A Poisson Regression Examination of the Relationship between Website Traffic and Search Engine Queries

Tierney, Heather L.R. and Pan, Bing

College of Charleston

5 November 2009

A Poisson Regression Examination of the Relationship between Website Traffic and Search Engine Queries

By

Heather L.R. Tierney* and Bing Pan

January 2010

Abstract

A new area of research involves the use of Google data, which has been normalized and scaled to predict economic activity. This new source of data holds both many advantages as well as disadvantages, which are discussed through the use of daily and weekly data. Daily and weekly data are employed to show the effect of aggregation as it pertains to Google data, which can lead to contradictory findings. In this paper, Poisson regressions are used to explore the relationship between the online traffic to a specific website and the search volumes for certain keyword search queries, along with the rankings of that specific website for those queries. The purpose of this paper is to point out the benefits and the pitfalls of a potential new source of data that lacks transparency in regards to the original level data, which is due to the normalization and scaling procedures utilized by Google.

KEYWORDS: Poisson Regression, Search Engine, Google Insights, Aggregation, Normalization Effects, Scaling Effects

JEL Classification Codes: C25, C43, D83

Contact author: Heather L.R. Tierney, School of Business and Economics, College of Charleston; 5 Liberty Street, Charleston, SC 29424, email: hlrtierney@yahoo.com; phone: (843) 953-7070; fax: (843) 953-5697, and Bing Pan, Department of Hospitality and Tourism Management, School of Business, College of Charleston; 5 Liberty Street, Charleston, SC 29424, phone: (843) 953-2025; fax: (843) 953-5697; email: bingpan@gmail.com. The authors would like to thank the following people in alphabetical order for their very helpful comments: Penny Goldberg, Thomas Lemieux, Steve Pischke, and Martin Ridout.

1. Introduction

A burgeoning field of research involves using data from Google Insights or Google Trends to predict economic activity. The lure of using Google data, which has already been normalized, scaled, and aggregated, is that one has instantaneous access to worldwide internet activity. This boon also has some problems with the main one being a lack of transparency in regards to the original level data although Google is transparent in the fact that it *only* provides data that has been normalized and scaled.[1] The main reason that Google normalizes and then scales *all* of their data obtained from Google Insights or Google Trends is due to privacy concerns such as the recent controversies of revealing a user's search history through search log data (Barbaro and Zeller 2006). This necessary protection of the privacy of the user of a search engine can be problematic to the researcher due to potentially limiting the sample size of a given data-frequency, which makes studying long-run trends virtually impossible (Rapach 2003, Gagnon 2008). In addition, the interpretation of the regression results based upon normalized and scaled data is not as straight forward, which has important policy implications.

In regards to scaling and aggregation, studies in other fields that use scaled variables such as in Geographic Information Systems (GIS) and aggregated variables such as in time series econometrics have found that models involving scaling or data aggregation can introduce a loss of information. The loss will be larger with larger scales and greater levels of aggregation (Rossana and Seater 1995, Bian 1997). Furthermore, scaling a variable in regards to range also has the possibility of producing distortions (Pyle 1999). In general, the effect of normalization, scaling, and aggregation in regards to website traffic and search engine queries has not been directly examined, and this paper investigates the potential pitfalls of using such data.

Internet activity has been used to predict economic activity and even possible flu epidemics (Ginsberg, Mohebbi, Patel, Brammer, Smolinski, and Brilliant 2009). In regards to internet activity with respect to economic activity, Azar (2009)

[1] Google does not use the terms normalizing and scaling interchangeably.

finds a negative relationship between oil prices and shocks to Google searches for electric cars in a Bayesian Vector Autoregression (BVAR). Askitas and Zimmermann (2009) observe a strong correlation between certain keywords such as *unemployment office or agency*, *unemployment rate*, *Personnel Consultant*, and *most popular job search engines in German* and the monthly German unemployment rate using Engle and Granger's (1987) error correction model. In a technical paper, Choi and Varian (2009b) state that the out-of-sample fit of U.S. initial unemployment claims is better explained with the inclusion of data from Google Trends in an ARIMA framework. Using daily and weekly data from Google Trends, Choi and Varian (2009a) also look at the relationship of retail sales and automotive sales using a seasonal autoregressive (AR) model, home sales in an AR model, and travel, with respect to visitor arrival in Hong Kong, in a fixed effects model in an earlier technical paper.

The purpose of this paper is two-fold. The first purpose is to investigate the effect that normalized, scaled, and aggregated variables of internet activity have on the empirical results, which has an important bearing on a new frontier of research that involves Google data. The second purpose is to understand the relationship between website traffic to a given website and keyword search queries as well as the rankings of that specific site for those queries. The understanding and modeling of website traffic could have important implications in terms of the gathering of predictive variables from external business environments, which could help in predicting revenue generation for individual businesses.

All of the data obtained from Google Insights or Google Trends is first normalized and then scaled, which truncates the data. Without loss of generality, daily and weekly data obtained from Google, as it pertains to the Charleston Area Convention and Visitors Bureau (CACVB) website, is used to demonstrate the effect of normalization, scaling, and then the aggregation.[2]

---

[2] The web address of the CACVB website is as follows: www.charlestoncvb.com.

A Poisson regression model with a conditional exponential mean function is adopted to model the website traffic of the Charleston Area Convention and Visitors Bureau (CACVB) website. Following the reasoning of Michener and Tighe (1992), the Poisson model is used since website traffic is a non-negative integer variable (Cameron and Trivedi 1998). In addition, the use of the Poisson regression permits the elimination of the interpretation problem of the regression coefficients obtained from data that has been normalized and then scaled. The nature of Poisson regression permits the interpretation of regression coefficients as elasticities or semi-elasticities, which is not automatically the case with all regression models.

The most important finding of this paper is that the frequency of Google data used in the regressions can greatly impact the empirical findings in terms of the magnitude of the estimated coefficients and even possibly with respect to the statistical significance of an estimated coefficient. This paper also finds that search volumes for certain keyword search queries has a larger impact on website traffic than does the ranking of a website especially with respect to the CACVB website.

The structure of this paper is as follows: Section 2 presents the theoretical model, Section 3 discusses the empirical findings, and Section 4 concludes.

2. Theoretical Model

The data used as the regressands are various forms of website traffic to the CACVB website, which is a count variable that takes on only non-negative integer values, $\left\{ y_t \in \mathbb{Z}^+ : y_t \geq 0 \right\}$ with $t = \left\{1, \cdots, T\right\}$. Given the nature of the data, this paper assumes that y_t is a Poisson distributed variable and is modeled using the Poisson generalized linear model (GLM) with the conditional mean being equal to the conditional variance, which is as follows:

$$y_t = m\left(\mathbf{x}_t, \beta\right) = \exp\left(\mathbf{x}_t'\beta\right) = \exp\left(\beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt}\right) \quad (1)$$

where $\mathbf{x}_t$ is the regressor matrix (Cameron and Trivedi 1998, Wooldridge 1999).

Since the assumption of the conditional mean and conditional variance being equal are too stringent and unrealistic for estimation purposes, the Poisson quasimaximum likelihood estimator (QMLE) is used since it relaxes the assumption that the conditional variance and conditional mean are the same. For this paper, the quadratic hill climbing method is implemented to estimate the Poisson QMLEs.

The main benefit of using the Poisson QMLE is that it permits for the inclusion of randomness with the addition of asymptotically normal error terms. This transforms y_t into a compound Poisson variable, which is needed to relax the assumption of the conditional mean and conditional variance being equal while being able to produce non-negative values of y_t (Cameron and Trivedi 1986, Wooldridge 1993). Another benefit of using the Poisson QMLE is that the parameters are consistent even if the underlying distribution is incorrectly specified. In addition, the Poisson QMLE is also relatively efficient with robust standard errors, which take into account heteroskedasticity and is of the following form:

$$y_t = \exp\left(\mathbf{x}_t'\beta + \varepsilon_t\right) = \exp\left(\beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + \varepsilon_t\right) \tag{2}$$

where the error term is $\varepsilon_t \sim N\left(0, \sigma^2\right)$ (Cameron and Trivedi 1998, Wooldridge 1999).[3] It should be noted that if σ^2 is zero, then Equation (2) collapses into Equation (1).

