ABATEMENT APPROXIMATIONS: USING ECONOMETRICS TO ESTIMATE COST OF ASBESTOS ABATEMENT

Kori Mayfield
mayf8550@bears.unco.edu

Follow this and additional works at: https://digscholarship.unco.edu/honors

Recommended Citation
Mayfield, Kori, "ABATEMENT APPROXIMATIONS: USING ECONOMETRICS TO ESTIMATE COST OF ASBESTOS ABATEMENT" (2019). Undergraduate Honors Theses. 20.
https://digscholarship.unco.edu/honors/20
ABATEMENT APPROXIMATIONS: USING ECONOMETRICS TO ESTIMATE COST OF ASBESTOS ABATEMENT

A Thesis/Capstone
Submitted in Partial Fulfilment for Graduation with Honors Distinction and the Degree of Bachelor of Arts

Kori Mayfield

College of Humanities and Social Sciences

AUGUST 2019
Abstract

The continued success and operation of a business requires that there be steady streams of income and that those incomes are greater than the business’s expenses. In the asbestos abatement industry, companies must provide accurate estimations of the price associated with a job in order to receive the contract to complete the work and maintain a profitable margin of invoiceable amount over labor and supply costs. Overestimating prices can result in lost work and underestimation can decrease or eliminate profit margins. For this reason, accuracy is extremely important. This is especially true for small businesses where profit margins are already quite tight. In these companies, a full schedule of work is required to keep the company running. Estimations are generally completed by human estimators drawing on personal field experience. This applied research project attempts to streamline this process by utilizing econometric analysis to develop a model to estimate these costs. Data for this project has been obtained by collecting information from 210 different jobs completed by a single firm located in the state of Colorado. The data has been run through estimation software and a model has been created using econometrics based linear regressions. This new model will be tested for statistical significance. The accuracy of the final model will be tested by comparing the model’s results with those of an estimating technique already in use at the firm. It is intended that the resulting model will be able to accurately estimate costs early on in the project process, thereby helping to improve consistency of profit margins and uniformity of estimations across human estimators.
Table of Contents

Introduction ........................................................................................................................................... 1

Literature Review ................................................................................................................................. 2

The Cost of Asbestos to Property Owners ......................................................................................... 4

Existing Asbestos Cost Research ......................................................................................................... 5

Variables Impacting Cost ..................................................................................................................... 6

Cost Prediction Models ......................................................................................................................... 9

Project Design ..................................................................................................................................... 11

Methodology ......................................................................................................................................... 11

Theoretical Model ................................................................................................................................. 13

Hypotheses ........................................................................................................................................... 14

Descriptive Statistics ............................................................................................................................. 16

Data for All Jobs ................................................................................................................................. 16

Material and Quantity Data ................................................................................................................ 17

Empirical Analysis ............................................................................................................................... 19

Testing for Multicollinearity ................................................................................................................. 19

Testing for Heteroskedasticity ............................................................................................................. 20

Establishing a Model ............................................................................................................................ 21

Discussion of Results ........................................................................................................................... 22
Introduction

For project-based industries, being able to properly estimate the cost of executing and completing a project can be integral to ensuring revenue and profitability. The asbestos abatement industry is one example of a project-based industry. When competing for business with other firms, a company often has only one opportunity to present a proposal including a cost bid to a prospective client. The proposal, if chosen, is then signed and becomes a contract between the client and the firm performing abatement work. For this reason, it is essential that initial cost estimations be accurate. Most clients will choose a firm based on the company that presents the smallest bid, so it does not serve a company to provide margin in their bid. However, firms must be careful to not underbid to the point of profit loss.

As firms and project managers gain experience, they often develop their own estimation systems. Built into these models are firm specific factors as well as years of experience. Each estimation model then is specific to the estimator that created it. However, all models evaluate the costs of the same set of core elements of an abatement job.

This applied research project attempts to create a cost estimation model for asbestos abatement jobs done by a single firm located in Colorado. The outcome of the project is expected to be a linear equation that can be used with proper and complete job data to predict overall cost. This research project will attempt to answer the following question: Can econometric regressions be used to estimate the cost of an asbestos abatement job?
In order to answer that question, a linear model will be created using econometric and statistical analysis. The construction of this model is based on data from the firm as well as from the other models and variables found through the literature review.

**Literature Review**

For many years, asbestos was used in a variety of materials as a way to strengthen, fireproof or soundproof. These materials include insulation, shingles, pipes, drywall, water supply lines, fire blankets, flooring, glues and several other elements (World Health Organization, 2014). During the 20th century, use of asbestos was common in many industries including construction and automotive manufacturing. As of 2010 there were over 3000 patented uses for asbestos in the European Union alone even though its use was banned in 1999 (Gregson, Watkins, & Calestani, 2010). Asbestos presents itself in a variety of forms the most common of which are chrysotile (white) and crocidolite (blue). Regardless of the type or use, all forms of asbestos have been found to be carcinogenic to humans. Exposure to asbestos can cause cancer in the lungs, larynx, and ovaries as well as mesothelioma, a specific cancer in the lining of the abdominal cavity and outer lungs. In addition to these cancers, other medical issues such as asbestosis can occur due to asbestos exposure (World Health Organization, 2014).

Current estimates suggest that over 100,000 people die each year from lung cancer, mesothelioma or asbestosis that can be traced back to occupational exposure to asbestos. Non-occupational exposure is responsible for nearly 400 deaths. “Exposure to asbestos occurs through inhalation of fibers primarily from contaminated air in the working environment, as well as from ambient air in the vicinity of point sources or indoor air in housing and buildings containing friable asbestos materials” (World Health Organization,
Asbestos materials used in building materials are generally not a risk factor when left alone, however the fibers become risk factors when they are disturbed due to maintenance, removal, alteration or demolition of buildings. Damage done to buildings during natural disasters can also cause previously stagnant asbestos to become friable and dangerous (World Health Organization, 2014).

