Leadership and Economic Growth: a Text Analytics Approach

Salfo Bikienga
Job Market Paper
November, 2017

Abstract
In recent years, a growing number of economists have come to recognize the importance of political leadership in promoting economic performance. However, without an agreed upon measure of leadership, formally demonstrating and testing this relationship remains elusive. This paper proposes identifying economic leadership by measuring the consistency with which leaders talk about economic issues. We employ a text analytics approach—Topic Modeling—to studying leaders’ discourses, and measure the relationship between these discourses and economic growth. Specifically, using the Latent Dirichlet Allocation (LDA) algorithm, we identified the topical content of U.S governors’ state of the state speeches from 2001 to 2013, constructed a consistency measure over these topics, and studied the relationship between the consistency of these topical content and the states’ real GDP growth. We find that the consistency with which governors address economic issues is strongly associated with economic growth. (JEL C40, H70, O40)

1 Introduction
In recent years, the importance of political leadership in promoting economic growth and performance has gained increasing recognition in the economics profession \cite{Brady and Spence 2010}. However, formally demonstrating and testing the relationship between leadership and growth remains elusive. Many economists remain skeptical of this seemly ephemeral concept, preferring to examine concrete policy actions. This paper demonstrates that text analytics is a viable approach for identifying leadership and testing its role in promoting economic growth. We identify economic leadership through textual analysis of their public speeches and then test whether leader’s
focus on relevant economic topics is related to subsequent economic performance. In particular, we analyze the “state of the state” speeches of U.S governors from 2001 to 2013. A governor is economically successful if his/her tenure was marked by a positive economic performance (e.g. average positive real GDP growth deviation from the U.S average). We find that governors who persistently talk about the economy significantly perform better than their counterparts.

This paper provides evidence that leadership matters and that leadership can be inferred from public speeches. Indeed, the role of leadership in the success of organizations is widely recognized in the management literature (Lieberson and O’Connor (1972); Thomas (1988); Jing and Gayle (2008)). However, with a few exceptions (Brady and Spence (2010); Jones and Olken (2005)), economists have typically remained skeptical of the role of leadership. To the extent that leader’s preferences and priorities shape both, direct policy actions and economic institutions, these priorities can be critical for economic growth, (Byman and Pollack (2001)). The leader’s priorities may be taken as proxies for the myriad actions taken by the leaders to encourage growth through appointments, setting the tone for governmental agencies and promoting legislation. These positive actions are difficult to measure directly and hence we seek to measure them indirectly through analysis of the leader’s priorities, as expressed in his/her speeches.

Textual analysis has become a standard tool in the social sciences, particularly in the political science literature (LAVER et al. (2003); Quinn et al. (2010); Wilkerson et al. (2015)), but is relatively new to the economics literature (Zubin et al. (2014); Baker et al. (2013); Alexopoulos and Cohen (2015)).

In this paper, we focus on the relationship between the priorities of U.S governors and state economic growth. Compared to national leaders in developing countries, U.S. governors have far less power to promote favorable economic policies. The main advantages of using U.S. governors is that most present a formal, annual, "state of the state" speech and the overwhelming majority of these speeches are readily available. We thus avoid many of language, cultural and political differences that make international comparisons problematic.

The remainder of the paper is organized as follows: section 2 reviews the debates about the role of political leaders for societal progress, and presents a framework informing on the usefulness of text analytics for studying the influence of political leaders. Section 3 explains the statistical method used in this paper. Section 4 describes the data, and section 5 presents the results. Section 6 concludes the paper.
2 Do political Leaders matter?

2.1 Leadership traits

Hermann et al. (2001) proposes a framework for analyzing political leaders. The framework identifies four types of leaders, in a two-dimensional scale: the crusader, the strategist, the pragmatist, and the opportunist. The two dimensions are: a) the way leaders challenge constraints, and b) the way they are open to new information. The goal-driven leaders (crusaders and strategists) interpret environmental constraints through a lens that is structured by their beliefs, motives, and passions. They see constraints as obstacles in their way, and must be overcome. Goals are to be achieved by all available means. Policy priorities are clearly defined and collaborators are chosen on the basis of their general belief and support of what the leader perceives to be best for all concerned parties. Information is filtered in accordance with the government’s policies, rather than their objectiveness. “We know what we want, and we only need information telling us how to get it”. They focus on achieving their goals; and because of their focus, they are more likely to be consistent in what they say. On the contrary, the more responsive leaders (pragmatists and opportunists) see life as "a theater where there are many roles to be played" and they avoid taking action unless the option chosen is supported by their constituencies. Principles and goals are sacrificed for the sake of consensus building. For such leaders, constraints set the parameters for action. Whereas a leader with agendas seeks information that reinforces their beliefs, the responsive leader is interested in what is possible under the current circumstances; and, as the saying goes, "runs an idea up the flagpole to see who salutes it." They are like chameleons, which change their stance according to the situation. Inaction is preferred to an action that may discontent constituents. They cannot be consistent in what they say, since they respond to circumstances. In this paper, we will identify the governors’ consistency (during their tenures) over a constructed list of topics, and analyze the relationship between the consistency measure and economic growth.

2.2 Text analytics and leaders’ discourses

Consistency in what one says is assessed through one’s discourse. Text analytics allows a systematic study of discourses, and it can be used to capture leaders’ expressed priorities. In fact, the desire to understand, and predict the behavior of political leaders has compelled political scientists to apply statistical methods on leaders’ discourses to characterize their leadership style (Hermann et al. (2001)), or to determine their expressed agenda (Grimmer (2010)). These methods, also referred as content analysis, are intended to extract leaders’ motives at distance (as opposed to surveying leaders). These methods are particularly useful for political leadership studies for several reasons: we cannot reach political leaders to administer surveys; we do not have clearly defined
leadership variables and data on these variables. The "one kind of data from political leaders that is produced and preserved in abundance" is their words (Winter (2005)). Political leaders communicate their agenda, mobilize followers, and research suggests that their public statements reflect what they want, and what they are pledging to be (Hermann (2008)). Thus, text analytics gives us a viable means for quantifying leaders’ expressed priorities, which then permits research to explore the relationships of these priorities to economic growth.

3 Methodology

3.1 Text Analytics

This paper uses text analytics methodology to infer topics covered in U.S. governors’ speeches, and analyze the correlation between these topics and the real GDP growth. Text analytics aims to extract useful information from text documents and do so in a formal, automatic manner. Text analytics consists of the application of statistical methodologies to textual sources (Solka (2008)). The idea of text analytics is to apply statistical methods, designed to analyze numbers, to words. One of the main tasks for text analytics is to transform unstructured texts into numerical data. As an example of how to convert words into a spreadsheet, assume the following two sentences constitute two documents (document 1 and 2):

[document1] "i love economics, but i hate math"
[document2] "he hates economics, and he hates math"

A spreadsheet of words count can be created as shown in table 1.

Once, the data table is created, the remaining of the analysis is just an application of several of traditional and modern statistical tools. By modern statistical tools, we refer to statistical learning or machine learning tools (Varian (2014)). Text analytics is widely used in social sciences, especially in political science to analyze political speeches and legislation (LÄVER et al. (2003); Quinn et al. (2010); Wilkerson et al. (2015)). A few examples of the use of text analytics can be found in economics too. Zubin et al. (2014) shows that political ideology influences economic research in the U.S by using "observed political behavior of economists and the phrases from their academic articles" to construct predictors "of political ideology by article, economists, school and journal." Baker et al. (2013) proposes a policy uncertainty index, based "on the frequency of newspaper

One challenge in text analytics is the dimensionality of the data; if words are the variables, and there are possibly thousands of different words for a given collection of documents, then the number of variables is unusual for traditional statistical tools (For example, OLS breaks down when $n \leq p$, $n$ being the number of observations and $p$, the number of variables, or words). Consequently, we resort to the use of dimensionality reduction methods. Text analytics is an extensive field, which ranges from key words finding and analysis (Romer and Romer, 2015), to inferring themes (or topics) from text documents. This paper uses topic modeling because it produces the relative importance (proportions) of topics covered in text documents. Thus, by using topic modeling, we can assess the relative importance of each topic for each governor over time.

