Investigating the Demand for Warehouse Space and the Spatial Attributes Influencing Those Decisions
Jamie N. Everett, Graduate Research Assistant
Dr. Eric L. Jessup, Assistant Professor
School of Economic Sciences, Washington State University

Introduction
The availability of space in warehouses is a key component to sustaining a thriving economy, especially for the trade-dependent state of Washington. Whether shipping goods to destinations within the state, across the country, or across the world, access to these facilities is essential for the economic vitality of Washington industries that utilize them and their down-stream customers. Supply chains are changing and accelerating; offering advanced technologies and other services to keep up with these shifting supply chains is paramount to ensure the continued efficiency and profits for both parties.

Although the demand of space in warehouses is such an important issue, estimating this demand can be a difficult task, primarily due to access limitations to specific data and information. Prior research efforts to estimate and determine the demand for warehouse space have utilized very simplistic models that have proven to be inaccurate and unreliable. Because these models use weak proxies and very broad information, the optimal location and square footage demanded is, at best, a rough estimation. The goal of this research is to develop an improved estimation model that will accurately determine the demand for warehouse space in the state of Washington, as dictated by specific indicators.

Literature
Traditional location theory describes the optimal placement of warehouses to be primarily determined by transportation, delivery, and opportunity costs. In the text *Location Theory and Decision Analysis*, Chan determines that the optimal location for warehouses is based on reordering needs. Probability and marginal analysis of cost relative to the demand function are employed. Although the optimal size of these warehouses is not determined with these methods, the location of warehouses is still an important aspect to study, as it may
be a determining factor in estimating the size of the facility. Proximity to multimodal transportation, commodities, markets, and employment are all possible determinants of the demand for warehouse space.

In *Transport Economics* (Button 1993), Weber’s (1929) model of industrial location is outlined. Weber notes the importance of good transportation when determining site location since raw materials, production, and consumers are all located in different areas. The location of an industrial firm is determined by focusing on the distance from where the raw materials are being picked up to the warehouse site to which they are being delivered, the amounts of raw materials being shipped, and the associated costs. Although this study uses distances to determine the location of industrial sites, this paper will study distances as a determinant of warehouse space.

Several other studies are cited in Button (1993). One is the Loschian model (1954) which states that different but interdependent product markets would lead to concentration of firms at certain locations. The market area is determined by the rate of transport costs. This research uses spatial statistical methods to examine patterns of concentration and spatial autocorrelation.

Another study cited by Button is the Greenhutt (1963) study. It states that transport costs are only important if they are a large part of total costs or if they widely differ according to different locations. However, Cook’s (1967) study shows that many firms in the area studied did not even know their transport costs. Button recognizes that transport costs are only one of the many determinants of industrial location. Although the variables used in the statistical analysis in this paper do not include any cost data, there are others that may be used as proxies.

Mansour and Christensen’s (2001) “An Alternative Determinant of Warehouse Space Demand: A Case Study” discusses a different approach to assess the demand for warehouse space. These

---

1 Location theory studies specific to industrial sites vary in their conclusions of what drives demand for warehouse space. As the model presented in *Transport Economics* is for industrial sites, it is not a precise model to use when determining warehouse space. However, some of its aspects may be incorporated into the model developed by this research paper.
authors cite several studies that have attempted to model the demand for warehouse space using models similar to those that are used to estimate the demand for office space. As cited by Mansour and Christensen, Wheaton and Torto (1990) use population and employment measures in manufacturing, distribution, and wholesale trade sectors to model demand for warehouse space. The study conducted by Rabianski and Black (1997) uses the type of analysis used to determine demand for office space to determine the demand for warehouse space. Mansour and Christensen note that there are “major differences between demand for office and demand for industrial space.” Wheaton and Torto (1990) agree that employment in a warehouse and an office widely vary. These measures may be too broad to accurately estimate the demand for warehouse space.

