ADDENDUM

As stated in Chapter 3, the vote counting methodology has certain limitations as an analytical tool, (e.g., mixing apples and oranges, combining studies that have been well done with some that have not, combining studies of different sample sizes. Grame Hodge (2000, p65) also says vote counting is statistically unreliable, can lead to false conclusions because it only uses part of the available information, and does not provide an estimate of effect size when the null hypothesis is false. In other words, vote counting uses data that lack statistical rigor, data that does not rise to the level of being useful for the accumulation of correlations procedures associated with more statistically rigorous procedures such as meta-analysis. However, that is not to say that the non-parametric data used in the vote counting procedure in this dissertation cannot have non-parametric statistical analysis applied to it, such as chi-square.

Without implying that this dissertation contains a statistical rigor and a higher level analysis than is really there, a chi-square analysis was conducted in order to determine whether the findings of the vote counting study just completed can be generalized to the broader population of outsourcing efforts. Two characteristics of the outsourcing efforts were analyzed: 1) whether the outsourced services were of a harder to measure professional nature, requiring much discretion, or were more easily measured discrete services of a more physical nature, and 2) what level of government was doing the outsourcing. **In summary, this analysis found that with 90% confidence, we can say that 5% of the savings can be explained by the nature of the outsourced job, and 27% of the savings can be explained by the level of the government doing the**

outsourcing. Also with 90% confidence, we can say that 15% of the service quality improvements can be explained by the nature of the outsourced job, and 35% of the service quality improvements can be explained by the level of the government doing the outsourcing.

Chi Square Analysis

Chi square is a non-parametric test of statistical significance for bivariate tabular analysis. Such a test can help determine the degree of confidence one can have in accepting or rejecting a hypothesis. In this case the hypothesis is that two or more different samples are different enough in some characteristic or aspect of their behavior that we can generalize from our samples that the populations from which our samples are drawn are also different in the characteristic or behavior.

Chi square is a rough estimate of confidence that accepts weaker, less accurate data as input than parametric tests (like t-tests and analysis of variance for example). Because of this limitation, chi square has less status than other statistical tests. However, this limitation is also its strength, because it is more forgiving in the data it will accept, and therefore can be used in wide variety of research, such as this vote-counting analysis of outsourcing effectiveness.

Bivariate tabular analysis is used when we are trying to summarize the intersections of independent and dependent variables and understand the relationship (if any) between those variables. The independent variable is the quality or characteristic that we

hypothesize will help explain some other quality or characteristic. In this case, we want

to know if there is any relationship between:

1) the independent variable such as human attributes required for a job (e.g., whether

    the job required more or less physical abilities), and the dependent variable,

    namely successful outsourcing of that job in terms of saving money and/or

    improving service quality;

2) the independent variable such as the government level of the service being

    outsourced (federal, state, or local) and the same dependent variable just

    mentioned.

The first step in determining these relationships was to re-sort the data along the

independent variable. Table A-1 (at the end of this addendum) was derived by re-sorting

Table 4-1 after adding another column which designated whether each of the 218 services

outsourced required the individual performing that service to use more or less physical

abilities, i.e., primarily his/her mental skills or his/her physical skills. Admittedly, all

jobs require some of both, but most require more of one than the other. For example, it is

safe to say that grounds-keeping requires more physical than mental attributes, while

providing legal aid to indigent arrestees requires the opposite. Using this as a guideline,

an educated determination about each service's attributes was made based on the title of

the service and any other information provided by the original source document. (Note

that Table 4-1 contains 222 entries while the Chart A-1 only has 207. This is because

four of the sources cited in the vote-counting contained multiple services whose job

attributes could not be determined either way, and another 11 of the individual services

included in the vote counting did not assess whether savings had been achieved through

outsourcing. Chart A-2 only has 141 observations because four of the sources cited in the

vote-counting contained multiple services whose job attributes could not be determined

either way, and another 83 of the individual services included in the vote counting did not

assess whether the quality of the service provided had improved through outsourcing).

