

Ask and Ye Shall Receive? Automated Text Mining of Michigan Capital Facility Finance Bond Election Proposals to Identify which Topics are Associated with Bond Passage and Voter Turnout¹

Alex J. Bowers² & Jingjing Chen
Teachers College, Columbia University

ABSTRACT:

The purpose of this study is to bring together recent innovations in the research literature around school district capital facility finance, municipal bond elections, statistical models of conditional time-varying outcomes, and data mining algorithms for automated text mining of election ballot proposals to examine the factors that influence the probability of school districts in the state of Michigan passing or failing capital facility finance bond elections from 1998-2014. Automated text mining is a data mining technique that identifies latent topics from a corpus of documents. We used an unsupervised correlated topic model to analyze the full text wording of all 1,210 school district capital facility bond election ballot proposals in Michigan over 16 years. The model identified 9 different latent topics across the bonds, including requests to purchase new buildings, renovations, and athletic facilities. Interestingly, equipment purchases appear to be a distinct category of bond proposal topics. We then examined the independent effect of the bond topics on the probability of passing the bond and voter turnout using modeling techniques and control variables from the recent literature. Bonds that focused exclusively on athletic facilities were 4.35 times less likely to pass than bonds that request new construction or omnibus ballot proposals. This work extends previous research to show that capital facility bond proposals that pass the most often include all facility needs in a single ballot proposal, are the first attempt at the polls, are at the top of the ballot, and request lower amounts of spending.

¹ This document is a pre-print of this manuscript, published in the *Journal of Education Finance*. Citation: Bowers, A.J., Chen, J. (2015) Ask and Ye Shall Receive? Automated Text Mining of Michigan Capital Facility Finance Bond Election Proposals to Identify which Topics are Associated with Bond Passage and Voter Turnout. *Journal of Education Finance*, 41(2), p.164-196.
<http://www.journalofeducationfinance.com/>

² Alex J. Bowers (bowers@tc.edu); Teachers College, Columbia University; Bowers@tc.edu; 525 W. 120th Street, New York, New York 10027. ORCID: 0000-0002-5140-6428, ResearcherID: C-1557-201

Bowers & Chen (2015)

BACKGROUND AND LITERATURE REVIEW:

The purpose of this study is to bring together recent innovations in the research literature around school district capital facility finance, municipal bond elections, statistical models of conditional time-varying outcomes, and data mining algorithms for automated text mining of election ballot proposals to examine the factors that influence the probability of school districts in the state of Michigan passing or failing capital facility finance bond elections from 1998-2014. Here, we bring together an emerging technique from the text data mining literature - automated text mining which is a data mining technique that identifies latent topics from a corpus of documents - with the recent literature on predicting school district capital facility finance bond election outcomes using a logistic regression framework that controls for the conditional dependence of floating or refloating failed bonds. Using this modeling framework, we test the current theory of bond passage from the literature on both bond election passage and voter turnout, focusing on the U.S. State of Michigan, as it provides an interesting context with long-term data on each ballot proposal and election result across its school districts. Across the states in the U.S., school building construction and renovation has historically been financed through local bond elections in which school districts propose taking on debt through the municipal bond market to finance capital facility construction while paying back the debt over time through a local property tax (Bowers, 2014; Sielke, 2003; Thompson, Crampton, & Wood, 2012). Over the past two decades, the school facilities and finance research literature has indicated that US schools have pressing unmet capital facility needs, with a large and concerning proportion of schools across the US in a state of disrepair (Crampton, 2003; Crampton, Thompson, & Hagey, 2001; Sielke, 2001; The Center for Green Schools, 2013). As noted in a recent National Center for Education Statistics (NCES) report on a survey of US school facilities in 2012-2013 “53 percent of public schools needed to spend money on repairs, renovations, and modernizations to put the school’s onsite buildings in good overall condition (p.3)” (Alexander,

Lewis, & Ralph, 2014). School facility construction and maintenance is an important issue given the need for local communities to build and maintain safe and adequate facilities for students, teachers and parents (Bowers & Urick, 2011; Cheng, English, & Filardo, 2011; Picus, Marion, Calvo, & Glenn, 2005; L. W. Roberts, 2009; Uline & Tschannen-Moran, 2008; Uline, Tschannen-Moran, & Wosley, 2009; Young, Green, Roehrich-Patrick, Joseph, & Gibson, 2003). In addition, recent research has demonstrated the positive impact of investing in school facility capital infrastructure as a means to increase local property values and thus build the overall tax base (Cellini, Ferreira, & Rothstein, 2010; Silverman, 2014), as well as improving local community capacity (Chung, 2002).

As school districts usually lack the substantial near-term funds needed for capital facility construction and renovations, long term school capital facility financing is traditionally funded in the US through local school districts issuing long term bonds on the municipal bond market, to be paid back with interest to bond holders through an increase in local property taxes (Sielke, 2003). As the federal government lacks a constitutionally-defined role in local school decisions, school facility financing has traditionally been a local and state issue (Ingle, Bowers, & Davis, 2014). For the majority of states, districts must put a capital facility bond proposal up for local election and gain the approval of the majority of voters before levying additional taxes to pay back the bond holders (Thompson, et al., 2012). Thus, for district administrators, how to lead district efforts to successfully pass facility bond elections is a major area of interest (Bowers & Lee, 2013; Frantz, 2014; Godown, 2010; Hiller & Spradlin, 2010; K. Roberts, Hannga, & Womack, 2012).

While there has historically been a robust set of normative literature with recommendations for school administrators on lessons learned for “how to pass your bond” (Bauscher, 1993; Boschee & Holt, 1999; L. Davis & Tyson, 2003; Holt, 2009; Lentz, 1999; Mathison, 1998) recent research has begun to describe the complex political nature and local strategies of bond and levy campaigns in districts (Ingle, Johnson, & Petroff, 2012). Through in-depth case study analysis, the research suggests that districts that mount a successful election campaign appear to focus on building an early comprehensive strategy for the campaign, involving and informing multiple community stakeholders as to the need and purpose of the funding request, focusing on cost effective get-out-the-vote efforts of stakeholders, and creating a sense of urgency in the community about the financial needs of the district (Holt, Wendt, & Smith, 2006; Ingle, et al., 2012; Ingle, Johnson, Ryan Givens, & Rampelt, 2013; Kraus, 2009). However, all of the research in this area notes that the local context of the community is paramount,

and so lessons learned from one district case may be difficult to apply in another community.

Informed by the case study literature, there is a strong set of quantitative research that investigates the variables most associated with passing or failing a school bond election. Researchers in this domain have worked to detail the variables that are most predictive of passing or failing a bond election, examining data from across multiple states and contexts, in an effort to provide local administrators and policymakers with generalizable recommendations as well as further information on which factors may be context dependent. Much of this work originally grew out of a large body of evidence from the 1950s and 1960s, summarized well in the in-depth literature review by Piele and Hall (Piele & Hall, 1973). This literature coincided with the rise in school enrollment and construction fueled by the Baby-Boom generation, culminating with Piele and Hall (1973). However, subsequent research interest trailed off for many years after the 1970s (Bowers & Lee, 2013). More recently, with the advent of richer and more complete state-level datasets around not only election and bond data, but also community and district variables, research on the factors most associated with passing or failing school capital facility finance bond elections has reemerged with multiple recent studies across multiple state contexts. Given that in the US, states on average together annually spend over \$40 billion on new school construction, equipment and renovations, over \$50 billion in maintenance and operations, and over \$17 billion in interest on debt (Cornman, 2015; NCEF, 2010), research in this area is an important concern for not only researchers, policymakers and administrators, but also parents, students and communities as well as states overall.

As a means to bridge between the findings and theory from the 1970s and the more contemporary research, Bowers and Lee (2013) recently provided a theory of school bond passage. In the Bowers and Lee model, the authors organized the findings to date from across the literature to serve as a means to promote further research in this area to understand how to model the dynamics within a community around a school district bond election. As opposed to political science median voter models or rational choice theory (Blais, 2000; Dunne, Reed, & Wilbanks, 1997; Fort, 1988; Fort & Bunn, 1998), the Bowers and Lee (2013) theory builds off of the long history of bond election research and takes the bond as the unit of analysis, as the concern of school administrators is on the specific outcome of individual bond election proposals. In the Bowers and Lee (2013) theory, in an attempt to provide a model that is useful and informative for district administrators and state policymakers attempting to pass bond elections for needed capital facility expenditures, they argue that strong evidence from the early work reviewed in Piele and Hall (1973) to the

present indicates that there are four main sets of variables that influence bond election outcomes, 1) the history of individual bond issues in a community, 2) the specifics of a bond proposal election, 3) the context and demographics of a community, 4) and voter participation. For this model, the key issue is the extent that each set of variables may be under the influence of district administrators, providing a means to structure recommendations. Across much of this work, findings focus on a few select states that have received attention in this research domain, namely Michigan, Oklahoma, and Texas, as the bond election outcome data is accessible in each of these states, and facility finance ballot measures are a separate issue from other school finance issues such as instructional and human resource expenditures. Research from states such as Ohio and New York, in which ballot measures may contain a combination of facility or more general school finance issues, also helps to provide some insights into the model, however as ballot measures in these states propose finance issues beyond facilities that may have different community reactions, such as teacher salaries and benefits (T. E. Davis, 2015; Ingle, Johnson, & Petroff, 2011; Ingle, et al., 2013), we include the relevant findings from this literature here, but focus primarily on the research specific to school facility bond passage.

The first main issue in the Bowers and Lee (2013) theory is the history of individual bond issues in a community. Historically, given the capital facility needs of school districts, if a bond election proposal fails at the polls, the district administration will put the same bond proposal back up for election within a few years, usually substantively unchanged in ballot wording and amount (Dunne, et al., 1997; Ehrenberg, Ehrenberg, Smith, & Zhang, 2004; Sielke, 1998), termed floating and refloating the bond (Bowers & Lee, 2013; Bowers, Metzger, & Militello, 2010a). Thus, in communities, ballot measures are remembered and can be voted on multiple times when they fail, and so districts will float and subsequently refloat failed bond proposals multiple times in the hopes of gaining community approval, as they usually have few other options to finance their facility needs (Bowers, Metzger, & Militello, 2010b). Nevertheless, while this tactic is employed by administrators quite regularly, the research across the domain clearly indicates that the first attempt at a bond proposal is the most likely to pass, with subsequent attempts failing at higher rates (Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b; Ehrenberg, et al., 2004; Ingle, et al., 2013). Additionally, from a statistical perspective, while the majority of studies in this domain model the probability of passing or failing a bond proposal using a logistic regression framework (Ingle, et al., 2013), the practice of districts floating and refloating a bond creates a conditional dependency issue within the model, as bonds that pass are not eligible to be refloated. To deal with this issue, the literature has turned to using a form of survival

Bowers & Chen (2015)

analysis to control for this conditional dependency, specifically discrete time hazard modeling (Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b). Nevertheless, modeling “attempts” or “floats” as a continuous variable or using a survival analysis framework demonstrates in multiple states that a district’s first attempt at a bond proposal election is the most likely to pass with subsequent attempts experiencing worse odds.