In the estimation of Equation (2), overdispersion is found to be present, which means that $\sigma^2 > 1$, and along with the Poisson GLM variance assumption, the conditional variance is permitted to be greater than the conditional mean:

$$\operatorname{var}\left(y|\mathbf{x}\right) = \sigma^2 m\left(\mathbf{x}, \beta\right). \tag{3}$$

For this paper, the Pearson error terms, which are weighted error terms, are used as a consistent estimator of σ^2 (Wooldridge 2002).

The interpretation of the regression parameters are not as straight forward as a linear regression model since the Poisson QMLE involves the exponential mean

[3] An alternative model to the Poisson QMLE is the two-step negative binomial QMLE, which does not produce robust findings, and hence is not used in this paper (Wooldridge, 1999).

function. The marginal effect of x_j on $E\left[y_t \middle| \mathbf{x}_t\right]$ is interpreted as the proportional change of $E\left[y_t \middle| \mathbf{x}_t\right]$ by the amount, β_j, i.e.

$$\frac{\partial E\left[y_t \middle| \mathbf{x}_t\right]}{\partial x_{jt}} = \exp\left(\beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt}\right) \times \beta_j = E\left[y_t \middle| \mathbf{x}_t\right] \times \beta_j \qquad (4)$$

where $j = \{1, \cdots, k\}$ (Cameron and Trivedi 1998).

Another way of interpreting the Poisson QMLE involves taking the log of Equation (2), which is:

$$\log\left(y_t\right) = \log\left(\exp\left(\mathbf{x}_t'\beta + \varepsilon_t\right)\right) = \mathbf{x}_t'\beta + \varepsilon_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + \varepsilon_t . \qquad (5)$$

Suppose x_j is not a log variable, then β_j can be interpreted as a semi-elasticity, meaning that a one unit change in x_j will change $E\left[y \middle| \mathbf{x}\right]$ by $100\beta_j$ (Wooldridge 2002). If the regressor x_j is a log variable, then β_j can be interpreted as an elasticity, meaning that β_j is the percentage change in $E\left[y \middle| \mathbf{x}\right]$ due to a percentage change in x_j (Michener and Tighe 1992, Cameron and Trivedi 1998).

3. Empirical Results

In this section, the results of the univariate and bivariate Poisson QMLE results are estimated with respect to five different sources of website traffic. A univariate model of each of the log of the search volumes for the seven keyword search queries is formed for each of the five regressands in order to isolate the effect of the keyword searches.[4] The relationship of the five different rank regressors to the five regressands of website traffic as well as their corresponding log transformation of the keyword search volume are also examined in a bivariate Poisson QMLE model. More detail about the data is presented in Subsection 3.1, and the univariate and bivariate Poisson QMLE results are presented in Subsection 3.2.

[4] The level variables for the search volume of keyword queries have also been used in model estimation, but the findings do not vary significantly.

The previously mentioned Poisson QMLE model regressions are estimated using both daily and weekly data to see whether aggregation has an effect on the empirical findings, which this paper does find and is discussed in Subsection 3.3.

3.1 Discussion of Data

When a visitor searches for a destination, he or she will be likely to type in a query in search engines, look through the returned results, and pick a webpage to follow. Thus, search volumes for certain keyword search queries and the website's ranking for those queries in major search engines will have a significant effect on the website's traffic (Pan, Litvin, and O'Donnell, 2007). This paper focuses on three variables, which are the ranking of a site for certain keyword search queries, search volumes for those queries, and online traffic for that specific website. The data used in the Poisson regressions comes from three different Google sources. The focus search engine is Google since it has dominated the market share during the period of the study, which is from January 2008 to March 2009.

The regressands of various forms of website traffic are obtained from Google Analytics of the CACVB website. Google Analytics uses a short Javascript on every page of a website to capture visitors' visitation behavior on the website. For this paper, five regressands are specifically analyzed in this paper as is listed in Tables 1A and 1B. The first regressand, entitled *all visits*, encompasses all the website traffic to the CACVB website. A sub-category of only first visits to the CACVB website (as indentified by new internet protocol (IP) addresses) forms the second regressand of *new visits*, and website traffic that are from only search traffic via a search engine is also examined and is referred to as *search traffic*, forms the third regressand. The fourth and fifth regressands are also sub-groups of website traffic from outside the local Charleston area and is entitled *nonlocal visits* and from the local Charleston area entitled *local visits*, respectively.

The rankings of the CACVB site for five search queries in Google were obtained through a custom-built program. A program has been designed to download daily search engine results as it pertains to five different keyword search queries: *charleston sc, travel charleston, charleston hotels, charleston restaurants*, and *charleston tourism*. These queries are the popular searches used by visitors to Charleston, SC, according to Google Keyword Tool.[5]

The regressors of normalized and scaled search volumes for the five queries, during the time period of January 2008 to March 2009, are obtained from Google Insights. Google Insights is a public search tool. Since January 2004, search volumes for specific searches for up to five queries can be obtained but only for a limited time period (three months) as it might be due to the normalization and scaling processes and the protection of the privacy of Google users. The normalizing and scaling methods used for daily data changes quarterly, which is why the daily data is examined only one quarter at a time for the sake of consistency.

Sometimes in statistics, the terms normalizing and scaling are used interchangeably, but this is not the case with Google data. Regarding normalization, according to Google Insights (2009a), the raw data is sorted by a common variable and is then normalized, i.e. divided by the website traffic in a given region. This normalization process could mean that if two regions have the same percentage in terms of the search volumes for a given search query, this does not automatically mean that the absolute search volumes are the same. The reason for the normalization process is to prevent regions with higher search volume activity from dominating the rankings as displayed by Google Insights. The scaling process occurs after the normalization process. The scale of the data as it pertains to search volumes is from 0 to 100 with each data point being divided by the highest point as 100 (Google Insights 2009b). Aside from the normalization and scaling processes, which can produce distortions, it is possible for the aggregation of data to have an effect on the empirical findings (Rossana and Seater 1995, Pyle 1999). For this paper, aggregation refers to the frequency of the data, meaning the transformation of daily data into weekly data, which is discussed in more detail in Subsections 3.2 and 3.3.

For the regressors, the search volumes for specific keywords searches in two categories of general overall type of searches is denoted as 'all' and a subcategory of

[5] The web address for Google Keyword Tool is as follows: https://adwords.google.com/select/KeywordToolExternal.

queries in just travel categories as specified by Google Trends, which is denoted as 'travel' is found in Table 1B. The search volume data for specific keyword queries is normalized and scaled by Google for both the daily data and the weekly data (Askitas and Zimmermann 2009, Google Trends 2009). The daily data has the same scale for three months, which limits the analysis of the daily datasets to only one quarter at a time. The weekly data has the same normalization and scaling regardless of quarter. The daily ranking of the CACVB website are captured using the afore-mentioned custom-built program, and the weekly data are the averages over a week's period of time.

As is shown in Table 1B, the log of the search volumes in the 'all' search category is as follows: *charleston hotels*, *charleston restaurants*, *charleston sc*, *charleston tourism*, and *charleston travel*, and the queries in the 'travel' search category is as follows: *charleston hotels*, and *charleston sc*. Google Trends did not provide the other three search traffic in the travel category due to their small volumes. The log of the scaled search volumes for the keywords searches in both the 'all' and 'travel' categories is used in the Poisson regression especially since the coefficients of the regressors can be interpreted as elasticities, which is analogous to Equation (4). The rank variables involve the rankings of *charleston sc*, *charleston hotel*, *charleston restaurants*, *charleston tourism*, and *charleston travel*. The coefficients of these regressors are interpreted as semi-elasticities as is defined in Equation (5).

