a shorter intermediation chain. To gauge the economic significance of these findings, we focus, consistent with the decomposition of differential markups in the previous section, on comparing the effects of specialized dealers in their specialized states with the effects of national dealers. We take the 90th percentile of DSC^N as the DSC^N of specialized dealers in their specialized states and the median DSC^N of dealers with a $Hellinger$ measure below the median as the DSC^N of national dealers in a typical state, and multiply the difference in DSC^N ($2.28 - (-0.34) = 2.62$) by the coefficient of DSC^N in each column to assess the effects of trading with specialized dealers relative to national dealers. In columns (1) and (2), the coefficients on DSC^N are positive and statistically significant at the 1% level. Focusing on the second column, the coefficient of 0.02 implies that specialized dealers are about five percentage points more likely to hold a bond overnight than national dealers. This is about 16% of the average probability of holding inventory overnight (which is about 33%).

In column (3), the coefficient on DSC^N is positive and statistically significant at the 1% level, while in column (4), it is significantly negative. By the same economic magnitude calculation as above, we find that for bonds issued in the states in which they are specialized, specialized dealers are about five percentage points more likely than national dealers to be the only dealer in the intermediation chain. This is about 9% of the average probability of having just one dealer in the chain (which is about 60%). Together, the results in columns (3) and (4) suggest that specialized dealers are better than national dealers at locating buyers for bonds within their states of specialization, and hence rely less on the dealer network to manage inventory.

While Li and Schürhoff (2019) present immediacy as the main justification for higher markups associated with central dealers, the economic effects that we document above for dealer specialization seem unlikely to fully explain the differential markups between specialized and national dealers. We argue that if immediacy is the only justification and dealers know at the time they take inventory whether they will hold it overnight, then the differential markups associated with DSC should disappear when we look within a sub-sample of chains that are completed within the same timeframe. We test whether this is the case by re-running the markup regression in equation (6) for sub-samples of same-day and overnight chains. Table VIII reports the results.

[Insert Table VIII]

In all columns of Table IX, the coefficient estimates for DSC^N are positive and statistically significant at the 1% level, and the magnitudes are almost the same as the corresponding estimates for institutional-sized trades in Table VI. Thus, relative to national dealers, the differential markups associated with specialized dealers are not only due to the fact that specialized dealers disproportionately take bonds into their inventory, allowing their customers to trade faster.

6.2 Allocation of New Bond Offerings

Given their local market connections, specialized dealers may also serve as or partner with local underwriters, providing their customers with access to underpriced new bond issues. Nikolova et al.

(2020) show that corporate bond underwriters reward their relationship customers with a higher allocation of new issues. Hendershott et al. (2020) also mention access to the primary market as one of the benefits of dealer-customer relationships. Customers may choose to trade with only one or two dealers, and pay more due to the lack of competition. In return, they are more likely to get an allocation of lucrative new bond offerings.

In our version of the MSRB data, we do not have dealer identities and therefore cannot directly observe their underwriting activities. We get around the problem in three ways, all of which provide consistent suggestive evidence. First, we assume that local dealers are more likely to work with local underwriters and therefore can provide more primary-market-related benefits in states in which local underwriters have greater market shares. If these benefits are valuable, then more benefits should translate to higher markups. We test our conjecture by adding a measure of local underwriter market shares and its interaction with DSC to the markup regression in equation (6). We measure local underwriters' market shares using the volume-weighted average underwriter $Hellinger$. Like dealer $Hellinger$, underwriter $Hellinger$ is calculated by equation (1) but with the distribution of trading volume of each dealer across states replaced by the distribution of each underwriter's issuance business. We collect data on municipal bond underwriters from SDC Platinum, and use a similar machine learning topic model to group states into regions based on the co-occurrence within the same underwriter.

[Insert Table IX]

The first column of Table IX reports the underwriter $Hellinger$ augmented results based on the specification in column (4) of Table VI. The coefficient on DSC^N is practically the same both in magnitude and significance as that in Table VI. The interaction of DSC^N and volume-weighted average underwriter $Hellinger$ also has a significantly positive coefficient, suggesting that the specialization premium is higher in states in which local underwriters have higher market shares. However, the economic magnitude seems small, especially given the variation in volume-weighted average underwriter $Hellinger$. In columns (2) and (3), we split the sample at the median of underwriter $Hellinger$ into the low and high group, and run a separate regression for each group (instead of using the interaction term). Consistent with the small interaction effect in column (1), the coefficients on DSC^N in the low and high underwriter $Hellinger$ groups are practically equal.

The weak association between underwriter $Hellinger$ and specialization premia at the state level may be due to the fact that there are many bonds issued within each state and some of these bonds are not underwritten by local underwriters. If the compensation for access to new bond offerings is directly attached to the local bonds that the specialized dealers help secure, then only the transactions in these bonds should be associated with higher markups. We test our conjecture, again, by adding a measure of underwriter local-ness for each bond and its interaction with DSC to the markup regression in equation (6). We measure bond-level underwriter local-ness using the equally-weighted average underwriter-state concentration, or USC. Like a dealer's DSC, an underwriter's USC is calculated by equation (2) but with the share of trading volume of each dealer in each state replaced by the share of each underwriter's issuance volume. We also normalize USC within a state to level its variation across small and larger states. We use the same region definitions as described for the calculation of underwriter $Hellinger$.

Column (4) of Table IX reports the results for the full sample with the interaction of DSC^N and USC^N capturing the incremental markups associated with access to new bond offerings. The interaction coefficient is positive and significant at the 1% level, confirming the results based on state-level underwriter local-ness. Columns (5) and (6) reports the regression estimates separately for the low and high USC groups (split at the median across bonds in each state). The coefficient on DSC^N in the high USC sub-sample is higher than that in the low sub-sample by about 0.01 (0.04 vs. 0.03). Overall, the results suggest that specialized dealers are compensated, in part, for helping customers secure an allocation of new bond offerings but such compensation may not be economically significant.

Another avenue is to look at quantities rather than prices. If local dealers work with local underwriters and can provide more primary-market-related benefits in states in which local underwriters have greater market shares, they should be able to draw more customers in those states. Thus, the market shares of specialized dealers and local underwriters should be positively correlated. As in our investigation of dealer markups, we test our quantity-based conjecture by adding the volume-weighted average underwriter $Hellinger$ to the state-level regression in equation (3). Table X Panel A reports the results based on the two specifications in Table III Panel A. The coefficient estimates for volume-weighted average underwriter $Hellinger$ are positive and statistically significant at the 1% level. In economic terms, a standard deviation increase in state-year average underwriter

$Hellinger$ is associated with an increase of 4.54×10^{-3} in state-year average dealer $Hellinger$ (11% of its standard deviation). That is, in states in which specialized underwriters have greater market shares in the primary market, specialized dealers also have greater market shares in the secondary market. This is suggestive of the existence of primary-market-related benefits for trading with specialized dealers.

[Insert Table X]

Our third avenue is to examine directly whether investors who trade more with specialized dealers are more likely to receive an allocation of new bond issues. To answer the question, we need to measure how much an investor trades with specialized dealers but we cannot do so for all customers in the academic MSRB due to the lack of unmasked customer identifiers. We therefore resort to the holdings data for mutual funds (Morningstar) and the holdings and trading data for insurance companies (NAIC Schedule D). On the former, we take a monthly change in position of a mutual fund, or, for the latter, a transaction reported by an insurance company, as a trade, and match it to a transaction in the MSRB data by CUSIP, trade size/direction, and trade time. To be conservative, we only consider unique matches, which represent 51% of all mutual funds' monthly position changes and 59% of insurance companies' reported transactions. Chernenko and Doan (2020) and Chernenko and Doan (2022) use a similar procedure to identify mutual funds' municipal bond trades, and report comparable match rates. From the matched transactions, we measure the degree of connection, denoted by $SConnect$, for each institution to specialized dealers in a particular state in a given year as the sum product, across all matched transactions in the year, of transaction size and the associated dealer's DSC^N in the state.

We assess whether an institution is more likely to receive an allocation of new bond issues if they trade more with specialized dealers by running the following regression:

$$Allocation_{i,s,y+1} = \beta_0 SConnect^N_{i,s,y} + \sum_{i=1}^{N} \beta_i X^i_{i,s,y} + \alpha_{s,y} + \epsilon_{i,s,y}, \tag{8}$$

where $Allocation_{i,s,y+1}$ measures the allocation of new bond issues of state s received by institution i in year $y+1$, and $SConnect^N_{i,s,y}$ is calculated as described above for year y but normalized within state s and year y. X is a vector of control variables, all measured in year y, which includes institution i's percentage holding of bonds issued by state s, the log of total assets under management, the log of holding of bonds issued by state s, the log of trading volume, the log of trading volume in bonds issued by state s, and a state fund dummy, if the institution is a mutual fund. We include state x year fixed effects, $\alpha_{s,y}$, to absorb the common variation within each state in each year, and double-cluster the standard errors by state and year.

Table X Panels B and C report the results for mutual funds and insurance companies, respectively. We measure $Allocation$ in two ways. In the first two columns of each panel, we use the dummy that equals one if the institution receives an allocation in a new bond issue of a particular state. In the latter two columns, we use the dollar allocation amounts of new bond issues of a particular state divided by the institution's assets under management.

For mutual funds (Panel B), the coefficients on $SConnect$ are positive and statistically significant at either the 1 or 5% levels in all columns. In columns (1) and (3), the magnitudes are more diluted as we include all states, even those in which the funds do not currently invest. The results are similar for insurance companies (Panel C). The coefficients on $SConnect$ are positive and statistically significant at the 1% level in all columns. The effects are also economically significant. Using the estimate in column (4) of Panel C, for example, a standard deviation increase in $SConnect$n is associated with an increase in the allocation of new bond issues in a particular state worth about 0.02% of an insurer's holding in municipal bonds. As a benchmark for comparison, on average, an insurer receives an allocation of new bond issues in a state worth about 0.13% of their holding of all municipal bonds, and so the effect of a standard deviation increase in $SConnect$ is about 15% of the average allocation. Overall, both mutual funds and insurance companies that trade more with specialized dealers in a particular state are more likely to receive an allocation of new bond issues in that state (or receive a relatively larger amount, on average). This is again consistent with specialized dealers helping their customers secure an allocation in lucrative new bond offerings.

6.3 Information and Expertise Sharing

By focusing on one or a few states in the same area, specialized dealers may develop a significant expertise in understanding and pricing of local municipal bonds. In addition, they may repeatedly deal with the same issuers and develop some informational advantage. Investors may pay more to trade with specialized dealers to compensate them for sharing the expertise and information. This is a form of soft dollar arrangement common in the asset management industry.

By exploring the cross section of bond complexity and opacity, we test whether information and expertise sharing is one of the explanations for specialized dealers posting higher markups. Specifically, we add interaction terms between DSC^N and various measures of bond complexity and opacity to the regression of markups in equation (6). Our candidate measures include a callable bond dummy, a sinkable bond dummy, an unrated bond dummy, a revenue bond dummy, and a bank-qualified bond dummy. We also use the count of complexity features, following Harris and Piwowar (2006). As we show in Section 3, specialized dealers enjoy higher market shares in complex bonds. If they also charge more for trading these bonds, then the differential markup is likely a form of compensation for additional informational benefits (otherwise, higher market shares in these complex and opaque bonds should imply that specialized dealers are relatively cheaper with whom to trade).

[Insert Table XI]

Table XI reports the results. The coefficients on the interactions of DSC^N and our various measures of bond complexity are positive and significant at the 1 or 5% levels except in column (5) where DSC^N is interacted with a bank-qualified bond dummy. The effects of complexity on specialization premia are economically meaningful. For example, the differential markups associated with dealer specialization almost double among callable, sinkable, revenue, or unrated bonds. The results are consistent with the notion that specialized dealers help clients overcome frictions in understanding complex and informationally opaque bonds.

6.4 Access to Local Network of Investors

About half of the outstanding municipal bonds are held directly by individuals or in private accounts. In addition, municipal bonds rarely trade, and so it is difficult to locate potential buyers of a bond being sold by a customer. By focusing on one or a few states in the same region, specialized dealers may develop connections with local investor clientele, and may be more effective than national dealers at locating potential buyers of local bonds. Connecting the seller to a right buyer who more highly values a particular bond means that the dealer may be able to share a larger surplus of the trade. The hypothesis thus predicts that specialized dealers should earn relatively larger markups in bonds whose natural investor clienteles are local.

We test the hypothesis by adding interactions of DSC^N and candidate measures of the importance of local investor clienteles to the regression of markups in equation (6). We consider two candidate measures: tax privilege and state fund holdings. As we show in Sections 3 and 6.2, specialized dealers have higher market shares in states with higher local ownership as proxied by tax privilege and state fund holdings (Babina et al. (2021)). If they also charge more for trading in these states, then the differential markup is likely a form of compensation for the additional surplus that they help generate (otherwise, in the more segmented states, higher market shares in these bonds should imply that specialized dealers are relatively cheaper with whom to trade).

[Insert Table XII]

Table XII reports the results. Columns (1) and (4) show that the coefficients on the interactions of DSC^N and the two measures of the importance of local investor clienteles are positive and significant at the 1 or 5% levels. In economic terms, the estimates in column (1) suggest that the effects of DSC on markups increase from 0.02 to 0.05 as we move from states with zero tax privilege to those with the state tax privilege at the 75th percentile (0.069). Using the change in DSC^N from the value associated with a median national dealer to the value associated with a specialized dealer in their local states, the differential markups increase by 8 basis points ($(2.28 - (-0.34)) \times (0.05 - 0.02)$). The sample split results in columns (2)-(3) and (5)-(6) confirm the economic significance of the interaction effects in columns (1) and (4). In states with higher than median tax privilege and those with higher than median state fund holdings, the effects of DSC on markups triple those in the other states. The results are consistent with the notion that specialized dealers help facilitate access to a local network of investors, who are likely to value local bonds more highly.