The second set of variables in the Bowers and Lee (2013) theory that are associated with passing or failing a capital facility finance bond election are local district and community characteristics, which are assumed to be constant in the community, and thus administrators must react and plan for these variables. These types of variables include district enrollment, enrollment trends, community demographics, student demographics, and urbanicity. For enrollment, overall district enrollment is either weakly or not significantly related to election outcomes across the states studied (Beckham & Maiden, 2003; Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b; Ingle, et al., 2013; Sielke, 1998; Zimmer, Buddin, Jones, & Liu, 2011; Zimmer & Jones, 2005), while enrollment growth was positively related to passing a bond in Texas (Bowers & Lee, 2013). Mixed evidence has been reported in the literature on community demographics as they are related to bond outcomes, with the percent of the population over age 65 shown to have a negative association with bond passage rates in Texas (Bowers & Lee, 2013) and Michigan in the 1990s (Zimmer & Jones, 2005) but not in Michigan from 1999-2001 (Zimmer, et al., 2011) and was positively associated with passing a school finance ballot initiative in California, Colorado and Minnesota (Shober, 2011). This corresponds to a debate in the research literature in general as to the perception of a negative influence of older voters on school finance decisions (Berkman & Plutzer, 2005; Button & Rosenbaum, 1989; Glass, 2008). In contrast to voter age, communities with larger proportions of residents with higher levels of education experience higher rates of passage in Michigan from 1999-2006 (Bowers, et al., 2010a, 2010b), New York (Ehrenberg, et al., 2004) and in Western states (Shober, 2011) but not in Michigan when the time period was focused only on 1999-2001 (Zimmer, et al., 2011) or in Texas (Bowers & Lee, 2013). Additionally, replicating the past research (Piele & Hall, 1973) student demographics, as a proxy for community demographics, have been shown to be related to bond outcomes, depending on the location. For example, increasing district student socio-economic status (SES) is positively related to bond passage in Michigan (Bowers, et al., 2010b; Sielke, 1998; Zimmer, et al., 2011; Zimmer & Jones, 2005) but was unrelated in Texas controlling for the other variables in the models (Bowers & Lee, 2013). Additionally while Zimmer et al. (2011) found no relationship between student ethnicity and bond passage, Bowers and Lee (2013) showed that in

Texas higher proportions of Asian and Hispanic students in a district were associated with a higher probability of passing a bond, all other variables being equal. And finally for community characteristics, urbanicity of the district appears to be context dependent, with rural and small town districts in Michigan having lower odds of passing a bond controlling for other variables in the models (Bowers, et al., 2010a, 2010b), while urbanicity appears to be unrelated to passage in Texas and Oklahoma (Bowers & Lee, 2013; C. Johnson & Maiden, 2010).

Throughout much of this research on capital facility finance bond elections, the state of Michigan has been studied extensively, and is the focus of the present study as well. Michigan is of interest in this research context for three main reasons of – an interesting capital facility finance context, a market choice environment, and long-term public reporting of elections and election outcomes. Michigan has been shown to have a large degree of variance in the capital facility stock across the state, especially in relation to district needs (T. E. Davis, 2015; Sielke, 1998). As has been noted in the past literature, facility construction is influenced by Michigan's market choice policies in which the state sets school and district-level funding formulas per student for curriculum and instruction expenditures, such that a student is able to move from district to district and take their state foundational formula funding with them (Arsen, Clay, Devaney, & Fulcher-Dawson, 2005; Arsen, Plank, & Sykes, 1999). In this market-choice environment, wealthy suburban districts with a large property tax base at times build what have been called "Taj Mahal" high schools (Arsen & Davis, 2006) that are large and attractive facilities with a wealth of services that serve to entice mobile students from urban centers to attend the suburban schools and bring their additional state foundation funding with them (Militello, Metzger, & Bowers, 2008), helping to offset funding issues in a state with decades-long school budget crises.

A third issue in the Bowers and Lee (2013) model is voter participation. Across most studies in bond elections, voter participation is operationalized as voter turnout, and the long-term research indicates that the larger the turnout at an election, the lower the probability of passing a bond (Piele & Hall, 1973). This effect is postulated to be due to a fairly constant turnout of voters with high school support attitudes, such as parents of school-aged children, but when overall voter turnout increases, such as during a presidential election, additional voters who are unaware of the school facility finance proposal, or who have low school support attitudes, make up a larger ratio of the voters at the polls which leads to lower probabilities of bond passage (Bowers, 2015; P. A. Johnson, 2015). Multiple studies to date that include voter turnout have shown a strong negative effect on election outcome across multiple contexts (Bowers, et al., Bowers & Chen (2015)

2010b; Dunne, et al., 1997; Holcombe & Kenny, 2007; Silverman, 2011). However, a small but growing body of literature has questioned this long-held belief about voter turnout (Gong & Rogers, 2014; Ingle, et al., 2013; Lentz, 1999). For example, Gong and Rogers (2014) recently examined the effects of percent voter turnout on the percent of yes votes on capital bond measures in Oklahoma, and found no evidence of a significant relationship when using an instrumental variable approach. Hence, more research is needed in this area.

The final set of variables in the Bowers and Lee (2013) model are the specifics of a bond proposal election. These include both the election itself, as well as ballot proposal characteristics, both of which Bowers and Lee (2013) argue are under the most influence by district administrators. For election characteristics, these include the location of the bond proposal on the ballot, the timing of the election and the long term debt of the district. As the first of these characteristics, one of the strongest predictors is proposal number of the ballot measure. Proposal number is an indication of the number of proposals on the ballot, and has been shown in both Michigan and Texas to be a strong negative predictor (Bowers & Lee, 2013; Bowers, et al., 2010b), with bonds with larger proposal numbers having much worse odds of passing controlling for other variables in the models. This indicates that if a bond is the first or only proposal on a ballot, it has much higher chances of passing. As one potential explanation, Bowers and Lee (2013) proposed a theory of *voter fatigue*, in which community voters are more willing to vote for a single omnibus ballot measure that includes all requests when it's at the top of the ballot, but as voters encounter other district or community issues on the ballot, such as issues pertaining to municipal infrastructure like the fire department or libraries, they are less likely to vote for the measure, controlling for what the proposal actually proposes. This "voter fatigue" theory aligns well with recent research from the political science domain (Augenblick & Nicholson, 2012; Conlin, Melnik, & Thompson, 2015) in which researchers have noted similar behavior across general elections in which voters are less likely to vote for subsequent ballot choices further down a ballot, termed voter "roll-off" (Bowler & Donovan, 1998; Bowler, Donovan, & Happ, 1992; Dubois & Feeney, 1998; Selb, 2008) or more recently "choice fatigue" (Augenblick & Nicholson, 2012).

Second in election characteristics, the timing of the election during the calendar year has received some attention since the early research in this area (Beckham & Maiden, 2003; Meredith, 2009; Piele & Hall, 1973). However, findings have been mixed, with research in Michigan indicating that elections later in the calendar year are more favorable times to propose bonds (Bowers, et al., 2010b), while there was no relationship in Texas to election timing (Bowers & Lee,

2013). Finally for the issue of election characteristics, the current long-term debt of the district has been shown in Michigan to be positively related to bond passage, an indication that communities that have previously been willing to tax themselves for increased school facility spending are more willing in the future to support similar proposals (Bowers, et al., 2010b; Zimmer, et al., 2011; Zimmer & Jones, 2005). This is in comparison to other indicators of community taxable value that would represent a proxy for the ability or likelihood of a community to support new taxation based on the local property values and current tax load, such as the tax rate or residential assessed value, which previous research would indicate may be related to bond issues (Ladd, 1975), but have been shown to be unrelated in the recent research (Bowers & Lee, 2013; Shober, 2011).

In addition to the issue of floating or refloating a bond proposal, there are two main aspects of the bond proposal itself which are related to bond passage, and to some extent under the influence of district administrators; the amount of the bond and the wording of the ballot proposal. While district administrators do not have full control over the amount and wording, as these are mostly guided by the needs of the district, the specifics are at the least somewhat under their discretion (Piele & Hall, 1973) and have been shown to be related to bond passage. First, research on the relationship of bond passage to the amount of the bond requested has been mixed, with both the early research (Piele & Hall, 1973) and more recent work showing that for some states, such as Oklahoma and Ohio, amount of the request was unrelated to passage (Beckham & Maiden, 2003; Ingle, et al., 2013) while there appears to be a strong relationship in Michigan and Texas (Bowers & Lee, 2013; Bowers, et al., 2010b; Zimmer, et al., 2011). For the second aspect of the characteristics of the bond proposal itself, research has examined the wording of the ballot measure as it relates to bond passage. As an example, Beckham and Maiden (2003) showed that bond proposals in Oklahoma which contained wording pertaining to technology passed more often. However, this result for technology was not replicated in Michigan (Bowers, et al., 2010b) or Texas (Bowers & Lee, 2013). Yet, for the Texas study, bonds that contained wording pertaining to renovations and refinancing had higher chances of passing, while bonds that contained wording pertaining to athletics were marginally less likely to pass (Bowers & Lee, 2013).

This issue of bond wording is of particular interest for the present study as it forms one of the central research questions described below. A critique of this past research on bond wording is that researchers have hand coded the full text of the ballot measures as either referencing a certain topic or not, such as coding for the presence of wording pertaining to technology, new construction, renovations, Bowers & Chen (2015)

refinancing, athletics, or art. However, the number of topics and extent that these topics may or may not exist in the proposal wording is an arbitrary decision by the researchers. To date, no research has empirically examined the extent that there are different topics requested across bonds, the number of topics, what the topics refer to, and the extent that these different topics are related to bond outcomes. As just one example in this area, wording pertaining to technology may not define a single topic, but in fact might relate to larger latent topics across many bond proposals, such as classroom instructional facilities or 21st century renovations. The issue is that currently there is very little known about what topics exist across bond proposals and the number of topics and the relationship of potential topics to bond passage. Said another way, this issue is concerned with the question of *what districts are asking for*, and *what voters are willing to buy*.