Regarding the weekly regressors, the entire period from January 2008 to March 2009 is examined simultaneously because the normalization and scaling methods are the same for the search volume variables. For the rank variables, the weekly data are the average of a seven-day week, but it should be noted that the rank variables are based upon search volume variables that have already been normalized and scaled. In regards to the Poisson regressions, both daily and weekly data are used. The number of observations for each Poisson regression is listed in Table 1A. There are some missing observations due to the temporary lack of connection to the Internet or the blockage of the custom-built program by the Google Server since the program was the suspect of spamming.

3.2 Interpretation of the Univariate and Bivariate Poisson QMLE Results

The benefit of using daily data organized quarter-by-quarter is that one can observe the peak quarter of website traffic to the CACVB website with respect to keyword queries. Regarding the daily data, the general relationship between the estimated coefficients of the log of the keyword search volumes and the regressand is a positive one as is shown in Tables 2A and 2B. For instance, in the general 'all' search category of keyword searches for *charleston hotels*, the quarter that produces the largest estimated elasticity is the third quarter for all five regressands which involves all website traffic to the CACVB website and the four subcategories of website traffic to the CACVB website. This is also true for the estimated coefficients involving the log of the keyword search volumes for *charleston restaurants*, *charleston tourism*, and *charleston travel* with the peak quarter being the fourth quarter. The only keyword search query whose peak quarter shows some variation in the five Poisson regressions is that of *charleston sc*. For the regressand of nonlocal visits, the peak quarter is in the first quarter and for the remaining four regressands, the peak quarter is in the second quarter.

As for the keyword searches for *charleston hotels*, in the 'travel' category, the third quarter produces the largest estimated elasticity for all five regressands as it does in the general overall category of keyword searches. Analogous to the general overall category of keyword searches, the keyword searches for *charleston sc*, in the 'travel' categories also produces different peak quarters with respect to the regressands. Having access to which keyword searches are going to have the biggest impact at a given time could help tourist boards and businesses to maximizing their advertising expense, which is just one benefit of using Google data.

It should be noted that a few of the coefficients are negative such as in the coefficient for the log of the keyword search volumes involving *charleston restaurants* in the second quarter, but these coefficients prove to be statistically insignificant. The coefficient for the log of the keyword search volumes involving *charleston tourism* are also negative and statistically insignificant, which could possibly be due to a missing data problem since the dataset is reduced to 58 observations for the first and fifth quarter and to 83 observations for the second quarter and to 86 observations for the third and fourth quarters.

A surprise finding of this paper is the impact of the regressors of rankings on website traffic. One might assume that the ranking of a website might help draw the attention of the search engine user to a higher ranked website, but this paper finds that most of the estimated coefficients of the rank regressors are statistically insignificant or very small as is shown in Tables 2A and 2B.

For the statistically significant estimate coefficients, there generally is a negative relationship between the regressor and the regressands. For instance, looking at the fifth quarter for the rank variable of *charleston travel* with respect to the regressand of *all visits*, Table 2 shows that the estimated coefficient is -0.197, which means that the log of website traffic decreases by 19.7% if the estimated coefficient is interpreted as a semi-elasticity. Alternatively, for the fifth quarter of the rank variable of *charleston hotel* with respect to the regressand of *all visits*, Table 2 shows that the estimated coefficient is 0.020, which means that the log of website traffic increases by 2.0%. It could be said that for certain variables, where a search engine user has something specific in mind such as a preferred hotel, rank does not help entice website traffic to a given website.

Working with rank variables is problematic, due to the lack of variability, which produced singular matrices in the Poisson QMLE results and is not reported in Tables 2A, 2B, 3, and 4 in order to conserve space. Furthermore, there is also a missing data problem, which reduced the dataset for the regressions for four quarters of daily data as well as the weekly data which involve the rank variables of *charleston sc* and *charleston hotel*. The estimated coefficients from the regressions with missing observations turned out to be statistically insignificant.

As is shown in Table 3, combining the log of the keyword search volumes with their respective ranking in the same regression did not greatly alter the empirical findings of the individual univariate regression results especially with respect to statistical significance.

3.3 Effects of Normalization, Scaling, and Aggregation

The most dramatic finding of this paper is with respect to the comparison of the estimated Poisson QMLE results using daily and weekly data, which has important implications in terms of policy implementation. Generally, one would expect that if the average at a lower frequency is rather consistent, then this ought to be reflected at the higher frequency. The same should hold for conditional means as well, but this is not the case when comparing the Poisson QMLEs when using daily and weekly data. For this comparison, the average of the regression estimates for the five quarters involving daily data is compared to the regression estimates of the weekly data, which is shown in Table 4.[6]

In some instances, depending on the frequency of data, there is a reversal of the sign of an estimated coefficient and even statistical significance, which is what occurs for the Poisson QMLE results that involve the regressand, *local visits*. When daily data is used, the sign for all seven regressors involving the log of keyword search volumes is positive with four out of the seven Poisson QMLE regressions producing statistically significant coefficients. Alternatively, when weekly data is used all seven estimated coefficients are statistically insignificant, and five out of the seven estimated coefficients become negative.

When examining the goodness of fit of a model, the higher the R-squared term and the lower the Akaike Information Criteria (AIC) the better. For the seven Poisson QMLE regressions that involve *local visits*, the regressions that use weekly data produce a slightly lower AIC, but the R-squared terms are much higher when daily data is used.[7] This is not the case for the remaining four regressands. For the other four remaining regressands, the R-squared terms are generally higher on average and the AICs are much higher when weekly data is as opposed to daily data as is shown in Table 4. Normally, one would expect a lower AIC with a higher R-squared term, but this anomaly is discussed in more detail a little further on in this subsection.

[6] It should be noted that the daily data is scaled differently for each quarter, which could have an impact on the comparison, but the general overall idea should hold.

[7] For evaluating the goodness of fit of a Poisson QMLE model, Wooldridge (2002) uses the R-squared terms, while Michener and Tighe (1992) use the AIC.

In regards to the interpretation of the estimated coefficients, a different story emerges depending on the frequency used in the regression. For instance, examining the regression involving the regressand, *new visits*, the Poisson QMLE regression using daily data produces a statistically significant estimated coefficient of 0.749 on average for the regressor concerning *charleston sc* in the travel category, which states that the conditional mean changes proportionally by 0.749. When weekly data is used, the conditional mean, for the same regressor, changes proportionally by 1.106, which is also statistically significant. The R-squared term is approximately 50% higher and the AIC is more than three times higher in the regression involving weekly data, which is listed in Table 4.

Another such example involves the regressand, *search traffic*, and the regressor *charleston sc* in the general overall search category as is also shown in Table 4. When daily data is used, the statistically significant estimated regression coefficient is 0.984 with an R-squared term of 0.34 and the statistically significant estimated regression coefficient is 1.450 with an R-squared term of 0.39 when weekly data is used. So, it appears that in terms of explaining the variability of the regressand, *search traffic*, both daily and weekly data capture approximately the same level of variability, but the AIC terms are vastly different. When daily data is used, the AIC is 61.85 and when weekly data is used, the AIC is 418.36. So, it would appear that based on the AIC, the Poisson QMLE regression model using daily data is a better fit.

When using weekly data, the regressions produce estimated coefficients that are generally larger on average, which naturally produce larger R-squared terms. So, in terms of goodness of fit, the R-squared terms might not be necessarily the best measure, which leaves the AIC as a measure of goodness of fit of the regression models for consideration. The AIC is generally lower when daily data is used, which indicates that the daily data might give a better picture of explaining the five various forms of website traffic discussed in this paper except in the case of *local visits*. This is not a prudent conclusion to draw due to the lack of transparency of the normalized and scaled regressors especially in regards to the keyword search volumes.

Hence, in the examination of the relationship between various types of website traffic and the regressor of the log of keyword search volumes, a different picture unfolds depending on the frequency of data used.

4. Conclusion

The benefit of using data from Google Insights and Google Trends for businesses is that it can better help businesses to target certain markets by providing instantaneous access to the most current data available at any given time. Using data from Google Insights and Google Trends is also helpful to researchers by providing a consumer-driven data source, i.e. internet activity, to help explain or predict economic activity without the time lag required for economic time series such as unemployment statistics.