6.5 Relative Importance of Different Benefits Provided by Specialized Dealers

The above results suggest that specialized dealers provide immediacy, access to new bond offerings, information and expertise useful for trading complex bonds, and access to local investor networks. In this section, we assess the relative importance of these benefits in explaining the specialization premium using the Blinder-Oaxaca decomposition. Specifically, we sort the institutional-sized transactions (of which the markups contain negligible dealer rents) into quintiles by DSC^N. The difference in markups between the top and bottom quintiles captures the total compensation for all benefits provided by specialized dealers. We then apply the full-blown Blinder-Oaxaca decomposition to split the total compensation into four components associated with the four benefits that we have examined, and within each, three distinct effects: endowment, coefficient, and interaction. We use overnight dummy to capture immediacy, bond-level average USC to capture access to new bond offerings, complexity count to capture information and expertise sharing, and tax privilege to capture access to local investor networks.

The endowment effects come from the difference in conditional average of each explanatory variable across the two comparison groups. For example, if transactions in the top DSC quintile are more likely to be held overnight than those in the bottom quintile and the overnight dummy is associated with higher markups, then the greater "endowment" of overnight transactions in the top DSC quintile must translate into a higher average markup for that group. The coefficient effects, on the other hand, result from the difference in "coefficient" of each explanatory variable in explaining the markups in the top and bottom DSC groups. For example, the overnight dummy may be associated with higher markups only in the top DSC quintile but not in the bottom, and thus the same representation of overnight transactions would translate to a higher average markup for the top group. Finally, since the endowment and coefficient effects are not orthogonal, they do not add up to the total effects of each explanatory variable. The residual is referred to as the "interaction" effects.

[Insert Figure 7]

Figure 7 illustrates the Blinder-Oaxaca decomposition results. After stripping out the effects of control variables, we find that the difference in full markups between transactions in the top and bottom DSC quintile is on average 10.3 basis points. Panel A shows that this difference is explained almost exclusively by the coefficient effects (9.4 out of 10.3 basis points). In Panel B, we further split the coefficient effects into the components associated with four explanatory variables representing four different benefits, and show that after correcting for the unexplained portion, privilege accounts for about two-thirds and complexity score accounts for about one-third of the coefficient effects. USC is also statistically significant but the economic magnitude is small (0.2 basis points). The overnight dummy is statistically and economically insignificant.[12] Overall, the results suggest that monopoly rents aside, specialization premia are largely compensations for access to local investor pools and assistance with opaque and complex bonds.

7. Conclusion

In a market generally characterized by geographic segmentation, we find that the municipal bond dealer network is organized to match investor clienteles in a manner more nuanced than the core-periphery structure that serves as the basis for many models. While some national dealers spread their business proportionally across the U.S., there are a number of important specialized dealers who focus on a few geographically adjacent states. These specialized dealers enjoy larger market shares in states with greater local ownership and in local bonds with more complex features.

Given this more nuanced network, we uncover that trades intermediated by specialized dealers are associated with higher costs. For small retail-sized trades, specialized dealers' markups are higher for bonds issued both by the states in which they are specialized and those issued by other states, suggesting that they capture clients with high search costs and can extract rents. For institution-sized trades, specialized dealers' markups are higher only for bonds issued by the states in which they are specialized. Taken together, our results suggest that while specialized dealers appear to possess some monopoly power, they also provide important differentiated services. We find that specialized dealers provide immediacy, reward customers with an allocation of new bond offerings, help customers overcome information frictions, and facilitate access to local customers, with the latter two explaining practically all of the specialization premium in institutional-sized trades.

[12]This suggests that the price of immediacy is the same for both specialized and non-specialized dealers. In unreported results, we find that the endowment effects associated with the overnight dummy are positive but statistically insignificant, consistent with the slightly higher representation of overnight transactions in the top DSC quintile.

References

Alok, Shashwat, Nitin Kumar, and Russ Wermers, 2020, Do fund managers misestimate climatic disaster risk, *Review of Financial Studies* 33, 1146–1183.

Babina, Tania, Chotibhak Jotikasthira, Christian Lundblad, and Tarun Ramadorai, 2021, Heterogeneous taxes and limited risk sharing: Evidence from municipal bonds, *Review of Financial Studies* 34, 509–568.

Bailey, Warren, and Julapa Jagtiani, 1994, Foreign ownership restrictions and stock prices in the Thai capital market, *Journal of Financial Economics* 36, 57–87.

Bena, Jan, Miguel A Ferreira, Pedro Matos, and Pedro Pires, 2017, Are foreign investors locusts? The long-term effects of foreign institutional ownership, *Journal of Financial Economics* 126, 122–146.

Bernile, Gennaro, Alok Kumar, and Johan Sulaeman, 2015, Home away from home: Geography of information and local investors, *Review of Financial Studies* 28, 2009–2049.

Blei, David M, and Michael I Jordan, 2006, Variational inference for Dirichlet process mixtures, *Bayesian Analysis* 1, 121–143.

Blei, David M, Andrew Y Ng, and Michael I Jordan, 2003, Latent Dirichlet allocation, *Journal of Machine Learning Research* 3, 993–1022.

Brennan, Michael J, and H Henry Cao, 1997, International portfolio investment flows, *Journal of Finance* 52, 1851–1880.

Brennan, Michael J, H Henry Cao, Norman Strong, and Xinzhong Xu, 2005, The dynamics of international equity market expectations, *Journal of Financial Economics* 77, 257–288.

Chalmers, John, Yu Steve Liu, and Z Jay Wang, 2021, The difference a day makes: Timely disclosure and trading efficiency in the muni market, *Journal of Financial Economics* 139, 313–335.

Chan, Kalok, Vicentiu Covrig, and Lilian Ng, 2005, What determines the domestic bias and foreign bias? Evidence from mutual fund equity allocations worldwide, *Journal of Finance* 60, 1495–1534.

Chang, Briana, and Shengxing Zhang, 2021, Endogenous market making and network formation, *Available at SSRN 2600242* .

Chernenko, Sergey, and Viet-Dung Doan, 2020, Forced sales and dealer choice in OTC markets, *Available at SSRN 3750171* .

Chernenko, Sergey, and Viet-Dung Doan, 2022, Mutual fund liquidity creation, *Available at SSRN 4019567* .

Choe, Hyuk, Bong-Chan Kho, and René M Stulz, 2005, Do domestic investors have an edge? The trading experience of foreign investors in Korea, *Review of Financial Studies* 18, 795–829.

Coval, Joshua D, and Tobias J Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *Journal of Finance* 54, 2045–2073.

Coval, Joshua D, and Tobias J Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811–841.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song, 2017, The value of trading relations in turbulent times, *Journal of Financial Economics* 124, 266–284.

Domowitz, Ian, Jack Glen, and Ananth Madhavan, 1997, Market segmentation and stock prices: Evidence from an emerging market, *Journal of Finance* 52, 1059–1085.

Dvořák, Tomáš, 2005, Do domestic investors have an information advantage? Evidence from Indonesia, *Journal of Finance* 60, 817–839.

Edwards, Amy K, Lawrence E Harris, and Michael S Piwowar, 2007, Corporate bond market transaction costs and transparency, *Journal of Finance* 62, 1421–1451.

Eichengreen, Barry, Romain Lafarguette, and Arnaud Mehl, 2016, Cables, sharks and servers: Technology and the geography of the foreign exchange market, *NBER Working Papers 21884* .

Eisfeldt, Andrea L, Bernard Herskovic, Sriram Rajan, and Emil Siriwardane, 2023, OTC intermediaries, *Review of Financial Studies* 36, 615–677.

Ferreira, Miguel A, Pedro Matos, Joao Pedro Pereira, and Pedro Pires, 2017, Do locals know better? A comparison of the performance of local and foreign institutional investors, *Journal of Banking & Finance* 82, 151–164.

French, Kenneth R., and James M. Poterba, 1991, Investor diversification and international equity markets, *American Economic Review* 81, 222–226.

Friewald, Nils, and Florian Nagler, 2019, Over-the-counter market frictions and yield spread changes, *Journal of Finance* 74, 3217–3257.

Gao, Pengjie, Chang Lee, and Dermot Murphy, 2020, Financing dies in darkness? The impact of newspaper closures on public finance, *Journal of Financial Economics* 135, 445–467.

Gore, Angela K, Kevin Sachs, and Charles Trzcinka, 2004, Financial disclosure and bond insurance, *Journal of Law and Economics* 47, 275–306.

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, 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?, *Journal of Finance* 65, 1669–1702.

Griffin, John M, Nicholas Hirschey, and Samuel Kruger, 2022, Do municipal bond dealers give their customers' fair and reasonable'pricing?, *Journal of Finance, Forthcoming* .

Harris, Lawrence E, and Michael S Piwowar, 2006, Secondary trading costs in the municipal bond market, *The Journal of Finance* 61, 1361–1397.

Hendershott, Terrence, Dan Li, Dmitry Livdan, and Norman Schürhoff, 2020, Relationship trading in over-the-counter markets, *Journal of Finance* 75, 683–734.

Hoffman, Matthew, Francis Bach, and David Blei, 2010, Online learning for latent Dirichlet allocation, *Advances in Neural Information Processing Systems* 23.

Hollifield, Burton, Artem Neklyudov, and Chester Spatt, 2017, Bid-ask spreads, trading networks, and the pricing of securitizations, *Review of Financial Studies* 30, 3048–3085.

Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai, 2012, Asset fire sales and purchases and the international transmission of funding shocks, *Journal of Finance* 67, 2015–2050.

Kang, Jun-Koo, and Rene M Stulz, 1997, Is bank-centered corporate governance worth it? A cross-sectional analysis of the performance of Japanese firms during the asset price deflation, *NBER Working Papers 6238* .

Li, Dan, and Norman Schürhoff, 2019, Dealer networks, *Journal of Finance* 74, 91–144.

Lundblad, Christian T, Donghui Shi, Xiaoyan Zhang, and Zijian Zhang, 2023, Are foreign investors informed? Trading experiences of foreign investors in China, *Available at SSRN 4112287* .

Malamud, Semyon, and Marzena Rostek, 2017, Decentralized exchange, *American Economic Review* 107, 3320–3362.

Munyan, Benjamin, and Sumudu W Watugala, 2018, What makes dealers central? Evidence from credit interdealer networks, *Available at SSRN 3031668* .

Nikolova, Stanislava, Liying Wang, and Juan Julie Wu, 2020, Institutional allocations in the primary market for corporate bonds, *Journal of Financial Economics* 137, 470–490.

Pirinsky, Christo A, and Qinghai Wang, 2011, Market segmentation and the cost of capital in a domestic market: Evidence from municipal bonds, *Financial Management* 40, 455–481.

Pollard, David, 2002, *A user's guide to measure theoretic probability*, number 8 (Cambridge University Press).

Sambalaibat, Batchimeg, 2022, A theory of liquidity spillover between bond and CDS markets, *Review of Financial Studies* 35, 2525–2569.

Schultz, Paul, 2012, The market for new issues of municipal bonds: The roles of transparency and limited access to retail investors, *Journal of Financial Economics* 106, 492–512.

Schultz, Paul, 2013, State taxes, local bias and differences in municipal bond yields across states, *Paper presented at the Brandeis-Bond Buyer Muni Conference* .

Üslü, Semih, 2019, Pricing and liquidity in decentralized asset markets, *Econometrica* 87, 2079–2140.

Van Nieuwerburgh, Stijn, and Laura Veldkamp, 2009, Information immobility and the home bias puzzle, *Journal of Finance* 64, 1187–1215.

Wongswan, Jon, 2006, Transmission of information across international equity markets, *Review of Financial Studies* 19, 1157–1189.

Appendix

A. Variable Definitions

A.1 State Characteristics

Value-weighted dealer Hellinger: Average Hellinger of dealers, weighted by the total dealer trading volume during the year

Privilege: The highest state income tax rate applied to income from municipal bonds issued by other states minus the highest state income tax rate applied to income from the state-issued municipal bonds

State fund holding (SFH): The amount of state-issued municipal bonds held by state municipal bond funds, presented as a percentage of the amount of state-issued municipal bonds held by all municipal bond funds.

Debt/Income: Debt/income is the total amount of the state's outstanding debt (from the Census Bureau) divided by the total income of state residents (from the U.S. Internal Revenue Service)

Volume-weighted average dealer centrality: The average centrality scores of dealers weighted by the total dealer trading volume during the year

Value-weighted underwriter Hellinger: The average Hellinger of underwriters, weighted by the total amount of state-issued bonds attributed to the underwriter in the year. When multiple underwriters facilitate a primary market sale, we assume each underwriter gets an equal amount from the deal.

Institutional holding: The amount of state-issued municipal bonds held by municipal bond funds and insurance companies as a percentage of the total amount of the state's outstanding debt

ln(Real GDP): ln(Real GDP) is the GDP with the 2015 U.S. dollar.

ln(Trading volume/Debt): The logarithm of the ratio of the annual trading volume, measured as the total par amount of state-issued bond transactions in MSRB, over the state's outstanding debt.

ln(Number of trades): The logarithm of the yearly total number of transactions of state-issued bonds.

Unemployment rate: State unemployment rates come from the Bureau of Labor Statistics.

Proactive: A dummy variable that equals one if the state is classified by Gao, Lee, and Murphy (2019) as being "proactive" in the bankruptcy of their agencies and municipalities and zero otherwise.

A.2 Dealer Characteristics

ln(Dealer size): The logarithm of the total par amount of dealer-customer transactions in a year

Dealer centrality: The eigenvector centrality, normalized between 0 and 1 by an empirical cumulative distribution function (Li and Shurhoff, 2019, pp109)

A.3 Transaction Characteristics

Round-trip markups: The customer purchase price minus the customer sale price in the chain as a percentage of the customer sale price.

1st dealer markups: 1st dealer markup is the price difference that the 1st dealer in the chain buys from a customer and sells to a customer or another dealer as a percentage of the customer sale price.

ln(Par): The logarithm of the par value of the customer sale transaction in the chain

ln(Chain length): The total number of dealers intermediating the round-trip chains

Sameday: A dummy equals one if the roundtrip transaction is completed within a day.

A.4 Bond Characteristics

ln(Remaining days to maturity): The logarithm of the number of days between the date of the customer sale transaction in the chain and the bond maturity date

ln(Days since issue date): The logarithm of the number of days between the bond issue date and the date of the customer sale transaction in the chain

ln(Issue size): The logarithm of the principal amount at issuance

General obligation: A dummy equals one if the bond is a general obligations bond

Callable: A dummy equals one if the bond is callable

Sinkable: A dummy equals one if the bond is sinkable

Bank qualified: A dummy equals one if the bond's interest payments are tax-exempt for banks

Insured: A dummy equals one if the bond is insured

Rating: The credit rating collected from Mergent (1 = AAA, 2 = AA+, etc.).