Unlike these previous models, Mansour and Christensen focus on warehouse inventory, using total shipment volume as a proxy, to model demand for warehouse space. This study focuses on distribution centers located in Dallas, Los Angeles, and Seattle. They concluded that by including freight shipments in model estimations, a better measure of demand is achieved for warehouse space than if using manufacturing employment, which may or may not affect demand depending on the area. However, their definition of warehouse employment presents the same weakness as the study conducted by Wheaton and Torto (1990). It may be too broad to accurately estimate the demand for warehouse space alone. Mansour and Christensen note that the growth in warehouse employment has not kept up with the increase of occupied warehouse stock. They believe that the negative sign of the coefficient for warehouse employment in the regression for the Los Angeles area is the wrong sign. It appears that they are ignoring advances in technology that make it possible to have an increase in warehouse stock without a proportional increase in employment, as the negative relationship in their regression shows.

In the statistical results for the Seattle area, warehouse employment is more significantly correlated with occupied warehouse space. An interesting note is that in the regression results in this research paper, the number of employees is found to be a statistically significant determinant of demand for warehouse space. Because this study concerns only warehouses in the state of Everett.
Washington, these results may be specific to this region while differing in other states and markets.

Another limitation in Mansour and Christensen’s approach is using total shipment volume as a proxy for warehouse inventory. This research uses specific, first hand data on specific commodities, inbound and outbound volume, and number of employees at warehouses located throughout the state of Washington.

Mueller and Laposa’s (1994) “The Path of Goods Movement” points out the weakness in previous industrial demand models: all industrial building types are viewed as one group instead of differentiating them by type of facility. It is important to differentiate these facilities because they serve different groups and their specific functions. They also make the distinction between regional and local warehouse markets, as different factors play a role in determining the demand for square footage of these facilities.

They introduce the theory that industrial warehouse demand can be better modeled with “The Path of Goods Movement.” The path of goods movement shows the routes in the United States that are used the most to transport goods, measured by the tonnage of the shipment. These routes are used because of efficient transportation and therefore, minimized costs. Goods are not evenly distributed throughout the country because people are not evenly distributed throughout the country. These goods must be stored in an efficient way so they can be readily accessible to consumers. By using the weight of the goods rather than the value, they find that demand for warehouse space is highly correlated with the movement of these goods. Highly valued goods are often small (i.e. diamonds), which doesn’t give an accurate estimate of the space needed to store them. Efficiency and customer service in the transportation sector have increased since its deregulation in 1980, which in turn has had an affect on the warehousing industry. Ordering and delivery have become highly efficient due to advances in technology, which reduces the time goods need to be stored in warehouses. Today’s consumers have grown accustomed to ‘instant gratification,’ thus increasing the importance of timely deliveries. Therefore, these goods must be stored in a place where they can be easily and quickly transported and that is close to the retail location and the consumer. However, having too many storage facilities is inefficient. Instead,
these goods are stored in warehouses on the path of goods movement. For that reason, population and overall employment growth prospects are better predictors than industrial employment growth prospects.

This study also indicates that regional markets have higher concentrations of large warehouses, while local markets tend to be smaller. Although the distinction between regional and local markets in this paper is not explicit, spatial statistics will be used study the concentration of warehouses by size, therefore determining if they serve a local or regional market.

“Design of warehousing and distribution systems: an object model of facilities, functions and information” by Govindaraj, Blanco, Bodner, Goetschalckx, McGinnis, and Sharp (2000), hereafter Govindaraj (2000), attempts to model warehouse design based on certain functions of a warehouse. Analysis of the warehouse’s clients is used to gather specific information concerning their usage of the warehouse. The designer then determines the type of technology and other functions that will be needed in order to accommodate these specific criteria. The designer uses “intuition, experience, and judgment” to create the design for the warehouse that will also be as cost effective as possible. One of the main problems pointed out by Govindaraj is the fact that “in current practice, in the early stages of design, the understanding of functional requirements and costs is intuitive.” This research effort will use specific attribute data from warehouses in the state of Washington to discover if these are accurate determinants for the demand for warehouse space.

In “Solution procedures for sizing of warehouses,” Rao and Rao (1998) discuss the problem of determining the size of a single warehouse for a firm producing a single process. In the past, the static problem has been solved using linear programming. Rao and Rao offer another method for solving the static problem by taking into consideration many other factors, including costs that vary over time and economies of scale. This research looks at a variety of warehouses and products, not a singular model like Rao and Rao.