Google is very transparent in the fact that they *only* provide data that already has been transformed through normalization and scaling procedures, but it is not transparent in regards to the raw data itself due to privacy issues, which is problematic to the researcher wishing to deal with the original level data. For the researcher as well as for policy implementation, it is difficult to see the direct effect of aggregation when normalized and scaled data is used because the normalization and scaling are different for each level of aggregation. As is demonstrated in this paper, the empirical Poisson QMLE results can be drastically different when using daily and weekly data. These differences could be due to the effects of normalization, scaling, and/or aggregation, but the lack of transparency in regards to the raw data itself makes it difficult to identify.

Furthermore, the use of data from Google Insights and Google Trends automatically means that the size of the dataset is going to be limited since the normalization and scaling is not uniform across time periods in regards to a given data frequency. For instance in this paper, the regressions involving daily data are limited to only one quarter at a time because of the normalization and scaling procedures used by Google, which hinders the examination of long-run trends.

Another potential problem of using Google data involves the interpreting of regression coefficients since normalized and scaled variables are used. In this paper, the problem of interpreting the regression coefficients is averted through the nature of the Poisson regression, which permits the interpretation of regression coefficients as elasticities or semi-elasticities.

A suggestion for future research involving Google data would be to work with the raw data, while maintaining the privacy of the user of a given search engine, which suggests a change in the data collection methods at Google Insights and Google Trends. In doing so, a researcher could still have the problem of aggregation in regards to the empirical results, but the potential doubts as it pertains to the empirical findings caused by normalizing and scaling would at least be removed. Balancing the great potential for research, which could benefit the consumers versus the protection of user privacy, will be the key to the future development in this research area.

References

Askitas, N. and Zimmerman, K.F. (2009), "Google Econometrics and Unemployment Forecasting," *Applied Economics Quarterly*, 55:2, 107-120.

Azar, J. (2009), "Oil Prices and Electric Cars", *Princeton University Working Paper*.

Barbaro, M. and Zeller, T. (2006) "A Face Is Exposed for AOL Searcher No. 4417749", New York Times, August 9, accessed online at http://www.nytimes.com.

Bian, L. (1997), "Multiscale Nature of Spatial Data in Scaling Up Environmental Models," *Scale in Remote Sensing and GIS*, D.A. Quattrochi and M.F. Goodchild, eds., London: CRC Press.

Cameron, A.C. and Trivedi, P.K (1986), "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests," *Journal of Applied Econometrics*, 1:1, 29-53.

Cameron, A.C. and Trivedi, P.K. (1998), *Regression Analysis of Count Data*, Cambridge: Cambridge University Press.

Choi, H. and Varian, H. (2009a), "Predicting the Present with Google Trends," *Google Technical Report*.

Choi, H. and Varian, H. (2009b), "Predicting Initial Claims for Unemployment Benefits," *Google Technical Report*.

Engle, R. F. and Granger, C. W. J. (1987), "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55:2, 251-276.

Gagnon, J.E. (2008), "Inflation Regimes and Inflation Expectations," *Federal Reserve Bank of St. Louis Review,* 90:3-Part 2.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009), "Detecting Influenza Epidemics using Search Engine Query Data," *Nature*, 457, 1012 –1014.

-----Google (2009), "About Google Trends,"
http://www.google.com/intl/en/trends/about.html (accessed August 30, 2009).

-----Google Insights (2009a), "How is the data normalized?"
http://www.google.com/support/insights/bin/bin/bin/answer.py?answer=87284&topic=13975 (accessed August 30, 2009).

-----Google Insights (2009b), "How is the data scaled?"

http://www.google.com/support/insights/bin/bin/answer.py?answer=87282 (accessed August 30, 2009).

Pan, B., Litvin, S.W. and O'Donnell, T. (2007), "Accommodation Search Query Formulation: Implications for Search Engine Marketing," *Journal of Vacation Marketing*, 13:4, 371-381.

Pyle, D. (1999), "Data Preparation for Data Mining (The Morgan Kaufmann Series in Data Management Systems)," San Francisco: Morgan Kaufmann.

Rapach, D. (2003), "International Evidence on the Long-Run Impact of Inflation," *Journal of Money Credit and Banking*, 35:1, 23-45.

Rossana, R.J. and Seater, J.J. (1995) "Temporal Aggregation and Economic Time Series," American Statistical Association, 13:4, 441-451.

Michener, R. and Tighe, C. (1992), "A Poisson Regression Model of Highway Fatalities," *The American Economic Review*, 82:2, 452-456.

Wooldridge, J.M. (1999), "Quasi-Likelihood Methods for Count Data," *Handbook of Applied Econometrics, Volume II: Microeconomics*, M.H. Pesaran and P. Schmidt, eds., Chichester: Blackwell.

Wooldridge, J.M. (2002), "*Econometric Analysis of Cross Section and Panel Data*," Cambridge: MIT Press.

Table 1A-Number of Observations							
Regressands	**Regressors**	**Daily-Q1**	**Daily-Q2**	**Daily-Q3**	**Daily-Q4**	**Daily-Q5**	**Weekly**
all visits	"charleston hotels" (all)	83	91	92	92	75	60
new visits	"charleston hotels" (travel)	83	91	92	92	75	60
search traffic	"charleston restaurants" (all)	83	91	92	92	75	60
nonlocal visits	"charleston sc" (all)	83	91	92	92	75	60
local visits	"charleston sc" (travel)	83	91	92	92	75	60
	"charleston tourism" (all)	58	83	86	86	56	60
	"charleston travel" (all)	83	91	92	92	75	60
	"charleston" (rank)	83	83	N/A	N/A	N/A	38
	"charleston hotel" (rank)	42	17	47	68	75	44
	"charleston restaurants" (rank)	83	91	92	92	75	60
	"charleston tourism" (rank)	83	91	92	92	75	60
	"charleston travel" (rank)	83	91	92	92	75	60

Table 1B-Legend				
Regressands	**Regressors**	**Abbreviations of Regressors**	**Type of Variable**	**Type of Searches**
all visits	"charleston hotels" (all)	ch hotels (all)	Log of search volume	All Category
new visits	"charleston hotels" (travel)	ch hotels (travel)	Log of search volume	Travel Category
search traffic	"charleston restaurants" (all)	ch restaurants (all)	Log of search volume	All Category
nonlocal visits	"charleston sc" (all)	ch sc (all)	Log of search volume	All Category
local visits	"charleston sc" (travel)	ch sc (travel)	Log of search volume	Travel Category
	"charleston tourism" (all)	ch tourism (all)	Log of search volume	All Category
	"charleston travel" (all)	ch travel (all)	Log of search volume	All Category
	"charleston" (rank)	ch (rank)	Rank Variable	All Category
	"charleston hotel" (rank)	ch hotel (rank)	Rank Variable	All Category
	"charleston restaurants" (rank)	ch restaurants (rank)	Rank Variable	All Category
	"charleston tourism" (rank)	ch tourism (rank)	Rank Variable	All Category
	"charleston travel" (rank)	ch travel (rank)	Rank Variable	All Category