Unrated: A dummy equals one if the bond is not rated

High yield: A dummy equals one if the bond is a high-yield bond (Rating>=11, BB+)

Complex: Complex is an integer ranging from 0 to 6, counting for the number of complex features following Harris and Piwowar (2006), including callable, sinkable, extraordinary call provisions, coupon payment schedule other than semi-annual, nonstandard accrual basis or credit enhancement.

Underwriter state concentration (USC): USC is the bond-level average underwriters' state concentration in the issue state. Underwriters' state concentration is calculated from a LDA with algorithm. Underwriter information is obtained from SDC platimum, available after 2005.

A.5 Financial Institution Characteristics

State-s new allocation: State-s new allocation is the total new municipal bond allocation in state s. New allocations are proxied by the holding increases and transactions in recently issued bonds. For mutual funds, we use holding increases in bonds with issue dates after the beginning of the portfolio holding period in Morningstar mutual fund holdings data. Recently issued bonds for insurance companies are inferred from NAIC transactions; specifically, bonds transacted within 30 days from issuance and with zero accrued interest rate when purchased.

Connection to specialized dealers: Connection to specialized dealers measures the ties between the financial institution and the specialized dealers at the institution-state-year level. We match transactions in seasoned bonds (at least 90 days after issuance) from MSRB to mutual fund holdings change and transactions by insurance companies. For all transactions of bonds with a given financial institution in state s in a given year, Connection to specialized dealers is the logarithm of the sum of the product of transaction size and the associated dealer-state concentration in states for all transactions.

State-s fund (insurer) holdings / AUM: Total market value of state-s-issued bonds over the total municipal bond holdings by the institution

ln(AUM): Total municipal bond holdings by the institution

ln(trading volume - fund): Total trading volume by the mutual fund, calculated from share changes in the Morningstar holdings database

ln(trading volume - fund × state): State-s-issued bonds trading volume by the mutual fund, calculated from share changes in the Morningstar holdings database

ln(trading volume - insurer): Total trading volume by the insurance company, calculated from transactions in NAIC

ln(trading volume - insurer × state): State-s-issued bonds trading volume by the insurance company, calculated from transactions in NAIC

State fund: A dummy equals one if the fund is a single state municipal bond fund and zero if the fund is a national municipal bond fund

B. Latent Dirichlet Allocation

B.1 Motivation

We apply the Latent Dirichlet Allocation (LDA), a well-developed approach in machine learning literature (e.g., Blei, Ng, and Jordan 2003, Blei and Jordan 2006). Our exact estimation approach follows Hoffman, Bach, and Blei (2010). In this section, we briefly introduce LDA and describe the intuition of our application in the context of dealer networks.

LDA is a generative probabilities model for collections of discrete data. One famous application is in topic modeling, which provides methods for automatically organizing, understanding, and summarizing large electronic archives. Provided with a large set of documents, topic modeling groups the words based on their "co-occurrence" among the documents, thus generating topics for distinct contents.

The academic historical transaction data by MSRB provides masked identifiers for each dealer that facilitates the transactions. We investigate the geographic business patterns of broker-dealers. Specifically, we are interested in addressing the following questions via LDA.

- Does a dealer trade extensively in one state tend to also trade in neighboring states?
- If so, can we group such neighboring states into regions, and what are the boundaries of the regions?

We attempt to cluster states into different regions and classify dealers into these regions. The challenge is that some dealers possibly conduct business in multiple regions, and meanwhile, a single state could be associated with multiple regions. For example, New York bonds are traded by national dealers, New York and New Jersey-based dealers, Florida and New York-based dealers, and Tennessee-based dealers.

We adopt the LDA method to resolve this challenge. Analogous to topic modeling, we treat the full trading transcript for each dealer in a given year as a document in which each word corresponds to the state of the bond in a transaction. Thus, LDA can cluster subsets of the states into regions if many dealers specialize in dealing with bonds in a collection of states.

B.2 Model

Formally, we relate the following terms in the context of municipal bond trading in related to the conventional terminologies in topic modeling:

- w_n is the basic unit of discrete data, such as a "word" in a topic mode. In our context, the basic unit is a bond transaction in MSRB. w is a 51×1 unit-based vector representing the 50 U.S. states and Washington, D.C. We use w to record the state-s issued bond being traded in a transaction with $w_n^s = 1$ and $w_n^{-s} = 0$. Superscript s denotes the s component in w, and $-s$ is the complement set; that is the other 50 elements except for w^s.
- A "document", $\boldsymbol{w_d} = \left(w_1, w_2, \dots, w_{N_d}\right)$, is a sequence of words. In our context, $\boldsymbol{w_d}$ is the trading transcript that records the sequence of state-issued for bonds that a dealer trades each year. N measures the length of $\boldsymbol{w_d}$; that is, the number of transactions that the dealer conducts each year.
- A "corpus" is the collection of documents $D = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_D}\}$. In our analysis, D is the collection of all dealer-year pairs of $\boldsymbol{w_d}$. Thus, D contains all transactions from the MSRB.

For each $\boldsymbol{w_d}$, the following generative process with total K topics applies.

a) Choose a variable $\theta_d \sim Dirichlet(\alpha)$. θ_d is the characteristic of a $\boldsymbol{w_d}$.

b) Draw the length of $\boldsymbol{w_d} = \left(w_1, w_2, \ldots, w_{N_d}\right)$: choose a variable $N_d \sim \text{Poisson}(\xi)$,

c) For each of the element in $\boldsymbol{w_d}$, denoted as w_n, draw the issue state of the bond:

   1. Choose a topic $z_{d,n} \sim Multinomial(\theta_d)$. The dealer characteristic θ_d determines the types of $z_{d,n}$ drawn.

   2. Choose a state w_n from $p\left(w_n \mid z_{d,n}, \beta\right)$, which is a multinomial probability distribution conditioned on the chosen topic $z_{d,n}$.

$\theta_d = \left(\theta_d^1, \theta_d^2, \ldots, \theta_d^K\right)$ is a K-dimensional random variable drawn from a Dirichlet distribution governed by a K-dimensional vector α, such that $\theta_d^k \geq 0$ and $\sum_{k=1}^{K} \theta_d^k = 1$. β is the $K \times 51$ conditional probability parameters such that $\beta_{k,s} = p\left(w_n^s = 1 \mid \mathrm{z}_\mathrm{n}^k = 1\right)$. A region k overweights a state s if $p\left(w_n^s = 1 \mid \mathrm{z}_\mathrm{n}^k = 1\right)$ is high. Our research focus on the estimated θ_d and β from the model. In the paper, we denote

$$p_k(state = s) = \beta_{k,s} \text{ and } p_d(region = k) = \theta_d^k.$$

Step a. We rely on θ_d to model the trading behavior of a dealer in a given year, that is, whether the dealer is a national dealer or a specialized dealer.[1] Loosely speaking, θ_d^k represents the market share of the dealer-year observation $\boldsymbol{w_d}$ in the region k. A national $\boldsymbol{w_d}$ is with every θ_d^k close to the average market share ($\bar{\theta}^k$) across all $\boldsymbol{w_d}$. A specialized $\boldsymbol{w_d}$ is with θ_d^k higher than $\bar{\theta}^k$ in the specialized region(s).

Step b. The total number of transactions conducted by the dealer (N_d) each year can be drawn from a Poisson distribution with the parameter ξ.

Step c. w_n is drawn to construct the trading transcript $\boldsymbol{w_d} = \left(w_1, w_2, \ldots, w_{N_d}\right)$. Given the parameters θ_d and N_d, step c iterates over $n = 1, 2, \ldots, N_d$. In each iteration, step c.1 draws z_n from the multinominal distribution with parameter θ_d. z_n is a K-dimensional unit-based vector representing that the kth region is drawn with $\mathrm{z}_\mathrm{n}^k = 1$ and $\mathrm{z}_\mathrm{n}^{-k} = 0$. Superscript $-k$ is the complement set: that is, the other K-1 elements except for z_n^k. Step c.2 draws w_n following a multinominal distribution $p(w_n \mid z_n, \beta)$.

---

[1] θ has the following probability density:

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{k=1}^{K} \alpha_k)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \theta_1^{\alpha_1 - 1} \ldots \theta_K^{\alpha_K - 1},$$

where α, a K-dimensional vector with $\alpha_k > 0$, governs the Dirichlet distribution, and $\Gamma(.)$ is the Gamma function.

For a specialized $\boldsymbol{w_d}$ in region k, θ_d^k is higher such that $\mathrm{z}_{\mathrm{n}}^k = 1$ is more likely to be drawn for some k. Thus, the realization of $\boldsymbol{w_d}$ will be concentrated in those states that are overweighted in region k. For a national $\boldsymbol{w}$, in contrast, all θ_k are close to $\bar{\theta}_k$ such that $\boldsymbol{w_d}$ will not overweight any states.

B.3 Parameter Estimation

A "corpus" $D = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \dots, \boldsymbol{w_D}\}$ contains all dealer-year pairs of $\boldsymbol{w_d}$ observed from MSRB data. The goal is to estimate the parameters α and β, such that the log likelihood of the data is minimized:

$$l(\alpha, \beta) = \sum_{d=1}^{D} \log p(\boldsymbol{w}_d \mid \alpha, \beta),$$

in which (Blei, Ng, and Jordan, 2003, p1003)

$$p(\boldsymbol{w} \mid \alpha, \beta) = \frac{\Gamma(\sum_{k=1}^{K} \alpha_k)}{\prod_{k=1}^{K} \Gamma(\alpha_k)} \int \left(\prod_{i=1}^{k} \theta_i^{\alpha_i - 1}\right) \left(\prod_{n=1}^{N} \sum_{k=1}^{K} \prod_{s=1}^{51} \left(\theta_i \beta_{k,s}\right)^{w_n^s}\right) d\theta.$$

$p(\boldsymbol{w} \mid \alpha, \beta)$ is intractable, but a variational inference provides a tractable lower bound on the log-likelihood. Then, α and β are selected to maximize the lower bound.

Table I

Dealer level summary statistics over time

This table presents summary statistics of municipal bond dealers over time. The sample period is from 2006 to 2017. At the dealer-year level, Panel A presents the number of dealers, measured by the unique masked dealer identifier from MSRB, and reports the mean, the standard deviation, the 25th, the 50th, and the 75th percentiles of the dealer Hellinger measure. Panel A also reports the pairwise correlations for three dealer-level measures: Hellinger, Centrality, and Dealer size. Dealer size is the total par amount of dealer-customer transactions in a year. Centrality is the eigenvector centrality calculated from all interdealer transactions. Panel B presents the summary statistics of the dealer state concentration by year. The sample is split into the "Low Hellinger dealers" group and the "High Hellinger dealers" group based on the median Hellinger each year. Reported are the number of dealer-year-state observations in each group and the mean, the standard deviation, the 25th, the 50th, the 75th, and the 95th percentiles.

Panel A: Hellinger

							Pairwise correlation		
Year	# Dealer	Mean	SD	P25	P50	P75	Hellinger w/t ln(Dealer size)	Hellinger w/t Centrality	Centrality w/t ln(Dealer size)
2006	881	0.43	0.19	0.27	0.45	0.59	-0.27	-0.22	0.80
2007	930	0.43	0.19	0.27	0.46	0.59	-0.29	-0.27	0.80
2008	885	0.44	0.20	0.28	0.46	0.60	-0.30	-0.29	0.82
2009	870	0.43	0.20	0.26	0.44	0.60	-0.31	-0.29	0.83
2010	869	0.41	0.20	0.25	0.41	0.57	-0.30	-0.25	0.84
2011	856	0.40	0.20	0.23	0.41	0.57	-0.32	-0.29	0.85
2012	829	0.39	0.20	0.21	0.39	0.56	-0.36	-0.32	0.86
2013	784	0.39	0.20	0.21	0.39	0.56	-0.38	-0.37	0.85
2014	726	0.38	0.21	0.19	0.37	0.55	-0.38	-0.34	0.83
2015	686	0.37	0.20	0.19	0.36	0.54	-0.38	-0.35	0.85
2016	617	0.36	0.21	0.19	0.36	0.54	-0.41	-0.38	0.83
2017	555	0.36	0.20	0.18	0.35	0.51	-0.42	-0.37	0.84

Table I, cont'd

Panel B: Dealer state concentration

	Low Hellinger dealers						
Year	# Dealer × State	Mean	SD	P25	P50	P75	P95
2006	22,389	1.05	0.61	0.65	0.95	1.31	2.21
2007	23,715	1.05	0.62	0.66	0.95	1.29	2.27
2008	22,542	1.02	0.60	0.63	0.92	1.28	2.21
2009	22,185	1.02	0.60	0.64	0.92	1.25	2.21
2010	22,083	1.02	0.55	0.67	0.93	1.24	2.10
2011	21,828	1.02	0.52	0.70	0.95	1.23	2.02
2012	21,114	1.00	0.49	0.70	0.94	1.21	1.95
2013	19,890	1.02	0.49	0.72	0.96	1.21	1.96
2014	18,411	1.03	0.46	0.75	0.97	1.21	1.89
2015	17,442	1.02	0.45	0.77	0.98	1.20	1.90
2016	15,708	1.03	0.43	0.79	1.00	1.21	1.85
2017	14,076	1.04	0.43	0.80	0.98	1.21	1.83

	High Hellinger dealers						
Year	# Dealer × State	Mean	SD	P25	P50	P75	P95
2006	22,491	1.06	1.21	0.14	0.60	1.58	3.94
2007	23,613	1.06	1.23	0.14	0.58	1.55	4.00
2008	22,491	1.06	1.24	0.14	0.58	1.55	4.04
2009	22,083	1.07	1.24	0.14	0.59	1.60	4.01
2010	22,134	1.07	1.22	0.17	0.60	1.59	3.96
2011	21,777	1.07	1.21	0.17	0.60	1.60	3.94
2012	21,063	1.07	1.20	0.18	0.60	1.57	3.88
2013	19,890	1.08	1.20	0.18	0.61	1.58	3.90
2014	18,462	1.07	1.17	0.19	0.62	1.56	3.80
2015	17,442	1.08	1.16	0.21	0.65	1.59	3.81
2016	15,708	1.09	1.15	0.22	0.67	1.59	3.76
2017	14,127	1.08	1.13	0.23	0.66	1.57	3.66

Table II

State, transaction, and institution-level summary statistics

This table presents summary statistics for the round-trip transactions (Panel A), the characteristics of states (Panel B), and the financial institutions (Panel C). The sample period is from 2006 to 2017. Panel A reports the summary statistics at the round-trip chains level. Round-trip chains start with customer sales to dealers and end with customer purchasing from dealers. The chains could involve one or more dealers to intermediate. The state sample in Panel B includes observations at the state-year level for the 50 US states and Washington, DC. Panel C includes observations at the institution-state-month level for municipal bond mutual funds and insurance companies. Appendix A defines all variables.