To mitigate the risk of exposure, asbestos abatement is performed to remove asbestos containing materials from buildings and other infrastructure. The removal of asbestos containing materials (ACM) can often make the materials friable and introduce a greater risk of exposure than before work began, for this reason, research suggests that abatement only be performed when absolutely necessary. Abatement workers are highly susceptible to being exposed to asbestos fibers and therefore must undergo a lengthy and costly training process. The research suggests that this training is not always effective and quickly forgotten. Further argument suggests that improper removal practices can lead to environmental exposures that can be harmful to future tenants of abated buildings. This creates long run societal health problems that may be avoided by leaving ACM alone rather than removing it (Esmen, 1991). This report on the quality if removal practices was published in 1991 shortly after the implementation of the Asbestos Hazard Emergency Response Act of 1986 also known as AHERA. AHERA outlined training and qualification practices for asbestos inspectors, workers, planners, supervisors and project designers. In order to become a certified abatement professional of any type one must undergo an initial training as well as yearly refresher courses. Courses must be approved and accredited by the EPA and the state in which they are conducted. In addition to courses there are examinations done on the state level to ensure the
competence of an individual before they are licensed to work in the industry (Asbestos hazard emergency response act of 1986, 1986). Perhaps the issues suggested in Esmen’s report were the result of kinks being worked out in this new legislation or were developed based on work done before the legislation was enacted.

The Cost of Asbestos to Property Owners

The presence of ACM in buildings can be costly for property owners. Average losses to homeowners trying to sell a home containing asbestos, as found in an Alabama based study are about 13.44%. The properties examined in the study showed those cost to average out to about $25,300 per property (Affuso, Cummings, & Le, 2018). This is the amount homeowners generally have to reduce asking price by when selling a home with ACMs. The study did not factor in the actual cost of abatement. Presumably, this cost will be assumed by the future owner upon purchase, however if the future owner does not do any renovations, or experience any disasters such as fire or flood, they will likely not have the materials removed. When homeowners do consider whether or not to pursue a project that will require professional abatement, they weigh the cost of the project including the abatement against the value it will add to the home. When asbestos remediation costs are compared with the effect of asbestos on property values, it is found that remediation is most optimally performed at certain times. Specifically, if waiting is an option, it is best for homeowners to wait until the value of the property is 23% higher with remediation than without (Espinoza & Luccioni, 2007). Proper timing of a project relies then on having accurate cost information in regard to the abatement process. Thus, it is important be able to properly estimate the cost of the asbestos removal from the early stages of a project.
Existing Asbestos Cost Research

Existing research into asbestos abatement costs have been comparative in nature. Two studies into the costs of asbestos abatement in public buildings in Erie County, Pennsylvania compared the costs of different projects to establish trends and patterns in the abatement industry. The original study examined data from 1992-1994 and found unsurprisingly that larger jobs, where the quantity of asbestos to be removed was higher, lead to larger costs. They also saw an increase in the number of jobs being done suggesting industry growth (Lange, J. H. et al., 1996). The second study revisited the same area looking at abatement jobs done between 1996 and 1999. The data collected confirmed previous findings that larger jobs lead to larger costs. This time however, it seemed that the quantity of projects being done was decreasing (Lange, Bules, Lindquist, Gray, & Ivarone, 2000). While this comparison can be beneficial in discovering trends, it does little to tell us the exact components that impact the costs of asbestos abatement. It compared jobs only based on the amount of material to be used and looked only at a small county in Pennsylvania. This produced a small data set of only 35 jobs total over the two studies. This small sample size combined with the age of the studies makes it hard to expand Lange’s finding to use in estimating cost in 2019.

The same researcher who spearheaded the studies into industry patterns and trends, also analyzed bid data for publicly funded projects in the same geographic area. He compared 15 different project bids and found that while the bids for one project were normally distributed, if compared across projects, bids have a non-normal distribution. The study deals with bids for public projects as that data is a matter of public record and unlike private industry, cost are not adjustable once a bid has been accepted (Lange,
Thus, the bid here is equivalent to the final cost of the project. Importantly, this study suggests that costs are fairly standard across firms for a single project, lending to the idea that a model can be created to estimate cost. It also indicates variation from one job to another as caused by differences in project details.

**Variables Impacting Cost**

There are a wide variety of variables that impact the cost of asbestos remediation. While little research has been done in regard to asbestos specifically, some of the variables used in general construction cost estimation can be carried over. They type of project, material cost, likelihood of changes, environmental conditions, size of project, and type of client are all factors to be considered for construction project costs. These factors are all project specific, they are will vary from project to project but are generally easy to measure and observe. There is another class of factors that influence pricing that are not as easy to quantify. These are what researchers consider estimator specific factors. This category of factors includes all pricing decisions made by the specific estimator based on experience, expectation, relationships, or regionally accepted pricing culture (Elfaki, Alatawi, & Abushandi, 2014). These same conditions and factors found in the estimation for construction jobs can also be found in asbestos jobs.

Another factor used in cost estimation in every industry is inflation ratios. These ratios make it possible to more accurately use past data to predict costs for projects in the present and future. Inflation ratios allow for standardization of monetary values across different time period by accounting for economic changes (Remer & Mattos, 2003). The more cases used to create a model, the more accurate it will be, this can require that a
model is produced using cases from varying time periods. Incorporating inflation ratios can reduce any bias that may be inserted into the model by external economic changes.

The time that it takes to complete a project is integral to cost estimation as man hours are one of the more expensive components of cost calculation. While true in all construction subsects, this is particularly true in asbestos removal as workers have to be highly trained and therefor garner high wages (Gregson et al., 2010). In one particular instance by creatively constructing work areas so that abatement could safely occur during business hours, thus shortening total work time and avoiding overtime, the cost of a New York Subway stop renovation was reduced by over $35 million (Zanoni & Singh, 2003). This massive savings displays the impact that man hours have on the total cost of an abatement project.

Other variables to consider come from the processes directly involved with abatement. This includes costs related to the materials used to package ACM and the work area containments used to increase safety during remediation. Standard practice requires that the materials be wetted and packaged securely then disposed of in specific and approved locations. To maintain environmental health standards, workers must be highly trained and don protective yet disposable clothing and filtration systems. (Gregson et al., 2010). The amount to materials used in packaging waste and protecting workers therefore an important element to consider in the process of estimating abatement costs. The protective materials required to complete asbestos abatement can quickly add up contributing largely to overall cost.

The most important variable is the quantity and type of ACM to be removed. Lange’s previously sited comparative studies look at the cost implications of quantity of
materials. In both the 1996 and 2000 studies it was found that projects in which there were greater amounts of material to be removed the overall cost of the abatement was larger (Lange et al., 1996). Lange’s reports however do not address the specific type of materials removed.