Framework of the present study:

Thus, the framework of the present study is to expand upon this recent capital facility finance bond election research, in an effort to not only inform district administrator decision making when considering a bond proposal, but also continue to investigate the theory of bond passage reviewed above, using the context of the state of Michigan. Here, our primary focus is to delve into the issue of examining the content of capital facility bond election proposals, namely through empirically identifying the latent topics across a large set of ballot proposals, identifying the topics that are most related to not only bond passage, but also voter turnout. The purpose of the present study is to address the following research question: Controlling for other variables associated with school bond election passage or failure, to what extent are the topics that bond proposals request associated with bond passage and voter turnout?

METHOD:

Dataset and Variables

This study is a secondary data analysis of the publically accessible State of Michigan Department of Treasury School Bond Qualification and Loan Program (SBQLP) database of all qualified Michigan school bond elections (MIDT, 2014). The Michigan qualified school loan program (Bowers, et al., 2010b; State of Michigan, 2005) provides school districts in the state with the ability to use the state's traditionally high bond rating on the municipal debt market in exchange for oversight of the bond proposal through the Michigan Department of Treasury (MIDT, 2014; Militello, et al., 2008). The dataset includes the school district, date of election, *amount of the bond*, *ballot proposal number*, number of yes/no voters, *election outcome*, and the full text

of the bond proposal. Following past recommendations on studying full longitudinal state datasets that include a single large metropolitan district (Bowers & White, 2014) we excluded the single Detroit bond proposal from the dataset, as the city of Detroit represents a very different context from the rest of the state (Sugrue, 1996). Using the previous literature and theory noted above to inform our selection of variables in this study, we merged this dataset with the US Department of Education National Center for Education Statistics (NCES) Common Core of Data (CCD) (NCES, n.d.) including district context, demographic and financial variables. These variables included *urbanicity*, coded as urban, suburban, small town, or rural; district *enrollment* (in thousands), the *outstanding long term debt* at the start of the previous year (in millions), *percent free and reduced price lunch students*, and percent *African American*, *Asian*, *Native American*, and *Hispanic* students. We also included the calculation of *average annual enrollment change* for each year for each district. Following the recommendations of previous studies in this domain (Bowers & Lee, 2013; Gong & Rogers, 2014; Ingle, et al., 2013) we relied on the US Census of 2000 data included in the NCES CCD for *percent population age 65 or over*, and *percent population with a college degree*, and calculated *voter turnout* by dividing the total number of recorded votes cast from the Michigan dataset by the US Census estimate of total population in the district age 18 or older. Additionally, we coded each bond as either a *first float*, or a *second float*, or *third float* if the district had a bond election fail and put a substantively similar bond back up for election within three years of the previous failed election. We examined all bonds proposed from 1998-2014 with complete data, n=1,210 bond proposals. This 16 year dataset represents one of the most comprehensive long-term capital facility finance bond datasets analyzed in the literature to date.

Analysis:

We conducted the analysis for this study in two phases. In the first phase, to identify the topics requested across the full text of the bond proposal ballot measures, we used recent innovations in latent topic modeling from the data and text mining literature. While automated text mining is a relatively new field in education data mining and data science (Baker & Yacef, 2009; Schutt & O'Neil, 2013) and has not received much attention to date in the education finance literature, we selected this technique as it is superior to the past research reviewed above in which ballots have been hand coded for inclusion of certain words or not that the author deems important. Unsupervised automated text mining allows us for the first time in this domain to empirically identify the number and content of latent topics across all bond ballot proposals in a state that may either be proposed in the literature to date or unknown.

Here, documents are considered to have a finite set of latent topics defined through the correlation of words across topics and documents, in which a form of unsupervised Bayesian matrix factorization is applied to a sparse matrix of documents by words (Grün & Hornik, 2011; Steyvers & Griffiths, 2007). For an in-depth review of this technique, please refer to Grün and Hornik (2011). Here, briefly, this technique is known as a “bag-of-words” model as each document (here, the full text of each bond ballot proposal) is broken down into the set of unique words in the document, which then forms the sparse document by words matrix. Application of matrix factorization then results in a matrix of topics by words and documents by topics in which each cell of the documents by topics matrix contains the probability of a document containing the latent topic identified (Steyvers & Griffiths, 2007). Here, following the recommendations of the methodological literature in this area, we used a Correlated Topic Model (CTM) (Blei, 2012; Blei & Lafferty, 2007, 2009; Ponweiser, 2012) in the statistical package R (R Development Core Team, 2014) and the R statistical software package RTextTools (Collingwood & Jurka, 2011; Jurka, Collingwood, Boydston, Grossman, & van Atteveldt, 2012) to automatically identify latent topics as features across the bond proposals, using 10-fold cross validation (Arlot & Celisse, 2010; Breiman & Spector, 1992; Kuhn, 2008) to identify the total number of topics across the bonds. We then categorized each bond by its most likely topic 0/1, assigning each bond uniquely to the mutually exclusive latent topic categories. We then used these topic categories as independent variables in the subsequent models.

In the second phase of the analysis, following the recommendations of the previous research on bond election outcomes (Beckham & Maiden, 2003; Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b; Ingle, et al., 2013; Zimmer, et al., 2011; Zimmer & Jones, 2005), we modeled the probability of passing a capital facility finance bond in Michigan over this time period as the dependent variable in a logistic regression framework controlling for the other variables noted above. We controlled for the conditional dependency of districts refloating failed bonds by using a modified form of discrete time hazard modeling (Singer & Willett, 2003), as is recommended (Bowers & Lee, 2013; Bowers, et al., 2010b). Additionally we analyzed a second set of models with voter turnout as the dependent variable, using the same variables and an OLS regression framework.

RESULTS:

The purpose of this study is to identify the multiple different topics requested for funding across all capital facility finance bond proposals in the state of Michigan from 1998-2014 and examine the extent to which these topics are related to bond election passage and voter turnout, controlling for other factors known to be associated with bond passage. As the first application of automated text mining in this domain, we first present our findings from the latent topic model, describing nine topics identified across the 1,210 bond proposals, and detail the longitudinal change in topics across the state as a means to understand what school districts in Michigan were requesting in these facility finance bonds over this time period. We then use the identified topics as independent variables to examine the relationship between topics and bond passage as well as any effect on voter turnout. We then turn to a discussion of the findings and the potential usefulness of text mining and latent topic models in education finance research.

A Correlated Latent Topic Model of Michigan Bond Proposals

As described in the methods, we applied an unsupervised Correlated Topic Model (CTM) to the document term matrix of all $n=1,210$ full text bond ballot proposals, identifying 1,298 unique terms across the bonds. Following the recommendations for probabilistic latent text mining methods (Grün & Hornik, 2011; Steyvers & Griffiths, 2007) we first calculated an optimal $tf-idf = 0.05$, which is the product of the overall term frequency (tf) by the inverse document frequency (idf), allowing us to exclude from the analysis both high frequency words such as “and”, “or”, “the” as well as extremely low frequency words such as the names of individual school buildings. We then applied a 10-fold cross validation and examined the topic by perplexity plots, identifying $k=9$ as the optimal solution for the number of latent topics, as this model had the lowest average perplexity out of a full set of models from $k=2$ to $k=15$.

For the first time in this domain, Table 1 provides the term lists from the $k=9$ correlated latent topic model, listing the first ten terms for each of the nine latent topics identified, the number of bond proposals classified into each topic, and our description of each of the nine latent topics, given the terms and the bonds they relate to. Overall, across this 16 year time period, the topic referred to the most across the bonds is Topic 7, Debt Issue Requirements, with 30.2% of the proposals classified as this topic, including terms such as “bond”, “valuation”, “per”, “rate”, and “expect”. The second most frequent topic was Topic 6, Facility Equipment, with 22.6% of the bond proposals, which included terms such as “facilities”, “acquiring”, “refurnishing”, and “equipping”. Topic 5, Non-Athletics

Outdoor, was the next most frequent topic at 15.6% of the bonds, with terms related to non-athletic specific improvements and outdoor renovations such as “middle”, “exist”, “outdoor”, “physic*”, “construct”, and “playground”. In comparison, Topic 8, Athletics, which only included 5.5% of the bonds, included terms such as “field”, “construct”, “track” and “gymnasium”. Surprisingly, given the past literature on new school construction, Topic 1, New School Construction, made up only 11.2% of the bonds with terms such as “buildings”, “build”, “purchase”, “facility”, and “developing”. Additionally, while the model identified Topic 9, Non-instructional, as a topic across the bonds, no bonds were classified as Topic 9 as the probability of most likely topic did not exceed any of the other eight topic probabilities for any bond. Of note across the topics identified, terms used in the prior literature to classify bonds, such as “technology”, are included in multiple term lists across the topics. This is expected in these types of model outcomes (Grün & Hornik, 2011) as the full term list is used to build a story about the latent topic included across the documents, rather than depending on strict frequency counts for individual terms, which we argue here, is an improvement over past methods for identifying topics across bond proposals.