Panel A: Round-trip transaction level

	N	Mean	SD	P25	P50	P75
Round-trip markups	7,301,333	1.53	1.53	0.42	1.21	2.36
1st dealer markups	7,301,333	1.03	1.22	0.23	0.62	1.58
Par ($1)	7,301,333	144,976	868,932	15,000	25,000	50,000
Dealer size ($1million)	7,301,333	62,890	122,895	3,375	12,209	43,116
Dealer centrality	7,301,333	0.95	0.08	0.94	0.98	0.99
Chain length	7,301,333	1.81	1.11	1.00	1.00	3.00
Remaining days to maturity	7,301,333	4,040	2,817	1,708	3,390	5,946
Days since issue date	7,301,333	2,062	1,408	1,004	1,909	2,907
Issue size ($1million)	7,301,333	28.48	80.84	1.86	6.35	22.03
Revenue	7,301,333	0.63	0.48	0.00	1.00	1.00
Callable	7,301,333	0.70	0.46	0.00	1.00	1.00
Sinkable	7,301,333	0.48	0.50	0.00	0.00	1.00
Bank qualified	7,301,333	0.06	0.24	0.00	0.00	0.00
Insured	7,301,333	0.47	0.50	0.00	0.00	1.00
Rating	7,301,333	2.68	2.39	1.00	2.00	4.00
Unrated	7,301,333	0.20	0.40	0.00	0.00	0.00
HY	7,301,333	0.00	0.07	0.00	0.00	0.00
Sameday	7,301,317	0.39	0.49	0.00	0.00	1.00
Complex	7,301,317	1.70	1.04	2.00	1.00	2.00
Underwriter state concentration	4,360,952	0.09	0.10	0.05	0.02	0.15

Table II, cont'd

Panel B: State-year level

	N	Mean	SD	P25	P50	P75
Value-weighted dealer Hellinger	612	0.20	0.04	0.16	0.19	0.22
Privilege	612	0.05	0.03	0.03	0.06	0.07
State fund holding (SFH)	612	0.21	0.19	0.04	0.15	0.34
Debt/Income	612	0.10	0.02	0.08	0.10	0.11
Volume-weighted average dealer centrality	612	0.96	0.01	0.96	0.96	0.97
Value-weighted underwriter Hellinger	612	0.28	0.08	0.23	0.27	0.32
Institutional holding	612	0.23	0.07	0.19	0.22	0.26
ln(Real GDP)	612	12.14	1.02	11.21	12.13	12.92
ln(Trading volume/Debt)	612	-3.64	0.47	-3.90	-3.56	-3.33
ln(Number of trades)	612	8.65	1.28	7.73	8.77	9.48
Unemployment rate	612	5.94	2.16	4.30	5.50	7.35
Proactive	612	0.16	0.36	0.00	0.00	0.00

Panel C: Institution-state-month level

	Municipal bond mutual funds (N = 356,286)		Insurance companies (N=260,559)	
	Mean	SD	Mean	SD
State-s new allocation (year+1)>0	0.12	0.32	0.10	0.30
State-s new allocation (year+1) / AUM (×100)	0.16	0.70	0.09	0.40
Connection to specialized dealers	16.03	1.76	14.90	2.24
AUM ($Million)	802	2090	781	2970
State-s institution holding / AUM	0.00	0.01	0.02	0.03
ln(Trading volume - institution)	16.46	2.96	17.08	2.14
ln(Trading volume - institution × state)	3.88	6.30	5.49	7.09
State fund	0.60	0.49		

Table III

Average dealer Hellinger across states and bonds

Panel A reports results from OLS panel regressions of value-weighted dealer Hellinger on State fund holding (SFH) in Model (1) and on Privilege in Model (2). Observations are state-years. Value-weighted dealer Hellinger is the average Hellinger of dealers weighted by dealer size in the state. SFH is the amount of state-issued municipal bonds held by state municipal bond funds, presented as a percentage of the amount of state-issued municipal bonds held by all municipal bond funds. Privilege is the highest state income tax rate applied to income from municipal bonds issued by other states minus the highest state income tax rate applied to income from the state-issued municipal bonds. The *t*-statistics, clustered by year, are in parentheses. Panel B reports results from OLS panel regressions of value-weighted dealer Hellinger, at bond-year level, which is the average Hellinger of dealers weighted by trading volume in the bond. Complex is an integer ranging from 0 to 6, counting for the number of complex features following Harris and Piwowar (2006), including callable, sinkable, extraordinary call provisions, coupon payment schedule other than semi-annual, nonstandard accrual basis or credit enhancement. The *t*-statistics, clustered by year and bond, are in parentheses. *p<.1; **p<.05; ***p<.01. Appendix A defines all control variables.

Panel A: State-level regressions

Dep var: Value-weighted dealer Hellinger (State-year)		
	(1)	(2)
State fund holding (SFH)	0.02***	
	(6.13)	
Privilege		0.09***
		(4.05)
Debt/Income	0.03	0.05
	(0.91)	(1.65)
Volume-weighted average dealer centrality	-2.88***	-2.93***
	(-18.69)	(-18.08)
Institutional holding	-0.09***	-0.09***
	(-5.54)	(-5.59)
ln(Real GDP)	0.01**	0.01***
	(2.67)	(4.35)
ln(Trading volume/Debt)	0.01**	0.01**
	(2.41)	(2.55)
ln(Number of trades)	-0.01***	-0.01***
	(-3.39)	(-4.22)
Unemployment rate	0.00	0.00
	(1.26)	(0.99)
Proactive	-0.01***	-0.01***
	(-11.59)	(-11.37)
Year FE	Yes	Yes
Observations	612	612
Adj R-squared	0.84	0.83

Table III, cont'd

Panel B: Bond-level regressions

Dep var: Value-weighted dealer Hellinger (Bond-year)	(1)	(2)
Complex		0.02***
		(10.49)
Callable	-0.00	
	(-0.10)	
Sinkable	0.05***	
	(8.11)	
Rev	0.04***	0.04***
	(8.49)	(8.52)
Bank qualified	0.50***	0.50***
	(21.86)	(22.21)
Insured	-0.13***	-0.14***
	(-10.89)	(-13.86)
Rating	0.01***	0.01***
	(11.23)	(10.53)
Unrated	0.16***	0.16***
	(16.20)	(15.17)
High yield	0.02	0.01
	(0.73)	(0.57)
ln(Remaining days to maturity)	0.01	0.00
	(0.76)	(0.18)
ln(Days since issue date)	-0.01	-0.01
	(-0.68)	(-0.95)
ln(Issue size)	-0.02***	-0.02***
	(-10.48)	(-9.60)
Year × State FE	Yes	Yes
Observations	2,583,547	2,583,547
Adj R-squared	0.09	0.09

Table IV

Trading costs sorted by Centrality and Hellinger

This table reports the average round-trip markup for dealers with different Centrality and Hellinger. We sort our transaction sample into three terciles (Low, Medium, and High) based on the monthly distribution of Centrality. Within each Centrality tercile, we conduct a second sort based on the monthly distribution of Hellinger. Reported are the average round-trip markups for each double-sorted tercile group over the sample period from 2006 to 2017. The last column reports the difference between the average markups from the High Hellinger and Low Hellinger terciles under the same Centrality tercile. To test whether the differences are significantly different from zero, we run OLS panel regressions $Markup_c = \beta_0 + \beta_1 High_c + \beta_2 Low_c + \epsilon_c$, and conduct the Wald test: $\beta_1 - \beta_2 = 0$. $High_c$ (Low_c) is a dummy equal to one if the dealer in the transaction is in the High (Low) Hellinger tercile. Standard errors are double clustered by month and dealer. *p<.1; **p<.05; ***p<.01.

		Average round-trip dealer markup			
		2nd stage sort: Hellinger			
		Low	Medium	High	High - Low
1st stage sort:	Low	1.46	1.36	1.69	0.23***
Centrality	Medium	1.59	1.50	1.77	0.19***
	High	1.27	1.51	1.68	0.42***

Table V
Trading costs and dealer specialization

This table reports results from OLS panel regressions of round-trip trading markups. Observations are no-split round trip transaction level. The explanatory variables include dealer characteristics, transaction characteristics, and issue characteristics described in Appendix A. Round-trip dealer markup is the total round trip markup charged by all dealers in a round-trip. 1st dealer markup is the markup charged by the first dealer in a round-trip. C(N)DC sample includes round-trips intermediated by one or more dealers, and dealer characteristics are for the 1st dealer in round-trips. CDC sample includes round-trips intermediated by a single dealer. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep var:	Round-trip dealer markup					1st dealer markup	
	C(N)DC	C(N)DC	C(N)DC	CDC	CDC	C(N)DC	C(N)DC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dealer-state concentration			0.04**		0.08***		0.06***
			(2.30)		(3.26)		(2.82)
Hellinger		0.13***	0.11***	0.12**	0.07	0.14***	0.10***
		(3.69)	(2.94)	(2.42)	(1.44)	(3.73)	(2.71)
ln(Dealer size)	0.01	0.03*	0.03*	0.05*	0.05*	0.04**	0.04***
	(0.44)	(1.84)	(1.86)	(1.91)	(1.92)	(2.61)	(2.62)
Dealer centrality	1.15*	1.48***	1.51***	0.84	0.90	1.45**	1.50**
	(1.81)	(2.69)	(2.74)	(1.15)	(1.22)	(2.45)	(2.51)
Dealer centrality × ln(Chain length)	-2.18***	-2.43***	-2.45***			-1.90***	-1.93***
	(-5.71)	(-6.89)	(-6.97)			(-5.14)	(-5.23)
ln(Chain length)	2.55***	2.82***	2.84***			1.15***	1.18***
	(7.45)	(9.21)	(9.32)			(3.65)	(3.77)
ln(Par)	-0.31***	-0.32***	-0.32***	-0.30***	-0.30***	-0.26***	-0.25***
	(-22.32)	(-23.42)	(-23.82)	(-19.05)	(-19.14)	(-18.72)	(-19.00)
ln(Remaining days to maturity)	0.61***	0.60***	0.60***	0.57***	0.56***	0.43***	0.42***
	(21.36)	(19.35)	(19.28)	(14.18)	(14.04)	(13.23)	(13.19)
ln(Days since issue date)	-0.07**	-0.08***	-0.08***	-0.09**	-0.09**	-0.04	-0.04
	(-2.25)	(-2.61)	(-2.64)	(-2.09)	(-2.16)	(-1.20)	(-1.22)
ln(Issue size)	-0.05***	-0.05***	-0.05***	-0.05***	-0.05***	-0.04***	-0.04***
	(-8.96)	(-7.21)	(-7.06)	(-5.20)	(-4.97)	(-5.82)	(-5.68)
Revenue	0.10***	0.10***	0.10***	0.06***	0.05***	0.04***	0.04***
	(7.22)	(7.84)	(7.97)	(3.11)	(3.22)	(3.21)	(3.29)

Cont'd next page

Table V, cont'd

	Round-trip dealer markup					1st dealer markup	
	C(N)DC	C(N)DC	C(N)DC	CDC	CDC	C(N)DC	C(N)DC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Callable	-0.05***	-0.06***	-0.06***	-0.02	-0.03	-0.04**	-0.04**
	(-3.18)	(-3.23)	(-3.27)	(-0.94)	(-0.98)	(-2.27)	(-2.32)
Sinkable	0.09***	0.08***	0.08***	0.09***	0.09***	0.06***	0.06***
	(5.56)	(5.45)	(5.50)	(3.93)	(3.94)	(4.68)	(4.68)
Bank qualified	0.03	-0.01	-0.01	0.04	0.03	0.03	0.03
	(1.42)	(-0.49)	(-0.63)	(1.22)	(1.02)	(1.46)	(1.29)
Insured	0.07***	0.07***	0.07***	0.08***	0.08***	0.06***	0.06***
	(4.61)	(5.51)	(5.61)	(3.48)	(3.52)	(3.33)	(3.43)
Rating	0.07***	0.07***	0.07***	0.07***	0.07***	0.05***	0.05***
	(18.54)	(18.12)	(18.35)	(13.43)	(13.74)	(11.42)	(11.60)
Unrated	0.37***	0.35***	0.35***	0.28***	0.27***	0.19***	0.19***
	(10.13)	(10.55)	(10.82)	(6.26)	(6.50)	(6.12)	(6.32)
High yield	0.47***	0.48***	0.49***	0.20*	0.20*	0.13	0.14*
	(6.95)	(6.93)	(6.98)	(1.78)	(1.85)	(1.60)	(1.67)
Month × State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,301,321	7,301,321	7,301,317	4,179,660	4,179,658	7,301,321	7,301,317
Adj R-squared	0.32	0.33	0.33	0.35	0.35	0.32	0.33