Chart A-1
Savings vs Job Attributes
(Observed Frequencies)

Job Attribute	Savings (Yes)	Savings (No)	Total
More Mental	44	10	54
More Physical	117	36	153
Total	161	46	207

Chart A-2
Quality Improvements vs Job Attributes
(Observed Frequencies)

Job Attribute	Improved Quality (Yes)	Improved Quality (No)	Total
More Mental	27	18	45
More Physical	41	55	96
Total	68	73	141

Interpreting these tables is most easily done by converting the raw frequencies in

each cell into percentages of each cell within the categories of the independent variable.

Percentages basically standardize cell frequencies as if there were 100 observations in

each category of the independent variable. Charts A-3 and A-4 show this conversion.

Chart A-3
Savings vs Job Attributes
(Observed Frequencies in Percentages)

Job Attribute	Savings (Yes) %	Savings (No) %	Sample Size
More Mental	81	19	54
More Physical	76	24	153

Chart A-4
Quality Improvements vs Job Attributes
(Observed Frequencies in percentages)

Job Attribute	Improved Quality (Yes) %	Improved Quality (No) %	Sample Size
More Mental	60	40	45
More Physical	43	57	96

Based on this information, it appears that it really doesn't matter whether a job requires mostly mental or physical attributes in order for it to result in savings between 76% and 81% of the time. However, the difference between metal and physical jobs is more pronounced when it comes to improving service quality. While 60% of the services requiring more mental activities that were contracted out resulted in quality improvements, only 43% of the more physical services contracted out resulted in quality improvements.

This exercise has allowed us to see more easily patterns in the data, but it tells us nothing about how confidently we can generalize our findings to the larger population from which the data was drawn. To accomplish this, we need to do a test of statistical significance such as chi square which compares our actual observed frequencies in our sample with the frequencies we would expect if their was no relationship at all between the two variables in the population from which the sample was drawn.

As stated earlier, the chi square test is forgiving in that it is more lenient about what data is useable. It also is non-parametric meaning that it does not require the sample data to be more or less normally distributed (like a t-test does), although it does rely on the assumption that the variable is normally distributed in the population from which the sample is drawn. However, chi square does have the following requirements:

1) <u>The sample data must be randomly drawn from the population</u>. With regard to this vote-counting study, it could be argued that the studies chosen were not randomly selected but were rather carefully selected (see the "winnowing process" of how studies were selected in Chapter 3 of this dissertation). While that is true, once the sources were found that met the vote-counting criteria, <u>all</u> of them were included in the study, regardless of whether they did or did not demonstrate savings or service quality improvements. So while the sample used in this vote-counting study comprised the entire population of legitimate, fact based, methodologically defensible sources that were found, the results shown in those sources were not random. Although this precludes the findings of this vote counting analysis from being generalizable to the general population of privatization/outsourcing literature in the strictest sense, using the criteria of only including methodologically defensible studies as the sample selection criteria should not present any systematic bias in the results.

2) <u>Data must be reported in raw frequencies (not percentages)</u>. All of the data in this

vote counting analysis was reported in raw frequencies. This requirement must be

met because if the data was already reported in percentages, we would be

standardizing already standardized measurements. Charts A-1, 2, 9, and 10

complete this step.

3) <u>Measured variables must be independent</u>. Any observation must fall into only

one category. That is true with this vote-counting analysis. An outsourced

service was evaluated to be either requiring more mental <u>or</u> more physical

attributes. It could not by definition be both. Further, an outsourced service could

only be attributable to one level of government for this part of the analysis.

3) <u>Categories of independent and dependent variables must be mutually exclusive</u>.

That is true with this vote counting analysis. No category is dependent or

influenced by another. For example, one outsourced service that is primarily

physical in nature does not influence whether or not the next outsourced service in

the study resulted in savings or improved quality.

4) <u>Observed frequencies cannot be too small.</u> Ideally, chi square analysis should be

conducted on data in its uncollapsed form. However, if cells in bivariate tables

show very low observed frequencies (5 or below) the expected frequencies may

be too low for chi square to be appropriately used. In this vote counting study, for

example, because the original data was sorted into 26 different types of jobs (in

order to prevent the mixing of apples and oranges), it resulted in some categories having too few data points (e.g., only one zoo-keeping data point).   Thus, in order for chi square analysis to be workable, the data had to be rearranged into fewer categories such as was done in Charts A-1, 2, 9, and 10.