Overall across the topics, as the first test of an automated text mining algorithm in this domain, the latent topic model performed well, identifying useful topics across the bonds, many of which have been proposed in the past literature (Bowers & Lee, 2013; Piele & Hall, 1973), while providing the first empirical evidence of the proportions of the different requested bond topics across a state over a long period of time. For example, the terms most associated with Topic 8, Athletics, as noted in Table 1, included terms such as “field, construct, track, gymnasium, center, football, stadium, and run”. The full text of the bond proposals identified as this Athletics Topic 8 by the model included proposals with wording such as the following: “*Erecting, furnishing and equipping a multipurpose gymnasium and fitness center addition to the high school; erecting, furnishing and equipping a new support facility for the athletic fields; and preparing, developing, and improving sites*”. A second example of the Athletics topic (Topic 8) includes “*Constructing and equipping a new track at the high school and developing and improving the site.*” In comparison, Topic 5 bond proposals referred to Non-Athletic Outdoor which included terms such as “outdoor, physic*, construct, and playground”. An example of the full text of a Topic 5 Non-Athletic Outdoor bond proposal with wording referring to non-specific athletic and outdoor issues was “*Erecting, furnishing and equipping additions to and remodeling the school building; and developing and improving playgrounds and the site.*”

Table 1: Correlated Latent Topic Model Outcome, with First Ten Terms and Proportion of Bonds Classified to Each Topic

Topic	Topic 1	Topic 2	Topic 3
<i>Term list</i>	Buildings Build Building Purchas* Buses Facil* Preparing System Reequipping Developing	Bond Per Value Build Equal Which First Pay Preparing Proposit*	District Improvements Include Relat* Remodel Middl* Replac* Transport* Properti* Together
<i>Proportion of Bonds</i>	<i>n=135 (11.2%)</i>	<i>n=35 (2.9%)</i>	<i>n=49 (4.0%)</i>
<i>Description</i>	<i>New School Construction</i>	<i>Non-specific</i>	<i>Renovations</i>
Topic	Topic 4	Topic 5	Topic 6
<i>Term list</i>	District Include Buildings Bond Exist Issu* Per Value Furnishing Construct	Middle Exist Outdoor Physic* Technology Construct Playground Purposes Bus Schools	Facilities Acquiring Refurnishing Constructing Equipping Buses Play Playgrounds Instruct Field
<i>Proportion of Bonds</i>	<i>n=96 (8.0%)</i>	<i>n=189 (15.6%)</i>	<i>n=274 (22.6%)</i>
<i>Description</i>	<i>Multi-Topic</i>	<i>Non-Athletics Outdoor</i>	<i>Facility Equipment</i>
Topic	Topic 7	Topic 8	Topic 9
<i>Term list</i>	Bond Valuation Per Rate Expect Issu* Exist Period Technology Increas*	Field Construct Include Track Gymnasium Center Football Stadium Run Junior	Improvements Buildings Playgrounds Field Facilities Auditorium Building Includ* Fields Playground
<i>Proportion of Bonds</i>	<i>n=366 (30.2%)</i>	<i>n=66 (5.5%)</i>	<i>n=0</i>
<i>Description</i>	<i>Debt Issue Requirements</i>	<i>Athletics</i>	<i>Non-instructional</i>

The topic model identified Renovations (Topic 3) as a separate topic across the bond proposals from Facility Equipment (Topic 6), which we report here for the first time in this literature. As noted in Table 1, the Renovations topic (Topic 3) included terms such as “improvements, remodel, and replac*”. An example of the full text of a Renovations topic bond proposal is “*Remodeling and replacing the Elementary school heating and ventilating systems, including related electrical and roofing modifications*”. A second example of a Renovations (Topic 3) bond proposal: “*Remodeling, furnishing, refurbishing, equipping and reequipping School District buildings, including boiler replacement, window and door replacement, energy conservation improvements, lighting and site improvements*”. The Renovations topic proposals are in comparison to the Facility Equipment (Topic 6) bond proposals, which included terms such as “facilities, acquiring, refurbishing, equipping, and buses” which include example ballot language such as “*Purchasing school buses*” and “*Remodeling, equipping and re-equipping the school building, in part for security purposes; and purchasing school buses*” as well as “*Acquiring, installing and equipping instructional technology for school facilities; partially remodeling, furnishing and refurbishing, equipping and re-equipping school facilities, in part, for technology and security; purchasing school buses; and developing and improving sites*”. All bond proposals that exclusively proposed to purchase school buses were categorized as Topic 6.

Interestingly, Topic 7, Debt Issue Requirements, was an unexpected outcome of the latent topic model. As noted in Table 1, this topic contains multiple words pertaining to the specifics of the bond issue such as “bond, valuation, per, rate, expect, issu*, and period”. These terms correspond to text in a large percentage of the bond proposals such as the following:

The maximum number of years the bonds may be outstanding, exclusive of any refunding, will not exceed thirty (30) years. The estimated millage that will be levied for the proposed bonds in the year 2001 is 2.01 mills (\$2.01 per \$1,000 of taxable valuation) for a total year 2001 tax rate of 7.20 mills (\$7.20 per \$1,000 of taxable valuation) which is a 1.70 mill (\$1.70 per \$1,000 of taxable valuation) increase from the year 2000 tax rate. The estimated simple average annual millage anticipated to be required to retire this bond debt is 3.07 mills (\$3.07 per \$1,000 of taxable valuation).

This latent topic of topic 7 is thus the terms and conditions of the bond issuance. From a text mining perspective, these

high frequency terms could be candidates for exclusion from the model. However, in examining the latent topic model solution over the 16 years of data, we found that this Debt Issue Requirements topic, Topic 7, occurred only during years 2000 through 2005. Interestingly, this time period corresponds precisely with the two significant changes in Michigan law on school district capital facility finance ballot proposals of these 16 years. In 2000, the state of Michigan passed Michigan Public Act 290 the School Qualification and Loan Act (State of Michigan, 2000) which stipulated inclusion of precise language pertaining to the bond valuation, issuance, debt repayment and local taxation changes in all school bond ballot proposals. In 2005, the state passed the Michigan School Bond Qualification, Approval and Loan Act (State of Michigan, 2005) which provided a greater role for the state department of the treasury to review and preapprove the bond issuance specifics, and thus this language was no longer required to be present next to the wording for the purpose of the funding request on the ballot.

Of interest in the present study, as one of the first applications of text mining to education finance issues, given that the latent topic model was unsupervised, and thus naïve to these structural changes in ballot language stipulated by the changes in the Michigan state law, this Topic 7 was an unexpected finding. Yet, as one of the first attempts to use text mining to explore the possible topics in bond proposals, the point that the latent topic model automatically detected a legal change over time using only the bond proposal language, is an intriguing finding that we believe has multiple applications for future research as a possible means to use ballot language as a way to identify significant legal shifts by states, especially in the domain of education finance. We discuss this issue at greater length below in the discussion. Nevertheless, while Topic 7 corresponds to state ballot wording requirements over a specific time period, and thus was outside the control of district administrators, to provide an initial example study of the possible use of text mining in this domain to identify relevant latent topics of capital facility finance bond proposals and examine the association between topics requested and election outcomes, we retained the final nine topic model in the subsequent analyses, which we turn to next.

Descriptive Statistics

Table 2 provides the descriptive statistics for all variables included in the models, including the mean, standard deviation, minimum and maximum values for each variable. The final bond proposal database included 1,210 bonds over years 1998-2014. As the district and community characteristics variables were obtained through merging the

Table 2: Descriptives for Variables Included in the Models

	<i>Mean</i>	<i>Stdev</i>	<i>Min</i>	<i>Max</i>
<i>Bond Characteristics</i>				
First float	0.71	0.455	0	1
Second float	0.20	0.398	0	1
Third float	0.03	0.178	0	1
Bond amount (in millions \$)	21.93	32.645	0.11	455.40
Topic 1 <i>New Construction</i>	0.11	0.318	0	1
Topic 2 <i>Non-Specific</i>	0.03	0.167	0	1
Topic 3 <i>Renovations</i>	0.04	0.196	0	1
Topic 4 <i>Multi-topic</i>	0.08	0.270	0	1
Topic 5 <i>Non-Athletics Outdoor</i>	0.15	0.362	0	1
Topic 6 <i>Facility Equipment</i>	0.23	0.421	0	1
Topic 7 <i>Debt Issue Requirements</i>	0.30	0.458	0	1
Topic 8 <i>Athletics</i>	0.05	0.266	0	1
Sample size from 1998-2014	1,210			
<i>District and community characteristics</i>				
Urban	0.05	0.219	0	1
Suburban	0.28	0.453	0	1
Small Town	0.21	0.406	0	1
Rural	0.45	0.498	0	1
Enrollment (in thousands)	2.88	3.294	0.05	29.36
Avg annual enrollment change	-0.01	0.077	-0.28	1.77
% population age 65 or over	0.23	0.040	0.05	0.29
% population with a college degree	0.11	0.068	0.02	0.45
% Free and reduced price lunch students	0.32	0.167	0	0.87
% Native American students	0.01	0.050	0	0.57
% Asian students	0.01	0.022	0	0.26
% Hispanic students	0.04	0.051	0	0.58
% African American students	0.05	0.112	0	1.00
<i>Election characteristics</i>				
Outstanding long-term debt (in millions \$)	20.90	36.200	0	331.19
Election held after July 1	0.42	0.493	0	1
Proposition 1	0.82	0.387	0	1
Proposition 2	0.13	0.340	0	1
Proposition 3 or greater	0.05	0.219	0	1
% Voter turnout	0.21	0.095	0.01	0.69
Sample size 2000-2011	779			

Table 3: Michigan Bond Election Characteristics By Year: Number Proposed, Percent Passed, and Descriptives for the amount of Bonds Proposed per Year (in Millions of \$) including Mean, Standard Deviation, Minimum, Maximum and Sum

<i>Year</i>	<i>Number</i>	<i>% Pass</i>	<i>Mean \$</i>	<i>SD \$</i>	<i>Min \$</i>	<i>Max \$</i>	<i>Sum \$</i>
1998	108	40.74%	20.733	40.296	0.635	396.000	2239.135
1999	118	47.46%	16.373	18.177	0.305	114.715	1931.960
2000	117	48.72%	21.178	28.400	0.720	175.000	2477.835
2001	108	62.96%	21.675	40.987	0.430	388.480	2340.900
2002	85	57.65%	21.003	23.213	0.260	114.470	1785.215
2003	70	37.14%	35.350	64.503	0.590	455.400	2474.505
2004	73	61.64%	34.724	40.084	0.105	205.465	2534.850
2005	58	39.66%	24.085	33.741	1.820	210.585	1396.935
2006	59	44.07%	27.616	30.606	0.335	112.075	1629.320
2007	68	47.06%	20.864	20.418	0.435	93.000	1418.720
2008	44	56.82%	19.299	18.009	0.775	79.040	849.170
2009	49	69.39%	20.495	24.687	0.302	112.500	1004.265
2010	50	66.00%	22.368	28.730	0.960	167.665	1118.395
2011	50	46.00%	20.245	22.102	1.050	103.560	1012.240
2012	49	51.02%	12.502	16.065	0.465	88.455	612.575
2013	44	75.00%	16.421	23.578	1.185	125.000	722.505
2014	63	69.84%	14.994	19.828	0.610	89.950	944.640
Totals	1213	54.19%	21.760	29.025	0.646	178.315	1558.421

bond database with data from the NCES CCD, the final sample size for full models is $n=779$, as some variables are not reported prior to 2000, such as percent free and reduced price lunch students (prior years record just free lunch status), as that the most recent year at the time of this writing for the long-term debt finance variable was 2011.