Table VI

Trading costs and dealer specialization: Retail- vs. institution-sized trades

This table reports results from OLS panel regressions of trading markups for retail-sized (<$100K) and institution-sized transactions (≥$100K). Observations are no-split round trip transaction level. C(N)DC sample includes round-trip transactions intermediated by one or more dealers, and dealer characteristics are for the 1st dealer in round-trips. CDC sample includes round-trips intermediated by a single dealer. Full markup is the total round-trip markup charged by all dealers, and 1st dealer is the markup charged by the first dealer in a round-trip. The independent variables are described in Appendix A. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep var: Dealer full markup	Retail (<$100K)			Institution (>=$100K)		
	Full markup		1st dealer	Full markup		1st dealer
	C(N)DC	CDC	C(N)DC	C(N)DC	CDC	C(N)DC
	(1)	(2)	(3)	(4)	(5)	(6)
Dealer-state concentration	0.04	0.08**	0.06**	0.04***	0.05***	0.04***
	(1.55)	(2.49)	(2.14)	(6.52)	(7.63)	(6.49)
Hellinger	0.13***	0.09	0.13***	0.01	-0.01	0.01
	(3.19)	(1.54)	(2.89)	(0.31)	(-0.27)	(0.67)
ln(Dealer size)	0.02	0.04	0.03*	0.06***	0.09***	0.07***
	(1.06)	(1.38)	(1.76)	(4.31)	(3.96)	(4.20)
Dealer centrality	1.80***	1.14	1.76***	0.21	-0.31	0.18
	(3.06)	(1.45)	(2.80)	(0.49)	(-0.56)	(0.38)
Dealer centrality × ln(Chain length)	-2.51***		-2.04***	-1.70***		-1.38***
	(-6.53)		(-5.43)	(-5.73)		(-3.73)
ln(Chain length)	2.94***		1.21***	2.02***		1.06***
	(8.96)		(3.74)	(7.71)		(3.35)
ln(Par)	-0.29***	-0.26***	-0.21***	-0.25***	-0.24***	-0.20***
	(-9.60)	(-7.36)	(-7.98)	(-9.77)	(-8.00)	(-7.88)
ln(Remaining days to maturity)	0.66***	0.64***	0.46***	0.37***	0.31***	0.26***
	(20.63)	(15.26)	(13.84)	(12.42)	(8.55)	(9.18)
ln(Days since issue date)	-0.08**	-0.11**	-0.04	-0.07***	-0.05***	-0.02
	(-2.47)	(-2.19)	(-1.27)	(-5.62)	(-3.06)	(-1.41)
ln(Issue size)	-0.06***	-0.06***	-0.04***	-0.02***	-0.02***	-0.02***
	(-8.56)	(-6.42)	(-6.34)	(-6.35)	(-3.46)	(-6.51)
Revenue	0.11***	0.07***	0.05***	0.02**	-0.01	-0.01
	(8.10)	(3.64)	(3.86)	(2.15)	(-1.07)	(-0.68)
Callable	-0.06***	-0.03	-0.05***	-0.04***	0.00	-0.01
	(-2.89)	(-1.09)	(-2.63)	(-2.99)	(0.34)	(-1.42)
Sinkable	0.08***	0.09***	0.06***	0.04***	0.05***	0.03***
	(4.85)	(3.36)	(4.08)	(5.42)	(4.81)	(4.88)
Bank qualified	-0.01	0.01	-0.01	0.15***	0.25***	0.16***
	(-0.44)	(0.23)	(-0.55)	(6.07)	(6.82)	(6.30)
Insured	0.08***	0.09***	0.06***	0.06***	0.06***	0.04***
	(5.74)	(3.50)	(3.26)	(6.45)	(5.28)	(3.84)
Rating	0.08***	0.07***	0.05***	0.05***	0.05***	0.03***
	(18.00)	(13.60)	(11.28)	(11.52)	(8.12)	(8.83)
Unrated	0.37***	0.28***	0.20***	0.20***	0.17***	0.13***
	(10.31)	(6.26)	(6.02)	(10.35)	(7.37)	(7.89)
High yield	0.65***	0.45***	0.23***	-0.12	-0.33***	-0.14*
	(8.68)	(4.36)	(2.77)	(-1.27)	(-3.10)	(-1.69)
Month × State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,876,986	3,318,845	5,876,986	1,424,281	860,725	1,424,281
Adj R-squared	0.29	0.30	0.31	0.30	0.27	0.21

Table VII

Dealer specialization, immediacy, and search efficiency

This table reports results from OLS panel regressions for institution-sized transactions (≥$100K). Observations are no-split round-trip transactions. The dependent variable in Model (1) is a dummy that equals one if a bond is held overnight by any dealer(s) in a round-trip transaction. In Model (2), the dependent variable is a dummy that equals one if a bond is held overnight by the first dealer in a round-trip transaction. In Model (3), the dependent variable is a dummy that equals one if the round-trip transaction is intermediated by a single dealer (CDC). In Model (4), the dependent variable is the logarithm of the total number of dealers intermediating the round-trip chains. The independent variables are described in Appendix A. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep vars:	Overnight		Chain length=1	ln(Chain length)
	Any dealer	1st dealer		
	(1)	(2)	(3)	(4)
Dealer-state concentration	0.01***	0.02***	0.02***	-0.02***
	(2.75)	(3.84)	(7.42)	(-7.51)
Hellinger	0.00	0.02	0.02	-0.01
	(0.01)	(1.37)	(1.45)	(-1.21)
ln(Dealer size)	0.05***	0.05***	0.04***	-0.03***
	(4.49)	(4.23)	(4.05)	(-3.98)
Dealer centrality	-0.52**	0.06	-0.57*	0.40
	(-2.43)	(0.24)	(-1.80)	(1.53)
ln(Par)	-0.09***	-0.06***	0.03***	-0.02***
	(-8.26)	(-6.56)	(4.10)	(-3.65)
ln(Remaining days to maturity)	0.04***	0.00	-0.00	0.02***
	(5.18)	(0.37)	(-0.45)	(2.92)
ln(Days since issue date)	0.01	-0.01	0.00	-0.01
	(1.46)	(-1.11)	(0.50)	(-1.47)
ln(Issue size)	-0.01***	-0.00	0.01***	-0.01***
	(-6.62)	(-1.18)	(8.47)	(-9.35)
Revenue	-0.01**	-0.01	0.03***	-0.03***
	(-2.27)	(-1.04)	(5.32)	(-5.55)
Callable	-0.09***	-0.09***	-0.00	-0.00
	(-8.44)	(-8.21)	(-0.19)	(-0.40)
Sinkable	0.00	-0.01	0.00	-0.00
	(0.08)	(-1.58)	(0.56)	(-0.92)
Bank qualified	0.09***	-0.01	0.01	-0.01
	(6.33)	(-0.54)	(0.72)	(-0.71)
Insured	-0.01	-0.02***	-0.02**	0.02**
	(-1.09)	(-3.06)	(-2.49)	(2.57)
Rating	0.00***	-0.00	-0.00	0.00
	(3.47)	(-0.22)	(-0.30)	(1.11)
Unrated	0.01	-0.00	0.02**	-0.01**
	(1.64)	(-0.59)	(2.29)	(-2.04)
High yield	-0.01	0.02	0.03	-0.03
	(-0.31)	(0.65)	(1.26)	(-1.21)
Month × State FE	Yes	Yes	Yes	Yes
Observations	1,424,281	1,424,281	482,406	482,406
Adj R-squared	0.07	0.07	0.03	0.03

Table VIII

Trading costs and dealer specialization: Overnight vs. same-day transactions

This table reports results from OLS panel regressions of trading markups for institution-sized transactions (≥$100K). Observations are no-split round-trip transactions. Model (1) includes same-day round-trip transactions in which the dealer(s) does not hold the inventory overnight. Model (2) includes overnight round-trip transactions in which the bond is held overnight by any of the dealer(s) in a round-trip transaction. Model (3) includes both same-day and overnight round-trip transactions. All models include control variables following the specification in Model (4) in Table VI. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep var: Dealer full markup	Same day (1)	Overnight (2)	All (3)
Dealer-state concentration	0.04***	0.04***	0.04***
	(8.45)	(3.71)	(7.07)
Overnight			0.25***
			(5.91)
Dealer-state concentration × Overnight			-0.02
			(-0.91)
Hellinger	0.01	-0.01	0.00
	(0.65)	(-0.36)	(0.18)
Controls	Yes	Yes	Yes
Month × State FE	Yes	Yes	Yes
Observations	786,971	637,191	1,424,281
Adj R-squared	0.31	0.30	0.30

Table IX

Trading costs and dealer specialization: Underwriter Hellinger and state concentration

This table reports results from OLS panel regressions of trading markups for institution-sized transactions (≥$100K). Avg. underwriter Hel. is the value-weighted average underwriter Hellinger at state year level. Underwriter state concentration (USC) is the bond-level average underwriters' state concentration in the issue state. The "Low" and "High" subsamples are grouped by the monthly sample median average underwriter Hellinger and USC, respectively. All models include control variables following the specification in Model (4) in Table VI. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep var: Dealer full markup						
	Avg. underwriter Hel.			USC		
	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Dealer-state concentration	0.04***	0.04***	0.04***	0.04***	0.03***	0.04***
	(6.65)	(5.06)	(7.24)	(6.38)	(5.23)	(6.95)
Dealer-state concentration	0.01***					
× Avg. underwriter Hel.	(2.75)					
Dealer-state concentration				0.01***		
× USC				(3.55)		
USC				0.01**		
				(2.58)		
Hellinger	0.01	0.00	0.01	0.01	0.01	0.02
	(0.27)	(0.03)	(0.61)	(0.67)	(0.59)	(0.77)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month × State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,424,281	775,924	648,357	926,115	463,156	462,677
Adj R-squared	0.30	0.30	0.28	0.33	0.32	0.34

Table X

Allocation of new bond offerings: Mutual funds and insurance companies

Panel A presents the results from OLS panel regressions of value-weighted dealer Hellinger on value-weighted underwriter Hellinger. The latter is the average Hellinger of underwriters, weighted by the total amount of state-issued bonds attributed to the underwriter in the year. The other control variables follow the specifications in Panel A of Table III. The *t*-statistics, clustered by year and bond, are in parentheses. Panels B and C report results from OLS panel regressions of allocation of new bond offerings to financial institutions: municipal bond mutual funds and insurance companies. Observations are institution-state-year. New allocation in s is the total new municipal bond allocation in state s, proxied by the holding increases from the Morningstar mutual fund holdings data and the National Association of Insurance Commissioners (NAIC) transactions for recently issued bonds. The dependent variable in Models (1) and (2), State-s new allocation (year+1)>0, is a dummy equal to one if the financial institution receives at least one new allocation in state s. The dependent variables in Models (3) and (4) are state-s new allocations divided by the AUM at the beginning of the year. Connection to specialized dealers measures the ties between the financial institution and the specialized dealers at the institution-state-year level. Models (1) and (3) present the results with all states for an institution–year pair, and Models (2) and (4) focus on the states that the institution has state-issued bonds in the previous year. We lag all dependent variables by one year. See Appendix A for detailed descriptions of dependent and independent variables. The *t*-statistics, double clustered by state and institutions, are in parentheses. *p<.1; **p<.05; ***p<.01.

Panel A: State-month level

Dep var: Value-weighted dealer Hellinger (State-year)		
	(1)	(2)
Value-weighted underwriter Hellinger	0.05***	0.05***
	(4.58)	(4.46)
State fund holding (SFH)	0.02***	
	(6.21)	
Privilege		0.09***
		(4.07)
Controls	Yes	Yes
Year FE	Yes	Yes
Observations	612	612
Adj R-squared	0.85	0.84

Table X, cont'd

Panel B: Municipal bond mutual funds

Dependent variables:	State-s new allocation (year+1)>0		State-s new allocation (year+1) / AUM	
	All states	State-s fund holding>0	All states	State-s fund holding>0
	(1)	(2)	(3)	(4)
Connection to specialized dealers	0.01**	0.02**	0.03***	0.09***
	(2.32)	(2.59)	(4.21)	(4.52)
State-s fund holding / AUM	8.20***	8.30***	46.51***	44.19***
	(28.02)	(17.74)	(33.52)	(28.22)
ln(AUM)	0.01***	0.03***	-0.03***	-0.08***
	(4.79)	(4.76)	(-4.68)	(-4.78)
ln(Trading volume - fund)	0.00***	0.01***	0.01***	0.02***
	(3.39)	(6.15)	(4.51)	(4.48)
ln(Trading volume - fund × state)	0.01***	0.01***	0.01***	0.01***
	(12.92)	(10.73)	(3.44)	(3.87)
State fund	-0.10***	-0.17***	-0.06***	-0.05*
	(-11.40)	(-10.22)	(-5.98)	(-1.68)
Year × State FE	Yes	Yes	Yes	Yes
Observations	356,286	117,379	356,286	117,379
Adj R-squared	0.365	0.300	0.487	0.492

Panel C: Insurance companies

Dependent variables:	State-s new allocation (year+1)>0		State-s new allocation (year+1) / AUM	
	All states	State-s insurer holding>0	All states	State-s insurer holding>0
	(1)	(2)	(3)	(4)
Connection to specialized dealers	0.02***	0.02***	0.02***	0.02***
	(3.31)	(2.84)	(3.31)	(3.18)
State-s insurer holding / AUM	0.83***	1.08***	1.86***	1.85***
	(6.02)	(8.59)	(7.44)	(6.89)
ln(AUM)	0.02***	0.04***	-0.01***	-0.02***
	(7.91)	(8.39)	(-3.28)	(-3.81)
ln(Trading volume - insurer)	0.00	0.00	0.01**	0.02***
	(0.47)	(0.60)	(2.35)	(3.19)
ln(Trading volume - insurer × state)	0.00***	0.00***	0.00***	0.00***
	(12.76)	(11.74)	(4.07)	(3.62)
Year × State FE	Yes	Yes	Yes	Yes
Observations	260,559	140,965	260,559	140,965
Adj R-squared	0.166	0.164	0.0863	0.0974

Table XI

Trading costs and dealer specialization: Complex bond

This table reports results from OLS panel regressions of trading markups for institution-sized transactions (≥$100K). The dependent variable is the total round trip markup charged by all dealers in a round-trip. In Models (1) - (5), we interact Dealer-state concentration with five dummy proxies for bond complexity: callable, revenue, unrated bonds, sinkable, and bank qualified. Complex in Model (6) counts the number of complex features, ranging from 0 to 6, following Harris and Piwowar (2006). All models include control variables following the specification in Model (4) in Table VI. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep var: Dealer full markup						
	Bond Characteristic					
	Callable	Revenue	Unrated	Sinkable	Bank	Complex
	(1)	(2)	(3)	(4)	(5)	(6)
Dealer-state concentration	0.02***	0.03***	0.03***	0.03***	0.04***	0.02***
	(4.28)	(4.35)	(5.67)	(5.47)	(5.63)	(3.59)
Dealer-state concentration	0.03***	0.02***	0.03***	0.02**	0.02	0.01**
× Bond characteristic	(3.17)	(2.82)	(3.66)	(2.60)	(1.45)	(2.11)
Complex						0.15***
						(12.51)
Hellinger	0.01	0.01	0.01	0.01	0.01	0.01
	(0.28)	(0.32)	(0.28)	(0.28)	(0.28)	(0.30)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month × State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,424,281	1,424,281	1,424,281	1,424,281	1,424,281	1,424,281
Adj R-squared	0.295	0.295	0.295	0.295	0.295	0.300

Table XII

Trading costs and dealer specialization: local investors

This table reports results from OLS panel regressions of trading markups for institution-sized transactions (≥$100K). The dependent variable is the total round-trip markup charged by all dealers in a round-trip. In models (1) and (2), we interact Dealer-state concentration with two proxies for investor localness in each state: privilege and state fund holding (SFH). The "Low" and "High" subsamples in Models (2), (3), (5), and (6) are split by the monthly sample median. All models include control variables following the specification in Model (4) in Table VI. The *t*-statistics, double clustered by month and dealer, are in parentheses. *p<.1; **p<.05; ***p<.01.