### 3.2 Topic models

Topic modeling derives from Latent Semantic Allocation (LSA), which is a linguistics theory of meaning that uses linear algebra to collapse words in a collection of documents into clusters of words (Landauer et al. (2007), chap1&2). The clusters of words are meant to represent themes in the documents. LSA postulates that meaning stems from words co-occurrence regardless of syntax. Thus, by a matrix factorization, from a matrix of thousands of words (variables), it is possible to reduce it to a matrix of a few topics, a matrix of their relative importance represented by their eigenvalues, and a new words-matrix. The new words-matrix provides clues for naming the topics. This paper uses a Bayesian matrix factorization algorithm known as Latent Dirichlet Allocation (Blei et al. (2003)). See Bikienga et al. (2017) for an extended exposition of topic modeling, and its suitedness as a tool for measuring leaders’ priorities. For the sake of exposition, the matrix decomposition below (a reduced form) is given for illustration.

\[
\begin{align*}
\begin{pmatrix}
  w_1 & w_2 & \ldots & w_V \\
  n_{1,1} & n_{1,2} & \ldots & n_{1,V} \\
  n_{2,1} & n_{2,2} & \ldots & n_{2,V} \\
  \vdots & \vdots & \ddots & \vdots \\
  n_{D,1} & n_{D,2} & \ldots & n_{D,V}
\end{pmatrix}
& \approx
\begin{pmatrix}
  t_1 & t_2 \\
  \theta_{1,1} & \theta_{1,2} \\
  \theta_{2,1} & \theta_{2,2} \\
  \theta_{3,1} & \theta_{3,2} \\
  \vdots & \vdots \\
  \theta_{D,1} & \theta_{D,2}
\end{pmatrix}
\begin{pmatrix}
  w_1 & w_2 & \ldots & w_V \\
  \phi_{1,1} & \phi_{1,2} & \ldots & \phi_{1,V} \\
  \phi_{2,1} & \phi_{2,2} & \ldots & \phi_{2,V}
\end{pmatrix}
\end{align*}
\]

The first matrix (left side) is a matrix of word frequencies for a collection of $D$ documents, and a...
list of V words. \( n_{1,1} \) is the frequency of word \( w_1 \) in the document \( d_1 \). The second matrix (middle) is the matrix of topic distributions in the documents. For example, \( \theta_{1,1} \) is the proportion of topic 1 (\( t_1 \)) in document \( d_1 \). This matrix (the matrix of the \( \theta \)s) preserves almost all essential information about the documents that the words in the first matrix contain. For example, by using the two dimensions, \( t_1 \) and \( t_2 \) of this matrix, we can analyze document similarities with a simple scatterplot. The third matrix (right side) gives the words distributions in each topic. For example, \( \phi_{1,1} \) gives the relative importance of word \( w_1 \) for the topic \( t_1 \). To name topic 1, we need to sort the first row of the third matrix by decreasing order. Ideally, the first few words will identify a recognizable concept. For example, if the first few words of topic \( t_1 \) are: education, college, tuition, teacher, we may conclude that topic \( t_1 \) is about education. And \( \theta_{1,1} \) is the relative importance of education in document \( d_1 \).

Intuitively, the goal of the LDA algorithm is to iteratively try different values for the \( \theta \)s and \( \phi \)s untill their joint product is highest. Thus, the algorithm searches for the \( \theta \)s and \( \phi \)s for which the likelihood of observing the given collection of documents is highest. The usefulness of topic modeling for the current paper is its ability to automatically provide the topics distributions \( \theta_{d,k} \), \( d \) being a governor’s speech, and \( k \) being a topic. Topic modeling informs on the topics, and their relative importance in every leader’s speech. Knowing the relative importance of each topic in each leader’s speech, and how the importance of topics changes over time can inform on the priorities of a leader.

4 Data

4.1 Text data

We choose to use the state of the state speeches of U.S governors from 2001 to 2013 for two reasons. Most of them can be accessed online, and they are given at a specific time of the year. Therefore, they can be used to compare governors. Most of the speeches (500 of the 598 speeches) were automatically collected (scrapped) from the state of the state website. Most of the remaining speeches were collected from The Pew Charitable Trusts website. A few of the speeches were collected from the governors’ websites. Once the speeches are gathered, the next step consists of preprocessing the data:

- Convert all words to lower cases, to avoid two identical words being considered different because one of them uses a capital letter;
- Remove stop words, which are words such as a, to, for, and, ...; they do not add content to texts.
- Strip white spaces, which is to remove unnecessary spaces and tabs in a text;
• Drop words of less than four characters; most of them do not add content to texts.

• Remove punctuation and numbers;

• Take words stems, i.e. take the roots of words to avoid, for example, economy, economics, economical to be considered as three different words.

Once the preprocessing is done, we create the Document-Term-Matrix (DTM), that is, our data matrix. The text documents (or speeches) are now converted to a spreadsheet with words counts in the cells. The DTM is then fed into the LDA algorithm to get the $\theta$ and $\phi$ matrices. The $\theta$s are used for the remainder of the analysis. The $\phi$s are used to interpret the topics.

### 4.2 Economic data

We use the state average real GDP growth rate deviation from the U.S growth rate (period 2001 to 2014) as our dependent variable in the analysis. The data were downloaded from the Bureau of Economic Analysis (BEA) website.

The real GDP percent growth variable is measured at the end of the year; the speeches are delivered at the beginning of the year, usually in January. Thus, the speech and the growth rate can be seen as one year apart. We may speculate that if a speech informs on what a leader intends to do, the effects of the leader’s actions are only visible after several months. We further use two-lead period in the analysis.

### 4.3 Ideology data

The government ideology data, compiled and maintained by [Berry et al. (1998)](Bibliography) was also collected to study the interaction effect of the state’s government ideology and the economic agenda variable constructed from the text data. The data set spans from 1960 to 2014 and provide an annual index of the ideology of each state government. The index ranges from 0 to 100, with 0 representing the most conservative government (i.e. the legislative and executive power is completely controlled by Republicans), and 100 the most liberal government. This paper converts the continuous index into three categories, under the assumption that an ideological index of 60 or above confers the Democrat governor a great political power. Similarly, an index of 40 or less confers the republican governor a great power. An index between 40 and 60 confers the governor a moderate power.

Table 2 shows that most state governments are ideologically controlled by either democrats or republicans (77 of 102 cases).
4.4 Variables construction

One goal of this paper is to show that text analytics is a viable tool for studying political leadership and economic growth. To do so, we study the correlation between governors’ consistency over certain topics (which is assumed to be a proxy for governors’ priorities) and economic growth. The inverse of the coefficient of variation (CV) of a topic is used as our consistency measure \( C_{i,j} \). Formally, the consistency over a topic \( j \) is given by:

\[
C_{i,j} = \frac{X_{i,j}'}{s_{i,j}},
\]

where \( X_{i,j}' \) is the average proportion of topic \( j \) in the speeches of a given governor \( i \), and \( s_{i,j} \) is the standard deviation of that topic for that governor. Intuitively, a governor who talks profusely and consistently about a topic should have a high average and low variance for that topic. We limit the data to governors with at least three speeches; the choice of three speeches is to assure we have enough observations to compute meaningful means and variances. 102 governors satisfy this condition (i.e. 102 observations). The consistency measures are the independent variables, and there are 5 topics. Based on a ten folds cross validation approach, 5 topics is the optimum number of topics for our data (see Appendix B for the details). Next, we compute the state’ real GDP growth rate deviation from the U.S, followed by their averages by governor.

\[
\bar{g}_{govrnor_{i}} = \frac{1}{tenure} \sum_{l=1}^{tenure} (g_{state_{i,l}} - g_{US_{i}}),
\]

where \( \bar{g}_{govrnor_{i}} \) is the average state growth \( (g_{state}) \) deviation from the US growth rate \( (g_{US}) \). The averages are computed by governor’s tenure.

Table 3 presents the summary of the final variables. Topic.1 and Topic.5 have a few outliers. Removing them from the analysis does not change the main results.

5 Results

In sum, the text data is converted into a matrix of words counts, which is then used to generate clusters of words that represent the topics (5 topics). Each topic is then converted into a consistency
Table 3: Data Summary Table

<table>
<thead>
<tr>
<th>Statistic</th>
<th>N</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic.1</td>
<td>102</td>
<td>5.561</td>
<td>5.611</td>
<td>1.081</td>
<td>43.573</td>
</tr>
<tr>
<td>Topic.2</td>
<td>102</td>
<td>3.560</td>
<td>2.802</td>
<td>1.235</td>
<td>17.029</td>
</tr>
<tr>
<td>Topic.3</td>
<td>102</td>
<td>4.142</td>
<td>2.558</td>
<td>1.426</td>
<td>15.059</td>
</tr>
<tr>
<td>Topic.4</td>
<td>102</td>
<td>3.814</td>
<td>2.787</td>
<td>1.004</td>
<td>18.338</td>
</tr>
<tr>
<td>Topic.5</td>
<td>102</td>
<td>3.961</td>
<td>3.845</td>
<td>1.165</td>
<td>34.427</td>
</tr>
<tr>
<td>gdp_r_perc_changeDv1</td>
<td>102</td>
<td>−0.137</td>
<td>1.166</td>
<td>−3.233</td>
<td>4.733</td>
</tr>
<tr>
<td>gdp_r_perc_changeDv2</td>
<td>102</td>
<td>−0.126</td>
<td>1.145</td>
<td>−2.800</td>
<td>3.300</td>
</tr>
</tbody>
</table>

measure \( C_{i,j} = \frac{\bar{X}_{i,j}}{s_{i,j}} \). Each governor performance is capture by \( \bar{g}_{\text{governor}} \). The final regression equation is:

\[
\bar{g}_{\text{governor}} = \beta_0 + \sum_{j=1}^{5} \beta_j C_{i,j} + \varepsilon_i,
\]

\( C_{i,j} \) being the topic \( j \) consistency measure for governor \( i \).