Information about type and quantity of material is most commonly acquired through asbestos surveys which present challenges of their own. Inaccuracy on asbestos surveys can lead to increased costs through the detection of previously unidentified ACM, adding to scope of work. This can slow down projects and have great impact on cost. In times of economic fluctuation, this imprecision of surveys can lead to not only small profit decreases but in complete loss or negative profit to property owners and contractors in the business of demolition for parts (Gregson et al., 2010). A study conducted in 2007 and published in the journal Cost Engineering looked at the relationship between the type of initial survey completed and the variation between estimation and final cost for 20 abatement projects done at “nondomestic” properties managed by a single property management firm in London. The survey included 10 projects that started with a Semi-Intrusive Survey and 10 that started with an Intrusive survey, noting that the intrusive survey was more expensive at the outset but supplied greater amounts of initial information. The study found not only that the presence of greater amounts of initial information as provided by more in-depth surveys could lower overall asbestos removal costs but also that in projects started based on the semi-intrusive surveys there was an average of 12.7% variation between initial estimates and final costs. There was no variation between estimate and final cost for any of the jobs done based on
an intrusive study (Kupakuwana, 2007). The findings of this study exemplify the need for proper and thorough studies in cost estimation.

Cost Prediction Models

A variety of models have been used to estimate costs in industries related to asbestos abatement. Many studies present and evaluate different cost estimation models for other construction related jobs. This evaluation helps to recognize the effectiveness of different methodologies and identify those that will be beneficial in the process of estimating the cost of an asbestos abatement job.

Recent advancements in technology have led to the emergence of artificial intelligence-based cost estimation models. There are a variety of these intelligent techniques used in the construction industry. Most commonly used are Knowledge Based and Machine Learning techniques. These techniques are effective and accurate based on their ability to continuously improve upon themselves from their increased use. This means that the more a specific intelligence-based model is used, the more accurate it becomes. (Elfaki et al., 2014). One particular machine learning model was created to examine deconstruction and demolition. It was built to read a database of past projects and estimate costs. It then built its estimates into its database and continued adapting and evolving getting increasingly accurate as it estimated more jobs (Tatiya, Zhao, Syal, Berghorn, & LaMore, 2018). The accuracy of these models is notable making them an intriguing exploration. However, the cost and technology associated with developing and programing an artificial intelligence-based system are not justifiable to the small businesses that mainly perform this type of work.
Artificial intelligence isn’t the only method in use in the academic evaluation of models used to estimate job costs. In an effort to better estimate the costs of road construction projects, one group of researchers used a parametric model that transformed data using logarithms and reciprocals. This allowed researchers to minimize heteroskedasticity and bias while maintaining the use of a linear estimation. The specific variables that researchers transformed were the bid quantity of concrete, and number of bidders for a project. Their study also included dummy variables used to represent different locations. (Swei, Gregory, & Kirchain, 2017). This linear regression model showed strong results.

To estimate the costs of tunnel production, researchers compiled data from past projects into a database and then used statistical analysis to create linear models to estimate the costs of different kinds of tunnel production. They were then able to program their results into estimation software where estimators can input the length and diameter of a tunnel project as well as the type of excavation and application of the tunnel and will receive back an approximate cost. While their model is found to be accurate for midsize jobs, the researchers suggest caution be exercised in regard to extremely large or extremely small jobs (Rostami, Sepehrmanesh, Gharahbagh, & Mojtabai, 2013). Notably this method is similar to that currently in use to estimate asbestos costs in terms of input variables needing to be project specific. The linear models created through statistical analysis are not currently in use in the asbestos industry but worth exploring.

In Summary, there are many different methodologies at play in cost estimation. No matter which model is in use, the most important element is correctly identifying and measuring the different components of job cost. These factors will have direct impact on
cost and if they are improperly explored and specified, any model created will be ineffective. Once the components are correctly outlined the importance falls to the relationships between those components and the monetary cost of the project.

Little research has been published in the field of asbestos cost estimation specifically. The studies that do evaluate asbestos costs, look at the costs comparatively rather than at creating a model to estimate them. There is a need in the field for a study that fully examines the variables associated with abatement and how they contribute to overall cost.

**Project Design**

The purpose of this project is to create a linear regression model to accurately estimate the cost of an asbestos abatement model. The model will be constructed using methods from econometric analysis. The model will then be tested for statistical significance and evaluated for overall fit and accuracy.

**Methodology**

A large portion of the work to develop a cost prediction model will involve the use econometric analysis to establish the trends in costs. The econometric analysis will consist of completing an OLS regression analysis on a data set of different project costs. This type of regression will be an effective tool for cost analysis because it can break the overall cost down into the degree by which individual pieces contribute to that cost.

Data collected was cross sectional in nature. It was collected through thorough examination of job files of previous jobs done by the company and stored in an excel spreadsheet saved to the company’s cloud storage system so that was accessible both in
the office where data will be collected and outside the office for analysis. The data was also backed up to a flash drive kept with the researcher in case of a digital system failure.

The model evaluates 209 cases. All relevant cases were considered but cases with incomplete information for all variables were not included so as not to bias the results of the analysis. The database of job information was built by starting with the newest possible cases and was then added to in reverse chronological order so as to create the largest possible set without having too much cost variation based on time or inflation.

The independent variable in the study will be the cost of the job. The final list of dependent variables includes the asbestos containing material to be removed, the quantity of that material, and the amount of time needed to complete the project. Other variables included can be found below in the theoretical model.

The control variable in the analysis is that all the cases come from the same company. This eliminates any difference in profit margin or worker pay rates across firms. This also standardizes for the processes done during the job as there are general health and safety practices that establish general procedure, but specifics may vary across firms.

The data found was run through a series of tests to ensure that the proposed model met the classical assumptions. It was tested for multicollinearity and heteroskedasticity. By ensuring that the model meets the classical assumptions that ensures that the OLS estimator will be the best estimator and will be unbiased.

Once the model was estimated each estimated coefficient was rigorously tested for statistical significance. This is done through t-tests, theory comparison, and evaluation
of the coefficient’s effect on the overall fit of the model. Once the estimations are shown to be statistically relevant, a final model will be constructed.