Dep var: Dealer full markup						
	Privilege			SFH		
	All	Low	High	All	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)
Dealer-state concentration	0.02	0.02*	0.07***	0.04***	0.02	0.07***
	(1.62)	(1.97)	(7.41)	(7.32)	(1.42)	(7.25)
Dealer-state concentration × Privilege	0.46**					
	(2.55)					
Dealer-state concentration × SFH				0.02***		
				(2.95)		
Hellinger	0.00	0.02	-0.01	0.00	0.02	-0.01
	(0.18)	(0.93)	(-0.43)	(0.21)	(1.03)	(-0.39)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Month × State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,424,281	747,132	677,149	1,424,281	733,309	690,972
Adj R-squared	0.296	0.295	0.296	0.296	0.298	0.295

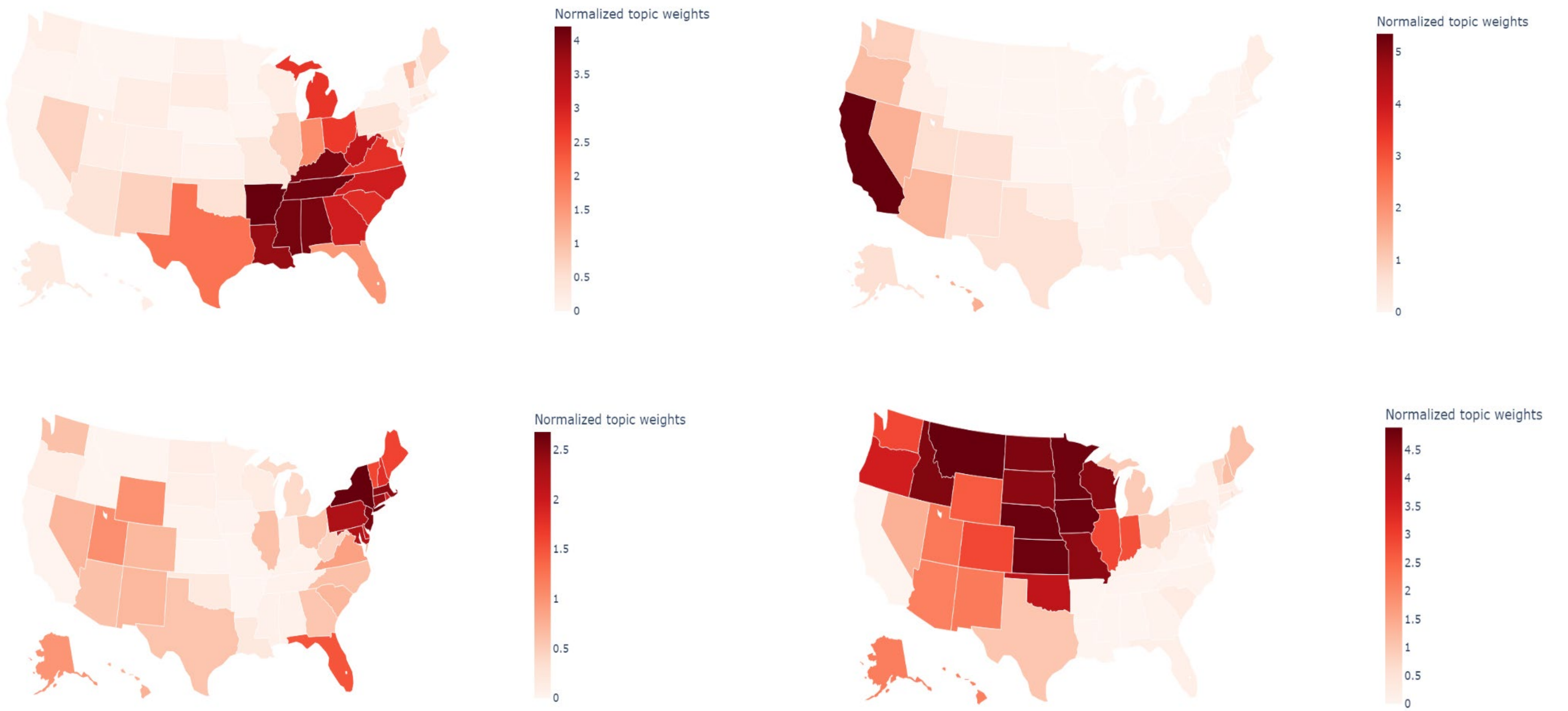
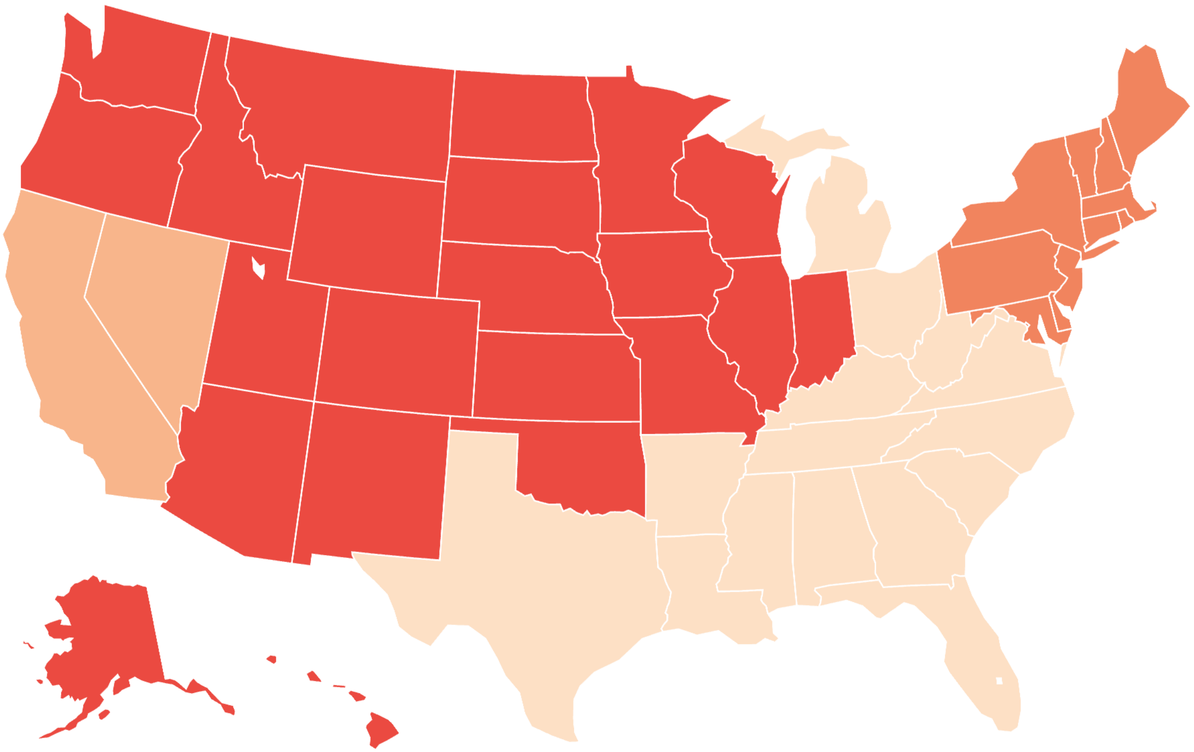


Figure 1. Normalized state weights for each topic. This figure presents the normalized weights in the states for each topic. The raw weights are the probability drawing state s in topic k, i.e., $p_k(state = s)$. We standardize by the unconditional probability of drawing state s: $p(state = s)$, estimated by the frequency of state-issued municipal bonds from MSRB transactions. $\frac{p_k(state=s)}{p(state=s)}$ measures the relative over(under)weight in the probability of drawing state s.



.

Figure 1b. Groups of states by the maximum probability. This figure describes the 50 U.S. states and Washington, D.C. with the topic of the maximum probability. For each of the states, let $k = \arg\max_k \{p_k(state = s)\}$ for k=1, 2, 3, and 4. Then, we fill a unique color for all states with the highest probability in topic k.

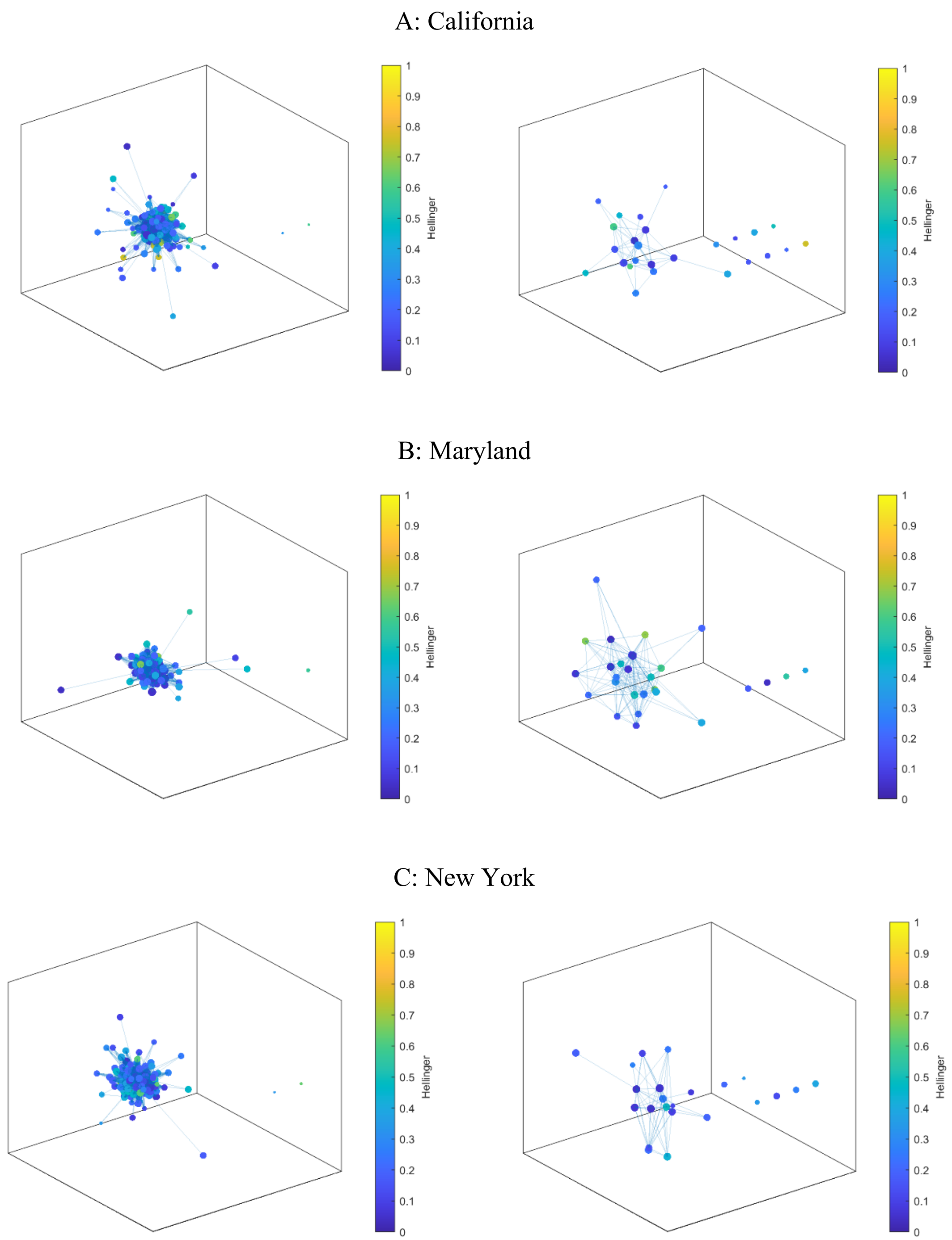


Figure 2. Dealer network and Hellinger. This figure presents the dealer network using trading in 2016 for bonds issued by California, Maryland, and New York. Each dot is a dealer: the size (color) represents Centrality (Hellinger). Figures in the left (right) column include all (top 30) dealers.