5.1 One period lead growth variable

The dependent variables used are the one and two-period leads of the state average real GDP growth rate, deviation from the U.S average real GDP growth rate. The following chart illustrates the one and two-period leads idea. For instance, assuming a one term governor with speeches from beginning 2001 to the beginning 2004, his/her agenda (i.e. the governor’s consistency measure over a topic) is matched with the average growth rate of the end of 2001 to the end of 2004 for the one period lead dependent variable, and the average growth rate of the end of 2002 to the end of 2005 for the two-period lead dependent variable.
5.1.1 Economic agenda and economic growth

Table 4 column (1) shows the result of the OLS regression of the five topics on the one period lead of the state real GDP average deviation from US growth ($\bar{g}_{\text{governor}}$). Only one topic (topic.4) appears significant. We further use the LASSO (Least Absolute Shrinkage and Selection Operator) method to drop the least relevant topics. The LASSO is a constrained OLS. The constraint is such that it sets the parameters of the non-relevant exogenous variables to zero, yielding a sparse model. It is one of the most used and reliable variable selection methods. LASSO can be seen as an alternative to the stepwise regression method; however it is more principled than the stepwise approach, as it is a model optimization based method.

Of the 5 topics, the LASSO regression picked 2 topics, that is, we can safely ignore three topics in our final regression model. By fitting an OLS on the selected 2 topics, the results are as shown in the second column of Table 4. Note that dropping the three topics from the model does not change the results, confirming that these three topics are not needed in the regression model.

Topic.4 is a variable of interest because it is about governors’ economic agenda as stated in their speeches, which we suspected might affect the state growth rate. The reason Topic.4 is called governor’s economic agenda is explained in the next section. Before then, column 3 of Table 4 shows the results for a different definition of the consistency variable (the standard deviation of the topic). The idea is that too much variation on a particular topic is a sign of lack of focus (or lack of agenda). The Topic.4 coefficient is negative and statistically significant suggesting that the lack of focus on the governor economic agenda is associated with negative economic growth.

A quick graphical look at the data (scatter plot of growth variable and Topic.4) shows two outliers values for Topic.4. To check whether the significance of Topic.4 is affected by the presence of these outliers, the regression model is run (1) using the log of the topics (column 2 of Table 5), (2) dropping the two outliers from the data and running the regular model (column 3). Topic.4 remains statistically significant under these two specifications.

It appears that Topic.4 is statistically significant however we look at its relationship with the growth variable. This topic is about economic development, and is positively related to the growth variable. It can be speculated that a governor with an economic development agenda is more likely to achieve a positive economic growth. Why Topic.4 can be understood as the economic agenda topic?

5.1.2 Interpreting Topic.4 as Economic Agenda Topic

Traditionally, topics from topic modeling are interpreted by looking at the most frequent words that constitute a topic, usually the first 30 words. We rely on three different approaches to interpreting the topics in this paper.

First, by looking at the first 40 words (Table 13 Appendix C), which are ranked by their
Table 4: OLS with 5 topics (1), with 2 topics picked by LASSO (2), and using topics’ standard deviation (3)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variable:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One period lead of state average real GDP growth (deviation from US)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic.1</td>
<td>−0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic.2</td>
<td>0.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic.3</td>
<td>−0.059</td>
<td>−0.057</td>
<td>−2.483</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.043)</td>
<td>(3.801)</td>
</tr>
<tr>
<td>Topic.4</td>
<td>0.134***</td>
<td>0.135***</td>
<td>−8.375**</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(4.090)</td>
</tr>
<tr>
<td>Topic.5</td>
<td>0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>−0.460</td>
<td>−0.418</td>
<td>0.524</td>
</tr>
<tr>
<td></td>
<td>(0.314)</td>
<td>(0.255)</td>
<td>(0.361)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>102</td>
<td>102</td>
<td>102</td>
</tr>
<tr>
<td><strong>R^2</strong></td>
<td>0.120</td>
<td>0.118</td>
<td>0.046</td>
</tr>
<tr>
<td><strong>Adjusted R^2</strong></td>
<td>0.075</td>
<td>0.100</td>
<td>0.026</td>
</tr>
<tr>
<td><strong>Residual Std. Error</strong></td>
<td>1.122 (df = 96)</td>
<td>1.107 (df = 99)</td>
<td>1.151 (df = 99)</td>
</tr>
<tr>
<td><strong>F Statistic</strong></td>
<td>2.627** (df = 5; 96)</td>
<td>6.603*** (df = 2; 99)</td>
<td>2.371* (df = 2; 99)</td>
</tr>
</tbody>
</table>

*Note:* *p<0.1; **p<0.05; ***p<0.01
relative importance for the topic, it can be seen that economically related words are highly ranked for Topic.4. Words such as business, work, create, energy, develop, company, invest, industry, company... are highly ranked.

Second, we constructed a web application aimed at highlighting topic’s keywords within the speeches that are highest on a given topic (Figure 2). Again, a speech is a distribution over topics. Some speeches are very high on a particular topic (for instance, over 60% of the speech is about a particular topic), while others are low on that particular topic. Thus, for a given topic, by identifying the speeches with the highest proportions of that topic and highlighting the topic’s keywords in those speeches, we may be able to fully identify the topic. The top five documents for Topic.4 are filled with "Economic development" phrases. One such phrase is: "aggressive economic development". It should be noted that the top five documents are from a single state (North Dakota). Figure 1 is an excerpt from the 2009 state of the state speech of the governor of North Dakota. This excerpt shows a clear delineation of economic agenda.

It is true that North Dakota had an oil boom that started in 2006, but governor John Hoeven economic agenda was prominent even in his earlier speeches. The word “develop” and its variants were used more than ten times in his 2001 state of the state speech. The phrase “New economy” was used 9 times in the same speech. Further, Topic.4 has been high in all of his speeches, ranging
Last, we introduced external speeches (speeches we know should be high on the economic topic) to our collection of governors’ speeches. We expect the algorithm to assign high proportions to the economic topic (Topic.4) in these speeches. Our interpretation of Topic.4 is validated if the topic modeling algorithm assigns high proportion to the economic topic in these speeches. Ten economic policy speeches by four different politicians (George W. Bush (3), Obama (3), Romney (1), McCain (1), Jeb Bush (1), Hilary Clinton (1)) were added to our collection of speeches. All these speeches are fairly high on Topic.4 (Table 6), supporting the interpretation that Topic.4 is indeed about economic agenda. It suffices to look at the totals rows of Table 6 to see that Topic.4 and Topic.1 are high for these ten speeches. The totals of the remaining topics are fairly low.

Topic.1 is what we will call a residual topic, that is, a topic made up of mostly contentless words. Topic.2 is about education, and Topic.5 is about the state budget. Topic.3 is a mix of education, budget, economy with a reform tone. A disproportionate space was given in this paper for the interpretation of Topic.4 because it is the one which is statistically significant however we
look at the relationship between the growth variable and the consistency measure of the topics.

Table 6: Topics distributions in documents labeled by their authors as speeches on the economy

<table>
<thead>
<tr>
<th></th>
<th>Topic.1</th>
<th>Topic.2</th>
<th>Topic.3</th>
<th>Topic.4</th>
<th>Topic.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_Obama_2010</td>
<td>0.272</td>
<td>0.079</td>
<td>0.203</td>
<td>0.376</td>
<td>0.071</td>
</tr>
<tr>
<td>B_Obama_2012</td>
<td>0.429</td>
<td>0.044</td>
<td>0.148</td>
<td>0.345</td>
<td>0.035</td>
</tr>
<tr>
<td>B_Obama_2013</td>
<td>0.422</td>
<td>0.121</td>
<td>0.104</td>
<td>0.309</td>
<td>0.044</td>
</tr>
<tr>
<td>GW_Bush_2003</td>
<td>0.297</td>
<td>0.023</td>
<td>0.221</td>
<td>0.367</td>
<td>0.092</td>
</tr>
<tr>
<td>GW_Bush_2006</td>
<td>0.514</td>
<td>0.103</td>
<td>0.049</td>
<td>0.268</td>
<td>0.066</td>
</tr>
<tr>
<td>GW_Bush_2008</td>
<td>0.202</td>
<td>0.040</td>
<td>0.217</td>
<td>0.329</td>
<td>0.212</td>
</tr>
<tr>
<td>H_Clinton_2015</td>
<td>0.349</td>
<td>0.081</td>
<td>0.170</td>
<td>0.356</td>
<td>0.043</td>
</tr>
<tr>
<td>Jeb_Bush_2015</td>
<td>0.395</td>
<td>0.062</td>
<td>0.095</td>
<td>0.364</td>
<td>0.085</td>
</tr>
<tr>
<td>J_McCain_2008</td>
<td>0.355</td>
<td>0.035</td>
<td>0.130</td>
<td>0.331</td>
<td>0.149</td>
</tr>
<tr>
<td>M_Romney_2012</td>
<td>0.384</td>
<td>0.068</td>
<td>0.287</td>
<td>0.221</td>
<td>0.041</td>
</tr>
<tr>
<td>Totals</td>
<td>3.617</td>
<td>0.655</td>
<td>1.624</td>
<td>3.266</td>
<td>0.837</td>
</tr>
</tbody>
</table>

It should be noted that the topics are not defined ex-ante. Before the analysis, we had no knowledge of the topics covered in the speeches, and the number of topics that should be considered. The analysis suggested we collapse the words into only five topics for analyzing the growth variable, and it turned out that the topic that mostly matters for economic growth is the economic development topic. This strongly suggests that the governor’s economic agenda matters for economic growth.