Table 3 provides the bond election characteristics in Michigan by year from 1998 through 2014, including the total number of proposals, the pass rate, the mean amount requested, standard deviation, minimum, maximum and sum

in billions of dollars. In comparison to other states recently examined in the literature such as Texas (Bowers & Lee, 2013), Michigan has an overall low bond election passage rate, at 54% overall across these 16 years, almost a 50-50 split. For district administrators in Michigan looking to finance much needed capital facility construction, renovations and equipment purchases, Michigan provides a fairly difficult context in which to propose a bond. The years 2003 and 2005 were especially problematic, with less than 40% of the bonds passing. In contrast, 2013 was an especially good year for bond proposals, with 75% passing.

However, the overall trend in Michigan appears to be that districts are proposing fewer bonds and overall across the time period, the sum of the requests is decreasing.

Analysis to Predict Bond Passage and Voter Turnout

Table 4 presents the findings from the final models of estimating the relationship between topics identified in the latent topic model and the probability of passing a bond proposal and percent voter turnout, controlling for variables previously identified in the literature. As a means to inform the theory of bond passage described above in the review of literature, through this analysis we wished to examine the relationship between what the districts asked for, what the voters agreed to fund, and the percentage of voters who came to the polls as a proxy for voter participation and interest. Our aim with this model is to examine the independent effect of each latent topic identified from the automatic text mining on both the probability of passing a bond and voter turnout to provide actionable information for school administrators looking to pass their bonds on which aspects of bond wording are associated with these two outcomes.

Table 4 provides three models. As noted in the methods, in the first model, Model A, we used a logistic regression discrete time hazard model to estimate the probability of passing a bond in Michigan from 1998-2014, controlling for the conditional dependency of floating and refloating bonds and examining the relationship of topic to election outcome. Model A covers the full 16 year time period. In Model A, as noted in the previous literature on the practice of floating and then refloating failed bonds (Bowers & Lee, 2013; Bowers, et al., 2010b), the first attempt at a bond has the highest chances of passing in comparison to second and third or more floats. As the focus of the present study on the content of the bond proposals, Model A demonstrates that two latent topics were significantly related to bond passage in the minimal model, Non-Athletic Outdoor (Topic 5) and Athletics (Topic 8). Both topics were negative and significant on the probability of passing a bond proposal. As noted above, Topic 5 does not exclude athletic facilities, but rather the latent topic model identified Topic 5 as including not only athletics but many other different requests in this area such as playgrounds and parking lots, whereas Topic 8, Athletics, identified bond proposals that were primarily concerned with athletic facilities, such as stadiums, gymnasiums and tracks. Past research in this domain that has relied on anecdotal evidence from district administrators as well as hand coding bond measures for the inclusion of athletics (Bowers & Lee, 2013; Piele & Hall, 1973) suggests that exclusive requests for athletic facilities are generally not favored by voters at the polls. Here in this study for the first time in this domain we confirm this result with the first

empirical evidence using a latent topic model to automatically identify topics from ballot language that are significant predictors of bond passage.

The second model in Table 4, Model B (Table 4, center column), provides the results from our full final model estimating the probability of passing a bond in Michigan, from 2000-2011. Following the recommendations of the past literature (Bowers & Lee, 2013; Ingle, et al., 2013), we separate the variables tested into sections that relate to the overall theory of bond passage reviewed above. In the first section of Model B, replicating the past literature, first float is a significant positive predictor of bond outcome, with first floats 2.484 times more likely to pass than refloats. Additionally, increasing bond amount is a strong negative predictor of passing a bond election. Of direct interest for the present study, in the final full model in Model B, Topic 8, Athletics, is a significant negative predictor of passing a bond election. As odds below 1.0 are difficult to interpret, inverting the odds of 0.230 suggests that bond proposals that focus on requesting funds to build athletic facilities are 4.348 times less likely to pass than bonds referring to the Topic 1 reference group, New Construction.

As noted in Model B in Table 4, few district and community characteristics were significant on bond passage, with only percentage population age 65 or over negative and significant in the model, replicating recent findings in Texas (Bowers & Lee, 2013). And finally for Model B, for election characteristics, outstanding long term debt was not significant while elections held after July 1, proposition number, and percent voter turnout were all significant, which is expected given previous findings in Michigan (Bowers, et al., 2010a, 2010b). Interestingly, controlling for the other variables in the model, bond proposal elections held after July 1 in a year were 1.768 times more likely to pass. Conversely, a bond proposal that was second on the ballot was 4.651 times less likely (1/0.215) to pass than a bond proposal that was at the top of the ballot or the only proposal on the ballot. Proposals that were third or lower had even worse chances of passing controlling for the other variables in the model, in which they were 14.493 times less likely to pass. Overall, Model B fit the data well. As variance explained equations are only able to estimate the R^2 in logistic regression (Borooah, 2002), we follow previous recommendations and report a conservative estimate (Bowers & Lee, 2013; Bowers, et al., 2010b; Ingle, et al., 2013), Cox and Snell, as well as a liberal estimate, Nagelkerke. Here, the full final Model B explains about 20% of the variance in bond election outcomes, which is comparable to previous studies in this area, especially in the Michigan case (Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b; Sielke, 1998; Zimmer, et al., 2011).

Table 4: Regression Model Parameter Estimation of Passing a Michigan School District Capital Facilities Finance Bond and Election Voter Turnout

	Model A			Model B			Model C		
	Logistic on Election Outcome			Logistic on Election Outcome			Linear on % Voter Turnout		
	Coeff.	SE	Odds	Coeff.	SE	Odds	Coeff.	SE	β
<i>Bond Characteristics</i>									
First float	0.629 ***	0.172	1.876	0.910 **	0.331	2.484	---		
Second float	0.051	0.205		0.645	0.355		0.032 ***	0.007	0.144
Third float	-0.221	0.372		0.483	0.481		0.056 ***	0.012	0.133
Bond amount (in millions \$) ^a	-0.003	0.047		-0.492 ***	0.111	0.611	0.029 ***	0.003	0.397
Topic 2 (ref. topic 1) Non-specific	-0.345	0.374		0.537	0.493		-0.023	0.016	
Topic 3 (ref. topic 1) Renovations	-0.253	0.322		-0.371	0.531		0.018	0.018	
Topic 4 (ref. topic 1) Multi-topic	-0.114	0.259		0.226	0.391		0.002	0.013	
Topic 5 (ref. topic 1) Non-Ath. Outdoor	-0.850 ***	0.206	0.427	-0.442	0.390		-0.005	0.013	
Topic 6 (ref. topic 1) Facility Equipment	-0.137	0.191		-0.080	0.326		-0.002	0.011	
Topic 7 (ref. topic 1) Debt Issue Reqs.	-0.248	0.173		0.095	0.303		-0.023 *	0.010	-0.123
Topic 8 (ref. topic 1) Athletics	-1.124 ***	0.298	0.325	-1.469 **	0.543	0.230	-0.003	0.016	
<i>District and community characteristics</i>									
Urban (ref. suburban)				0.286	0.426		-0.017	0.014	
Small Town (ref. suburban)				0.157	0.272		0.006	0.009	
Rural (ref. suburban)				-0.104	0.273		0.025 **	0.009	0.139
Enrollment (in thousands) ^a				0.192	0.187		-0.047 ***	0.006	-0.448
Avg annual enrollment change				-0.126	0.859		0.028	0.029	
% population age 65 or over ^b				-0.913 **	0.272	0.401	0.058 ***	0.010	0.194
% population with a college degree ^b				0.269	0.242		0.037 ***	0.008	0.203
% Free and reduced price lunch students ^b				-0.334	0.808		-0.079 **	0.028	-0.137
% Native American students ^b				-0.191	1.076		0.108 **	0.034	0.095
% Asian students ^b				2.536	1.857		-0.071	0.060	
% Hispanic students ^b				-0.887	0.942		0.044	0.031	
% African American students ^b				-0.790	0.741		-0.058 *	0.024	-0.019
<i>Election characteristics</i>									
Outstanding long-term debt (in millions \$)				0.003	0.003		0.001	0.001	
Election held after July 1				0.570 **	0.172	1.768	0.029 ***	0.005	0.156
Proposition 2 (ref. prop 1)				-1.537 ***	0.295	0.215	0.041 ***	0.009	0.155
Proposition 3 or greater (ref. prop 1)				-2.680 ***	0.497	0.069	0.060 ***	0.014	0.146
% Voter turnout				-2.907 **	1.026	0.055	---		
<i>Goodness-of-fit</i>									
- 2 log likelihood	1618.30			926.91					
Cox and Snell R ²	0.048			0.178					
Nagelkerke R ²	0.064			0.238					
R ²							0.376		
First year included	1998			2000			2000		
Last year included	2014			2011			2011		
n	1,210			779			779		

Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

^a indicates variable was natural log transformed

^b indicates variable was square root transformed

Model A: Discrete-time hazard model with election outcome as the dependent variable

Model B: Discrete-time hazard model with election outcome as the dependent variable

Model C: OLS Regression with percent voter turnout as the dependent variable

Additionally, given the Bowers and Lee (2013) theory of bond passage reviewed above, we wished to examine the influence of bond topics on voter turnout as a potentially interesting means to begin to capture and examine a proxy of community school support and participation. Table 4 Model C provides the results from an OLS regression in which we estimate the percent voter turnout for each bond election using the same variables as with Models A and B. Here, we show for the first time in the literature that bond topics overall appear to not be significantly related to voter turnout, controlling for the other variables in the model. Topic 7, Debt Issue Requirements, is marginally significant and negative. However we posit that this effect is most likely due to the time dependency of Topic 7 noted above which coincides with some of the worst years for bond outcomes in Michigan in this dataset. Thus, our model suggests that voters in Michigan over this time period do not come to the polls to vote on specific topics. The remaining variables replicate and extend previous literature in this area (Gong & Rogers, 2014; Piele & Hall, 1973), indicating that for Michigan, controlling for the other variables in the model, voter turnout is significantly higher for rural districts, districts with higher percentages of population age 65 or over and with a college degree, as well as districts with higher proportions of Native American students, as well as for bond elections after July 1. Voter turnout was significantly negatively related to larger district enrollments, and higher proportions of free and reduced price lunch students. Interestingly, voter turnout was positively related to higher proposition numbers, in that as there are more propositions on the ballot, more voters turn out. This finding suggests a link between the number of requests on the ballot, voter turnout, and election outcome, which has previously been postulated in the literature (Piele & Hall, 1973), but has not been modeled in this way in the past. We turn next to a discussion of the findings.