Traffic World magazine’s June 5, 2006 issue published the article “Warehouses Stack Up” by William Hoffman. This article indicates the push for warehouse operators to offer more value added services to their customers. Supply chains are changing; consumers are demanding faster delivery, which is in turn driving demand for
shorter leases in public warehouses. Mergers are leading to warehouses being consolidated in order to accommodate this trend. The article quotes Richard Armstrong of Armstrong & Associates, who expects “growth rates to be in the neighborhood of 10 percent (annually) over the next five years.” This growth along with the change in supply chains is making long-term contracts with warehouses less attractive to shippers. Technology that can keep up with fast paced supply chains, along with value added services, must be offered in order to draw customers to these warehouses. Public warehouses are also starting to offer shorter lease contracts. This research will examine several variables that this article says are becoming increasingly important to warehouse customers to determine if they have an impact on the demand for warehouse space: value added services (labeling, ticketing, pooling, pick and pack, assembly/consolidation, etc.), inventory controls (just-in-time delivery (JIT) and electronic data interface (EDI)), and the type of warehouse (public, private, distribution center, cold storage).

Research Questions

These previous research efforts, while seeking improved estimation and modeling approaches, have generally relied upon rather simplistic techniques for estimating the demand for warehouse space. Additionally, they do not attempt to incorporate key spatial features of these warehouses. This research paper focuses on the specific functional characteristics and spatial attributes of warehouses in Washington and seeks to empirically identify and estimate those factors which have the greatest impact on the demand for space. Particularly, 1) what determines the demand for warehouse/distribution center space? 2) What are the determinants that influence the size and location of warehouses? 3) Are the existing models still valid when these additional factors are in play?

In addition to those variables used in prior studies, this effort will include the different types of warehouse facilities, the value-added services offered at these different sites, the technology utilized in these warehouses, and the average length of haul of shipments to and from these warehouses. Additionally, spatial attributes and characteristics related to proximity to transportation infrastructure will also be tested.
Data

Firm level warehouse characteristics utilized in this study were gathered from a primary survey conducted by the Transportation Research Group at Washington State University in 2004. In this study, questionnaires were mailed to 973 warehouses and distribution centers in Washington. As per Pike (2005), the “sample used in this study was compiled from local, county, state, and federal agencies. Collected information was categorized by the Standard Industrial Classification (SIC) or North American Industry Classification System (NAICS) codes associated with industries that operate in the storage and distribution of freight.” Warehouses were contacted through telephone calls to follow up in order to increase the response rate. For the original study, a total of 142 companies (16%) returned useable survey responses. For this study, 63 survey responses (6%) contained the unique warehouse characteristics that were needed to conduct this research.

The summary statistics\(^2\) for the square footage of warehouses show that the data are not normally distributed. This is an exploratory study, therefore, research continued despite the skewed nature of the data.

In addition to the operational characteristics collected from the warehouse survey, spatial attribute data were created using numerous GIS shapefiles from the Washington State Department of Transportation. These files were used to identify the geographical location of each warehouse used in this study, to calculate the distances from these warehouses to multimodal transportation, and to calculate the population densities surrounding the warehouses at two, five, and ten mile radii.

Ordinary Least Squares Regressions

Before any statistical analysis was conducted, the variables that were being considered for the ordinary least squares regression equation were tested for multicollinearity by examining the correlation coefficient matrix. Two variables, namely the outbound number of trucks and ten mile population density, were found to be

---

\(^2\) Summary statistics, rate of occurrence in the sample, normality tests, and the correlation coefficient matrix for each of these variables are available by request from the author.
highly correlated and thus, were removed from the list of possible variables to be used.