We further perform the analysis by interacting the economic agenda variable with the governor’s political affiliation and find that both, democrats and republicans economic agenda correlate positively with economic growth (Table 7); however the correlation is stronger for republicans than for democrats. Column 1, of Table 7 shows the results when no particular treatment of the outliers is considered, and column 2 shows the results when we take the log of the topics as a way to temper the effects of the outliers on the results.

Table 12 in appendix A shows that the results remain qualitatively the same when we control for regions, and for state human capital level, suggesting that the governors’ focus on the economy strongly correlates with higher economic growth.

5.1.3 Economic Agenda, Political Power and Economic Growth

Does ideological control of the government influences the effect of governor’s economic agenda? It can be argued that the governor’s agenda will be difficult to implement without an ideological control of the government. For instance, a Democrat governor operating within a context in which Democrats control the legislative branch of the government has more leeway in implementing
Table 7: Interaction with Governor’s Political Affiliation. OLS without consideration of outliers (1), OLS using the log of Topic.4 (2)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable:</td>
<td>One period lead of state average real GDP growth (deviation from US)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic.3</td>
<td>−0.054 (0.044)</td>
<td>−0.212 (0.226)</td>
</tr>
<tr>
<td>Topic.4:PartyD</td>
<td>0.094** (0.045)</td>
<td>0.519** (0.222)</td>
</tr>
<tr>
<td>Topic.4:PartyR</td>
<td>0.177*** (0.046)</td>
<td>0.697*** (0.229)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.426* (0.254)</td>
<td>−0.569 (0.381)</td>
</tr>
</tbody>
</table>

|                  |                    |                    |
| Observations     | 101                | 101                |
| R²               | 0.143              | 0.093              |
| Adjusted R²      | 0.116              | 0.065              |
| Residual Std. Error (df = 97) | 1.099 | 1.130 |
| F Statistic (df = 3; 97) | 5.394*** | 3.325** |

Note: *p<0.1; **p<0.05; ***p<0.01
his/her agenda. Table 8 presents an analysis of the economic agenda effect on economic growth, by political party interacted with ideological control of the government.

Table 8: Interaction with Governor’s Political Affiliation and Ideology. OLS without a special treatment of the outliers (1), OLS after removing the two outliers (2), OLS after taking the log of Topic.4 (3)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One period lead of state average real GDP growth (deviation from US)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic.3</td>
<td>−0.040</td>
<td>−0.052</td>
<td>−0.091</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.042)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_D:PartyD</td>
<td>0.057</td>
<td>0.026</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.061)</td>
<td>(0.225)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_N:PartyD</td>
<td>0.247**</td>
<td>0.198*</td>
<td>0.916***</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.102)</td>
<td>(0.345)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_R:PartyD</td>
<td>0.232</td>
<td>0.189</td>
<td>0.836</td>
</tr>
<tr>
<td></td>
<td>(0.226)</td>
<td>(0.222)</td>
<td>(0.725)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_D:PartyR</td>
<td>0.116</td>
<td>0.044</td>
<td>0.526</td>
</tr>
<tr>
<td></td>
<td>(0.225)</td>
<td>(0.224)</td>
<td>(0.674)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_N:PartyR</td>
<td>−0.073</td>
<td>−0.124</td>
<td>−0.059</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.111)</td>
<td>(0.374)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_R:PartyR</td>
<td>0.201***</td>
<td>0.117**</td>
<td>0.802***</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.056)</td>
<td>(0.224)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.430*</td>
<td>−0.204</td>
<td>−0.650*</td>
</tr>
<tr>
<td></td>
<td>(0.258)</td>
<td>(0.279)</td>
<td>(0.379)</td>
</tr>
</tbody>
</table>

| Observations                   | 101              | 99               | 101              |
| Adjusted R²                    | 0.184            | 0.143            | 0.187            |
| Residual Std. Error            | 1.056 (df = 93)  | 1.030 (df = 91)  | 1.093 (df = 93)  |
| F Statistic                    | 4.222*** (df = 7; 93) | 2.163** (df = 7; 91) | 3.064*** (df = 7; 93) |

Note: *p<0.1; **p<0.05; ***p<0.01

The results suggest that the economic agenda of a Republican governor whose political party controls the government is positively associated with economic growth. The result is not significant for a Democrat governor whose political party controls the government, though the relationship
remains positive. By ideological control, we mean at least 60% of the state legislative and executive power is held by one political party. The data suggests that Democrat governors’ economic agenda matters when they do not have ideological control of the government. The results hold under different specifications (removing outliers values from the dataset (column 2), and using the log values of Topic.3 and Topic.4 column (3)). This analysis uses 101 observations instead of 102 governors because one governor is independent. The scatterplots (Figure 2) show the relationship between economic growth and economic agenda by governor political party, and the ideological dominance of the state government.

The scatterplots offer a visual prospect of the relationships shown in the regression table (Table 8).

### 5.2 Two-period lead growth variable

The analysis in section 5.1 shows that a governor’s economic agenda matters for economic growth by using a one lead growth rate. Next, we perform the same analysis with a two-period lead growth variable. The result of the OLS regression using $K = 5$ topics is given in Table 9, column (1).

Again, Topic.4 seems to be strongly relevant for the two-period lead economic growth variable. We further use LASSO regression to drop the least relevant topics, and LASSO picked two topics of which the economic development topic (Topic.4), and the budget topic, Topic.5 are statistically significant (Table 9, column (2)). Column 3 of Table 9 shows the regression results when the two outliers values of Topic.4 are removed from the data set; and column 4 shows the results when we take the log of the dependent variable.

Table 9 shows that the economic development agenda is statistically significant and positive however we deal with the outliers values. This suggests that the significance of Topic.4 must be meaningful. Also, it can be said that the governor’s economic agenda has a lasting effect on economic growth.

#### 5.2.1 Economic Agenda, Political Power and Economic growth

We further the analysis by taking into consideration the governor political party and the ideological control of the government. The results are similar to the results performed using the one lead growth rate as the dependent variable. The results in Table 10 are for the regular model (column 1), removing the two outliers values of Topic.4 (column 2), and taking the log of the topics (column 3).

Again, as shown in section 5.1.3, the governor’s economic agenda matter for economic growth for a Republican governor whose political party controls the government. It also matters for a Democrat governor whose political party do not control the government. A Democrat governor
Figure 2: Relationship between the consistency with which governors talk about the economy and the state real GDP average deviation from U.S growth rate. The economic agenda of republican governors whose party control the legislative power tend to achieve higher economic growth than their counterparts. The relationship seems weak for Democrat governors with political power.
Table 9: OLS with 5 topics (1) and 2 topics picked by LASSO (2), after removing the two outliers of Topic.4 (3), after taking the log of Topic.4 (4)

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
<th>Two-period lead of state average real GDP growth (deviation from US)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3) (4)</td>
<td></td>
</tr>
<tr>
<td>Topic.1</td>
<td>−0.014</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Topic.2</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td></td>
</tr>
<tr>
<td>Topic.3</td>
<td>−0.081*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td></td>
</tr>
<tr>
<td>Topic.4</td>
<td>0.110***</td>
<td>0.110***</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Topic.5</td>
<td>0.055*</td>
<td>0.049*</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.475</td>
<td>−0.740***</td>
</tr>
<tr>
<td></td>
<td>(0.306)</td>
<td>(0.214)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Observations: 102</th>
<th>102</th>
<th>100</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.137</td>
<td>0.102</td>
<td>0.067</td>
<td>0.101</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.092</td>
<td>0.084</td>
<td>0.048</td>
<td>0.083</td>
</tr>
<tr>
<td>Residual Std. Error</td>
<td>1.091 (df = 96)</td>
<td>1.096 (df = 99)</td>
<td>1.103 (df = 97)</td>
<td>1.097 (df = 99)</td>
</tr>
<tr>
<td>F Statistic</td>
<td>3.054** (df = 5; 96)</td>
<td>5.613*** (df = 2; 99)</td>
<td>3.494** (df = 2; 97)</td>
<td>5.561*** (df = 2; 99)</td>
</tr>
</tbody>
</table>

Note: *p<0.1; **p<0.05; ***p<0.01
Table 10: Interaction with Governor’s Political Affiliation and Ideology. OLS without a special treatment of the outliers (1), OLS after removing the two outliers (2), OLS after taking the log of Topic.4 (3)