Table 2A--Univariate Poisson Regressions for Regressands: All Visits, New Visits, and Search Traffic																			
		Regressand--All Visits						Regressand--New Visits						Regressand--Search Traffic					
Time	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC
Q1	ch hotels	0.573	0.094	6.126	**0.00**	0.32	135.56	0.528	0.096	5.530	**0.00**	0.28	90.79	0.487	0.096	5.068	**0.00**	0.24	67.22
Q2	(all)	0.604	0.090	6.705	**0.00**	0.33	149.65	0.586	0.088	6.648	**0.00**	0.33	91.74	0.536	0.099	5.400	**0.00**	0.24	79.71
Q3		0.828	0.095	8.748	**0.00**	0.46	139.04	0.729	0.098	7.426	**0.00**	0.38	97.67	0.735	0.094	7.805	**0.00**	0.41	61.30
Q4		0.543	0.059	9.172	**0.00**	0.48	85.16	0.523	0.058	9.069	**0.00**	0.47	55.69	0.512	0.056	9.147	**0.00**	0.47	38.27
Q5		0.727	0.103	7.051	**0.00**	0.40	123.34	0.654	0.097	6.711	**0.00**	0.38	70.61	0.692	0.100	6.950	**0.00**	0.39	54.12
W		0.746	0.053	14.183	**0.00**	0.78	196.75	0.825	0.069	12.007	**0.00**	0.71	222.34	0.714	0.061	11.781	**0.00**	0.70	201.83
Q1	ch restaurants	0.107	0.082	1.317	0.19	0.02	190.61	0.065	0.081	0.810	0.42	0.01	121.58	0.037	0.080	0.467	0.64	0.00	86.35
Q2	(all)	-0.114	0.085	-1.339	0.18	0.02	215.70	-0.103	0.084	-1.234	0.22	0.02	129.99	-0.054	0.090	-0.599	0.55	0.00	102.21
Q3		0.089	0.093	0.953	0.34	0.01	245.54	0.075	0.090	0.831	0.41	0.01	149.95	0.125	0.088	1.426	0.15	0.02	94.56
Q4		0.192	0.079	2.432	**0.02**	0.06	146.78	0.165	0.077	2.145	**0.03**	0.05	93.67	0.208	0.075	2.788	**0.01**	0.08	60.71
Q5		0.012	0.080	0.151	0.88	0.00	200.22	0.000	0.074	0.001	1.00	0.00	108.29	0.016	0.077	0.204	0.84	0.00	83.58
W		0.806	0.138	5.820	**0.00**	0.35	573.72	0.921	0.159	5.799	**0.00**	0.35	509.43	0.782	0.143	5.477	**0.00**	0.33	462.24
Q1	ch sc	1.290	0.171	7.544	**0.00**	0.43	118.03	1.239	0.173	7.179	**0.00**	0.40	77.89	1.086	0.180	6.033	**0.00**	0.32	61.94
Q2	(all)	1.473	0.139	10.570	**0.00**	0.59	100.27	1.462	0.133	10.975	**0.00**	0.61	60.11	1.264	0.169	7.496	**0.00**	0.41	65.71
Q3		0.870	0.185	4.700	**0.00**	0.17	201.85	0.932	0.173	5.390	**0.00**	0.20	117.64	0.695	0.182	3.818	**0.00**	0.12	84.47
Q4		1.018	0.084	12.175	**0.00**	0.62	64.09	1.053	0.071	14.754	**0.00**	0.70	34.96	0.900	0.086	10.424	**0.00**	0.54	34.16
Q5		1.214	0.178	6.805	**0.00**	0.40	125.67	1.165	0.161	7.219	**0.00**	0.43	66.78	0.977	0.186	5.248	**0.00**	0.28	62.97
W		1.353	0.239	5.665	**0.00**	0.34	581.02	1.420	0.285	4.976	**0.00**	0.29	559.89	1.450	0.230	6.297	**0.00**	0.39	418.36
Q1	ch tourism	-0.215	0.097	-2.231	**0.03**	0.08	158.83	-0.198	0.092	-2.159	**0.03**	0.07	94.41	-0.220	0.088	-2.510	**0.01**	0.10	63.80
Q2	(all)	-0.139	0.075	-1.851	0.06	0.04	217.25	-0.140	0.073	-1.904	0.06	0.05	129.85	-0.179	0.079	-2.277	**0.02**	0.06	100.50
Q3		-0.187	0.145	-1.283	0.20	0.02	244.78	-0.175	0.140	-1.245	0.21	0.02	149.83	-0.222	0.136	-1.630	0.10	0.03	93.21
Q4		0.550	0.106	5.210	**0.00**	0.26	124.40	0.522	0.102	5.102	**0.00**	0.25	79.11	0.525	0.100	5.273	**0.00**	0.26	52.54
Q5		0.044	0.124	0.355	0.72	0.00	160.94	0.018	0.115	0.158	0.87	0.00	88.20	0.106	0.121	0.877	0.38	0.01	70.04
W		0.457	0.042	10.777	**0.00**	0.65	303.76	0.540	0.045	11.983	**0.00**	0.69	234.03	0.417	0.049	8.448	**0.00**	0.52	315.62
Q1	ch travel	-0.126	0.096	-1.315	0.19	0.02	190.62	-0.128	0.094	-1.356	0.18	0.02	119.99	-0.095	0.093	-1.024	0.31	0.01	85.56
Q2	(all)	0.064	0.083	0.773	0.44	0.01	218.32	0.046	0.081	0.561	0.57	0.00	131.55	0.083	0.087	0.955	0.34	0.01	101.66
Q3		0.112	0.090	1.248	0.21	0.02	246.38	0.059	0.087	0.673	0.50	0.01	151.82	0.142	0.084	1.686	0.09	0.03	94.61
Q4		0.239	0.096	2.477	**0.01**	0.07	124.11	0.216	0.093	2.319	**0.02**	0.06	78.16	0.237	0.094	2.523	**0.01**	0.08	54.78
Q5		0.151	0.084	1.797	0.07	0.05	183.28	0.148	0.078	1.911	0.06	0.05	98.98	0.129	0.082	1.573	0.12	0.03	79.15
W		0.750	0.075	9.986	**0.00**	0.62	328.19	0.843	0.089	9.419	**0.00**	0.59	311.65	0.693	0.084	8.203	**0.00**	0.52	321.40
Q1	ch hotels	0.485	0.089	5.439	**0.00**	0.27	144.89	0.458	0.090	5.117	**0.00**	0.25	94.51	0.427	0.089	4.773	**0.00**	0.22	69.12
Q2	(travel)	0.526	0.077	6.867	**0.00**	0.34	147.41	0.516	0.074	6.940	**0.00**	0.34	89.43	0.516	0.082	6.307	**0.00**	0.30	73.84
Q3		0.677	0.106	6.379	**0.00**	0.31	174.31	0.628	0.104	6.020	**0.00**	0.29	110.95	0.594	0.103	5.784	**0.00**	0.27	73.12
Q4		0.474	0.059	7.973	**0.00**	0.41	95.21	0.456	0.058	7.889	**0.00**	0.41	61.77	0.440	0.057	7.739	**0.00**	0.40	42.76

Table 2A (Continued)--Univariate Poisson Regressions for Regressands: All Visits, New Visits, and Search Traffic

