1974

The Statistical Development of an Index of Office Performance

William (Bill) H. Williams
CUNY Hunter College

Hwei Chen
Bell Laboratories

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/hc_pubs
Part of the Business Commons, and the Social and Behavioral Sciences Commons

Recommended Citation
https://academicworks.cuny.edu/hc_pubs/585

This Article is brought to you for free and open access by the Hunter College at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
The Statistical Development of an Index of Office Performance

W. H. Williams¹ and H. J. Chen²
¹ Hunter College, New York, NY
² AT&T Bell Laboratories, Murray Hill, NJ

ABSTRACT

In many organizations, similar operations are carried out at different locations. This is certainly true of plant operations, and for many corporations, is also true for their business offices. For example, in the AT&T System, the everyday contact with customers is performed in approximately twenty-one hundred different business office locations. These offices all perform very similar work functions. Other organizations such as the utilities and various government agencies also have this characteristic in that many different office locations perform essentially the same work tasks.

This situation leads to consideration of formal measurement schemes which can be used to compare the performance of these offices in an objective way. There are actually two types of important comparisons. The first is the comparison of an office with its own earlier performances. The second is the comparison of two different offices. The same index of office performance should not necessarily be used for both types of comparisons. This aspect of the measurement of office performance does not seem to have been discussed by earlier authors and we believe the procedures used to adjust indices for environmental differences are new. It is important to notice that these procedures do not involve expensive work sampling.

The problem is discussed with specific emphasis on the statistical models used to develop both intra- and inter-office comparisons. The models described were developed for use in the AT&T System, however, they are described with emphasis on development rather than the implications of their use in the AT&T System.
The Statistical Development of an Index of Office Performance

W. H. Williams¹ and H. J. Chen²
¹ Hunter College, New York, NY
² AT&T Bell Laboratories, Murray Hill, NJ

1. INTRODUCTION

In many businesses, similar operations are carried out at different locations. For example, in the AT&T System, the everyday contact with customers is carried out in approximately twenty-one hundred different business offices. Such a situation often leads to consideration of formal measurement schemes which can be used to compare the performance of the various business units in an objective way. Although such schemes can be very useful, they can also be very troublesome and even detrimental.

The first of these troubles relates to the operational definition of performance. It should not be so broad that it is meaningless or misleading, nor should it be overly narrow. Additionally, this definition must reflect the real characteristics that are associated with performance. While this problem of the meaning and measurement of performance efficiency is obvious, it is not necessarily easy to solve. On the other hand, the problems which come about as a result of the influence of the measurement scheme on the business itself are usually neither easy to solve nor obvious. For example, if a scheme is not carefully evaluated, it can modify the activities of the business operation in undesirable ways. To illustrate, if inappropriate subcategories of work are measured for use in a final efficiency ranking, the result may be that these subcategories become ends in themselves, simply to get the credit associated with them. This same characteristic may also work against reorganizations which are designed to take advantage of local characteristics. An office unit able to automate certain operations may not feel inclined to do so if it will eliminate categories of work for which "credit" is given. However, it should be noted that interactions such as this can be potentially useful for inducing desired objectives. Of course, this kind of "tinkering" requires a good deal of knowledge about the unit and the interaction.
Another difficulty with performance measurement schemes is that there may be real reasons why a task is more difficult to perform in one part of the country than another. That is, there may be real external influences on a particular location which ensure that its task is harder (or easier) to perform than it is in other locations. Such factors must be accounted for in the measurement scheme.

A final, and perhaps tangential, difficulty with any analytic measurement scheme is that it will not separate the offices into those which are "efficient" and those which are "inefficient." Such a separation is usually achieved by a comparison with external norms which may be obtained either from statistical studies or from theoretical considerations. However, at some stage the separation must always require the judgement of management.

In summary, to be useful to management, a measurement scheme must relate to and measure some understandable characteristics of the work performance in such a way that it is informative, and not potentially misleading. At the same time, it must not interact with the actual work procedure in such a way that it invites less efficiency. And it must allow the local managers to be flexible. It naturally follows that if meaningful measurements can be constructed, they would be very helpful to the immediate office management as well as to the higher level personnel.

This paper describes the statistical development of a model which was designed to possess as many as possible of the characteristics just discussed. The data used are from a sample of forty-six AT&T System business offices. These offices are very much alike in that they perform most of the functions required in the daily contact with the customers of the AT&T System. For example, they handle the orders for new service, toll inquires and complaints. In this paper, the statistical procedures and models are emphasized rather than the details of the use of these models in the AT&T System.
THE FORMULATION OF A MODEL

2.1 The Single Office Model

It is possible to undertake studies of this kind either in terms of time expenditure or in terms of cost data. Even though cost analysis is a natural and important aspect of any overall efficiency study, no cost data are presented in this paper. In this case, the vast geographical area covered ensures that cost factors are highly variable and require careful, separate consideration. Consequently, it was decided for this study to use time expenditure as the basis, and to restrict the study to performance, where performance is measured in terms of time usage without reference to cost factors.