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable: Two-period lead of state average real GDP growth(deviation from US)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Topic.5</td>
<td>0.048*</td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_D:PartyD</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_N:PartyD</td>
<td>0.241**</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_R:PartyD</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_D:PartyR</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_N:PartyR</td>
<td>−0.081</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
</tr>
<tr>
<td>Topic.4:PowerPower_R:PartyR</td>
<td>0.146***</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
</tr>
<tr>
<td>Constant</td>
<td>−0.690***</td>
</tr>
<tr>
<td></td>
<td>(0.227)</td>
</tr>
</tbody>
</table>

|                  | (2)                                                                                     |
| Topic.5          | 0.049*                                                                                 |
|                  | (0.028)                                                                                 |
| Topic.4:PowerPower_D:PartyD | 0.043                                                                                     |
|                  | (0.065)                                                                                 |
| Topic.4:PowerPower_N:PartyD | 0.227**                                                                                 |
|                  | (0.108)                                                                                 |
| Topic.4:PowerPower_R:PartyD | 0.130                                                                                     |
|                  | (0.234)                                                                                 |
| Topic.4:PowerPower_D:PartyR | 0.029                                                                                     |
|                  | (0.237)                                                                                 |
| Topic.4:PowerPower_N:PartyR | −0.096                                                                                     |
|                  | (0.118)                                                                                 |
| Topic.4:PowerPower_R:PartyR | 0.132**                                                                                 |
|                  | (0.060)                                                                                 |
| Constant         | −0.639**                                                                               |
|                  | (0.256)                                                                                 |

|                  | (3)                                                                                     |
| Topic.5          | 0.379*                                                                                 |
|                  | (0.221)                                                                                 |
| Topic.4:PowerPower_D:PartyD | 0.334                                                                                     |
|                  | (0.223)                                                                                 |
| Topic.4:PowerPower_N:PartyD | 0.844**                                                                                 |
|                  | (0.343)                                                                                 |
| Topic.4:PowerPower_R:PartyD | 0.448                                                                                     |
|                  | (0.720)                                                                                 |
| Topic.4:PowerPower_D:PartyR | 0.268                                                                                     |
|                  | (0.669)                                                                                 |
| Topic.4:PowerPower_N:PartyR | −0.125                                                                                  |
|                  | (0.370)                                                                                 |
| Topic.4:PowerPower_R:PartyR | 0.593**                                                                                 |
|                  | (0.226)                                                                                 |
| Constant         | −1.089***                                                                               |
|                  | (0.351)                                                                                 |

Observations | 101 | 99 | 101 |
R² | 0.180 | 0.144 | 0.172 |
Adjusted R² | 0.119 | 0.078 | 0.109 |
Residual Std. Error | 1.079 (df = 93) | 1.090 (df = 91) | 1.085 (df = 93) |
F Statistic | 2.921*** (df = 7; 93) | 2.191** (df = 7; 91) | 2.751** (df = 7; 93) |

Note: *p<0.1; **p<0.05; ***p<0.01
whose political party control the government can affect the economy positively. However, this relationship is not statistically significant. A Republican governor economic agenda effect on economic growth is not statistically significant when his political party does not control the state government.

5.3 Robustness Check

5.3.1 A different consistency measure

A different consistency measure is used for robustness check. That consistency measure is the log of the inverse of the variance. The cross validation method suggests the use of 16 topics when the consistency measure of choice is the log of the inverse of the variance. The Table summarizes the regression results.

Column 1 and 3 shows the results when using all 16 topics. Column 2 and 4 shows the regression results after using a LASSO method to reduce the number of topics. It is worthwhile noting that the topics that are significant when the dependent variable is the one-period lead growth rate are also significant when the dependent variable is the two-period lead growth rate.

Topic.15 is the economic topic when we consider a 16 topics set (Table Appendix C). Topic.15 is significant at 5% significance level for the one-period lead of the growth variable (column 1 and 2), and highly significant for the two-period lead growth variable (column 3 and 4). From Table it can be seen that Topic.15 is very similar to Topic.4 in term of the ranking of their most important words. That confirms that the economic agenda of the governor is positively associated with economic growth. Topic.7, which is about reforms is also positively associated with economic growth. Note that Topic.3 in section 5.1, which has a reform tone to it, was found to be negatively associated with economic growth. There is an apparent contradiction between the finding in section 5.1 and the robustness check finding with respect to the reform topic. But, it is clearer that Topic.7 is about reforms, when Topic.3 in section 5.1 is a mix of several topics with a reform tone. It is also safe to say that Topic.3 in section 5.1 is not specific enough.

Another significant topic is Topic.8, which is clearly about education (see Table Appendix C); it is also positively associated with economic growth. This suggests that the state governor’s education agenda is positively associated with economic growth. There are a few canals where state education policy can induce economic growth in the relatively short run. First, it sends a signal to potential investors that the state is serious about improving its education system and company can expect to tap into well-trained workers in the coming years. Second, it helps the state attract talents because highly qualified workers may be concerned about the education their children get. The education topic was found to be positively associated with economic growth in section 5.1 and 5.2, where the education topic where referred as Topic.2 (see Table column 1, and Table column 1).
Table 11: OLS with 16 topics (1 and 3); and OLS with topics picked by LASSO (2 and 4)

<table>
<thead>
<tr>
<th>Topic</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic.1</td>
<td>$-0.074$</td>
<td>$0.063$</td>
<td>$0.063$</td>
<td>$0.093$</td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.098)</td>
<td>(0.093)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>Topic.2</td>
<td>$-0.072$</td>
<td>$-0.100$</td>
<td>$-0.041$</td>
<td>$-0.061$</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.091)</td>
<td>(0.093)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Topic.3</td>
<td>$-0.100$</td>
<td>$-0.119$</td>
<td>$-0.048$</td>
<td>$-0.061$</td>
</tr>
<tr>
<td></td>
<td>(0.085)</td>
<td>(0.079)</td>
<td>(0.079)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Topic.4</td>
<td>$0.063$</td>
<td>$-0.072$</td>
<td>$-0.100$</td>
<td>$0.063$</td>
</tr>
<tr>
<td></td>
<td>(0.098)</td>
<td>(0.072)</td>
<td>(0.091)</td>
<td>(0.085)</td>
</tr>
<tr>
<td>Topic.5</td>
<td>$-0.041$</td>
<td>$-0.119$</td>
<td>$-0.048$</td>
<td>$-0.061$</td>
</tr>
<tr>
<td></td>
<td>(0.093)</td>
<td>(0.085)</td>
<td>(0.079)</td>
<td>(0.083)</td>
</tr>
<tr>
<td>Topic.6</td>
<td>$0.271^{***}$</td>
<td>$0.244^{***}$</td>
<td>$0.263^{***}$</td>
<td>$0.219^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.076)</td>
<td>(0.079)</td>
<td>(0.072)</td>
</tr>
<tr>
<td>Topic.7</td>
<td>$0.185^{**}$</td>
<td>$0.182^{**}$</td>
<td>$0.163^{*}$</td>
<td>$0.183^{**}$</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.081)</td>
<td>(0.083)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Topic.8</td>
<td>$0.055$</td>
<td>$0.055$</td>
<td>$0.026$</td>
<td>$0.026$</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.079)</td>
<td>(0.076)</td>
<td>(0.076)</td>
</tr>
<tr>
<td>Topic.9</td>
<td>$-0.164^{*}$</td>
<td>$-0.136$</td>
<td>$-0.071$</td>
<td>$-0.096$</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.088)</td>
<td>(0.093)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Topic.10</td>
<td>$0.093$</td>
<td>$-0.137$</td>
<td>$-0.118$</td>
<td>$-0.137$</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.084)</td>
<td>(0.077)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>Topic.11</td>
<td>$0.048$</td>
<td>$0.048$</td>
<td>$0.021$</td>
<td>$0.048$</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.094)</td>
<td>(0.090)</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Topic.12</td>
<td>$0.008$</td>
<td>$0.008$</td>
<td>$-0.009$</td>
<td>$0.008$</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.100)</td>
<td>(0.096)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Topic.13</td>
<td>$0.259^{**}$</td>
<td>$0.198^{**}$</td>
<td>$0.329^{***}$</td>
<td>$0.284^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.107)</td>
<td>(0.096)</td>
<td>(0.102)</td>
<td>(0.093)</td>
</tr>
<tr>
<td>Topic.14</td>
<td>$0.021$</td>
<td>$0.021$</td>
<td>$0.045$</td>
<td>$0.021$</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.110)</td>
<td>(0.105)</td>
<td>(0.105)</td>
</tr>
<tr>
<td>Constant</td>
<td>$-2.490$</td>
<td>$-2.490$</td>
<td>$-2.784$</td>
<td>$-3.514^{***}$</td>
</tr>
<tr>
<td></td>
<td>(2.287)</td>
<td>(2.287)</td>
<td>(2.194)</td>
<td>(1.153)</td>
</tr>
</tbody>
</table>

Dependent variable: One-period lead (1 and 2) and Two-period lead (3 and 4) of state average real GDP growth (deviation from US)

<table>
<thead>
<tr>
<th>Observations</th>
<th>102</th>
<th>102</th>
<th>102</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.276</td>
<td>0.239</td>
<td>0.310</td>
<td>0.248</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.140</td>
<td>0.182</td>
<td>0.180</td>
<td>0.209</td>
</tr>
<tr>
<td>Residual Std. Error</td>
<td>1.082 (df = 85)</td>
<td>1.055 (df = 94)</td>
<td>1.037 (df = 85)</td>
<td>1.019 (df = 96)</td>
</tr>
<tr>
<td>F Statistic</td>
<td>2.027^{**} (df = 16; 85)</td>
<td>4.211^{***} (df = 7; 94)</td>
<td>2.384^{***} (df = 16; 85)</td>
<td>6.326^{***} (df = 5; 96)</td>
</tr>
</tbody>
</table>

Note: "p<0.1; **p<0.05; ***p<0.01
In sum, the results obtained using only a five topics construct with the inverse of the coefficient of variation as our consistency measure and the results obtained using a sixteen topics construct with the log of the inverse of the variance are similar. The governor’s focus on economic issues matters for economic growth.