5) <u>Minimum frequency thresholds should be obeyed.</u> For a 1 X 2 or 2 X 2 table, expected frequencies in each cell should be at least 5. By sorting the data according to only one dimension of job attributes and by level of government, this criterion was met in practically all cases.

The first step in chi square analysis is to determine what threshold of tolerance for error we are willing to accept that we are wrong in generalizing from the results in our sample to the population it represents. Since no one's life depends on the interpretation of these results, a probability of error threshold of 1 in 10 (90% confidence level) will be set.

The next step is to calculate the expected frequencies in each cell which is the product of that cell's row total multiplied by that cell's column total, divided by the sum total of all observations. For example, the expected frequency in the 'more mental job' resulting in saving is (54 X 161)/207 or 42. The expected frequencies for Savings and Quality Improvements are shown in Charts A-5 and A-6 respectively.

Chart A-5
Savings vs Job Attributes
(Expected Frequencies)

Job Attribute	Savings (Yes)	Savings (No)	Total
More Mental	42	12	54
More Physical	119	34	153
Total	161	46	207

Chart A-6
Quality Improvements vs Job Attributes
(Expected Frequencies)

Job Attribute	Improved Quality (Yes)	Improved Quality (No)	Total
More Mental	22	23	45
More Physical	46	50	96
Total	68	73	141

We now have a comparison between the observed results versus the results we would expect if the null hypothesis were true, which states that there is no relationship between savings or quality improvements in service whether a service involves more or less physical work than mental work. But in order to determine whether we can reject the null hypothesis, and whether or not we would be making a mistake in generalizing from our sample results to the larger population of outsourcing experiences, we need to calculate the difference between the observed and expected frequency in each cell, square that difference, and then divide that product by the difference itself. Squaring the difference ensures a positive number so that we end up with an absolute value of differences. The last step is to add up the products which gives us the chi square value. In this case:

Chart A-7
Savings vs Job Attributes
(Observed vs Expected Frequencies)

Job Attribute	Savings (Yes)	Savings (No)	Total
More Mental (Obs)	44	10	54
More Mental (Exp)	42	12	54
More Physical(Obs)	117	36	153
More Physical(Exp)	119	34	153

(44 Mental/Yes Obs - 42 Mental/Yes Exp) = 2 squared = 4 divided by expected 42 = .095
(10 Mental/No Obs - 12 Mental/No Exp) = -2 squared = 4 divided by expected 12 = .333
(117 Phys/Yes Obs – 119 Phys/Yes Exp) = -2 squared = 4 divided by expected119 =.034
(36 Phys/No   Obs–   34 Phys/ No Exp)= 2 squared = 4 divided by expected 34= .<u>118</u>
Chart A-7 Total Chi square =   .580

Chart A-8
Quality Improvements vs Job Attributes
(Observed vs Expected Frequencies)

Job Attribute	Improved Quality <u>(Yes)</u>	Improved Quality <u>(No)</u>	Total
More Mental (Obs)	27	18	45
More Mental (Exp)	22	23	45
More Physical(Obs)	41	55	96
More Physical(Exp)	46	50	96

(27 Mental/Yes Obs – 22 Mental/Yes Ex) =5 squared = 25 divided by expected 22= 1.14
(18 Mental/No Obs – 23 Mental/No Ex) = -5 squared = 25 divided by expected 23= 1.09
(41 Phys/Yes Obs – 46 Phys/Yes Ex) = -5 squared = 25 divided by expected 46 = 0.54
(55 Phys/No   Obs – 50 Phys/No Ex) =   5 squared = 25 divided by expected 50= <u>0.50</u>
Chart A-8 Total Chi square =   3.27

Interpreting the Chi Square Value

Now we need to know how much larger than zero (the absolute chi square value

of the null hypothesis) our chi square value must be before we can confidently reject the

null hypothesis. This requires us to first determine the degrees of freedom using the

following formula: df = (r – 1) (c – 1) where 'r' is the number of rows in the table and 'c'

is the number of columns. Getting back to the original data as shown in Table 7-1 we get

the following: df = (2-1)(2-1) = 1(1) = 1.