Next, using the stepwise regression model selection technique, many of the variables that were expected to play a role in the determination of the demand for warehouse square feet were found to be insignificant. The statistically significant independent variables were identified, resulting in the following OLS regression equation:

\[
\text{Square Footage} = 284 \text{ Average Length of Haul} + 834 \text{ Number of Employees} + 193790 \text{ Other Services} - 162881 \text{ Frozen Foods} - 117377 \text{ Public Warehouse} - 94884 \text{ Other Control Programs} + 0.0146 \text{ Inbound Volume} + 75990 \text{ Paper and Lumber}
\]

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>No constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Length of Haul</td>
<td>284.18</td>
<td>34.57</td>
<td>8.22</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of Employees</td>
<td>834.0</td>
<td>171.4</td>
<td>4.86</td>
<td>0.000</td>
</tr>
<tr>
<td>Other Services</td>
<td>193790</td>
<td>64573</td>
<td>3.00</td>
<td>0.004</td>
</tr>
<tr>
<td>Frozen Foods</td>
<td>-162881</td>
<td>49686</td>
<td>-3.28</td>
<td>0.002</td>
</tr>
<tr>
<td>Public Warehouse</td>
<td>-117377</td>
<td>49803</td>
<td>-2.36</td>
<td>0.022</td>
</tr>
<tr>
<td>Other Control Programs</td>
<td>-94884</td>
<td>40454</td>
<td>-2.35</td>
<td>0.023</td>
</tr>
<tr>
<td>Inbound Volume</td>
<td>0.014560</td>
<td>0.006960</td>
<td>2.09</td>
<td>0.041</td>
</tr>
<tr>
<td>Paper and Lumber</td>
<td>75990</td>
<td>40240</td>
<td>1.89</td>
<td>0.064</td>
</tr>
</tbody>
</table>

\[S = 155583 \quad R^2 = 79.21\% \quad R^2 (adj) = 74.75\%\]

Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>8</td>
<td>5.07372E+12</td>
<td>6.34216E+11</td>
<td>26.2</td>
<td>0.000</td>
</tr>
<tr>
<td>Residual Error</td>
<td>55</td>
<td>1.33134E+12</td>
<td>24206220505</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td></td>
<td>6.40507E+12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>Seq SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Length of Haul</td>
<td>1</td>
<td>3.08344E+12</td>
</tr>
<tr>
<td>Number of Employees</td>
<td>1</td>
<td>1.20315E+12</td>
</tr>
<tr>
<td>Other Services</td>
<td>1</td>
<td>1.47381E+11</td>
</tr>
<tr>
<td>Frozen Foods</td>
<td>1</td>
<td>2.57268E+11</td>
</tr>
<tr>
<td>Public Warehouse</td>
<td>1</td>
<td>1.06921E+11</td>
</tr>
</tbody>
</table>

\(^3\) After determining the best model for the data, another correlation coefficient matrix was created which indicated no additional problems with multicollinearity.
A one unit increase in the average length of haul will increase the size of the warehouse by 284 square feet. This implies that warehouses providing services for long distance, cross-country shipments require larger inventory/storage space. It is also likely that commodities that do not have to travel as far to their ultimate destination have considerably higher turnover ratios and therefore, may be kept in the warehouse for shorter periods of time. Companies may want to keep their stock in local warehouses so it is on hand for quick restocking at nearby retail stores. Also, goods that have a short shelf life must be transported quickly to close locations. Examples of this may include food and grocery items. Many of these items are perishable and cannot sit in a warehouse for a long period of time and have to be readily available. However, the variables for food and grocery commodities and cold storage were not statistically significant.

The two commodities that did have explanatory power when estimating warehouse square footage were frozen foods and paper and lumber. Frozen foods have a large, negative coefficient which may be because frozen goods can be stored for longer periods of time at their origin and in grocery stores, cutting down on the time they must be stored in another location.

On the other hand, paper and lumber has a large, positive coefficient which may be a seasonal effect as lumber is largely used in building. Typically, construction’s busy season is during the spring and summer months. Lumber needs to be stored through the off-season until demand increases again in the spring.

Although the specific value added variables were found to be statistically insignificant, the variable for “Other Services” was highly significant and has a large positive coefficient. When these services are offered, the demand for warehouse square footage increases by 193,790.