5.3.2 Governor’s Agenda and state spending

Part of the argument that leadership matters for economic growth is that leaders design, or at least decide on economic, educational, and social policies. The policy options have a bearing on economic outcomes since they affect how society’s scarce resources are allotted. For instance, we would expect a governor with an education agenda to devote a high share of state spendings to education. Indeed, education tends to be prominent in U.S governors’ state of the state speeches. Of the five topics we constructed, Topic.2, the easiest to interpret of all the five topics, is about education. Does this topic correlate with the share of state spending devoted to education? The following scatter plot shows that the consistency with which governors talk about education is positively correlated with state spending on education. That is, the higher education is a priority for a governor, the higher the share of his/her state spendings devoted to education during that governor’s tenure.

Figure 3: The consistency with which governors talk about education is positively correlated with the average share of state spendings on education.
The positive relation is valid both for Democrats and Republicans as shown on the graph (Figure 4).\

We further plot the correlations of all the topics and the share of state spending on education (Figure 5).

The plots show the correlations and their strength (as measured by their p-values) between the five topics and the education share of the state spendings. The first row of the panel shows the relationship when using annual data (i.e., the annual speech is the unit of observation). The second row of the panel shows the relationship when using the constructed consistency measure (i.e., the governor is the unit of observation). The bars in the first column represent the correlation values, and the bars in the second column represent the p-values of the correlations. Topic 2 (the education topic) is emphasized with a darker bar. The first row of the plot suggests a strong relationship between what the governor says in his/her speech (at the beginning of the year) and the share of the state spendings devoted to education (in that year).

The second row suggests that the governor consistency over the education topic is fairly correlated, positively, with the average share of state spending on education.

For both cases (i.e., using annual speeches as the unit of observation, or state governor as the

---

3 We dropped the observations for which the education consistency measure is greater than 6 to highlight the spread of the observations. By dropping the 9 observations with values greater than 6, the qualitative trend of the data does not change.
Figure 5: The correlation between the importance of the topics and the share of state spending on education (first row) is highest for the education topic (Topic.2). The correlation is highly significant with a p-value close to zero. Using the constructed variables, where the governor is the unit of observation, the education topic has the highest correlation value (second row). The correlation is marginally significant, with a p-value = 0.071.
unit of observation), the education topic is the most correlated with the share of state spendings on education. That suggests that the governor’s education agenda matters for the resources the state devotes towards education\textsuperscript{4}.

5.4 Goal driven leaders and economic growth

The main take away from the robustness check is that the relevant topics found using the inverse of the coefficient of variation as the consistency measure are also relevant when we use the log of the inverse of the variance as the consistency measure. This suggests that being consistent on certain topics matters for economic growth; however we measure consistency. This finding supports the main thesis of this paper, which postulates that the goal-driven leader is persistent on his/her agenda; and persistence yields results. To capture the persistence of the governor, we constructed a consistency measure which captures governors’ consistency over certain topics. Being persistent on a topic over several years is a sign of commitment to an agenda. An example of a commitment to an agenda can be seen by looking at the speeches of the governor of West Virginia from 2001 to 2004\textsuperscript{5}. It is safe to say that the governor is an education governor. We arrive at this conclusion because this governor scores high on our education consistency measure. Then, by looking at his speeches over four years period, it is apparent that this governor main focus was in improving West Virginian education system. Another example is the North Dakota governor of 2001 to 2009 whose economic development agenda rank among the top in our consistency measure\textsuperscript{6}.

5.5 Text analytics and leadership studies

The main goal of the present paper was to study US governors’ professed agenda and economic growth. We define professed agenda the consistency with which governors talk about certain topics in their state of the state speeches. Talking about a topic consistently is perceived as a sign of commitment to an agenda. Thus, we first define a consistency measure which is the inverse of the coefficient of variation \(C_{i,j} = \frac{x_{i,j}}{s_{i,j}}\). The main idea of this definition is that high variance is a sign of lack of focus, and high mean is a sign of importance. The means and variances are computed from the proportion of the speeches devoted to a particular topic. By defining consistency as the inverse of the coefficient of variation, a high score indicates that the governor shows low variability and high proportion of a given topic in his/her speeches.

\textsuperscript{4}It should be noted that four outliers valued where dropped before computing and plotting the correlations seen on the second row of the panel.

\textsuperscript{5}http://stateofthestate.com/content.aspx?state=WV&date=02/14/2001

\textsuperscript{6}http://stateofthestate.com/content.aspx?state=ND&date=01/09/2001
One main challenge of text analytics is separating signals from noise. We rely on cross-validation methods to identify a limited set of topics that are correlated with economic growth rate. It turns out that the identified topics bear economic significance. For example, we found that the economic development topic is always statistically, significantly related to economic growth for all the different specifications we used. Using a different definition of the consistency measure, the log of the inverse of the variance, the topics that are statistically related to economic growth are those expected when indeed the governors do what they profess to be doing. Furthermore, the same topics were identified when using the inverse of the coefficient of variation as the consistency measure. These findings suggest that there are signals in the governors’ speeches. The findings further show that text analytics methods applied to leaders’ speeches can provide measures of leaders’ professed agenda. This provides an avenue for studying political leadership and economic growth. We do not have a measure of what they do, but we can measure what to say they are doing. Consequently, this paper makes the case that text analytics, applied to leaders’ speeches, is a viable approach to studying political leadership and economic growth.

6 Conclusion

This paper demonstrates that text analytics methods, particularly topic modeling, applied to leaders’ speeches, is a viable tool for studying political leadership and economic performance. Using a corpus of 598 U.S governors’ state of the state speeches, topic modeling allowed us to identify 5 topics covered in these speeches. Of the 5 topics, the LASSO method was used to identify the topics which co-vary most with real GDP growth rate (state growth average deviation from U.S growth rate). Then, simple OLS regression using the most important topics suggested that only one topic (the economic development topic) matters for economic growth under the diverse specifications used in this paper.

The paper demonstrates the possibility of studying U.S governors in terms of what they say, rather than their ideological affiliation to a political party, as is often seen in the economics literature (Beland (2015)).

The paper used mainly 5 topics for a corpus of 598 speeches. Five topics, though enough for the current paper’s purpose, may appear too small. For example, Topic.3 appears to be a mix of several topics; and had we chosen a larger set of topics (maybe 10) this mix-up topic could be separated to well-defined topics. Indeed, by redefining the consistency measure as the log of the inverse of the variance, the reform topic (which is apparent in Topic.3 when using 5 topics set) appeared as a meaningful topic. The governor’s education agenda is also positively related to economic growth. From our findings, it appears that the governor’s economic development, education, and reform agendas are strongly associated with economic growth. A future research would be to look at
governor’s professed agendas and action taken by the state to achieve these agenda. For instance, are governor’s with professed education agenda spend more on education? Are governor’s with reform agenda pass more reforms legislation?

References


A  Regression with control variables

Our focus in the main text was to assess the ability to capture political leaders’ professed agendas, and analyze the relationship between these agendas and economic growth. Here, we add control variables to assess the strength of the relationship. We consider the region and human capital as control variables. We use the Bureau of Economic Analysis regions (8 regions), and the proportion of adults age 25 and over with a bachelor degree or more. Table 12 indicates that the governors’ consistency over economics issues is positively associated with higher economic growth.

B  Choosing the optimum K

Topic modeling algorithms require that the number of topics K be provided. To avoid picking K arbitrarily, we use a ten folds cross-validation approach to picking the optimum K; that is, we iterate the regression through different values for K. For each K, an OLS is run on 90% of the observations; then the 10% remaining observations are used to predict the remaining values for the dependent variable. Using the predicted values, we compute the errors sum of squares (sse) of the dependent variable for the 10% of the observations. The process is repeated 10 times, each using a different set of observations to be predicted. For each K, a single value for the sse is computed. Then the K for which the sse is small is considered the optimum number of topics to be used.