**Theoretical Model**

\[ COST_i = \beta_0 + \beta_1 TXT_i + \beta_2 FLR_i + \beta_3 VAT_i + \beta_4 MAST_i + \beta_5 L_i + \beta_6 S_i + \beta_7 C_i + \beta_8 DAY_i \]
\[ + \beta_9 EXT_i + \beta_{10} SCHL_i + \beta_{11} EMER_i + \beta_{12} ES1_i + \beta_{13} ES2_i + \beta_{14} ES3_i \]
\[ + \beta_{15} ES4_i + \beta_{16} NCO_i \]  
(Eq. 1)

*Where:*

- \( COST_i \) = The total invoice cost, in dollars, of the \( i \)th asbestos abatement job
- \( TXT_i \) = The quantity, in square yards of texture or drywall to be removed during the \( i \)th asbestos abatement job
- \( FLR_i \) = The quantity, in square yards, of sheet flooring (vinyl or laminate) to be removed during the \( i \)th asbestos abatement job
- \( VAT_i \) = The quantity, in square yards, of vinyl asbestos tiling to be removed during the \( i \)th asbestos abatement job
- \( MAST_i \) = The quantity, in square yards, of mastic or other adhesive to be removed during the \( i \)th asbestos abatement job
- \( L_i \) = The quantity, in linear yards, of other material to be removed during the \( i \)th asbestos abatement job
- \( S_i \) = The quantity, in square yards, of other material to be removed during the \( i \)th asbestos abatement job
- \( C_i \) = The quantity, in cubic yards, of other material to be removed during the \( i \)th asbestos abatement job
- \( DAY_i \) = The number of days required to complete the \( i \)th asbestos abatement job
\[ EXT_i = \begin{cases} 1 & \text{if any portion of the } i\text{th asbestos} \\
& \text{abatement job is performed on the exterior of a building and zero otherwise} \\
0 & \text{otherwise} 
\end{cases} \]

\[ SCHL_i = \begin{cases} 1 & \text{if the } i\text{th asbestos abatement job is} \\
& \text{performed for a school district or other educational institution and zero if otherwise.} \\
0 & \text{otherwise.} 
\end{cases} \]

\[ EMER_i = \begin{cases} 1 & \text{if any portion the } i\text{th asbestos abatement job is defined as an emergency job and zero if otherwise} \\
0 & \text{otherwise.} 
\end{cases} \]

\[ ES1_i = \begin{cases} 1 & \text{if the job was originally quoted by} \\
& \text{Estimator 1 and zero if otherwise.} \\
0 & \text{otherwise.} 
\end{cases} \]

\[ ES2_i = \begin{cases} 1 & \text{if the job was originally quoted by} \\
& \text{Estimator 2 and zero if otherwise.} \\
0 & \text{otherwise.} 
\end{cases} \]

\[ ES3_i = \begin{cases} 1 & \text{if the job was originally quoted by} \\
& \text{Estimator 3 and zero if otherwise.} \\
0 & \text{otherwise.} 
\end{cases} \]

\[ ES4_i = \begin{cases} 1 & \text{if the job was originally quoted by} \\
& \text{Estimator 4 and zero if otherwise.} \\
0 & \text{otherwise.} 
\end{cases} \]

\[ NCO_i = \begin{cases} 1 & \text{if the job was done at a location outside} \\
& \text{and zero if otherwise. of Colorado} \\
0 & \text{otherwise.} 
\end{cases} \]

**Hypotheses**

Based on the cost models presented in the literature the following hypotheses have been developed on how the variables will contribute to the cost of an abatement job.

Nearly all the variables in the model will increase the cost of asbestos abatement. This includes, \( TXT_i, FLR_i, VAT_i, MAST_i, L_i, S_i, C_i, DAY_i \), and \( EXT_i \). Previous
examination of asbestos abatement jobs suggests that greater quantities of material to be removed results in greater costs so for the different types of materials, the coefficients should be positive. It is fairly obvious that projects that take longer to complete will cost more. This is due to increasing labor costs, thus the variable representing days will also have a positive coefficient. Work occurring on the exterior of a building requires more complex containments this means that materials cost will increase. This encompasses the first nine variables as all hypothesized to be positive. They will be tested for significance at a 5% significance level using the following test.

\[
H_0: \beta_{1-9} \leq 0
\]
\[
H_A: \beta_{1-9} > 0
\]

Where \(H_0\) is the null hypothesis that the variable is either insignificant or not representing with the expected sign and \(H_A\) is the alternative hypothesis that the variable is positive and significant. The null hypothesis is rejected if the t-score associated with the variable is greater than the critical t-score calculated based on the size of the sample collected, and the coefficient is positive.

Jobs at schools require specific measures in regard to safety and record keeping this may increase costs. In addition to increased safety protocols, schools must be worked on during times when students are not present, therefor firms have to work harder to fit them into schedules and therefor charge a higher rate. However, the bidding process for school jobs is competitive, this may decrease costs in that firms have to bid lower to get the work. Emergency jobs are often smaller, but they require an expedited process so it is unclear whether the characteristic of being an emergency will increase or decrease costs. For those reasons, the \(SCHL_i\) and \(EMER_i\) will be tested using a two-tailed test where the
null hypothesis is that the variable is insignificant and the alternative hypothesis is that
the variable is significant in either the positive or negative direction. Similar tests will be
done for each of the variables that represent the different estimators as well as out of state
jobs.

$$H_0: \beta_{10-16} = 0$$

$$H_A: \beta_{10-16} \neq 0$$

The null hypothesis is rejected as long as the t-score of the variable is greater than the
critical t-value.

**Descriptive Statistics**

Because of the nature of the data, the descriptive statistics section has been broken into
two separate subsections, one with data for all jobs, and one with data for the materials
and estimators.

**Data for All Jobs**

The following table presents the data for variables that impact all jobs including
cost, number of days and all dummy variables.

*Table 1: Table of Variables in All Cases*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COST</strong></td>
<td>$9,452.47</td>
<td>$5,513.50</td>
<td>$87,822.00</td>
<td>$582.00</td>
<td>$87,240.00</td>
</tr>
<tr>
<td><strong>DAY</strong></td>
<td>2.97</td>
<td>2</td>
<td>15</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td><strong>EXT</strong></td>
<td>0.11</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>EMER</strong></td>
<td>0.05</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>SCHL</strong></td>
<td>0.19</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>NCO</strong></td>
<td>0.03</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>ES1</strong></td>
<td>0.11</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>ES2</strong></td>
<td>0.20</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>ES3</strong></td>
<td>0.45</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>ES4</strong></td>
<td>0.04</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

The average cost of projects examined for this study was $9,452.47 this is over
$3,900 more than the median of $5,513.50 suggesting that the data set is skewed to the
right with more higher cost jobs than lower cost. The highest invoice price for a job was $87,882 and the lowest price was $582. The resulting range of job costs is $87,240.

The longest job examined in this study was 15 days, there were 61 jobs in the data set that lasted only one day. The median number of days was 2 and the average was fairly close at 2.97. 89% of jobs included in this study lasted 5 or fewer days.