The relationship between demand for warehouse space and public warehouses is negative. Public warehouses can be rented on a short-term basis, which may decrease the amount of space demanded.
There is also a negative relationship between demand for warehouse square footage and other control programs. The variable ‘Other Control Programs’ includes inventory control programs except for JIT and EDI. If a facility offers ‘other’ control programs, the demand for space is decreased by 94,884 square feet. This may indicate that warehouse customers prefer technology that will insure fast and accurate service.

These findings are consistent with the article “Warehouse Stack Up” by William Hoffman. Public warehouses are becoming less attractive because their lease terms are often longer than their customers want, while value added services and technology-based inventory control programs are becoming increasingly important to accommodating accelerating supply chains.

In this equation, the inbound volume variable could be used as a proxy for the demand for warehouse space. This is similar to Mansour and Christensen (2001) who used total shipment volume as a proxy for the demand for warehouse space. The above equation shows that a one unit increase in inbound payload increases the demand for warehouse space by only 0.015 square feet. Although this is a very small coefficient, it is significant. This may indicate a very fast turnaround time. If the inbound volume increases by one unit and the square footage only increases by a fraction of that, it may be due to cross docking; inbound volume is unloaded and immediately loaded on to an outbound truck for transport to its next destination. However, the variable for cross docking was not statistically significant. Both inbound and outbound volumes were considered, but only inbound volume was statistically significant.

Unlike Mansour and Christensen (2001), this regression does find that the number of employees at a warehouse facility is significant in determining the facility size. For each additional employee, the facility increases in size by 834 square feet. Although Mansour and Christensen believe that the numbers for square footage and employment should be more closely related, this research notes the importance of technology and inventory programs. Therefore, the large increase in square footage with the addition of just one employee is realistic.

It is interesting to find that many of the variables that were believed to influence the size of the warehouse were statistically
insignificant. None of the spatial variables play a role in determining the square footage of a warehouse. Proximity to highways, railroads, ports did not matter. The population density at two, five, and ten mile radii around each warehouse was not significant. This could be because the goods are being exported rather than being used for local consumption. As in Mueller and Laposa (1994), this may indicate that these warehouses are regional, not local. Future research efforts can be focused on determining the latent or unmeasured variables that have an effect on the size of these warehouses. Data on local markets, the sources of the commodities, and their terminal markets would be valuable in this search.

**Causality**

When examining the relationships between the dependent and independent variables, possible concerns with causality/endogeneity need to be addressed. It is important to determine if the observed relationships are from correlation or causation.

By using Tetrad 3 software, it was determined that there are no direct causal paths between square footage and the following variables:\(^4\): Type of commodity, Inbound/Outbound Payload and Number of Trucks, Population Density at Two, Five, or Ten Miles, Type of Facility, Distance to Multimodal Transportation, Warehouse Services. When determining the possible causal relationships between the variables, the likelihood of latent common causes and correlated errors was considered. At a \( p = 0.05 \) significance level, there is a relationship between square footage and the number of employees and the number of loading bays, albeit ambiguous. There is not a direct path between these variables, but there may be an unmeasured variable that is influencing this relationship. It is not possible to determine the relationships.

\(^4\) These results hold only under certain criteria and when the standard assumptions used by the Tetrad 3 software are not violated. These factors are given in the Tetrad 3 User’s Manual: “1. The correctness of the background knowledge input to the algorithm. 2. Whether the recursiveness condition holds, i.e., that there are no feedback loops. 3. Whether the Causal Independence assumption holds. 4. Whether the Faithfulness assumption holds. 5. Whether the distributional assumptions made by the statistical tests hold. 6. The power of the statistical tests against alternatives. 7. The significance level used in the statistical tests.” For detailed explanations of these assumptions, please refer to the User’s Manual.
Geographic Characteristics

It was an interesting discovery to find that the geographic characteristics of each warehouse did not contain any explanatory power concerning the square footage of the facility. The studies of Weber (1929), Wheaton and Torto (1990), and Mueller and Laposa (1994) found that the proximity to multimodal transportation and population measures impacted the location and square footage of warehouses. Although these variables were not significant in this research, they may be telling another story. Loschian (1954), Button (1993), and Mueller and Laposa (1994) all discuss some form of concentration of firms. Spatial statistical methods are used in this section to determine if concentration based on any of the other characteristics of the warehouses is occurring. The first step in the type of analysis was to plot the 63 warehouses on a map of the state of Washington. After doing so, the geographic coordinates were obtained (as well as the population densities and proximity to multimodal transportation), which makes spatial inference possible.