The figure above shows the path of the log sses as K takes different values ($K = 5, 6, ..., 60$). The vertical blue line indicates the optimum number of topics K. K = 5 is the optimum K for the two dependent variables. The dependent variables are the state average deviation from the U.S growth rate of the real GDP (one-period lead, and two-period lead). Each independent variable is a consistency measure over a topic. We measure consistency as the inverse of the coefficient of variation (CV). We are interested in governors who talk a lot about a topic (which we capture by the mean $\bar{X}_i$), with a minimum variance over time (which we capture by the standard deviation $s_i$).
<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Real GDP (One period lead)</th>
<th>Real GDP (Two period lead)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLAK</td>
<td>$-0.841^{*}$</td>
<td>$-1.036^{**}$</td>
</tr>
<tr>
<td></td>
<td>(0.469)</td>
<td>(0.454)</td>
</tr>
<tr>
<td>MEST</td>
<td>$-0.550$</td>
<td>$-0.775$</td>
</tr>
<tr>
<td></td>
<td>(0.487)</td>
<td>(0.472)</td>
</tr>
<tr>
<td>NENG</td>
<td>$-0.866^{*}$</td>
<td>$-1.116^{**}$</td>
</tr>
<tr>
<td></td>
<td>(0.485)</td>
<td>(0.470)</td>
</tr>
<tr>
<td>PLNS</td>
<td>$-0.178$</td>
<td>$-0.423$</td>
</tr>
<tr>
<td></td>
<td>(0.430)</td>
<td>(0.416)</td>
</tr>
<tr>
<td>RKMT</td>
<td>0.321</td>
<td>0.279</td>
</tr>
<tr>
<td></td>
<td>(0.480)</td>
<td>(0.465)</td>
</tr>
<tr>
<td>SEST</td>
<td>$-0.541$</td>
<td>$-0.818^{**}$</td>
</tr>
<tr>
<td></td>
<td>(0.397)</td>
<td>(0.385)</td>
</tr>
<tr>
<td>SWST</td>
<td>$-0.114$</td>
<td>$-0.392$</td>
</tr>
<tr>
<td></td>
<td>(0.550)</td>
<td>(0.533)</td>
</tr>
<tr>
<td>Human Capital (education)</td>
<td>1.014</td>
<td>$-1.975$</td>
</tr>
<tr>
<td></td>
<td>(2.474)</td>
<td>(2.396)</td>
</tr>
<tr>
<td>Topic.4 (economics)</td>
<td>0.092^{**}</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.041)</td>
</tr>
<tr>
<td>Constant</td>
<td>$-0.376$</td>
<td>0.802</td>
</tr>
<tr>
<td></td>
<td>(0.797)</td>
<td>(0.771)</td>
</tr>
</tbody>
</table>

| Observations        | 102                        | 102                        |
| R$^2$               | 0.190                      | 0.213                      |
| Adjusted R$^2$      | 0.111                      | 0.136                      |
| Residual Std. Error (df = 92) | 1.100                  | 1.065                      |
| F Statistic (df = 9; 92) | 2.402^{**}             | 2.761^{***}                |

*Note:* $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$
Thus the inverse of the CV, that is $\bar{X}_i / S_i$ can be seen as a consistency measure.
Note: the observational unit is a governor.

C  Topics’ words table

The topics’ words table are often used to interpret the topics. The words are ordered in terms of their relative importance (proportion) with respect to the topic. The following table shows the keywords by topic. A more elaborate tool for interpreting the topics can be found at this link: https://salif.shinyapps.io/topic_context

Topic.1 and Topic.3 remains diffused and difficult to interpret.

D  Web links of economic speeches

http://www.americanrhetoric.com/speeches/gwbuseconomicgrowthpackage.htm
http://www.presidentialrhetoric.com/speeches/02.02.06.html
https://www.whitehouse.gov/the-press-office/2012/06/14/remarks-president-economy-cleveland
https://www.whitehouse.gov/the-press-office/2013/12/04/remarks-president-economic-mobility
http://www.americanrhetoric.com/speeches/barackobama/barackobamacarnegiemellon.htm
http://www.p2016.org/bush/bush090915sp.html
http://mittromneycentral.com/speeches/2012-speeches/102612-remarks-on-the-american-economy
Table 13: Topics’ words table

<table>
<thead>
<tr>
<th>Topic.1</th>
<th>Topic.2</th>
<th>Topic.3</th>
<th>Topic.4</th>
<th>Topic.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 state</td>
<td>school</td>
<td>will</td>
<td>will</td>
<td>state</td>
</tr>
<tr>
<td>2 peopl</td>
<td>educ</td>
<td>state</td>
<td>state</td>
<td>fund</td>
</tr>
<tr>
<td>3 year</td>
<td>year</td>
<td>must</td>
<td>work</td>
<td>will</td>
</tr>
<tr>
<td>4 want</td>
<td>student</td>
<td>govern</td>
<td>busi</td>
<td>year</td>
</tr>
<tr>
<td>5 know</td>
<td>children</td>
<td>time</td>
<td>creat</td>
<td>budget</td>
</tr>
<tr>
<td>6 just</td>
<td>teacher</td>
<td>make</td>
<td>energi</td>
<td>million</td>
</tr>
<tr>
<td>7 work</td>
<td>health</td>
<td>budget</td>
<td>develop</td>
<td>increas</td>
</tr>
<tr>
<td>8 thing</td>
<td>care</td>
<td>futur</td>
<td>also</td>
<td>propos</td>
</tr>
<tr>
<td>9 make</td>
<td>everi</td>
<td>work</td>
<td>nation</td>
<td>also</td>
</tr>
<tr>
<td>10 need</td>
<td>help</td>
<td>challeng</td>
<td>build</td>
<td>program</td>
</tr>
<tr>
<td>11 come</td>
<td>make</td>
<td>econom</td>
<td>compani</td>
<td>servic</td>
</tr>
<tr>
<td>12 time</td>
<td>will</td>
<td>economi</td>
<td>help</td>
<td>revenu</td>
</tr>
<tr>
<td>13 like</td>
<td>program</td>
<td>face</td>
<td>invest</td>
<td>provid</td>
</tr>
<tr>
<td>14 last</td>
<td>famili</td>
<td>reform</td>
<td>economi</td>
<td>public</td>
</tr>
<tr>
<td>15 south</td>
<td>high</td>
<td>public</td>
<td>peopl</td>
<td>spend</td>
</tr>
<tr>
<td>16 money</td>
<td>invest</td>
<td>today</td>
<td>opportun</td>
<td>continu</td>
</tr>
<tr>
<td>17 great</td>
<td>colleg</td>
<td>citizen</td>
<td>econom</td>
<td>need</td>
</tr>
<tr>
<td>18 thank</td>
<td>state</td>
<td>famili</td>
<td>make</td>
<td>percent</td>
</tr>
<tr>
<td>19 good</td>
<td>work</td>
<td>busi</td>
<td>industri</td>
<td>addit</td>
</tr>
<tr>
<td>20 back</td>
<td>need</td>
<td>togeth</td>
<td>year</td>
<td>dollar</td>
</tr>
<tr>
<td>21 think</td>
<td>last</td>
<td>protect</td>
<td>futur</td>
<td>educ</td>
</tr>
<tr>
<td>22 said</td>
<td>parent</td>
<td>need</td>
<td>provid</td>
<td>current</td>
</tr>
<tr>
<td>23 look</td>
<td>learn</td>
<td>nation</td>
<td>thank</td>
<td>feder</td>
</tr>
<tr>
<td>24 take</td>
<td>child</td>
<td>everi</td>
<td>communiti</td>
<td>legisl</td>
</tr>
<tr>
<td>25 chang</td>
<td>system</td>
<td>governor</td>
<td>effort</td>
<td>support</td>
</tr>
<tr>
<td>26 govern</td>
<td>communiti</td>
<td>peopl</td>
<td>west</td>
<td>depart</td>
</tr>
<tr>
<td>27 mani</td>
<td>first</td>
<td>educ</td>
<td>technolog</td>
<td>govern</td>
</tr>
<tr>
<td>28 first</td>
<td>provid</td>
<td>respons</td>
<td>plan</td>
<td>includ</td>
</tr>
<tr>
<td>29 give</td>
<td>insur</td>
<td>servic</td>
<td>million</td>
<td>project</td>
</tr>
<tr>
<td>30 right</td>
<td>nation</td>
<td>live</td>
<td>need</td>
<td>plan</td>
</tr>
<tr>
<td>31 believ</td>
<td>improv</td>
<td>leader</td>
<td>grow</td>
<td>issu</td>
</tr>
<tr>
<td>32 bill</td>
<td>better</td>
<td>serv</td>
<td>must</td>
<td>develop</td>
</tr>
<tr>
<td>33 import</td>
<td>must</td>
<td>come</td>
<td>home</td>
<td>system</td>
</tr>
<tr>
<td>34 better</td>
<td>qualiti</td>
<td>mani</td>
<td>univers</td>
<td>last</td>
</tr>
<tr>
<td>35 that</td>
<td>propos</td>
<td>children</td>
<td>like</td>
<td>cost</td>
</tr>
<tr>
<td>36 talk</td>
<td>read</td>
<td>great</td>
<td>north</td>
<td>administr</td>
</tr>
<tr>
<td>37 start</td>
<td>rais</td>
<td>system</td>
<td>resourc</td>
<td>address</td>
</tr>
<tr>
<td>38 much</td>
<td>opportun</td>
<td>address</td>
<td>today</td>
<td>improv</td>
</tr>
<tr>
<td>39 everi</td>
<td>life</td>
<td>choic</td>
<td>center</td>
<td>local</td>
</tr>
<tr>
<td>40 spend</td>
<td>increas</td>
<td>health</td>
<td>product</td>
<td>recommend</td>
</tr>
</tbody>
</table>
Table 14: Topics’ words table