The average value of the EXT variable is 0.11 suggesting that 11% of jobs included work on the exterior of the building. Other dummy variable averages can be used to describe the data set similarly. Based on these averages we see that only 5% of jobs were considered emergency jobs 19% of jobs were done at schools and only 3% of the jobs occurred at sites outside of Colorado.

The last four dummy variables are representative of project managers who have worked for the firm at any time during which jobs in the study were created. Their averages give the percentage of jobs bid by each of them. Estimator 4 has the lowest percentage at only 4% (9 jobs). Estimator 3 bid 94 of the jobs giving him 45% of the total. Estimator 2 who started at the company in early 2018 bid 11% of the jobs examined and Estimator 1 bid 20% with 23 and 43 jobs respectively. The omitted condition of this set represents the company owner who bid the remaining 20% of jobs.

**Material and Quantity Data**

Each job contains quantities of only selected materials. 46% of jobs involved the removal of more than one type of material. The following displays data for each type of material present in the database of jobs.

Texture or drywall was removed from 82 jobs in the study making up 39% of the total. The average amount removed from these jobs was 1147.32 square feet and the
median was 574 square feet indicating a skew to the right. The largest amount removed was 9600 square feet and the smallest was 1 square foot.

Sheet flooring was removed from 27% of jobs. The largest amount was 13,216 square feet and the smallest was 2 square feet. The median amount removed was 144 square feet and the mean was 448.09 square feet.

The maximum and minimum amounts for mastic and flooring are identical at 13,216 and 2 square feet and those numbers come from the same jobs. This makes sense as mastic is the adhesive used to secure flooring to a base. However, mastic is not always present with floor tile so the two amounts are not identical for every job. Interestingly, mastic is also present in 27% of jobs but they are not all the same jobs. The average amount of mastic removed is much larger than the average amount of sheet flooring at 1158.68 square feet and the median is 440 square feet.

Mastic can also be attached to tile though again one is not always indicative of the other. The average amount of tile removed was 898.48 square feet and the median was 383 square feet. The maximum amount removed from any one job was 5688 square feet and the minimum was 3 square feet.

Materials removed less frequently than the four above were combined into other materials in either linear, square, or cubic feet. 10% of jobs removed items measured in linear square feet (usually TSI or Thermal System Insulation). The minimum amount of linear feet removed was 1 and the maximum was 747. The median amount was 20 linear feet and the average was 83.10. Square feet of other material were most commonly transite siding panels and consist of 21% of all jobs. The mean amount of miscellaneous square foot materials was 852.06 square feet and the median was 126.5 square feet. There
was a range of 8197.8 square feet between the maximum of 8198 square feet and the minimum of 0.2 square feet. Material measured in cubic feet is not as commonly removed and therefore appears in only 5% of the jobs in this study. Most common material removed in cubic feet is contaminated soil. The median amount of material removed was 15 cubic yards while the mean was 272.80. The largest amount was 1125 cubic yards and the smallest was 1 cubic yard.

Table 2: Variables in Select Cases

<table>
<thead>
<tr>
<th>Variable</th>
<th>Count</th>
<th>Percentage of Total</th>
<th>Mean</th>
<th>Median</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>TXT</td>
<td>82</td>
<td>39%</td>
<td>1147.32</td>
<td>574</td>
<td>9600</td>
<td>1</td>
<td>9599</td>
</tr>
<tr>
<td>FLR</td>
<td>57</td>
<td>27%</td>
<td>448.09</td>
<td>144</td>
<td>13216</td>
<td>2</td>
<td>13124</td>
</tr>
<tr>
<td>MAST</td>
<td>57</td>
<td>27%</td>
<td>1158.68</td>
<td>440</td>
<td>13216</td>
<td>2</td>
<td>13214</td>
</tr>
<tr>
<td>VAT</td>
<td>66</td>
<td>31%</td>
<td>898.48</td>
<td>383</td>
<td>5688</td>
<td>3</td>
<td>5685</td>
</tr>
<tr>
<td>L</td>
<td>21</td>
<td>10%</td>
<td>83.10</td>
<td>20</td>
<td>747</td>
<td>1</td>
<td>746</td>
</tr>
<tr>
<td>S</td>
<td>44</td>
<td>21%</td>
<td>852.06</td>
<td>126.5</td>
<td>8198</td>
<td>0.2</td>
<td>8197.8</td>
</tr>
<tr>
<td>C</td>
<td>10</td>
<td>5%</td>
<td>272.80</td>
<td>15</td>
<td>1125</td>
<td>1</td>
<td>1124</td>
</tr>
</tbody>
</table>

**Empirical Analysis**

Before an estimation is made it should be confirmed that the model is suited to linear estimation. To do this the model is tested for multicollinearity and heteroskedasticity.

**Testing for Multicollinearity**

Initial test of multicollinearity utilizes a simple correlation analysis (Shown in Appendix A). The results of this correlation test show that there is no strong correlation at a 0.8 threshold. The strongest apparent correlation is 0.72 between cost and number of days.

The second test for multicollinearity is the variance inflation factor (VIF) test. This measures the effect of multicollinearity on the model’s estimated variance. VIF factors above 5 are generally viewed as representations of high multicollinearity. As seen
in the VIF table below, the mastic variable (MAST) has a VIF value of 5.41. This is something to keep a close eye on. All other VIF factors are free of concern.

<table>
<thead>
<tr>
<th>Variable</th>
<th>TXT</th>
<th>FLR</th>
<th>VAT</th>
<th>MAST</th>
<th>T</th>
<th>S</th>
<th>C</th>
<th>DAY</th>
<th>L</th>
<th>S</th>
<th>C</th>
<th>DAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIF</td>
<td>1.89</td>
<td>4.32</td>
<td>2.89</td>
<td>5.41</td>
<td>1.10</td>
<td>1.08</td>
<td>2.28</td>
<td>1.14</td>
<td>1.40</td>
<td>1.87</td>
<td>1.58</td>
<td>2.09</td>
</tr>
</tbody>
</table>

Testing for Heteroskedasticity

Next the model will be tested for heteroskedasticity to ensure that is well specified and there is little correlation between error terms. Some degree of heteroskedasticity if expected due to the cross sectional nature of the data set.