		Regressand--All Visits						Regressand--New Visits						Regressand--Search Traffic					
Time	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC
Q5	ch hotels	0.546	0.094	5.789	**0.00**	0.31	140.68	0.478	0.090	5.336	**0.00**	0.28	80.72	0.466	0.095	4.910	**0.00**	0.25	65.23
W	(travel)	0.765	0.046	16.692	**0.00**	0.83	156.10	0.875	0.055	15.897	**0.00**	0.81	149.09	0.729	0.056	13.079	**0.00**	0.74	176.89
Q1	ch sc	0.640	0.145	4.431	**0.00**	0.20	158.59	0.600	0.145	4.145	**0.00**	0.18	102.57	0.484	0.147	3.289	**0.00**	0.12	77.24
Q2	(travel)	0.872	0.105	8.273	**0.00**	0.44	128.75	0.859	0.102	8.399	**0.00**	0.45	78.04	0.896	0.111	8.084	**0.00**	0.43	62.95
Q3		0.745	0.149	5.006	**0.00**	0.22	195.85	0.760	0.141	5.382	**0.00**	0.24	116.81	0.655	0.143	4.582	**0.00**	0.19	79.96
Q4		0.707	0.062	11.419	**0.00**	0.60	69.48	0.674	0.062	10.958	**0.00**	0.57	47.45	0.667	0.058	11.431	**0.00**	0.59	32.20
Q5		0.910	0.129	7.076	**0.00**	0.41	122.97	0.850	0.118	7.176	**0.00**	0.41	67.58	0.837	0.127	6.600	**0.00**	0.37	55.92
W		0.969	0.095	10.151	**0.00**	0.62	328.65	1.106	0.110	10.028	**0.00**	0.61	294.88	0.976	0.096	10.156	**0.00**	0.63	254.22
Q1	ch sc	0.010	0.005	1.947	0.05	0.04	186.58	0.008	0.005	1.552	0.12	0.03	119.35	0.006	0.005	1.166	0.24	0.02	85.30
Q2	(rank)	-0.193	0.141	-1.374	0.17	0.02	215.71	-0.179	0.136	-1.322	0.19	0.02	127.85	-0.319	0.155	-2.055	**0.04**	0.05	99.51
W		0.031	0.018	1.684	0.09	0.07	302.32	0.042	0.023	1.813	0.07	0.08	352.34	0.020	0.021	0.970	0.33	0.03	284.30
Q1	ch hotel	0.000	0.001	-0.327	0.74	0.00	177.52	-0.001	0.001	-0.382	0.70	0.00	110.33	-0.001	0.001	-0.547	0.58	0.01	72.92
Q2	(rank)	0.002	0.006	0.352	0.72	0.01	125.30	0.002	0.005	0.392	0.70	0.01	70.22	0.001	0.006	0.228	0.82	0.00	64.27
Q3		0.002	0.001	1.681	0.09	0.06	214.90	0.002	0.001	1.655	0.10	0.06	127.31	0.001	0.001	1.226	0.22	0.03	85.90
Q4		0.000	0.001	0.166	0.87	0.00	134.60	0.000	0.001	0.397	0.69	0.00	82.51	0.000	0.001	0.043	0.97	0.00	51.32
Q5		0.020	0.006	3.216	**0.00**	0.13	177.32	0.018	0.006	3.117	**0.00**	0.12	97.01	0.021	0.006	3.551	**0.00**	0.15	72.93
W		-0.001	0.001	-0.914	0.36	0.02	986.01	-0.002	0.001	-1.230	0.22	0.04	830.89	-0.001	0.001	-0.956	0.34	0.02	756.68
Q1	ch restaurants	0.004	0.002	1.964	**0.05**	0.04	186.19	0.003	0.002	1.562	0.12	0.03	119.22	0.004	0.002	1.926	0.05	0.04	83.16
Q2	(rank)	-0.006	0.002	-2.702	**0.01**	0.08	204.72	-0.006	0.002	-2.681	**0.01**	0.08	123.35	-0.009	0.002	-3.839	**0.00**	0.15	89.92
Q3		0.002	0.003	0.784	0.43	0.01	246.29	0.003	0.003	0.887	0.37	0.01	149.79	0.001	0.003	0.278	0.78	0.00	96.41
Q4		0.006	0.004	1.545	0.12	0.03	151.96	0.007	0.004	1.744	0.08	0.03	95.03	0.004	0.004	0.934	0.35	0.01	64.67
Q5		0.003	0.004	0.727	0.47	0.01	198.96	0.002	0.004	0.634	0.53	0.01	107.77	0.003	0.004	0.855	0.39	0.01	82.91
W		0.004	0.003	1.428	0.15	0.03	861.78	0.002	0.003	0.691	0.49	0.01	786.48	0.003	0.003	1.001	0.32	0.02	678.93
Q1	ch tourism	-0.031	0.008	-3.730	**0.00**	0.15	166.53	-0.027	0.008	-3.240	**0.00**	0.12	109.10	-0.026	0.008	-3.166	**0.00**	0.11	77.65
Q3	(rank)	-0.203	0.045	-4.505	**0.00**	0.18	203.78	-0.171	0.044	-3.834	**0.00**	0.14	131.14	-0.174	0.043	-4.011	**0.00**	0.15	83.08
Q4		-0.092	0.040	-2.316	**0.02**	0.06	147.50	-0.101	0.038	-2.627	**0.01**	0.07	91.64	-0.086	0.038	-2.258	**0.02**	0.05	62.14
W		-0.015	0.027	-0.537	0.59	0.01	886.63	0.001	0.031	0.020	0.98	0.00	792.70	-0.033	0.028	-1.155	0.25	0.02	674.78
Q1	ch travel	-0.026	0.009	-2.763	**0.01**	0.09	178.14	-0.022	0.009	-2.410	**0.02**	0.07	114.75	-0.023	0.009	-2.495	**0.01**	0.07	80.83
Q2	(rank)	-0.580	0.233	-2.490	**0.01**	0.07	203.90	-0.547	0.225	-2.435	**0.01**	0.06	123.21	-0.647	0.251	-2.577	**0.01**	0.07	94.79
Q3		0.193	0.231	0.834	0.40	0.01	245.98	0.240	0.227	1.053	0.29	0.01	149.18	0.127	0.212	0.597	0.55	0.00	96.12
Q4		0.246	0.153	1.610	0.11	0.03	151.33	0.217	0.146	1.484	0.14	0.02	95.67	0.246	0.146	1.681	0.09	0.03	63.34
Q5		-0.197	0.031	-6.404	**0.00**	0.37	132.78	-0.173	0.029	-5.866	**0.00**	0.33	77.04	-0.184	0.030	-6.109	**0.00**	0.34	58.72
W		-0.041	0.034	-1.206	0.23	0.03	869.06	-0.041	0.039	-1.029	0.30	0.02	778.29	-0.054	0.035	-1.531	0.13	0.05	662.78

Table 2B--Univariate Poisson Regressions for Regressands: Nonlocal Visits and Local Visits													
		Regressand--Nonlocal Visits						Regressand--Local Visits					
Time	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-stat	PV	R-Sq	AIC
Q1	ch hotels	0.615	0.093	6.623	**0.00**	0.35	65.98	0.538	0.101	5.318	**0.00**	0.26	88.04
Q2	(all)	0.554	0.088	6.308	**0.00**	0.31	73.13	0.650	0.096	6.800	**0.00**	0.34	91.29
Q3		0.806	0.091	8.807	**0.00**	0.46	67.51	0.849	0.104	8.176	**0.00**	0.43	89.71
Q4		0.544	0.056	9.647	**0.00**	0.51	43.83	0.542	0.066	8.168	**0.00**	0.42	56.29
Q5		0.773	0.099	7.823	**0.00**	0.45	58.59	0.686	0.113	6.083	**0.00**	0.33	81.13
W		0.918	0.067	13.724	**0.00**	0.77	249.15	-0.047	0.065	-0.727	0.47	0.01	65.42
Q1	ch restaurants	0.149	0.083	1.801	0.07	0.04	93.43	0.073	0.085	0.860	0.39	0.01	115.43
Q2	(all)	-0.127	0.082	-1.559	0.12	0.03	99.02	-0.103	0.091	-1.126	0.26	0.01	131.84
Q3		0.111	0.090	1.227	0.22	0.02	115.26	0.068	0.099	0.691	0.49	0.01	148.39
Q4		0.194	0.077	2.527	**0.01**	0.07	74.84	0.190	0.084	2.247	**0.02**	0.05	86.89
Q5		0.020	0.080	0.246	0.81	0.00	99.60	0.005	0.083	0.062	0.95	0.00	117.24
W		0.950	0.171	5.557	**0.00**	0.33	734.14	0.039	0.112	0.352	0.72	0.00	65.80
Q1	ch sc	1.423	0.164	8.664	**0.00**	0.50	54.64	1.179	0.190	6.200	**0.00**	0.33	81.44
Q2	(all)	1.414	0.132	10.693	**0.00**	0.60	48.57	1.527	0.154	9.923	**0.00**	0.56	66.74
Q3		0.738	0.185	3.985	**0.00**	0.13	101.12	0.991	0.192	5.158	**0.00**	0.19	117.99
Q4		1.030	0.077	13.439	**0.00**	0.67	32.29	1.005	0.098	10.216	**0.00**	0.54	46.74
Q5		1.260	0.174	7.228	**0.00**	0.43	61.70	1.172	0.192	6.100	**0.00**	0.35	80.52
W		1.598	0.295	5.417	**0.00**	0.32	742.64	0.073	0.188	0.386	0.70	0.00	65.77
Q1	ch tourism	-0.216	0.103	-2.086	**0.04**	0.07	86.94	-0.215	0.095	-2.255	**0.02**	0.08	88.53
Q2	(all)	-0.100	0.072	-1.386	0.17	0.02	101.09	-0.175	0.080	-2.194	**0.03**	0.06	130.65
Q3		-0.198	0.141	-1.400	0.16	0.02	115.50	-0.176	0.154	-1.142	0.25	0.01	147.50
Q4		0.536	0.104	5.157	**0.00**	0.26	64.95	0.564	0.112	5.027	**0.00**	0.24	73.70
Q5		0.066	0.127	0.521	0.60	0.00	84.46	0.024	0.126	0.191	0.85	0.00	92.66
W		0.557	0.052	10.785	**0.00**	0.64	372.56	-0.066	0.045	-1.470	0.14	0.03	63.93
Q1	ch travel	-0.116	0.099	-1.176	0.24	0.02	95.32	-0.135	0.099	-1.362	0.17	0.02	114.01
Q2	(all)	0.072	0.079	0.913	0.36	0.01	100.54	0.056	0.088	0.639	0.52	0.00	132.97
Q3		0.096	0.088	1.093	0.27	0.01	116.61	0.128	0.095	1.341	0.18	0.02	147.89
Q4		0.203	0.095	2.132	**0.03**	0.06	65.52	0.275	0.102	2.683	**0.01**	0.09	73.59
Q5		0.150	0.084	1.779	0.08	0.04	92.21	0.152	0.087	1.742	0.08	0.04	107.80
W		0.909	0.094	9.642	**0.00**	0.60	419.17	-0.042	0.083	-0.511	0.61	0.01	65.61
Q1	ch hotels	0.509	0.090	5.648	**0.00**	0.29	71.98	0.465	0.095	4.904	**0.00**	0.23	91.53
Q2	(travel)	0.493	0.074	6.670	**0.00**	0.33	70.89	0.556	0.082	6.776	**0.00**	0.33	91.47
Q3		0.657	0.103	6.368	**0.00**	0.31	83.68	0.695	0.114	6.089	**0.00**	0.29	108.81
Q4		0.475	0.057	8.324	**0.00**	0.43	48.90	0.473	0.066	7.192	**0.00**	0.36	61.26