DISCUSSION:

In this study, we set out to examine the independent effect of bond proposal topics on capital facility finance bond election outcomes and voter turnout, examining all bond proposal elections in Michigan over an extended period of time as a means to inform current theories of capital facility finance bond elections. This study is a novel and innovative contribution to the literature in three ways. First, for the first time in the education finance literature, to identify bond proposal topics we used an automated text mining algorithm and identified 9 separate latent topics that varied across the bond proposals. Second, we then used these topics to examine the association between the topic of the ballot proposal and election outcome, all other things being equal in the model. Third, we did the same for voter turnout. The central finding of this study is that other than specific requests for Athletic facilities, the topic of a bond proposal

is not significantly related to the probability of passing the bond or voter turnout, controlling for the other variables in the model. Said another way, for districts looking to pass needed facility renovations, construction and improvements for their local communities, it's not so much what you ask for, but how you ask for it. In relation to voter turnout, our findings indicate that many past variables proposed as predictors in bond outcome models are significant on turnout. Interestingly, some of these effects are in opposite directions in relation to their effects on bond passage.

As noted in the previous research (Holt, 2009; Ingle, et al., 2012; Piele & Hall, 1973), qualitative studies of district voters have indicated that when going to the polls to vote for proposals, voters are concerned with questions of "if the number of students are growing, that most students would benefit, and that the dollar amount of the bond issue was reasonable" (p.49) (Kastory & Harrington, 1996). Here we replicate and extend this work into a large longitudinal state-wide sample. The results for bond passage replicate and extend the prior research in this area reviewed above (Beckham & Maiden, 2003; Bowers, 2014; Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b; Ingle, et al., 2013; Piele & Hall, 1973; Sielke, 1998; Zimmer, et al., 2011) to the large longitudinal Michigan dataset included here. We find that bonds that pass more often in Michigan are those that are the first attempt, request lower amounts of funding, are not focused on athletic facility requests, are proposed in communities with lower percentages of senior citizens, and in which the election is held after July 1, are at the top of the ballot or are the only proposal on the ballot, and have lower percentages of voter turnout.

As noted above, one of the unique contributions of this study is the use of data mining and latent topic modeling to empirically identify the total number of topics requested in the bonds across the state as well as defining the subject of the topics. While only one of the topics was a significant predictor of bond passage, Athletics with a negative effect, we argue here that as one of the first applications of latent topic modeling in this domain of education finance, this study lays the groundwork for future studies in this area. Automated text mining is a rapidly developing field (Blei, 2012; Grün & Hornik, 2011; Hofmann, 1999; Junqué de Fortuny, Martens, & Provost, 2013; Lee, Song, & Kim, 2010; Ponweiser, 2012; Steyvers & Griffiths, 2007), and the application demonstrated here is just one in a multitude of useful applications of the technique in the education finance research domain, a domain that could leverage the use of this technique to examine latent topics across a large corpus of relevant documents, from election ballots, to law, news reports, cases, budget documents, school district and state-level document analysis, and interviews and observations. Our finding here on the temporal dependency of the Debt Issue Requirements topic (Topic 7) which corresponds to

the legal changes in Michigan, provides a tantalizing example of the possibility of using the technique to examine ballot proposal wording as a means to study the impact of changes in the law and what goes before voters across the ten thousand districts in the US. Our aim in this study is to help provide an additional useful tool for document analysis which can provide an empirical means to test for the total number of topics across a lexical space and provide information on the latent subjects across the documents. As examples here of the use of this technique, as opposed to previous research on the relation of bond proposals and technology wording (Beckham & Maiden, 2003) we were able to identify that technology appears to not be a term that uniquely identifies topics in Michigan, but rather is used quite often across a broad range of topics, as show in Table 1. This perhaps helps to explain the difference in findings between the Oklahoma context and Michigan (Beckham & Maiden, 2003; Bowers, et al., 2010b). Additionally, across bond election studies, our finding that facility equipment is a unique topic that identifies over a fifth of the bonds in Michigan during this time period, has not previously been articulated as a possible unique topic in the past literature. Thus, while we found few effects of these topics in the final models, we encourage future research in this area as we believe that the topic identification strategy provides a new useful tool to researchers in this domain.

While this study was focused on the factors most associated with passing a facility finance bond as well as the factors associated with voter turnout, questions remain in the literature as to the effects of bond passage and subsequent spending on facilities on schooling and community outcomes. Historically, while the US spends billions per year on school facilities (Cornman, 2015), research on the effects of facility quality on student achievement has shown weak to non-existent effects (Bowers & Urick, 2011; Picus, et al., 2005; L. W. Roberts, 2009; M. Schneider, 2002). Thus, recent research has turned to examining the discontinuity in a community before and after a successful bond election or infusion of funds to build new schools in an attempt to model the effect of facilities on student and community outcomes. Most notably, Neilson and Zimmerman (2014) examined the New Haven Connecticut school district which worked over multiple years to demolish and rebuild the majority of its school facilities as new facilities (Neilson & Zimmerman, 2014). As the first study of its kind, they showed a strong effect on achievement. However, as a classic before and after study (B. Schneider, Carnoy, Kilpatrick, Schmidt, & Shavelson, 2007), the counterfactual is problematic as the authors were unable to control for a placebo or Hawthorn effect, and so the effects could be attributable not to the new facilities, but that the community felt that they were being paid attention to. In contrast, recent research that has examined effects of bond passage on student achievement in comparison to *Bowers & Chen (2015)*

communities that do not pass similar bonds has shown very weak to non-existent effects (Hong & Zimmer, 2014; Martorell, McFarlin, & Stange, 2015). These findings are in-line with the previous facilities research which used direct effects models, indicating that beyond basic shelter, heating and lighting needs, school facilities have little effect on student achievement (Bowers & Urick, 2011). Nevertheless, other recent studies have suggested that facility spending may be associated with other school and community outcomes, such as reducing overcrowding in urban schools (Welsh, Coghlan, Fuller, & Dauter, 2012) and increasing property values with a subsequent rise in the property tax base (Cellini, et al., 2010). Given the apparent increasing interest in this field, we encourage future research in this area, working to link the larger theory of school facility bond passage to community support, voter preference, improved school facilities, and ultimately increased school and community outcomes (Bowers & Lee, 2013).

In replication of the past literature on district bond elections (Bowers & Lee, 2013; Bowers, et al., 2010a, 2010b), our findings suggest strong negative effects of placing a ballot initiative farther down a ballot, as measured by proposal number. Here we found that bonds that were the first or only proposal on the ballot fared significantly better than bonds that were proposal number 2 or farther down, controlling for the other variables in the model. We attribute this finding to the “voter fatigue” theory (Bowers & Lee, 2013) (also referred to as “choice fatigue” and “roll-off” theory) from the previous literature noted in the literature review above (Augenblick & Nicholson, 2012; Bowler & Donovan, 1998; Bowler, et al., 1992; Conlin, et al., 2015; Dubois & Feeney, 1998; Selb, 2008). In this conception of how voters interact with ballots at the polls, the present study lends further support to the strong findings indicating that as voters are met with more and more choices to vote on, they either abstain or vote no on the choices that are further down the ballot as they tire of making selections. Thus, for district administrators, placement of the proposal on the ballot is an important consideration, as the research across the bond research, and more general election research, shows that the theory of voter fatigue is real. This is especially important given the results here, which shows a strong negative effect of a bond being further down the ballot, regardless of the topic the bond requests, the amount of the bond and all other variables in the model. Thus, as noted above, one of the central findings of the present study for administrators to consider when attempting to obtain needed funding to build safe and adequate facilities for the students of their community, it is not so much what they ask for, but how they ask for it.

In addition, while bond topic generally did not appear to be associated with either bond passage or voter turnout, the

number of propositions was strongly related to both bond passage and voter turnout, such that for ballots with larger proposition numbers, voters turned out in higher proportions. Thus, the present study provides strong evidence of a negative association between voter turnout and bond passage. This replicates a long history of research that has indicated that higher voter turnout usually leads to worse outcomes at the polls for school district finance issues (Bowers, 2014, 2015; Bowers, et al., 2010b; P. A. Johnson, 2015; Piele & Hall, 1973). Conversely, Gong and Rogers (2014) have recently challenged this view, in which they show little association between voter turnout and bond passage for elections in Oklahoma. However, they only considered the continuous variable percent voting yes on a ballot proposal as the outcome in their models, rather than the probability of passing or failing a bond. While the two are related, from a district's perspective as well as the probability of a bond passing, there is no difference between a bond barely passing by a required majority or receiving near 100% of the vote. In addition, Gong and Rogers (2014) examined voter turnout only as an independent predictor variable, rather than model it as an outcome variable. Thus, voter turnout is an active area of research, and we encourage further research in this area.

And finally, as one of the first times in the school facilities bond election literature in which voter turnout is modeled as an outcome variable, our findings indicate that voter turnout appears to be associated with many of the variables traditionally nominated as significant variables in the bond passage literature. We found that voters turn out more often for bonds that are refloats, are further down the ballot, and request larger sums of money. Our finding that voters turn out more often for elections held after July 1 is a good check on the model, as voters should turn out more often at the polls for national presidential and congressional elections held in November, independent of the other variables in the model. Additionally, as one of the first times reported in the literature, we show that voter turnout for school facility elections is strongly associated to the demographics and context of a community. Rural communities appear to have higher turnout, as do communities with higher proportions of senior citizens, and residents with a college degree. However, the strongest negative effect on voter turnout appeared to be school enrollment, while community poverty was also a strong negative predictor. These findings contribute important information for the overall Bowers and Lee (2013) theory of bond passage reviewed above in the literature review. Throughout this theory of bond passage, authors have noted that proxies for overall community school support as well as community and individual voter bond preference are difficult variables to acquire but should be sought out and modeled to help further explicate a full model of bond passage. In the present study, we set out to understand the extent that specific bond topics may motivate

Bowers & Chen (2015)

voters to turn out at the polls. While we found no major associations between topic and voter turnout, we posit that modeling voter turnout using the same variables used previously for bond passage provides an additional means to further understand community dynamics as they relate to school district bond passage.