The first test conducted is a Breusch-Pagan test which regresses the original independent variables but uses the squared residuals of the first estimation as the dependent variable. The Breusch-Pagan test for this model uses the following regression equation.

\[
e^{2}_i = \alpha_0 + \alpha_1 TXT_i + \alpha_2 FLR_i + \alpha_3 VAT_i + \alpha_4 MAST_i + \alpha_5 L_i + \alpha_6 S_i + \alpha_7 C_i + \alpha_8 DAY_i + \alpha_9 EXT_i + \alpha_{10} SCHL_i + \alpha_{11} EMER_i + \alpha_{12} ES1_i + \alpha_{13} ES2_i + \alpha_{14} ES3_i + \alpha_{15} ES4_i + \alpha_{16} NCO_i + u_i \quad (Eq. 2)
\]

Where:

- \(e^{2}_i\) = The squared residuals of the estimated regression equation derived from the data set and Eq. 1
- \(TXT_i - NCO_i\) = Variables as defined in Eq 1.

The hypothesis to be tested is

\[H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_9 = 0\]

\[H_A: H_0 \text{ is false}\]
This hypothesis is testing using a chi-squared test at the 5% significance level with a degree of freedom of 16. The critical chi-squared value is 26.296. From the regression of Eq 2 (Appendix B) we get an $NR^2$ value of 24.4321. Because the observed $NR^2$ value is less than the critical chi squared value, the null hypothesis is not rejected, and the model is assumed to be homoscedastic.

The second test for heteroskedasticity is the White Test. This test goes deeper than the previous test by regressing the residuals of the estimation using as independent variables, the explanatory variables, their squares and their cross products. It uses the same hypothesis as the Breusch-Pagan test but many more variables. The critical chi-squared for the white test on this model comes from 5% significance and 113 degrees of freedom. This chi-squared value is 138.811. The white test run in Stata (Appendix C) gives a calculated chi-squared of 101.76 so we do not reject the null hypothesis and assume a sufficient level of homoskedasticity.

**Establishing a Model**

Since the model has been tested for multicollinearity and heteroskedasticity and no major problems have been detected, an initial model has been constructed. Using the model specified in Equation 1 and the data explored above, the following estimated regression equation (Equation 3) was found. The rest of this paper will explore the meaning and significance of this result.
\begin{equation}
\text{COST}_i = -529.901 + 1.3417XT_i^{**} + 1.991FLR_i^{**} + 1.775VAT_i^{**} + 0.754MAST_i + 0.63L_i
\end{equation}

\begin{align*}
t & = & (2.10, 1.71, 1.76, 0.74, 0.07) \\
+ & 0.271S_i + 19.716C_i^* + 2970.14DAY_i^* + 1639.852EXT_i - 1979.65SCHL_i \\
& (9.6209, 0.7699, 5.0937, 333.7806, 1784.576) \\
& 0.35, 3.87, 8.90, 0.92, -1.23 \\
- & 472.8EMER_i - 2106.774ES1_i - 4712.856ES2_i^{**} + 699.8774ES3_i \\
& (2440.122, 1772.423, 2100.752, 1520.347) \\
& -0.19, -1.19, -2.24, 0.46 \\
+ & 1242.028ES4_i + 1906.028NCO_i \\
& (2911.14, 3001.18) \\
& 0.43, 0.64 \tag{Eq. 3}
\end{align*}

\begin{equation}
N = 209 \quad \bar{R}^2 = 0.6072
\end{equation}

**Discussion of Results**

**Overall Model Fit and Significance**

The estimated regression equation has an \( \bar{R}^2 \) of 0.6072. This means that if significant, the model can be used to estimate 60.72\% of variation in job costs. An F-test is employed to test the overall significance of the model. An F-test tests use the explained and residual sums of squares to test the significance of an estimated regression equation against the mean of the independent variable alone. The F-test uses and the F-Statistic distribution to test the following hypothesis:

\[ H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0 \]

\[ H_a: H_0 \text{ is not true} \]

The model has 16 explanatory variables and 193 degrees of freedom, so the critical F-statistic is approximately 1.7. Stata has calculated an F-statistic for the model to be 21.10. Since this is much larger than the critical F-Statistic, the null hypothesis is rejected, and the overall model is found to be significant.
Explanatory Variable Significance and Interpretation

Next each variable must be tested for its significance. To determine whether or not each variable should be really be included in the model, there are four things that must be looked at, what theory says about the variable, how the variable impacts the overall fit of the equation ($R^2$), the t-score hypothesis test, and whether exclusion of the variable biases the remaining explanatory variables. The theory behind why each variable was included in the model is presented in the hypotheses for each variable from Section III so this section will mainly explore the other three tests of variable significance.

The t-scores hypothesis test for all variables will be tested at the five percent significance level. For the single sided tests ($TXTi$, $FLRi$, $VATi$, $MASti$, $Li$, $Si$, $Ci$, $DAYi$, and $EXTi$) the critical t-value is 1.653. The remaining variables use a two-sided test and have a critical t-value of 1.972.

The t-value of the $TXTi$ variable is 2.10, meaning that the null hypothesis is rejected, and the variable is significant. Exclusion of the $TXTi$ variable does not bias the other variables, but it does decrease the $R^2$ value. The variable will remain in the model.

The $FLRi$ variable has a t-value of 1.71 making it significant. Excluding it from the regression decreases the $R^2$ value and biases one of the other variable. The variable should remain in the model.

The t-value of the $VATi$ variable is 1.76, meaning that the null hypothesis is rejected, and the variable is significant. Exclusion of the $VATi$ variable biases two of the other variables and decreases the $R^2$ value. The variable will remain in the model.

The $MASti$ variable has a t-value of 0.74 making it essentially zero based on the 5% significance level. Excluding it from the regression increases the $R^2$ value and does
not bias any of the other variables. The variable should be removed from the model. This is in-line with theory in that mastic is generally only present with tile or other flooring, so its impact is probably best represented in those numbers.

The t-value of the $L_i$ variable is 0.07, meaning that the null hypothesis is not rejected, and the variable is not significant. Exclusion of the $L_i$ variable does not bias any of the other variables and increases the $\bar{R}^2$ value. The lack of significance is probably due to the fact that only 10% of jobs involved the removal of linear feet of material. For this reason it is best to remove the variable and let any variation in cost be assumed by other variables.

The $S_i$ variable has a t-value of 0.35 making it essentially zero based on the 5% significance level. Excluding it from the regression increases the $\bar{R}^2$ value and does not bias any of the other variables. Nearly all the materials included in this group of materials are removed from the outside of buildings so this variable will be removed from the model and the variance should be explained through the $EXT_i$ variable.