To construct a first model for daily time expenditure in a single office, assume that time is used up, partly as a result of direct customer demands and partly by overhead time, see Equation (2.1).

\[
\text{time spent on all commercial office work} = \text{overhead time} + \text{time required for customer generated demands}
\]

(2.1)

Next, suppose that the time required to carry out a single customer contact in the jth work category is \(a_j\), and that it is performed \(F_{ij}\) times on day i. Then the total time spent that day on category j is \(F_{ij}a_j\) and the right bracket of the right side of Equation (2.1) could be written as

\[
\sum_{j=1}^{k} F_{ij}a_j
\]

where \(k\) is the total number of work categories. Thus, if \(a_0\) denotes overhead time, Equation (2.1) can be written as

\[
T_i = a_0 + \sum_{j=1}^{k} F_{ij}a_j
\]

(2.2)
where \( T_i \) is the total time expenditure on day \( i \), \( i = 1, 2, 3, \ldots, n \). It is doubtful that such exact relationships ever hold, so that an obvious next step is the introduction of the random component shown in the elementary statistical model of Equation (2.3).

\[
T_i = a_0 + \sum_{j=1}^{k} F_{ij} a_j + e_i, \quad i = 1, 2, \ldots, n.
\]  

Equation (2.3) has the form of a standard regression model but, as we shall see, cannot be used directly in time studies without some modification.

With the model of Equations (2.2) and (2.3) in mind, data were gathered from each of the 46 sample offices. First, daily frequency counts were obtained for \( k = 2 \) categories of work. One of the categories is the number of times in which a demand request (such as a new installation) was received by that office from a business customer. The second category is similar except that it applied to residence customers. These counts are routinely available in each office, as is the daily number of hours spent by the office staff on all customer demand work. Consequently, in the notation of Equation (2.2), \( T_i \) is the gross time on commercial operation on day \( i \), \( F_{i1} \) is the total number of business contacts on day \( i \), and \( F_{i2} \) is the total number of residence contacts on day \( i \). The observations were gathered for each of the 46 offices for \( i = 1, 2, 3, \ldots, n = 65 \) days. The simple regression model was then fitted to each of the 46 offices separately.

This initial analysis of the 46 offices revealed two factors, both of which suggested a log transformation of the data. First, it was found that the \( T \)'s and the \( F \)'s are non-linearly related; time expended does not increase in proportion to the number of customer contacts. This is a reflection of increasing productivity in that the more a particular item of work needs to be carried out, the faster it is done. Presumably, however, this relationship could not be extrapolated indefinitely. At some point further, time-efficiency gains would not be possible. The point at which such gains would no longer be possible was not evident in our data however.
The next characteristic which suggested a log transformation was that the standard deviation of the number of contacts (per office) is approximately linearly related to the mean number of contacts. Since both of these factors suggested a log transformation, the functional form of the model was modified and became,

\[ t_i = B_0 + \sum_{j=1}^{k} B_j f_{ij} \]  \hspace{1cm} (2.4)

where \( t_i = \log T_i \) and \( f_{ij} = \log F_{ij} \). Notice that the time parameters are now measures of the time required per log contact; in the discussion which follows this will not always be repeated but the meaning should be clear.

Equation (2.4) was then used as the basis for a simple separate regression model for each of the 46 offices. The associated assumptions were the standard ones for linear regression models. For illustration, some statistical details of this fit to data obtained from an office in Pittsburgh are presented in Table 2.1. The fitted model explains \((0.4081/0.6542) \times 100\) or 62 percent of the variance after the mean, i.e., a multiple \( R^2 \) of 0.62. As far as the other individual offices are concerned, some have considerable better fits than Pittsburgh, others are not as good. The number of observations available for each office fit is equivalent to the number of days (65) of the study. The important result of this part of the analysis is the development of estimates of the parameters \( (\theta_0, \theta_1, \theta_2) \) for each of the 46 sample offices. This permitted a statistical study of the relationship of these time estimates to the environmental factors which are possibly influencing them. This part of the study is described in the next sections.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimated Value</th>
<th>St. Dev. of Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.515</td>
<td>0.375</td>
</tr>
<tr>
<td>1</td>
<td>0.187</td>
<td>0.054</td>
</tr>
<tr>
<td>2</td>
<td>0.211</td>
<td>0.055</td>
</tr>
</tbody>
</table>