Table 2B(Continued)--Univariate Poisson Regressions for Regressands: Nonlocal Visits and Local Visits													
		Regressand--Nonlocal Visits						Regressand--Local Visits					
Time	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-stat	PV	R-Sq	AIC
Q5	ch hotels	0.589	0.091	6.485	**0.00**	0.37	66.85	0.507	0.102	4.973	**0.00**	0.25	90.14
W	(travel)	0.946	0.057	16.526	**0.00**	0.82	191.06	-0.066	0.064	-1.044	0.30	0.02	64.90
Q1	ch sc	0.702	0.145	4.848	**0.00**	0.23	77.28	0.590	0.154	3.834	**0.00**	0.16	99.77
Q2	(travel)	0.814	0.103	7.936	**0.00**	0.42	63.30	0.924	0.113	8.198	**0.00**	0.44	80.31
Q3		0.704	0.146	4.832	**0.00**	0.21	94.68	0.783	0.158	4.943	**0.00**	0.21	119.26
Q4		0.702	0.059	11.902	**0.00**	0.61	36.70	0.712	0.071	10.089	**0.00**	0.54	47.73
Q5		0.917	0.128	7.157	**0.00**	0.41	62.46	0.903	0.137	6.603	**0.00**	0.37	77.15
W		1.162	0.120	9.704	**0.00**	0.60	429.02	-0.026	0.099	-0.259	0.80	0.00	65.85
Q1	ch sc	0.011	0.005	2.077	**0.04**	0.05	92.56	0.009	0.005	1.735	0.08	0.04	112.70
Q2	(rank)	-0.161	0.133	-1.213	0.23	0.02	99.46	-0.224	0.014	-16.087	**0.00**	0.03	131.73
W		0.037	0.023	1.602	0.11	0.07	418.72	0.037	0.023	1.602	0.11	0.07	418.72
Q1	ch hotel	0.000	0.001	0.007	0.99	0.00	89.70	-0.001	0.001	-0.592	0.55	0.01	104.77
Q2	(rank)	0.000	0.006	0.083	0.93	0.00	71.06	0.003	0.006	0.608	0.54	0.02	67.59
Q3		0.002	0.001	1.485	0.14	0.05	104.60	0.002	0.001	1.807	0.07	0.07	125.90
Q4		0.000	0.001	0.282	0.78	0.00	68.27	0.000	0.001	0.097	0.92	0.00	81.48
Q5		0.020	0.006	3.264	**0.00**	0.13	88.47	0.019	0.006	3.044	**0.00**	0.11	105.49
W		-0.002	0.002	-1.000	0.32	0.02	1227.01	-0.002	0.002	-1.000	0.32	0.02	1227.01
Q1	ch restaurants	0.004	0.002	2.014	**0.04**	0.04	92.67	0.004	0.002	1.817	0.07	0.04	112.23
Q2	(rank)	-0.005	0.002	-2.284	**0.02**	0.06	96.55	-0.007	0.002	-2.984	**0.00**	0.09	122.95
Q3		0.002	0.003	0.698	0.49	0.01	116.45	0.003	0.003	0.831	0.41	0.01	148.05
Q4		0.006	0.004	1.453	0.15	0.02	77.85	0.007	0.004	1.572	0.12	0.03	89.04
Q5		0.003	0.004	0.622	0.53	0.01	99.22	0.003	0.004	0.792	0.43	0.01	116.37
W		0.004	0.003	1.232	0.22	0.03	1078.12	0.004	0.003	1.232	0.22	0.03	1078.12
Q1	ch tourism	-0.030	0.009	-3.539	**0.00**	0.14	84.68	-0.032	0.009	-3.663	**0.00**	0.15	100.57
Q3	(rank)	-0.211	0.043	-4.899	**0.00**	0.21	94.14	-0.195	0.049	-4.013	**0.00**	0.15	127.82
Q4		-0.095	0.039	-2.443	**0.01**	0.06	75.08	-0.090	0.043	-2.109	**0.04**	0.05	87.36
W		-0.015	0.033	-0.444	0.66	0.00	1101.60	-0.015	0.033	-0.444	0.66	0.00	1101.60
Q1	ch travel	-0.024	0.010	-2.440	**0.01**	0.07	90.64	-0.028	0.010	-2.878	**0.00**	0.09	106.13
Q2	(rank)	-0.516	0.217	-2.380	**0.02**	0.06	95.13	-0.641	0.255	-2.510	**0.01**	0.07	123.82
Q3		0.274	0.233	1.177	0.24	0.02	115.25	0.123	0.237	0.517	0.60	0.00	148.69
Q4		0.165	0.145	1.137	0.26	0.01	78.42	0.336	0.169	1.995	**0.05**	0.04	87.37
Q5		-0.210	0.030	-7.073	**0.00**	0.41	63.54	-0.186	0.034	-5.553	**0.00**	0.30	85.64
W		-0.052	0.043	-1.232	0.22	0.03	1076.55	-0.052	0.043	-1.232	0.22	0.03	1076.55