The first calculation was the Nearest Neighbor Index (NNI)\(^5\). The NNI for the warehouses used in this study is 0.33553, with \(Z = -10.2486\), which indicates clustering. There is significant clustering around the Seattle/Tacoma and Spokane areas, two areas that support much of the commerce and trade within the state. The location of the other warehouses in the state is consistent with the agricultural nature of the state’s economic activity. This is consistent with the previously mentioned studies that note concentration of firms. Spatial autocorrelation is another characteristic that can be measured using spatial statistics. Tobler’s Law (the first law of geography) is “everything is related to everything else, but near things are more related than distant things” Dezzani (2006). As noted by Chang (2002), Cliff and Ord (1973) state “spatial autocorrelation measures the relationship among values of a variable according to the

\(^5\) As defined in the help section of the Crimestat 3 software package, which was used to calculate the index, the NNI “provides an approximation about whether points are more clustered or dispersed than would be expected on the basis of chance. It compares the average distance of the nearest other point (nearest neighbor) with a spatially random expected distance by dividing the empirical average nearest neighbor distance by the expected random distance.” If the average distance is the same as the mean random distance, the index will equal 1.
spatial arrangement of the values.” More simply put, spatial autocorrelation is the same as traditional statistical autocorrelation, but with space as an added dimension. It is the “correlation of a variable with itself through space.” Spatial autocorrelation can be measured in a specific area using Local Indicators of Spatial Association, otherwise known as LISA statistics. The LISA statistic for each variable was examined to determine clustering (spatial autocorrelation).

The warehouses in this study are not clustered according to their respective square footages, the average length of haul, or warehouses that store frozen foods. They are randomly arranged.

Warehouses with a small number of employees are clustered in areas with other warehouses with a small number of employees; these warehouses are similar in that they employ a small amount of people.

There are varying types of clustering that occur with warehouses that store paper and lumber and facilities that offer other services. Clustering related to public warehouses cluster in only a few areas, and have opposite patterns. The clustering pattern associated with other control programs indicates that the target warehouse offers other control programs while the surrounding warehouses do not.

There is clustering in regards to the inbound volume of these warehouses. In some areas, warehouses that receive low inbound volumes are clustered. In other areas, the clustering of warehouses differs in their inbound volumes, the target warehouse has a high inbound volume and others have low. In one area, a warehouse that receives a low inbound volume is surrounded by warehouses that receive high inbound volumes.

---

6 The "working definition" from Dezzani (2006) states that spatial autocorrelation "can be interpreted as a spatially weighted Pearson’s correlation coefficient, with values closer to 1 indicating a spatially clustered pattern and values near 0 show no evidence of a spatially clustered pattern."
7 Calculated using Geoda software. The LISA cluster maps are available upon request from the author.
8 These are indicator variables, which is not the optimal type of variable to use in LISA analysis.
Conclusion

Although this is a small exploratory research effort, some strong determinants of the demand for warehouse space do emerge, while factors that were believed to be significant were found not to be so. The statistically significant determinants of the demand for warehouse space are average length of haul, number of employees, other services, frozen foods, public warehouse, other control programs, inbound volume, and paper and lumber. Some of the findings were on par with past research, while others were quite different. This may be a spatial issue; characteristics of warehouses in the state of Washington differ from warehouses throughout the United States. This is a possibility, as Washington has access to ocean routes for trade. There is not a clear picture concerning the possible problem of endogeneity. There are no direct causal paths, but the relationships between certain variables are ambiguous. Spatial attributes do not play a role in determining the demand for square footage in a warehouse facility, but examining the clustering patterns based on different attributes does give some information concerning location. Future research efforts will be able to determine the robustness of the results found in this paper.

References