<table>
<thead>
<tr>
<th>Topic.3</th>
<th>Topic.5</th>
<th>Topic.6</th>
<th>Topic.7</th>
<th>Topic.8</th>
<th>Topic.10</th>
<th>Topic.12</th>
<th>Topic.15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>peopl</td>
<td>must</td>
<td>energi</td>
<td>reform</td>
<td>school</td>
<td>year</td>
<td>health</td>
</tr>
<tr>
<td>2</td>
<td>know</td>
<td>educ</td>
<td>will</td>
<td>state</td>
<td>educ</td>
<td>state</td>
<td>care</td>
</tr>
<tr>
<td>3</td>
<td>thing</td>
<td>today</td>
<td>also</td>
<td>govern</td>
<td>teacher</td>
<td>last</td>
<td>children</td>
</tr>
<tr>
<td>4</td>
<td>just</td>
<td>futur</td>
<td>north</td>
<td>system</td>
<td>student</td>
<td>nation</td>
<td>insur</td>
</tr>
<tr>
<td>5</td>
<td>want</td>
<td>opportun</td>
<td>develop</td>
<td>spend</td>
<td>children</td>
<td>program</td>
<td>famili</td>
</tr>
<tr>
<td>6</td>
<td>come</td>
<td>invest</td>
<td>creat</td>
<td>school</td>
<td>everi</td>
<td>percent</td>
<td>cost</td>
</tr>
<tr>
<td>7</td>
<td>make</td>
<td>innov</td>
<td>peopl</td>
<td>will</td>
<td>need</td>
<td>million</td>
<td>must</td>
</tr>
<tr>
<td>8</td>
<td>time</td>
<td>work</td>
<td>million</td>
<td>chang</td>
<td>will</td>
<td>public</td>
<td>year</td>
</tr>
<tr>
<td>9</td>
<td>think</td>
<td>continu</td>
<td>build</td>
<td>need</td>
<td>learn</td>
<td>increas</td>
<td>work</td>
</tr>
<tr>
<td>10</td>
<td>work</td>
<td>univers</td>
<td>renew</td>
<td>cost</td>
<td>read</td>
<td>sinc</td>
<td>make</td>
</tr>
<tr>
<td>11</td>
<td>said</td>
<td>commit</td>
<td>power</td>
<td>year</td>
<td>high</td>
<td>rate</td>
<td>afford</td>
</tr>
<tr>
<td>12</td>
<td>great</td>
<td>growth</td>
<td>percent</td>
<td>taxpay</td>
<td>help</td>
<td>first</td>
<td>help</td>
</tr>
<tr>
<td>13</td>
<td>like</td>
<td>togeth</td>
<td>compani</td>
<td>incom</td>
<td>classroom</td>
<td>three</td>
<td>protect</td>
</tr>
<tr>
<td>14</td>
<td>need</td>
<td>build</td>
<td>plant</td>
<td>properti</td>
<td>parent</td>
<td>made</td>
<td>provid</td>
</tr>
<tr>
<td>15</td>
<td>back</td>
<td>qualiti</td>
<td>product</td>
<td>billion</td>
<td>child</td>
<td>even</td>
<td>will</td>
</tr>
<tr>
<td>16</td>
<td>give</td>
<td>effort</td>
<td>help</td>
<td>propos</td>
<td>teach</td>
<td>next</td>
<td>access</td>
</tr>
<tr>
<td>17</td>
<td>right</td>
<td>communiti</td>
<td>futur</td>
<td>save</td>
<td>grade</td>
<td>time</td>
<td>need</td>
</tr>
<tr>
<td>18</td>
<td>take</td>
<td>challeng</td>
<td>nation</td>
<td>educ</td>
<td>graduat</td>
<td>four</td>
<td>reduc</td>
</tr>
<tr>
<td>19</td>
<td>tell</td>
<td>economi</td>
<td>produc</td>
<td>better</td>
<td>today</td>
<td>medic</td>
<td>small</td>
</tr>
<tr>
<td>20</td>
<td>money</td>
<td>success</td>
<td>wind</td>
<td>competit</td>
<td>take</td>
<td>still</td>
<td>system</td>
</tr>
<tr>
<td>21</td>
<td>look</td>
<td>life</td>
<td>opportun</td>
<td>busi</td>
<td>first</td>
<td>report</td>
<td>citizen</td>
</tr>
<tr>
<td>22</td>
<td>that</td>
<td>research</td>
<td>industri</td>
<td>administr</td>
<td>program</td>
<td>higher</td>
<td>live</td>
</tr>
<tr>
<td>23</td>
<td>start</td>
<td>prosper</td>
<td>just</td>
<td>pass</td>
<td>provid</td>
<td>histori</td>
<td>drug</td>
</tr>
<tr>
<td>24</td>
<td>talk</td>
<td>center</td>
<td>generat</td>
<td>choic</td>
<td>improv</td>
<td>thank</td>
<td>expand</td>
</tr>
<tr>
<td>25</td>
<td>today</td>
<td>provid</td>
<td>plan</td>
<td>first</td>
<td>math</td>
<td>past</td>
<td>nation</td>
</tr>
<tr>
<td>26</td>
<td>someth</td>
<td>make</td>
<td>electr</td>
<td>lower</td>
<td>colleg</td>
<td>even</td>
<td>program</td>
</tr>
<tr>
<td>27</td>
<td>chang</td>
<td>colleg</td>
<td>school</td>
<td>growth</td>
<td>mani</td>
<td>month</td>
<td>legislatur</td>
</tr>
<tr>
<td>28</td>
<td>them</td>
<td>ensur</td>
<td>need</td>
<td>local</td>
<td>system</td>
<td>number</td>
<td>child</td>
</tr>
<tr>
<td>29</td>
<td>world</td>
<td>state</td>
<td>increas</td>
<td>good</td>
<td>propos</td>
<td>progress</td>
<td>senior</td>
</tr>
<tr>
<td>30</td>
<td>mean</td>
<td>live</td>
<td>fund</td>
<td>long</td>
<td>standard</td>
<td>anoth</td>
<td>communiti</td>
</tr>
<tr>
<td>31</td>
<td>littl</td>
<td>great</td>
<td>fuel</td>
<td>make</td>
<td>make</td>
<td>done</td>
<td>coverag</td>
</tr>
<tr>
<td>32</td>
<td>good</td>
<td>grow</td>
<td>state</td>
<td>pension</td>
<td>time</td>
<td>well</td>
<td>healthi</td>
</tr>
<tr>
<td>33</td>
<td>here</td>
<td>higher</td>
<td>home</td>
<td>privat</td>
<td>money</td>
<td>said</td>
<td>mani</td>
</tr>
<tr>
<td>34</td>
<td>place</td>
<td>technolog</td>
<td>market</td>
<td>rais</td>
<td>work</td>
<td>rank</td>
<td>made</td>
</tr>
<tr>
<td>35</td>
<td>this</td>
<td>econom</td>
<td>provid</td>
<td>real</td>
<td>good</td>
<td>mani</td>
<td>prescript</td>
</tr>
<tr>
<td>36</td>
<td>sure</td>
<td>also</td>
<td>resourc</td>
<td>control</td>
<td>account</td>
<td>high</td>
<td>home</td>
</tr>
<tr>
<td>37</td>
<td>import</td>
<td>best</td>
<td>research</td>
<td>plan</td>
<td>test</td>
<td>special</td>
<td>allow</td>
</tr>
<tr>
<td>38</td>
<td>believ</td>
<td>strong</td>
<td>take</td>
<td>rate</td>
<td>earli</td>
<td>enact</td>
<td>give</td>
</tr>
<tr>
<td>39</td>
<td>live</td>
<td>vision</td>
<td>center</td>
<td>district</td>
<td>rais</td>
<td>pass</td>
<td>increas</td>
</tr>
<tr>
<td>40</td>
<td>life</td>
<td>world</td>
<td>project</td>
<td>mani</td>
<td>meet</td>
<td>initi</td>
<td>challeng</td>
</tr>
</tbody>
</table>