Table 3--Bivariate Poisson Regressions for Regressand--All Visits												
Time	Regressors	Coef.	S.E.	Z-Stat	PV	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC
Q1	ch sc (all)	1.265	0.170	7.439	**0.00**	ch (rank)	0.007	0.004	1.790	0.07	0.45	114.02
Q2		1.478	0.146	10.107	**0.00**		-0.199	0.095	-2.100	**0.04**	0.61	97.23
W		0.787	0.239	3.289	**0.00**		0.036	0.016	2.219	**0.03**	0.30	235.67
Q1	ch hotels (all)	0.607	0.142	4.279	**0.00**	ch hotel (rank)	0.000	0.001	0.416	0.68	0.32	123.84
Q2		0.303	0.281	1.079	0.28		0.004	0.006	0.615	0.54	0.08	117.00
Q3		0.779	0.140	5.571	**0.00**		0.001	0.001	1.243	0.21	0.44	130.98
Q4		0.522	0.077	6.758	**0.00**		-0.001	0.001	-0.955	0.34	0.40	82.78
Q5		0.693	0.118	5.887	**0.00**		0.004	0.006	0.613	0.54	0.40	122.79
W		0.745	0.047	15.910	**0.00**		-0.001	0.000	-2.699	**0.01**	0.86	139.20
Q1	ch restaurants (all)	0.088	0.082	1.076	0.28	ch restaurants (rank)	0.004	0.002	1.805	0.07	0.06	183.69
Q2		-0.128	0.083	-1.542	0.12		-0.006	0.002	-2.802	**0.01**	0.10	199.83
Q3		0.090	0.093	0.964	0.34		0.002	0.003	0.802	0.42	0.02	243.87
Q4		0.206	0.078	2.644	**0.01**		0.007	0.004	1.869	0.06	0.09	141.67
Q5		0.004	0.081	0.053	0.96		0.003	0.004	0.709	0.48	0.01	198.98
W		0.790	0.141	5.590	**0.00**		0.002	0.002	0.863	0.39	0.36	566.66
Q1	ch tourism (all)	-0.165	0.087	-1.904	0.06	ch tourism (rank)	-0.142	0.034	-4.136	**0.00**	0.31	123.80
Q3		-0.182	0.134	-1.363	0.17		-0.200	0.045	-4.407	**0.00**	0.21	200.10
Q4		0.512	0.117	4.365	**0.00**		-0.032	0.042	-0.762	0.45	0.26	123.64
W		0.456	0.043	10.660	**0.00**		-0.006	0.016	-0.347	0.73	0.65	303.18
Q1	ch travel (all)	-0.150	0.092	-1.641	0.10	ch travel (rank)	-0.027	0.009	-2.928	**0.00**	0.12	172.74
Q2		0.039	0.081	0.484	0.63		-0.569	0.235	-2.422	**0.02**	0.07	203.43
Q3		0.104	0.091	1.148	0.25		0.162	0.233	0.695	0.49	0.02	245.05
Q4		0.237	0.094	2.523	**0.01**		0.269	0.135	1.988	***0.05***	0.12	117.99
Q5		0.039	0.072	0.539	0.59		-0.186	0.032	-5.896	**0.00**	0.36	127.16
W		0.756	0.072	10.501	**0.00**		-0.045	0.019	-2.376	**0.02**	0.66	297.79

Table 4: Comparing the Average Coefficients of Regressions involving Daily Data against Weekly Data																			
		Regressand--All Visits						Regressand--New Visits						Regressand--Search Traffic					
Time	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC
D	CH hotels	0.655	0.088	7.560	**0.00**	0.40	126.55	0.604	0.087	7.077	**0.00**	0.37	81.30	0.592	0.089	6.874	**0.00**	0.35	60.13
W	(all)	0.746	0.053	14.183	**0.00**	0.78	196.75	0.825	0.069	12.007	**0.00**	0.71	222.34	0.714	0.061	11.781	**0.00**	0.70	201.83
D	CH hotels	0.541	0.085	6.489	**0.00**	0.33	140.50	0.507	0.083	6.260	**0.00**	0.31	87.47	0.489	0.085	5.902	**0.00**	0.29	64.81
W	(travel)	0.765	0.046	16.692	**0.00**	0.83	156.10	0.875	0.055	15.897	**0.00**	0.81	149.09	0.729	0.056	13.079	**0.00**	0.74	176.89
D	CH rests	0.057	0.084	0.703	0.32	0.02	199.77	0.040	0.081	0.510	0.41	0.02	120.70	0.066	0.082	0.857	0.44	0.02	85.48
W	(all)	0.047	0.138	5.820	**0.00**	0.35	573.72	0.921	0.159	5.799	**0.00**	0.35	509.43	0.782	0.143	5.477	**0.00**	0.33	462.24
D	CH sc	1.173	0.151	8.359	**0.00**	0.44	121.98	1.170	0.142	9.104	**0.00**	0.47	71.48	0.984	0.161	6.604	**0.00**	0.33	61.85
W	(all)	1.353	0.239	5.665	**0.00**	0.34	581.02	1.420	0.285	4.976	**0.00**	0.29	559.89	1.450	0.230	6.297	**0.00**	0.39	418.36
D	CH sc	0.775	0.118	7.241	**0.00**	0.37	135.12	0.749	0.114	7.212	**0.00**	0.37	82.49	0.708	0.117	6.798	**0.00**	0.34	61.65
W	(travel)	0.969	0.095	10.151	**0.00**	0.62	328.65	1.106	0.110	10.028	**0.00**	0.61	294.88	0.976	0.096	10.156	**0.00**	0.63	254.22
D	CH tourism	0.011	0.109	0.040	0.20	0.08	181.24	0.005	0.105	-0.010	0.24	0.08	108.28	0.002	0.105	-0.053	0.10	0.09	76.02
W	(all)	0.457	0.042	10.777	**0.00**	0.65	303.76	0.540	0.045	11.983	**0.00**	0.69	234.03	0.417	0.049	8.448	**0.00**	0.52	315.62
D	CH travel	0.088	0.090	0.996	0.19	0.03	192.54	0.068	0.087	0.822	0.27	0.03	116.10	0.099	0.088	1.142	0.17	0.03	83.15
W	(all)	0.750	0.075	9.986	**0.00**	0.62	328.19	0.843	0.089	9.419	**0.00**	0.59	311.65	0.693	0.084	8.203	**0.00**	0.52	321.40

Table 4 (Continued): Comparing the Average Coefficients of Regressions involving Daily Data against Weekly Data													
		Regressand--Nonlocal Visits						Regressand--Local Visits					
Time	Regressors	Coef.	S.E.	Z-Stat	PV	R-Sq	AIC	Coef.	S.E.	Z-stat	PV	R-Sq	AIC
D	CH hotels	0.659	0.085	7.842	**0.00**	0.42	61.81	0.653	0.096	6.909	**0.00**	0.36	81.29
W	(all)	0.918	0.067	13.724	**0.00**	0.77	249.15	-0.047	0.065	-0.727	0.47	0.01	65.42
D	CH hotels	0.545	0.083	6.699	**0.00**	0.35	68.46	0.539	0.092	5.987	**0.00**	0.30	88.64
W	(travel)	0.946	0.057	16.526	**0.00**	0.82	191.06	-0.066	0.064	-1.044	0.30	0.02	64.90
D	CH rests	0.069	0.082	0.848	0.25	0.03	96.43	0.047	0.088	0.547	**0.42**	0.02	119.96
W	(all)	0.950	0.171	5.557	**0.00**	0.33	734.14	0.039	0.112	0.352	0.72	0.00	65.80
D	CH sc	1.173	0.147	8.802	**0.00**	0.46	59.67	1.175	0.165	7.520	**0.00**	0.39	78.69
W	(all)	1.598	0.295	5.417	**0.00**	0.32	742.64	0.073	0.188	0.386	0.70	0.00	65.77
D	CH sc	0.768	0.116	7.335	**0.00**	0.38	66.88	0.782	0.126	6.734	**0.00**	0.34	84.84
W	(travel)	1.162	0.120	9.704	**0.00**	0.60	429.02	-0.026	0.099	-0.259	0.80	0.00	65.85
D	CH tourism	0.018	0.110	0.161	0.19	0.08	90.59	0.004	0.114	-0.075	0.23	0.08	106.61
W	(all)	0.557	0.052	10.785	**0.00**	0.64	372.56	-0.066	0.045	-1.470	0.14	0.03	63.93
D	CH travel	0.081	0.089	0.948	0.20	0.03	94.04	0.095	0.094	1.008	0.19	0.04	115.25
W	(all)	0.909	0.094	9.642	**0.00**	0.60	419.17	-0.042	0.083	-0.511	0.61	0.01	65.61