Table 2.1: Estimated Parameters, Pittsburgh (log basis)

2.2 Difficulties with Office Comparisons

The model of Equation (2.4) is of the general form \( t = g(f_1, f_2, \ldots, f_k) \) which relates the demand put on an office and the time consumed by it. Given the separate office estimates, \((\theta_0, \theta_1, \theta_2)\). The model could then be used to give estimates of the time required to carry out a certain demand load, say \((f_0^1, f_0^2, \ldots, f_0^k)\). This estimate, or allowed time, could then be compared with the actual time used, to produce a performance index. A natural way to do this is as a ratio, \( P = \frac{\text{allowed time}}{\text{actual time}} \). These performance ratios could then be computed at regular intervals, say monthly, to follow the progress of the office.

The comparison of different offices is not so simple, however. If each office is fitted separately, and performance indices obtained for each office for each month, then the offices could presumably be compared on the basis of the \( P \) scores. However, the separate office model-fitting associates each office with its own individual set of parameters, and this procedure has two defects for use in inter-office comparisons.

The first defect is operational. If the number of offices is large, it would be very cumbersome. The second defect is more devastating, however, in that such a comparison would give an inefficient office a time allowance based solely on its own inefficient procedures. Similarly, an efficient office would be hurt in the comparison by only being given a time allowance based on its own efficient organization. This is not
what is wanted. The use of such a model as a basis for comparisons would be clearly most unfair.

This latter defect suggests fitting a model to all data in which the parameters are common to all offices. This would allow all offices the same standard overhead time and the same standard times for each work classification. Such a model would be tractable and would eliminate the defect of giving each office a time allowance based on its own procedures. However, consideration of the real operation usually makes it clear that there really may be valid reasons why one office should have different time allowances from another. Consequently, we seem to stand between two models, one which allows every office the same time allowances and one which gives every office different allowances based on individual office performances.

What is needed at this stage is a method and a model which gives offices a fair time allowance, based on the factors which actually influence the performance times. Operationally, this means relating the estimates, \( (\theta_0, \theta_1, \theta_2) \), for each office to exogenous variables which describe the environment of the office and building these relationships into the final multi-office model. This second stage of analysis is described in the next section.

2.3 The Multi-Office Model

2.3.1 The Adjustment of Overhead* Time

The next stage of the study was the analysis of the time estimates, \( (\theta_0, \theta_1, \theta_2) \), for possible relationships to environmental variables. To facilitate this, a profile survey was conducted to determine basic characteristics about both the offices and their environment. In this survey, more than two hundred measurements were obtained for each office which were then studied for possible relationships to the office estimates of time usage.
It was found that overhead time, $\theta_0$, is very much related to office size as measured by $A_i$, the monthly number of accounts carried by the office. Figure 1 shows a plot of $\log \theta_0$ against $\log \left( \frac{A_i}{100} \right)$, $i = 1, 2, \ldots, 46$. The strength of this relationship suggests that the overhead time allowance for each office be adjusted by the number of accounts carried by it. This is logical because an efficient large office may well have more overhead time associated with it than a poorly run small office. A regression model was fitted, but the details are not included in the paper because this fit mainly affects the decision to use $A_i$ as a scale variable for overhead time and only indirectly affects the final models. However the good linear relationship between $\log \theta_0$ and $\log \left( \frac{A_i}{100} \right)$ suggests modifying the model by inserting $\theta_0 A_i^i$ for the $\theta_0$. This modification was made and the details are given in section 2.3.3.

* Overhead time means time for which no frequency count can sensibly be made.

2.3.2 The Adjustment of the Business and Residence Contact Times

As has been pointed out, the time that it takes an office to carry out a business or residence contact may well be influenced by outside factors. The hope was that the estimates $\theta_1$ and $\theta_2$ would be related to variables that were included among the profile variables. A number of good relationships were found. Percent employment losses (Service Representatives) and number of main business telephones were found to be approximately logarithmically related to the business time parameter. Percent residence main and number of live customer bills handled were related to the residence parameter in a similar way. An example is given in Figure 2.

2.3.3 The Multi-Office Model
Using the relationships discussed in Section 2.3.2, the multi-office model was modified in a manner similar to that described for overhead time. Specifically, the multi-office model was put in the form,

\[ T = A F_1 \log C_1 + F_2 \log C_2 + \log C_3 \]  

(2.5)

where \( C_1 \), \( C_2 \), and \( C_3 \) are selected profile variables and the other variables, \( A \), \( T \), \( F_1 \) and \( F_2 \), are as previously specified. As we saw in the last section, useful profiles are \( C_1 \), percent service representative losses; \( C_2 \), percent business telephones; and \( C_3 \), the number of main telephones.