Limitations:

While we argue here that our findings are robust, our study was limited in the following ways. First, for the latent topic model, as with many data mining models (Blei, 2012; Bowers & Sprott, 2012a, 2012b; Steyvers & Griffiths, 2007) the number of latent topics is derived *a priori* empirically through iterative testing, selecting the appropriate term frequency cut-off values in an effort to identify the most relevant number of latent topics while keeping the most documents within the dataset. In the present study, we examined multiple topic models before running the cross-validation procedure and selected the $k=9$ model discussed above based on its ability to retain the most bonds. Selecting a higher cutoff on the term frequency resulted in losing a large fraction of the bonds (data not shown). As one of the first attempts of latent topic modeling in this domain, we chose to use the more conservative model as it appeared to not overfit the data, and while this most likely resulted in the model with some overlapping topics with only a few or no bonds categorized as those topics, such as here with Topic 4 "Multi-topic" or Topic 9 "Non-Instructional" (see Table 1), a more restrictive model would have excluded many bonds and we wished to be as inclusive with this first study as possible. Second, latent topic modeling traditionally is carried out on documents with much more lexical complexity, such as the full text of research journal articles or newspaper stories (Blei & Lafferty, 2007; Grün & Hornik, 2011). For the present study, Michigan bond election proposals are a fairly constrained lexical space. However, despite this issue, we identified strong latent topics that not only mirrored some of the previously nominated topics from researchers who only used hand coding, such as the "Athletics" topic, but we also identified topics which have not received much attention in the past literature, such as "Facility Equipment", as separate and distinct latent topics. We encourage more research in this area. Third, in this study we chose to use a correlated topic model, given its success with similar types of studies in other domains (Blei & Lafferty, 2007; Lee, et al., 2010). Automatic text data mining is a rapidly developing field of research, and so we encourage future work in this area, as recent developments in the field may model election proposal text in better ways. We look forward to future work in this area.

Suggested Citation Format:

Bowers, A.J., Chen, J. (2015) Ask and Ye Shall Receive? Automated Text Mining of Michigan Capital Facility Finance Bond Election Proposals to Identify which Topics are Associated with Bond Passage and Voter Turnout. *Journal of Education Finance*, 41(2), p.164-196. <http://www.journalofeducationfinance.com/>

REFERENCES:

- Alexander, D., Lewis, L., & Ralph, J. (2014). Condition of America's Public School Facilities: 2012-13 (NCES 2014-022). Washington, DC: U.S. Department of Education, Washington DC: National Center for Education Statistics.
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistical Surveys*, 4, 40-79.
- Arsen, D., Clay, T., Devaney, T., & Fulcher-Dawson, R. (2005). Adequacy, equity, and capital spending in Michigan schools: The unfinished business of Proposal A. East Lansing, MI: Michigan State University, Education Policy Center.
- Arsen, D., & Davis, T. (2006). Taj Mahals or decaying shacks: Patterns in local school capital stock and unmet capital need. *Peabody Journal of Education*, 81(4), 1-22.
- Arsen, D., Plank, D. N., & Sykes, G. (1999). School choice politics in Michigan: The rules matter. East Lansing: Michigan State University.
- Augenblick, N., & Nicholson, S. (2012). *Ballot Position, Choice Fatigue, and Voter Behavior*. University of California, Berkeley. Berkeley, CA. Retrieved from http://faculty.haas.berkeley.edu/ned/Choice_Fatigue.pdf
- Baker, R. S., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1(1), 3-16.
- Bauscher, R. H. (1993). Winning school bonds at the polls. *School and College*, January, 15-16.
- Beckham, J., & Maiden, J. (2003). The effects of technology inclusion on school bond election success in Oklahoma. *Journal of Education Finance*, 28(4), 557-574.
- Berkman, M. B., & Plutzer, E. (2005). *Ten thousand democracies: Politics and public opinion in America's school districts*. Washington, DC: Georgetown University Press.
- Blais, A. (2000). *To vote or not to vote? The merits and limits of rational choice theory*. Pittsburgh, PA: University of Pittsburgh Press.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Blei, D. M., & Lafferty, J. D. (2007). A Correlated Topic Model of Science. *The Annals of Applied Statistics*, 1(1), 17-35. doi: 10.2307/4537420
- Blei, D. M., & Lafferty, J. D. (2009). Topic models. *Text mining: classification, clustering, and applications*, 10, 71-86.
- Bowers & Chen (2015)
- Borooah, V. K. (2002). *Logit and probit: Ordered and multinomial models*. Thousand Oaks, CA: Sage Publications.
- Boschee, F., & Holt, C. R. (1999). *School bond success: A strategy for building America's schools*. Lancaster, PA: Technomic Publishing.
- Bowers, A. J. (2014). Bonds in School Financing. In D. J. Brewer & L. O. Picus (Eds.), *Encyclopedia of Education Economics and Finance*. Thousand Oaks, CA: Sage Publications.
- Bowers, A. J. (2015). Voter's Role in Deciding School Facility Issues. *UCEA Review*, 56(1), 20-21.
- Bowers, A. J., & Lee, J. (2013). Carried or Defeated? Examining the Factors Associated With Passing School District Bond Elections in Texas, 1997-2009. *Educational Administration Quarterly*, 49(5), 732-767. doi: 10.1177/0013161x13486278
- Bowers, A. J., Metzger, S. A., & Militello, M. (2010a). Knowing the odds: Parameters that predict passing or failing school district bonds. *Educational Policy*, 24(2), 398-420. doi: 10.1177/0895904808330169
- Bowers, A. J., Metzger, S. A., & Militello, M. (2010b). Knowing what matters: An expanded study of school bond elections in Michigan, 1998-2006. *Journal of Education Finance*, 35(4), 374-396. doi: 10.1353/jef.0.0024
- Bowers, A. J., & Sprott, R. (2012a). Examining the Multiple Trajectories Associated with Dropping Out of High School: A Growth Mixture Model Analysis. *Journal of Educational Research*, 105(3), 176-195. doi: 10.1080/00220671.2011.552075
- Bowers, A. J., & Sprott, R. (2012b). Why tenth graders fail to finish high school: A dropout typology latent class analysis. *Journal of Education for Students Placed at Risk*, 17(3), 129-148. doi: 10.1080/10824669.2012.692071
- Bowers, A. J., & Urick, A. (2011). Does high school facility quality affect student achievement? A 2-level hierarchical linear model. *Journal of Education Finance*, 37(1), 72-94.
- Bowers, A. J., & White, B. R. (2014). Do Principal Preparation and Teacher Qualifications Influence Different Types of School Growth Trajectories in Illinois? A Growth Mixture Model Analysis. *Journal of Educational Administration*, 52(5), 705-736. doi: 10.1108/JEA-12-2012-0134
- Bowler, S., & Donovan, T. (1998). *Demanding choices: Opinion, voting and direct democracy*. Ann Arbor, MI: University of Michigan Press.
- Bowler, S., Donovan, T., & Happ, T. (1992). Ballot Propositions and Information Costs: Direct Democracy and the Fatigued Voter. *The Western Political Quarterly*, 45(2), 559-568. doi: 10.2307/448727
- Breiman, L., & Spector, P. (1992). Submodel Selection and Evaluation in Regression. The X-Random Case.

- International Statistical Review / Revue Internationale de Statistique*, 60(3), 291-319. doi: 10.2307/1403680
- Button, J. W., & Rosenbaum, W. A. (1989). Seeing gray: School bond issues and the aging in Florida. *Research on Aging*, 11(2), 158-173. doi: 10.1177/0164027589112002
- Cellini, S. R., Ferreira, F., & Rothstein, J. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. *The Quarterly Journal of Economics*, 125(1), 215-261. doi: 10.1162/qjec.2010.125.1.215
- Cheng, G., English, S., & Filardo, M. (2011). Facilities: Fairness and Effects: Evidence and recommendations concerning the impact of school facilities on civil rights and student achievement. Washington, DC: 21st Century School Fund.
- Chung, C. (2002). Using Public Schools as Community-Development Tools: Strategies for Community-Based Developers. Cambridge, MA: Jount Center for Housing Studies, Harvard University.
- Collingwood, L., & Jurka, T. P. (2011). Getting Started with Latent Dirichlet Allocation using RTextTools + topicmodels. Retrieved from <http://www.rtexttools.com/1/post/2011/08/getting-started-with-latent-dirichlet-allocation-using-rtexttools-topicmodels.html>
- Conlin, M., Melnik, W., & Thompson, P. (2015). *Ballot Order, Ballot Roll-Off and Home Sales: Evidence from Local Referenda in Ohio*. Paper presented at the Annual meeting of the Association of Education Finance and Policy (AEFP), Washington, DC. <https://aefpweb.org/sites/default/files/webform/aefp40/aefp2015conlin.pdf>
- Cornman, S. Q. (2015). Revenues and Expenditures for Public Elementary and Secondary Education: School Year 2011–12 (Fiscal Year 2012) (NCES 2014-301). Washington, DC: National Center for Education Statistics, U.S. Department of Education.
- Crampton, F. E. (2003). Unmet school infrastructure funding need as a critical educational capacity issue. In F. E. Crampton & D. C. Thompson (Eds.), *Saving America's school infrastructure*: (pp. 3-26). Greenwich, CT: Information Age Publishing.
- Crampton, F. E., Thompson, D. C., & Hagey, J. M. (2001). Creating and sustaining school capacity in the twenty-first century: funding a physical environment conducive to student learning. *Journal of Education Finance*, 27(2), 633-652.
- Davis, L., & Tyson, B. (2003). Strategic bonding. *American School Board Journal*, 190(7), 34-36.
- Davis, T. E. (2015). State and Federal Policies for School Facility Construction: A Comparison of Michigan and Ohio. *Educational Administration Quarterly*, 51(1), 3-26. doi: 10.1177/0013161x13508773
- Dubois, P. L., & Feeney, F. (1998). *Lawmaking by initiative: issues, opinions, and comparisons*. New York: Agathon Press.
- Dunne, S., Reed, W. R., & Wilbanks, J. (1997). Endogenizing the median voter: Public choice goes to school. *Public Choice*, 93, 99-118.
- Ehrenberg, R. G., Ehrenberg, R. A., Smith, C. L., & Zhang, L. (2004). Why do school district budget referenda fail? *Educational Evaluation and Policy Analysis*, 26(2), 111-125. doi: 10.3102/01623737026002111
- Fort, R. (1988). The median voter, setters, and non-repeated construction bond issues. *Public Choice*, 56, 213-231.
- Fort, R., & Bunn, D. N. (1998). Whether one votes and how one votes. *Public Choice*, 95, 51-62. doi: 10.1023/A:1004934330814
- Frantz, J. T. (2014). The Bricks and Mortar Superintendency: How to Survive a Building Project. *Education*, 134(3), 384-394.
- Glass, G. V. (2008). *Fertilizers, pills and magnetic strips: The fate of public education in American*. Charlotte, NC: Information Age Publishing.
- Godown, M. P. (2010). *Factors Influencing the Successful Passage of a School Bond Referendum as Identified by New Jersey School Superintendents*. Ed.D, Widener University, Chester, PA.
- Gong, H., & Rogers, C. L. (2014). Does Voter Turnout Influence School Bond Elections? [Article]. *Southern Economic Journal*, 81(1), 247-262. doi: 10.4284/0038-4038-2013.061
- Grün, B., & Hornik, K. (2011). topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software*, 40(13), 1-30.
- Hiller, S. C., & Spradlin, T. E. (2010). School Referenda in Indiana (Vol. Volume 8, Number 2). Bloomington, IN: Center for Evaluation and Education Policy, Indiana University.
- Hofmann, T. (1999). *Probabilistic latent semantic indexing*. Paper presented at the Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, Berkeley, California, USA.
- Holcombe, R., & Kenny, L. (2007). Evidence on voter preferences from unrestricted choice referendums. *Public Choice*, 131(1-2), 197-215. doi: 10.1007/s11127-006-9113-1
- Holt, C. R. (2009). *School bond success: A strategy for building America's Schools* (3rd ed.): Rowman & Littlefield Education.
- Holt, C. R., Wendt, M. A., & Smith, R. M. (2006). School bond success: An exploratory case study. *Rural Educator*, 27(2), 11-18.
- Hong, K., & Zimmer, R. W. (2014). *Does Investing in School Capital Infrastructure Improve Student Achievement?* Paper presented at the Annual meeting of the Association for Education Finance and Policy