The t-value of the $C_i$ variable is 3.87, meaning that the null hypothesis is rejected, and the variable is significant. Exclusion of the $C_i$ variable does not bias any of the other variables but it decreases the $\bar{R}^2$ value. The variable will remain in the model.

The $DAY_i$ variable has a t-value of 8.90 making it significant even at the 1% significance level. Excluding it from the regression decreases the $\bar{R}^2$ value and biases six of the other variables. The variable should not be removed from the model.

The t-value of the $EXT_i$ variable is 0.92, meaning that the null hypothesis is not rejected, and the variable is not significant. Exclusion of the $EXT_i$ variable does not bias any of the other variables or change the $\bar{R}^2$ value. The variable will remain in the model.
because the processes to remove material from the exterior of a building are significantly enough different that theory suggests the cost impact is important.

The $SCHL_i$ variable has a t-value of -1.23 making it essentially zero based on the 5% significance level. Excluding it from the regression decreases the $\bar{R}^2$ value but it does not bias any of the other variables. Based on theory and impacts on overall fit, the variable should not be removed from the model.

The t-value of the $EMER_i$ variable is -0.19, meaning that the null hypothesis is not rejected, and the variable is not significant. Exclusion of the $EMER_i$ variable does not bias any other variables and increases the $\bar{R}^2$ value. The variable will not remain in the model.

The $ES1_i$ variable has a t-value of -1.19 making it essentially zero based on the 5% significance level. Excluding it from the regression increases the $\bar{R}^2$ value and does not bias any of the other variables. The variable however will not be removed from the model. This is due to the theory that costs are generally highly impacted by the estimator who bids a job.

The t-value of the $ES2_i$ variable is -2.24, meaning that the null hypothesis is rejected, and the variable is significant. Exclusion of the $ES2_i$ variable biases one of the other variables and decreases the $\bar{R}^2$ value.

The $ES3_i$ variable has a t-value of 0.46 making it essentially zero based on the 5% significance level. Excluding it from the regression increases the $\bar{R}^2$ value and does not bias any of the other variables. The variable will not be removed from the model for the same reason as keeping the $ES3_i$ variable.
Similarly to the $ES_1_i$ and $ES_3_i$ variables, the $ES_4_i$ variable is insignificant at the 5% significance level and its exclusion improves $\bar{R}^2$ without biasing other variables but will be kept in the model based on theory.

The $NCO_i$ variable has a t-value of 0.64 making it essentially zero based on the 5% significance level. Excluding it from the regression increases the $\bar{R}^2$ value and does not bias any of the other variables. The variable should be removed from the model.

**Final Model and Interpretation.**

Removing the variables seen to not fit the model, the following final estimated regression equation has been obtained.

$$\hat{COST}_i = -582.94 + 1.32TXT_i** + 2.73FLR_i* + 2.33VAT_i* + 19.83C_i* + 2987DAY_i*$$

$$t = \begin{array}{cccc}
(0.62) & (0.58) & (0.70) & (5.01) \\
2.12 & 4.70 & 3.31 & 3.95 & 9.23 \\
\end{array}$$

$$+ 1803.24EXT_i* - 1872.21SCHL_i - 2019.14N_i - 4664.17D_i**$$

$$\begin{array}{cccc}
(1687.45) & (1580.664) & (1747.26) & (2044.67) \\
1.07 & -1.18 & -1.16 & -2.28 \\
\end{array}$$

$$+ 770.47NA_i + 1292.33T_i$$

$$\begin{array}{cc}
(1497.831) & (2861.46) \\
0.51 & 0.45 \\
\end{array}$$

(Eq. 4)

$$N = 209 \quad \bar{R}^2 = 0.6147$$

This model has no VIF factors over 2.2 and still passes the Breusch-Pagan test (See Appendix D). Standard error is decreased in all remaining terms and t-scores are increased for all variables except $SCHL_i$ and $N_i$, neither of which had a significant t-score in the original estimation. The $\bar{R}^2$ for this model is 0.0075 higher than the previous model’s.

Interpreting the coefficients in the model we make the following conjectures. A one-square-foot increase in the quantity of texture or drywall to be removed will increase job costs by $1.32. A one-square-foot increase in the quantity of sheet flooring to be
removed will increase job costs by $2.73. A one-square-foot increase in the quantity of vinyl asbestos tile to be removed will increase job costs by $2.33. A one-cubic-foot increase in the quantity of asbestos containing material to be removed will increase job costs by $19.83. Each day that a job goes on adds $2987 to the total cost. Jobs where any work occurs on the exterior of a structure cost $1803.24 more than jobs done entirely on interiors. Jobs done at schools cost $1872.21 less than jobs that are the same otherwise but completed outside an educational institution. The coefficients attached to variables for each estimator represent the difference between the costs of a job bid by the business owner and one bid by that particular estimator. A job that Estimator 1 bids will cost $2019.14 less, a job that Estimator 2 bids will cost $4664.17 less, a job that Estimator 3 bids will cost $770.47 more, and a job that Estimator 4 bids will cost $1292.33 more than if the company owner had bid the exact same job.

Conclusion

The purpose of this paper was to develop a model to estimate the influence of certain factors on the cost of an asbestos abatement job. The model and analysis of factors that were and were not significant suggests that the most significant factor in the cost of a job is the number of days that the job takes to complete. This is supported by general business ideology that frames labor as the most expensive input to production. The only materials that were significant on their own were texture, flooring, and vinyl asbestos tiling. These are the main materials removed from buildings by the firm so it makes sense that their presence will contribute significantly. Though insignificant by t-score much of the variation in the model is due to the variables representing different cost estimator. This matches up with theory from other construction literature that estimations
are often specific to the human making them. The firm can use the information culled from this report to evaluate differences in estimation and try to minimize them for the variables within the firm’s control. Thus creating a more uniform estimation to improve consistency of estimation and bids across employees.
Appendix A: Variable Correlation.
Appendix B: Breusch Pagen Regression of Equation 3

\[
\text{. reg e2 TXT FLR VAT MAST L S C DAY EXT SCHL EMER N D NA T NCO}
\]