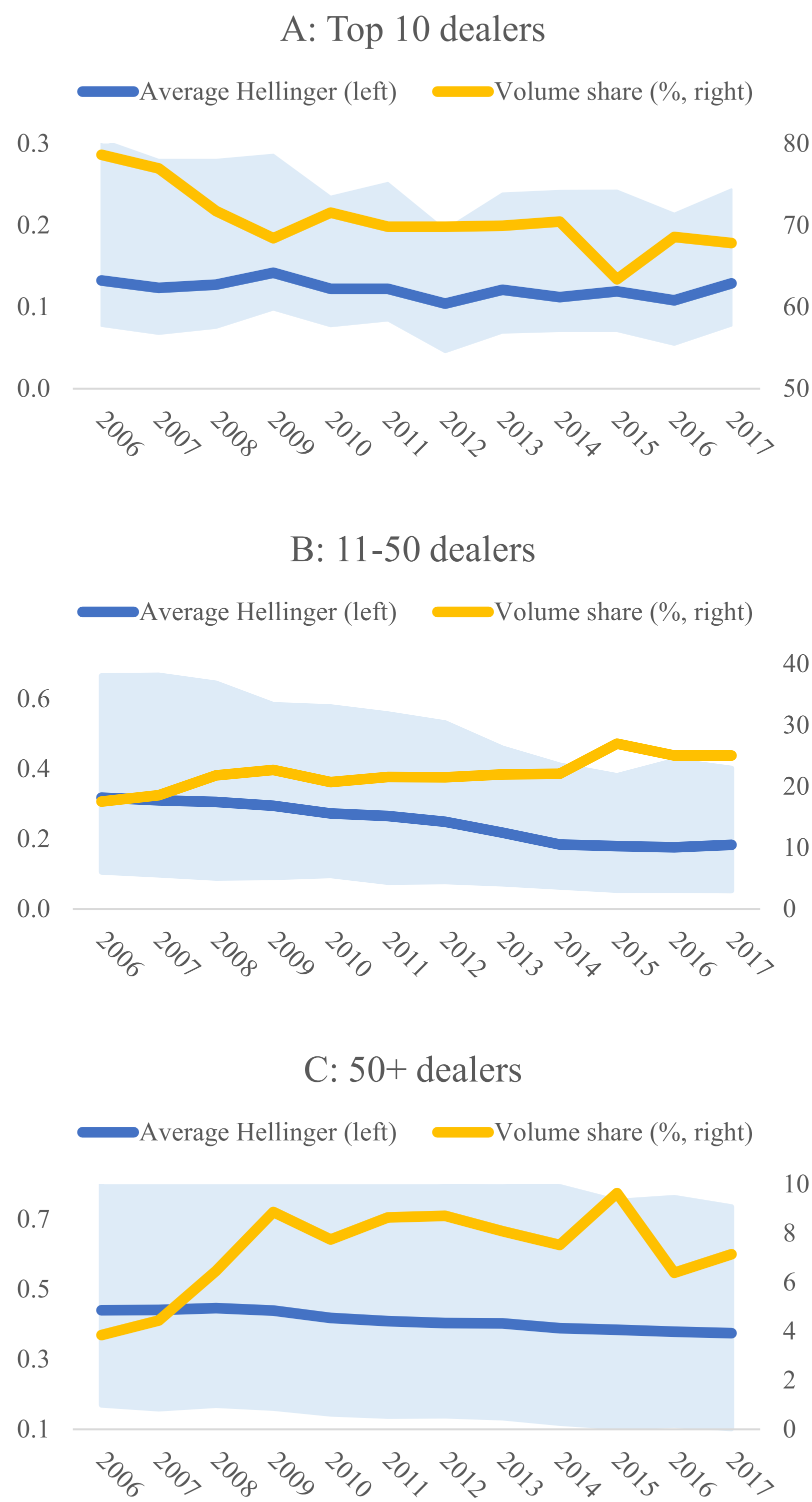


Figure 3. Average Hellinger and market share of municipal bond dealers by size group. This figure presents the average Hellinger (in blue) and the overall market share (in yellow) over the sample period from 2006 to 2017 for dealers of different sizes. The shaded blue area represents the 10th and 90th percentile of dealer Hellinger. Dealer size is the total par amount of dealer-customer transactions in a year. Panel A focuses on the top 10 dealers by size each year. Panel B includes the 11-50th dealers, and Panel C contains the 51st and smaller dealers.

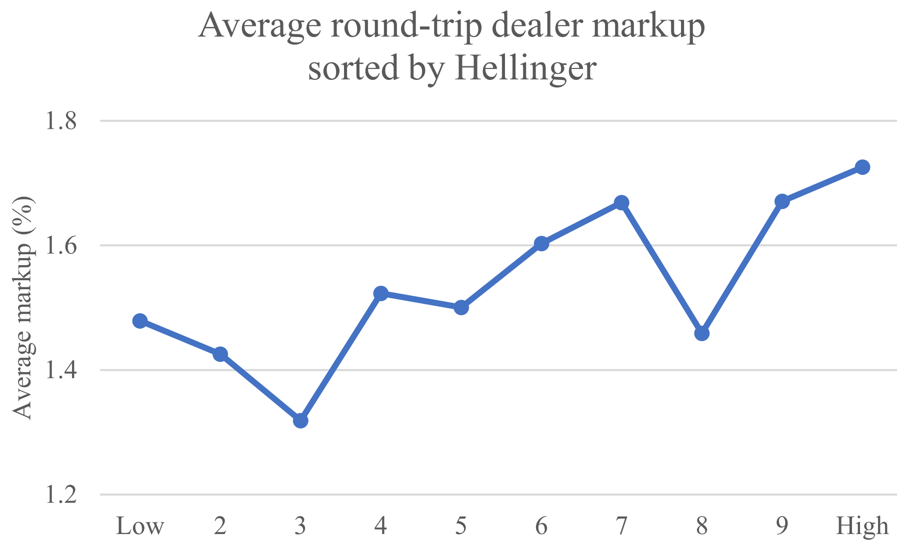


Figure 4. Average round-trip dealer markup by Hellinger decile groups. This figure presents the average round-trip dealer markup for national dealers versus specialized. We sort each round-trip chain into a decile group based on the monthly sample Hellinger deciles. Plots are the average round-trip markup for each decile group over the sample period from 2006 to 2017.

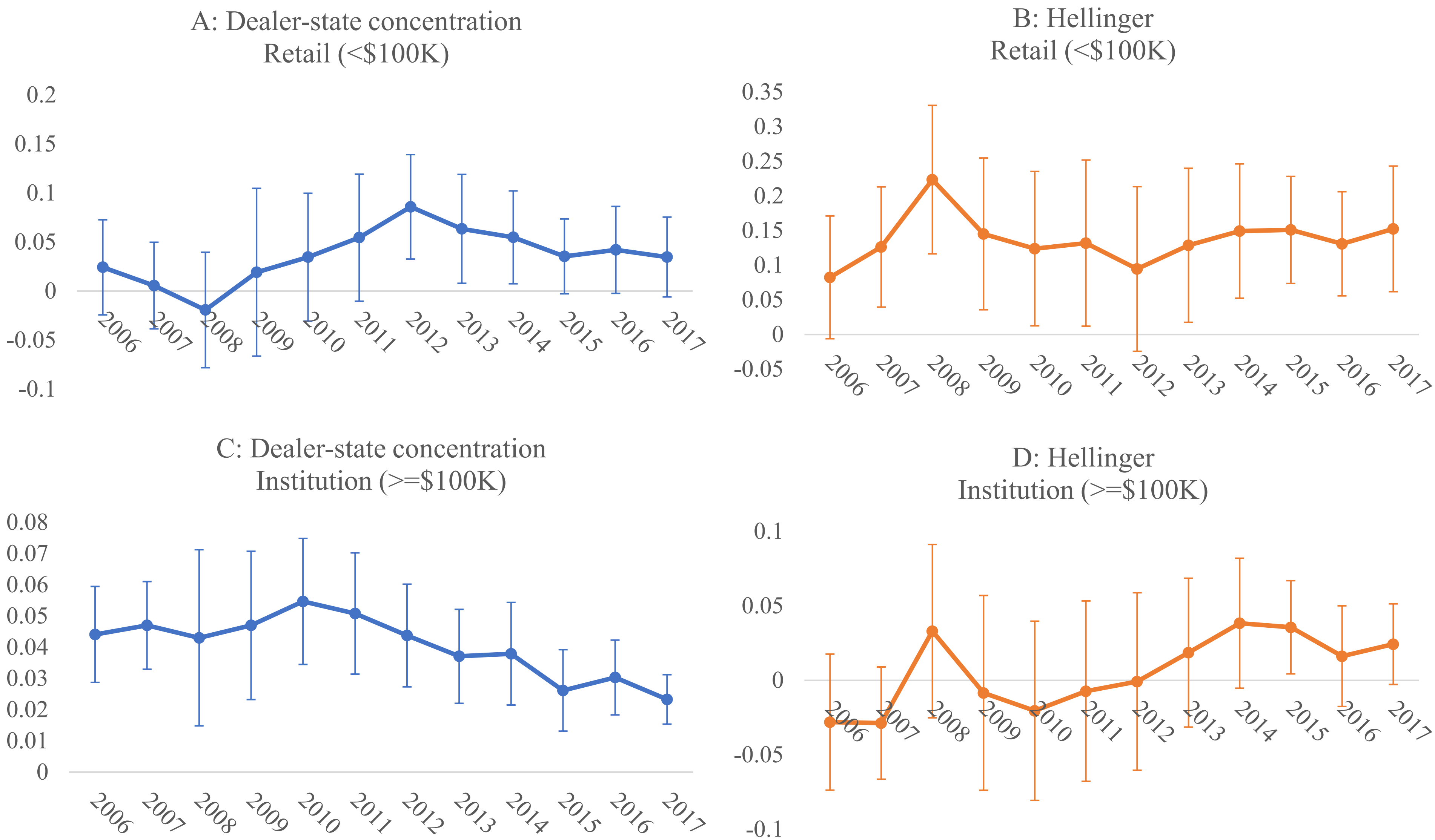
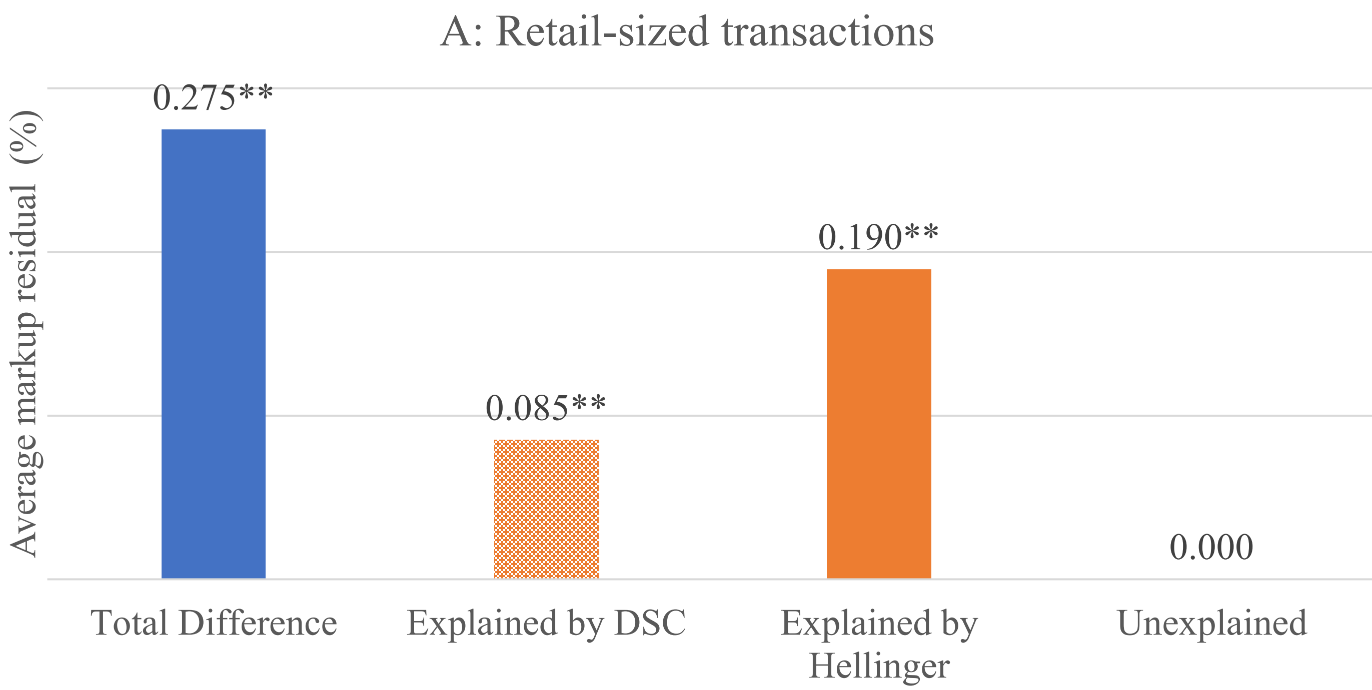


Figure 5. Trading costs and dealer specialization over time. This figure presents the estimated coefficients of Dealer-state concentration and Hellinger from OLS panel regressions of round-trip trading markups. The regression models follow Model (1) in Table VI for retail-sized transactions and Model (4) in Table VI for institution-sized transactions, and we estimate the models by year. Error bars represent the 95% confidence intervals with standard errors double clustered by month and dealer.