- (AEFP), San Antonio, TX. https://my.vanderbilt.edu/kaihong/files/2014/03/Hong_Zimmer_Capital_AEFP_2014.pdf
- Ingle, W. K., Bowers, A. J., & Davis, T. E. (2014). Which School Districts Qualified for Federal School Facility Funding Under ARRA? Evidence from Ohio. *Journal of Education Finance*, 40(1), 17-37.
- Ingle, W. K., Johnson, P. A., & Petroff, R. A. (2011). Estimating resource costs of levy campaigns in five Ohio school districts. *Journal of Education Finance*, 37(1), 52-71.
- Ingle, W. K., Johnson, P. A., & Petroff, R. A. (2012). "Hired guns" and "legitimate voices": The politics and participants of levy campaigns in five Ohio school districts. *Educational Administration Quarterly*, 48(5), 814-858. doi: 10.1177/0013161X12448251
- Ingle, W. K., Johnson, P. A., Ryan Givens, M., & Rampelt, J. (2013). Campaign Expenditures in School Levy Referenda and Their Relationship to Voter Approval: Evidence From Ohio, 2007–2010. *Leadership and Policy in Schools*, 12(1), 1-36. doi: 10.1080/15700763.2013.766347
- Johnson, C., & Maiden, J. (2010). An examination of capital outlay funding mechanisms in Oklahoma. *Journal of Education Finance*, 36(1), 1-21.
- Johnson, P. A. (2015). Voters' Role in Deciding School Levies. *UCEA Review*, 56(1), 21-23.
- Juqué de Fortuny, E., Martens, D., & Provost, F. (2013). Predictive Modeling With Big Data: Is Bigger Really Better? *Big Data*, 1(4), 215-226. doi: 10.1089/big.2013.0037
- Jurka, T. P., Collingwood, L., Boydston, A. E., Grossman, E., & van Atteveldt, W. (2012). RTextTools: Automatic Text Classification via Supervised Learning. R package version 1.3.9. Retrieved from <http://CRAN.R-project.org/package=RTextTools>
- Kastory, R. C., & Harrington, S. J. (1996). Voter perceptions are key to passing a school bond. *Educational Research Quarterly*, 20(1), 49-58.
- Kraus, B. W. (2009). *A descriptive analysis of selected community stakeholder opinions regarding potentially critical factors in school bond referenda success or failure in Kansas during the years 2004-2007*. Ed.D. Dissertation, Kansas State University, Manhattan, Kansas. Retrieved from <http://krex.k-state.edu/dspace/handle/2097/2170>
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1-26.
- Ladd, H. (1975). Local education expenditures, fiscal capacity, and the composition of the property tax base. *National Tax Journal*, 28(2), 145-158.
- Lee, S., Song, J., & Kim, Y. (2010). An empirical comparison of four text mining methods. *Journal of Computer Information Systems*, 51(1), 1-10.
- Bowers & Chen (2015)
- Lentz, C. (1999). Predicting school referenda outcomes: Answers from Illinois. *Journal of Education Finance*, 24(4), 459-482.
- Martorell, P., McFarlin, I., & Stange, K. (2015). *Investing in schools: Capital spending, facility conditions, and student achievement*. Paper presented at the Annual meeting of the Association for Education Finance and Policy (AEFP), Washington, DC. http://m.newyorkfed.org/research/education_seminar_series/Stange.pdf
- Mathison, T. R. (1998). Successful bond elections. *School Business Affairs*, 64(1), 30-33.
- Meredith, M. (2009). The strategic timing of direct democracy. *Economics & Politics*, 21(1), 159-177. doi: 10.1111/j.1468-0343.2008.00342.x
- MIDT. (2014). Qualified school bond loan program Retrieved October 20, 2014, from http://www.michigan.gov/treasury/0,1607,7-121-1753_56435---,00.html
- Militello, M., Metzger, S. A., & Bowers, A. J. (2008). The high school "space race": Implications of a school-choice market environment for a Michigan metropolitan region. *Education and Urban Society*, 41(1), 26-54. doi: 10.1177/0013124508321358
- NCEF. (2010). National Clearinghouse for Educational Facilities Retrieved Feb 24, 2010, from <http://www.edfacilities.org/ds/index.cfm>
- NCES. (n.d.). Common Core of Data, 2014, from <http://nces.ed.gov/ccd/>
- Neilson, C. A., & Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120(0), 18-31. doi: <http://dx.doi.org/10.1016/j.jpubeco.2014.08.002>
- Picus, L. O., Marion, S. F., Calvo, N., & Glenn, W. J. (2005). Understanding the relationship between student achievement and the quality of educational facilities: Evidence from Wyoming. *Peabody Journal of Education*, 80(3), 71-95.
- Piele, P. K., & Hall, J. S. (1973). *Budgets, bonds, and ballots: Voting behavior in school financial elections*. Lexington, Mass.: Lexington Books.
- Ponweiser, M. (2012). *Latent Dirichlet Allocation in R*. Vienna University of Business and Economics, Vienna. Retrieved from <http://epub.wu.ac.at/3558/1/main.pdf>
- R Development Core Team. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Roberts, K., Hannga, S. L., & Womack, S. T. (2012). *The First-Year Experiences of Successful Superintendents*. Plymouth, UK: Rowman & Littlefield.
- Roberts, L. W. (2009). Measuring school facility conditions: an illustration of the importance of purpose. *Journal of Educational Administration*, 47(3), 368-380.

- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). Estimating casual effects: Using experimental and observational designs (report from the Governing Board of the American Educational Research Association Grants Program). Washington, DC: American Educational Research Association.
- Schneider, M. (2002). Do school facilities affect academic outcomes? Washington, DC: National Clearinghouse for Educational Facilities.
- Schutt, R., & O'Neil, C. (2013). *Doing Data Science: Straight Talk from the Frontline*. Cambridge, MA: O'Reilly.
- Selb, P. (2008). Supersized votes: ballot length, uncertainty, and choice in direct legislation elections. *Public Choice*, 135(3-4), 319-336. doi: 10.1007/s11127-007-9265-7
- Shober, A. F. (2011). Attracting capital: Magnets, charters, and school referendum success. *Journal of School Choice*, 5(2). doi: 10.1080/15582159.2011.577676
- Sielke, C. C. (1998). Michigan school facilities, equity issues, and voter response to bond issues following finance reform. *Journal of Education Finance*, 23(3), 309-322.
- Sielke, C. C. (2001). Funding school infrastructure needs across the states. *Journal of Education Finance*, 27(2), 653-662.
- Sielke, C. C. (2003). Financing school infrastructure needs: An overview across the 50 states. In F. E. Crampton & D. C. Thompson (Eds.), *Saving America's school infrastructure* (pp. 27-52). Greenwich, CT: Information Age Publishing.
- Silverman, R. M. (2011). How unwavering is support for the local property tax?: Voting on school district budgets in New York, 2003-2010. *Journal of Education Finance*, 36(3), 294-311. doi: 10.1353/jef.2011.0003
- Silverman, R. M. (2014). Urban, Suburban, and Rural Contexts of School Districts and Neighborhood Revitalization Strategies: Rediscovering Equity in Education Policy and Urban Planning. *Leadership and Policy in Schools*, 13(1), 3-27. doi: 10.1080/15700763.2013.876051
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- School Qualification and Loan Act, Michigan Public Act 290 (2000).
- School bond qualification, approval, and loan act of 2005 § 388.1921-1939 (2005).
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. *Handbook of latent semantic analysis*, 427(7), 424-440.
- Sugrue, T. J. (1996). *The Origins of the Urban Crisis*. Princeton: Princeton University Press.
- The Center for Green Schools. (2013). State of our schools report. Washington, DC: The Center for Green Schools.
- Bowers & Chen (2015)
- Thompson, D. C., Crampton, F. E., & Wood, C. R. (2012). *Money and Schools* (5 ed.). New York, NY: Routledge.
- Uline, C. L., & Tschannen-Moran, M. (2008). The walls speak: the interplay of quality facilities, school climate, and student achievement *Journal of Educational Administration*, 46(1), 55-73.
- Uline, C. L., Tschannen-Moran, M., & Wosley, T. D. (2009). The walls still speak: the stories occupants tell. *Journal of Educational Administration*, 47(3), 400-426.
- Welsh, W., Coghlan, E., Fuller, B., & Dauter, L. (2012). New Schools, Overcrowding Relief, and Achievement Gains in Los Angeles – Strong Returns from a \$19.5 Billion Investment. Stanford, CA: School of Education, Stanford University.
- Young, E., Green, H. A., Roehrich-Patrick, L., Joseph, L., & Gibson, T. (2003). Do K-12 school facilities affect education outcomes? Staff information report. Nashville, TN: Tennessee Advisory Commission on Intergovernmental Relations.
- Zimmer, R. W., Buddin, R., Jones, J. T., & Liu, N. (2011). What types of school capital projects are voters willing to support? *Public Budgeting & Finance*, 31(1), 37-55. doi: 10.1111/j.1540-5850.2011.00973.x
- Zimmer, R. W., & Jones, J. T. (2005). Unintended consequence of centralized public school funding in Michigan education. *Southern Economic Journal*, 71(3), 534-544.