Next, using the sampling distribution of chi square (also known as 'critical values

of chi square') in the back of any statistics book, with one degree of freedom and a

confidence level of 90 percent that our data does represent the population about which we

are generalizing, reveals a critical value of 2.706. This can be interpreted to mean that

since the calculated chi square in Chart A-7 (.58) is less than the critical value, then we

cannot make the statement that there is a meaningful relationship between whether an

outsourced service is primarily physically or mentally demanding, and savings achieved

from that outsourcing. On the other hand, because the calculated chi square in Chart A-8

(3.27) is greater than the critical value, we can say with a 90% degree of confidence that

there is a relationship between whether an outsourced service is more physically or

mentally demanding, and whether or not the quality of that outsourced service improved

after outsourcing. This relationship can not be proven with chi square to be causal in

nature, but the relationship between variables described in these results indicates that the

results are not attributable to random error.

<u>Measure of Association</u>

In order to measure the strength of the relationship between these variables, and

because we're using a 2 X 2 table, it is necessary to calculate a correlation coefficient

called 'phi' as follows:

Savings phi = square root of the result of dividing the calculated chi square by the total number of observations or the square root of (.58/207) = square root of .00280 = **.0527**.

Quality phi = square root of the result of dividing the calculated chi square by the total number of observations or the square root of (3.27/141) = square root of .02319 = **.1523**.

In keeping with the earlier finding that there is little meaningful relationship between savings achieved from outsourcing services and whether or not those services are more mentally or physically challenging, the calculated savings phi of .0527 indicates that only 5% of those savings are explained by the physical or mental demands of the job. On the other hand, over 15% of improvements in outsourced service quality can be attributed to the physical or mental aspects of the outsourced activity. This means there are one or more variables which cumulatively account for and predict the remaining 95% of the influence on savings from outsourcing, and the remaining 85% of influence on service quality improvements from outsourcing. One of those other variables is now addressed.

Level of Government

Another factor that may influence the success of outsourcing efforts is the level of government that is conducting the effort. Based on the data provided earlier in Table 4-2 and discussed in Chapter 5 of this dissertation, the following bivariate tables result:

Chart A-9
Savings vs Government Level
(Observed and Expected Frequencies)

Gov't Level	Savings (Yes)	Savings (No)	Total
Federal (observed)	55	7	62
Federal (observed %)	89	11	100
Federal (expected)	49	13	62
State (observed)	92	22	114
State (observed %)	81	19	100
State (expected)	90	24	114
Local (observed)	18	15	33
Local observed %)	55	45	100
Local (expected)	26	7	33
Total	165	44	209

(55 Fed/Yes Obs - 49 Fed/Yes Exp = 6) squared =   36 div by expected 49 =   .7347
(7 Fed/No   Obs - 13 Fed/No Exp = -6) squared =   36 div by expected 13 = 2.7692
(92 State/Yes Obs - 90 State/Yes Exp =2) squared =   4 div by expected 90 =   . 0444
(22 State/No Obs - 24 State/No Exp = -2) squared =   4 div by expected 24 =   .1667
(18 Local/Yes Obs - 26 Local/Yes Exp =-8)squared =   64 div by expected 26 = 2.4615
(15 Local/No Obs - 7 Local/No Exp = 8) squared =   64 div by expected 7 =   9.1429
Chart A-9 Total Chi square =   15.3194

Using the degrees of freedom formula, $df = (r – 1) (c – 1)$ where 'r' is the number

of rows in the table and 'c' is the number of columns, we get the following: $df = (3-1)(2-1) = 2(1) = 2$.

The chi square critical values table at the back of any statistics book, with two

degrees of freedom and a confidence level of 90 percent that our data does represent the

population about which we are generalizing, reveals a critical value of 4.605. This can be

interpreted to mean that since the calculated chi square in Chart A-9 is greater than the

critical value, we can say with a 90% degree of confidence that there is a relationship

between whether an outsourced service is at the federal, state, or local level, and whether

or not savings were achieved as a result of that service being outsourced. This

relationship can not be proven with chi square to be causal in nature, but the relationship between variables described in these results is probably systematic in the larger population and probably is not attributable to random error.