Some statistical details of the fit of this model to the combined data of all offices are given in Tables 2.2 and 2.3.

The percentage variance removed by the parameters after that mean, i.e., \( R^2 \), is (983/1062) x 100 or approximately 93.0 percent. This is a very satisfactory result because the final model not only has logic to its functional form but, in addition, fits the data very well. This is very important, because any model which has no relationship to the statistical facts would be of little practical value.

In summary to this point, the time allowance for an office would be made up of two components as follows,

\[ \text{Time allowance} = \text{Overhead time allowance} + \text{Allowance for time generated by customer demands} \]  

(2.6)

The overhead time allowance is based on the size of the office, adjusted by the number of accounts carried. The time allowed for customer generated demands is based on the \((\log)\) number of contacts multiplied by an allowed time per contact. For business contacts, this allowed time is bigger for offices with higher service representative losses. For residence contacts, the per contact time allowance is higher for offices which have a higher percentage of business telephones and number of main station.
3. **SOME ACTUAL OFFICE COMPARISONS**

To make some actual office comparisons, the parameter estimates presented in the previous section were used along with actual time \( (T_i) \) and frequency counts \( (F_{i1}, F_{i2}) \) for a three-month period to produce monthly time allowances for each of the 46 offices. This allotted time was then compared with the actual time used as a percentage \( P = \frac{\text{allowed time}}{\text{actual time}} \times 100 \). Thus, at the end of each month, each office received a performance rating which allows two types of comparisons. The first is the month-to-month comparison of each office with itself, and the second is the comparison of offices with each other. It is important to notice that these are different comparisons and that an office may slip in comparison with itself from one month to the next but rank higher when compared with all other offices. A sample of the performance indices and rankings is given in Table 3.1.

Since each \( P \) value is estimated by regression methods, each one has a different standard error. For the \( j \)th month, the standard error of \( P_j = t_j/t_i \) can be approximated by the square root of the mean square error of the regression divided by \( t_j \). For example, the standard error of the \( P = 96.1 \) value of office 1 for month one is 0.80.

Finally, the \( P \) rankings must be checked to see if there is any obvious systematic behavior. The presence of such behavior might mean that the scheme is favoring certain offices. To check this, the relationship of the \( P \) values to the gross time used by the office and all variables used as inputs into the model were studied, but no evidence of any relationship was found. The interpretation of this is that, for these variables, the measurement scheme is not favoring any particular kind of office.

The residuals from the regression fits were also checked for time dependencies, both by graphical techniques and the formal Durbin Watson statistic, but no systematic behavior was found.

4. **SUMMARY**
In this paper we have discussed some of the procedures and difficulties associated with the development of statistical performance measures which can be used for the assessment of separate offices which perform essentially the same set of work operations. The procedures are developed and illustrated using a sample of 46 Bell System business offices. These offices are organized in such a way that the demands placed upon them are very similar. The techniques are general, however, and will be found useful in other large organizations, such as the utilities and the government. In fact, the procedures have been used by at least two government agencies.

For reasons discussed in the paper, it is not always appropriate to use the same performance index to compare the same office in different time periods as it is to compare two different offices in the same time period. However, the model presented is specifically designed to permit both intra- and inter-office performance comparisons. Models with this characteristic have not previously appeared in the literature.

Operationally, the performance indices were developed in different working stages. The first stage was a separate analysis of each of the 46 sample business offices. At this stage, estimates were obtained for each office of overhead time ($\square_0$), the average time required for each business contact ($\square_1$), and the average time required for each residence contact ($\square_2$).

Since the individual office time estimates might reasonably be expected to be related to environmental factors not yet considered in the model, the second stage of the analysis was to examine the 46 sets of estimates, ($\square_0, \square_1, \square_2$), for their relationship to these environmental variables. A number of such relationships were found for each of the time parameters, $\square_0$, $\square_1$, $\square_2$.

In the next stage, the developed relationships of the estimates, ($\square_0, \square_1, \square_2$), to the environmental variables were used to formulate a multi-office model with a functional form that could be reasonably used for both intra- and inter-office
comparisons. This model was then fitted to the entire set of original observations and was found to fit the data very well.

The fitted model was then used to construct performance indices for the sample of 46 offices for each of three consecutive months.

Finally, it is important to notice that the scheme requires no expensive work sampling.