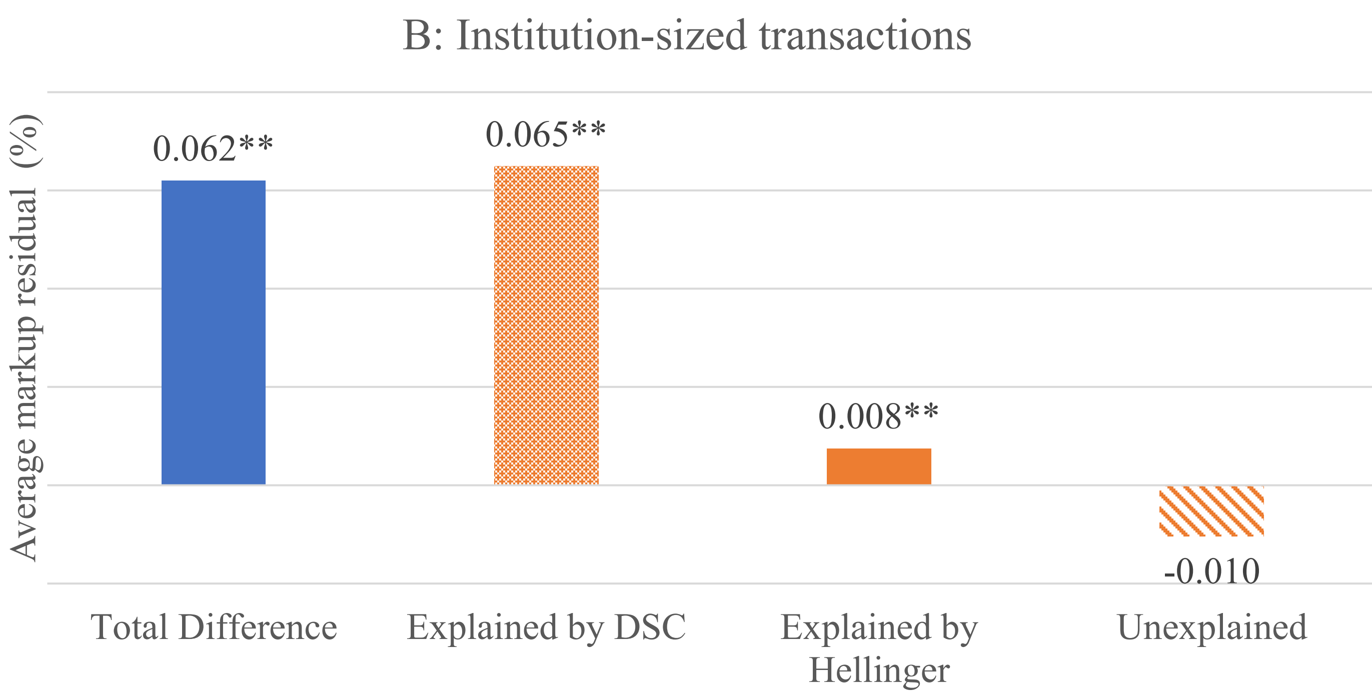


Figure 6. Oaxaca decomposition: Hellinger vs. Dealer-state concentration. The figure presents the results of Oaxaca decomposition for the impact of Hellinger and DSC on dealer markups for retail-sized (<100K) and institution-sized (≥$100K) transactions, respectively. We regress dealer markup on the control variables and month × state fixed effects, and obtain the residual dealer markup (e_c). The total difference (blue bar) is calculated by subtracting the average residual dealer markup by national dealers (bottom tercile in $Hellinger^N$; group A) from the average residual dealer markup by specialized dealers (top tercile in $Hellinger^N$ and top tercile in DSC^N; group B). We then regress e_c on $Hellinger^N$ and DSC^N: $e_c = \alpha + \beta_0 DSC^N + \beta_1 Hellinger^N + u_c$, using all retail-sized (<100K) and institution-sized (≥$100K) transactions, respectively. The component of total difference explained by DSC^N is calculated as $\beta_0(\overline{DSC_B^N} - \overline{DSC_A^N})$, where $\overline{DSC_B^N}$ and $\overline{DSC_A^N}$ are the average DSC^N for group B and group A, respectively. The component of total difference explained by $Hellinger^N$ is calculated as $\beta_1(\overline{Hellinger_B^N} - \overline{Hellinger_A^N})$, where $\overline{Hellinger_B^N}$ and $\overline{Hellinger_A^N}$ are the average $Hellinger^N$ for group B and group A, respectively. Standard errors are clustered by the dealer. *p<.1; **p<.05; ***p<.01.

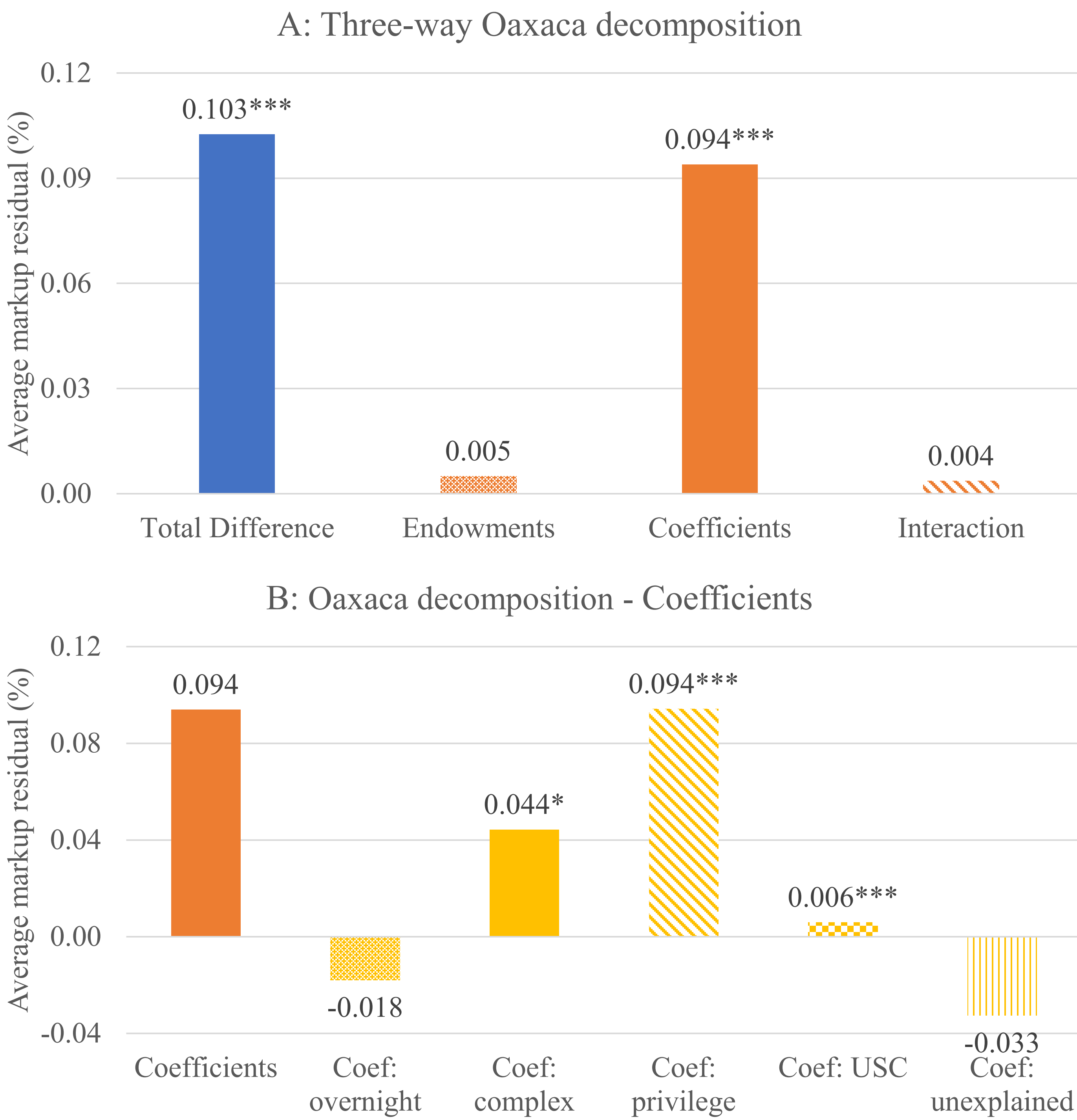


Figure 7. Three-way Oaxaca decomposition: Differential impact by DSC. The figure presents the results of Oaxaca decomposition for the differential impacts of DSC on dealer markups for institution (≥\$100K) sized transactions. We regress dealer markup on the control variables and month × state fixed effects, and obtain the residual dealer markup (e_c). The total difference (blue bar) is calculated by subtracting the average residual dealer markup by low DSC^N dealers (bottom quintile in DSC^N; group L) from the average residual dealer markup by high DSC dealers (top quintile in DSC^N; group H). We then regress e_c on four characteristics X_i (i.e., overnight, complex, privilege, and USC for $i = 1, \ldots, 4$): $e_{c,G} = \alpha_G + \sum_{i=1}^{4} \gamma_{i,G} X_{c,i} + u_{c,G}$, where G is group L or group H, respectively. In Panel A: we decompose the total difference into the component explained by endowments $\sum_{i=1}^{4} \gamma_{i,H}(\bar{X}_{i,H} - \bar{X}_{i,L})$, the component explained by coefficients $\sum_{i=1}^{4} \bar{X}_{i,H}(\gamma_{i,H} - \gamma_{i,L}) + \alpha_H - \alpha_L$, and the interaction terms. In Panel B, we present each coefficient component: $\bar{X}_{i,H}(\gamma_{i,H} - \gamma_{i,L})$, and the unexplained component is $\alpha_H - \alpha_L$. Standard errors are clustered by the dealer. *p<.1; **p<.05; ***p<.01.

Table A.1

Nuveen's top ten counterparties

This table presents the principal transaction amount with the top 10 counterparties for the six single state municipal bond funds from June 1st, 2020, to May 31st 2021. The N-CEN is filed on August 13th, 2021, with CIK 0001018972. For each fund, we collect the value in item C.17.b, which is the aggregate value of principal purchase/sale transactions of the fund during the reporting period, and present the value in the parenthesis next to the state name. The table reports the full name of the dealer (item C.17.a.i) and the fraction of principal transactions with each of the top 10 dealers, which is the ratio of C.17.a.vii over C.17.b. Item C.17.a.vii is the total purchase and sales value (excluding maturing securities) with the fund.

	Nuveen state municipal bond funds (Aggregate annual principal purchase/sale transactions)					
Top 10 counterparties	Maryland ($79 Mil)		Virginia ($134 Mil)		New Mexico ($17 Mil)	
1	BoA	40%	BoA	23%	Pershing	37%
2	Hilltop	15%	Barclays	15%	BoA	27%
3	Citigroup	14%	Citigroup	14%	Huntington	13%
4	JPMorgan	6%	JPMorgan	13%	Barclays	5%
5	Barclays	5%	Hilltop	9%	RBC	4%
6	Goldman Sachs	4%	Wells Fargo	8%	D.A. Davidson	4%
7	Wells Fargo	4%	Morgan Stanley	7%	Stifel	4%
8	Morgan Stanley	3%	Stifel	2%	Hilltop	3%
9	Pershing	2%	Pershing	2%	BNY	2%
10	RBC	2%	Janney	1%		
Sum of %		95%		95%		100%

Top 10 counterparties	Colorado ($125 Mil)		Arizona ($58 Mil)		Pennsylvania ($181 Mil)	
1	Stifel	18%	Pershing	24%	ICE	14%
2	RBC	16%	BoA	15%	RBC	11%
3	D.A. Davidson	11%	Stifel	11%	Barclays	10%
4	BoA	10%	RBC	10%	PNC	8%
5	Pershing	8%	JPMorgan	9%	BoA	7%
6	Barclays	6%	Morgan Stanley	7%	JPMorgan	7%
7	Morgan Stanley	4%	Wells Fargo	6%	Jefferies	6%
8	Jefferies	3%	Barclays	5%	Citigroup	5%
9	JPMorgan	3%	Goldman Sachs	2%	Janney	5%
10	Piper Jaffray	3%	Ziegler	2%	Wells Fargo	4%
Sum of %		84%		92%		79%

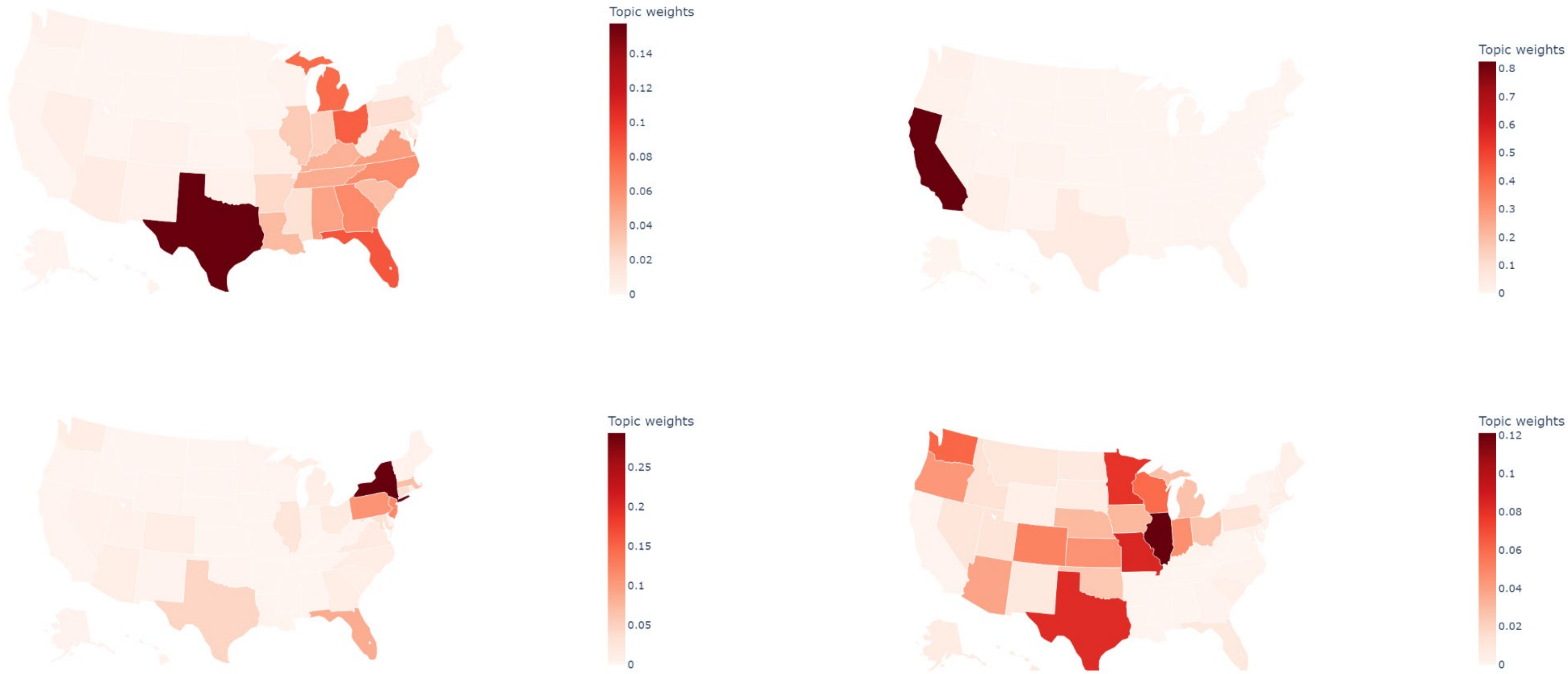


Figure A.1. State weights for each topic. This figure presents the weights in the states for each topic. The weights are the probability drawing state s in topic k, i.e., $p_k(state = s)$.