Measure of Association

In order to measure the strength of the relationship between these variables, and because we're using a 3 X 2 table, it is necessary to calculate a correlation coefficient called 'Cramer's phi' which = the square root of (chi square divided by (N times (k -1)) where N = the total number of observations, and k = the smaller of the number of rows or columns. So, for Table 7-9

1. N(k-1) = 209(2-1) = 209(1) = 209

2. chi square/209 = 15.3194/209 = .0733

3. square root of .0733 = .2707 (This product is interpreted as a Pearson r, that is, as a correlation coefficient.

The calculated Pearson r tells us that approximately 27% of the savings achieved from outsourcing efforts identified in this vote counting study are explained by the level of government doing the outsourcing. If we consider that there was no overlap in cases between the 27% of savings explained by the level of government and the 5% that can be explained by whether or not an outsourced job is more or less physically challenging, that would mean that at least two thirds (68%) of the savings are explained by other factors.

Chart A-10
Quality Improvement vs Government Level
(Observed and Expected Frequencies)

Government Level	Quality Improvement (Yes)	Quality Improvement (No)	Total
Federal (observed)	7	1	8
Federal (observed %)	88	12	100
Federal (expected)	4	4	8
State (observed)	58	52	110
State (observed %)	53	47	100
State (expected)	53	57	110
Local (observed)	3	19	22
Local observed %)	14	86	100
Local (expected)	11	11	22
Total	68	72	140

(7 Fed/Yes   Obs - 4 Fed/Yes Exp = 3) squared =   9 div by expected 4 =   2.2500
(1 Fed/No   Obs – 4 Fed/No Exp = -3) squared =   9 div by expected 4 = 2.2500
(58 State/Yes Obs - 53 State/Yes Exp = 5) squared = 25 div by expected 53 =   .4717
(52 State/No   Obs - 57 State/No Exp = -5) squared = 25 div by expected 57 =   .4386
 (3 Local/Yes Obs – 11 Local/Yes Exp = -8)squared = 64 div by expected 11 = 5.8181
(19 Local/No Obs - 11 Local/No Exp = 8) squared = 64 div by expected 11 = 5.8181
Chart A-10 Total Chi square =   17.0465

Using the degrees of freedom using the following formula, $df = (r – 1)(c – 1)$

where 'r' is the number of rows in the table and 'c' is the number of column, we get the

following: $df = (3-1)(2-1) = 2(1) = 2$.

The chi square critical values table at the back of any statistics book, with two

degrees of freedom and a confidence level of 90 percent that our data does represent the

population about which we are generalizing, reveals a critical value of 4.605.   This can

be interpreted to mean that since the calculated chi square in Chart A-10 is greater than

the critical value, we can say with a 90% degree of confidence that there is a relationship

between whether an outsourced service is at the federal, state, or local level, and whether

or not improvements on the quality of service were achieved as a result of that service being outsourced. This relationship can not be proven with chi square to be causal in nature, but the relationship between variables described in these results probably is systematic in the larger population and probably is not attributable to random error.

Measure of Association

Using the same procedures listed above for Table 7-9, Table 7-10 shows:

1) $N(k-1) = 140 (2-1) = 140 (1) = 140$

2) chi square/140 = $17.0465/140 = .1218$

3) square root of $.1218 = .3489$ (This product is interpreted as a Pearson r, that is, as a correlation coefficient.

The calculated Pearson r tells us that approximately 35% of the improvements in service quality achieved from outsourcing efforts identified in this vote counting study are explained by the level of government doing the outsourcing. If we consider that there was no overlap in cases between the 35% of savings explained by the level of government and the 15% that can be explained by whether or not an outsourced job is more or less physically challenging, that would mean that at least half (50%) of the improvements in service quality are explained by other factors.

(Note: The analysis in this section relies heavily on the "Chi Square Tutorial" published by Professor Jeff Connor-Linton from the Department of Linguistics at Georgetown University. The tutorial can be found at: www.georgetown.edu/faculty/ballc/webtools/web_chi_tut.html)