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 289</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F(16, 192) = 1.59</td>
</tr>
<tr>
<td>Model</td>
<td>2.8285e+18</td>
<td>16</td>
<td>1.2678e+17</td>
<td>Prob &gt; F = 0.0748</td>
</tr>
<tr>
<td>Residual</td>
<td>1.5321e+19</td>
<td>192</td>
<td>7.9795e+16</td>
<td>R-squared = 0.1169</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adj R-squared = 0.0433</td>
</tr>
<tr>
<td>Total</td>
<td>1.7349e+19</td>
<td>208</td>
<td>8.3409e+16</td>
<td>Root MSE = 2.5e+08</td>
</tr>
</tbody>
</table>

|       | Coef.    | Std. Err. | t     | P>|t|  | [95% Conf. Interval] |
|-------|----------|-----------|-------|-----|---------------------|
| TXT   | 25398.37 | 23839.73  | 1.07  | 0.288 | -21631.03 to 72411.77 |
| FLR   | 564.8961 | 43494.46  | 0.01  | 0.999 | -85223.47 to 86353.26 |
| VAT   | 61896.14 | 37751.56  | 1.64  | 0.103 | -12564.9 to 136357.2 |
| MAST  | -10218.2 | 38132.2   | -0.27 | 0.789 | -85429.62 to 64993.22 |
| L     | -94664.83| 359103.4  | -0.26 | 0.792 | -882959.31 to 613629.4 |
| S     | -22679.78| 28730.56  | -0.79 | 0.431 | -79363.62 to 34004.05 |
| C     | 215911.5 | 190124    | 1.14  | 0.258 | -159888.5 to 590911.5 |
| DAY   | 2.37e+07 | 1.25e+07  | 1.91  | 0.058 | -830392.7 to 4.03e+07 |
| EXT   | -3.25e+07| 6.66e+07  | -0.49 | 0.626 | -1.64e+08 to 9.88e+07 |
| SCHL  | -45648.64| 6.00e+07  | -0.00 | 0.999 | -1.18e+08 to 1.18e+08 |
| EMER  | -2.88e+07| 9.11e+07  | -0.32 | 0.752 | -2.08e+08 to 1.51e+08 |
| N     | -5.15e+07| 6.62e+07  | -0.78 | 0.438 | -1.82e+08 to 7.90e+07 |
| D     | 1.75e+07 | 7.64e+07  | 0.12  | 0.24 | -1.37e+08 to 1.72e+08 |
| NA    | 4.26e+07 | 5.67e+07  | 0.75  | 0.454 | -6.93e+07 to 1.55e+08 |
| T     | -3.80e+07| 1.09e+08  | -0.28 | 0.783 | -2.44e+08 to 1.84e+08 |
| NCO   | -2.73e+07| 1.12e+08  | -0.24 | 0.808 | -2.48e+08 to 1.94e+08 |
| _cons | -4.49e+07| 5.70e+07  | -0.79 | 0.432 | -1.57e+08 to 6.75e+07 |

Appendix C: White's Test for Heteroskedasticity

White's test for Ho: homoskedasticity
against Ha: unrestricted heteroskedasticity

\[
\chi^2(113) = 101.76
\]

\[
\text{Prob > } \chi^2 = 0.7671
\]

Cameron & Trivedi's decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>(\chi^2)</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>-12099.98</td>
<td>113</td>
<td>1.0000</td>
</tr>
<tr>
<td>Skewness</td>
<td>.</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>.</td>
<td>130</td>
<td></td>
</tr>
</tbody>
</table>

Appendix D: Breusch Pagan Test Equation 4

\[ R^2_i = \delta_0 + \delta_1 TXT_i + \delta_2 FLR_i + \delta_3 VAT_i + \delta_4 C_i + \delta_5 DAY_i + \delta_6 EXT_i + \delta_7 SCHL_i \]

\[ + \delta_8 ES_1_i + \delta_9 ES_2_i + \delta_{10} ES_3_i + \delta_{11} ES_4_i + \mu_i \]

Where:

\[ R^2_i = \] The squared residuals of the estimated regression equation derived from the data set and Eq. 4

\[ TXT_i - T_i = \] Variables as defined in theoretical model on pg.16

The hypothesis to be tested is

\[ H_0: \alpha_1 = \alpha_2 = \cdots = \alpha_9 = 0 \]

\[ H_A: H_0 \text{ is false} \]

The critical chi-squared value is 19.67 (11 degrees of freedom at 5% significance)

\[ . \text{reg} \ R^2 \ \text{TXT} \ \text{FLR} \ \text{VAT} \ \text{C} \ \text{DAY} \ \text{EXT} \ \text{SCHL} \ \text{N D NA T} \]

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs =</th>
<th>209</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1.924e+18</td>
<td>11</td>
<td>1.7492e+17</td>
<td>Prob &gt; F =</td>
<td>0.8122</td>
</tr>
<tr>
<td>Residual</td>
<td>1.5115e+19</td>
<td>197</td>
<td>7.6723e+16</td>
<td>R-squared =</td>
<td>0.1129</td>
</tr>
<tr>
<td>Total</td>
<td>1.7839e+19</td>
<td>208</td>
<td>8.1916e+16</td>
<td>Adj R-squared =</td>
<td>0.0634</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Root MSE =</td>
<td>2.8e+00</td>
</tr>
</tbody>
</table>

| R2    | Coef.  | Std. Err. | t     | P>|t|   | [95% Conf. Interval] |
|-------|--------|-----------|-------|-------|---------------------|
| TXT   | 31155.68 | 22978.74  | 1.36  | 0.177 | -14100.22            |
| FLR   | -7975.296 | 21471.09  | -0.37 | 0.711 | -50317.99            |
| VAT   | 51753.77  | 26838.74  | 1.99  | 0.048 | 419.887              |
| C     | 216420.3 | 185263.8  | 1.17  | 0.244 | -148934.8            |
| DAY   | 2.05e+07  | 1.20e+07  | 1.72  | 0.087 | -3028838            |
| EXT   | -4.34e+07 | 6.24e+07  | -0.70 | 0.487 | -1.66e+08            |
| SCHL  | -4150178  | 5.84e+07  | -0.07 | 0.943 | -1.19e+08            |
| N     | -5.88e+07 | 6.46e+07  | -0.79 | 0.433 | -1.78e+08            |
| D     | 1.48e+07  | 7.56e+07  | 0.20  | 0.845 | -1.34e+08            |
| NA    | 3.95e+07  | 5.54e+07  | 0.71  | 0.476 | -6.97e+07            |
| T     | -2.21e+07 | 1.96e+08  | -0.21 | 0.835 | -2.31e+08            |
| _cons | -4.11e+07 | 5.46e+07  | -0.75 | 0.453 | -1.49e+08            |

\[ NR^2 = 209 \times 0.0634 = 13.0604 \]
Resources


doi:10.1080/13504851.2017.1327114


