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

Tierney, Heather L.R. and Pan, Bing

College of Charleston, College of Charleston

1 September 2010

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

By

Heather L.R. Tierney[*] and Bing Pan

July 2011

Abstract

A new area of research involves the use of normalized and scaled Google search volume data to predict economic activity. This new source of data holds both many advantages as well as disadvantages. Daily and weekly data are employed to show the effect of aggregation in 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 search queries, along with the rankings of that 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 raw 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, College of Charleston; 5 Liberty Street, Charleston, SC 29424, email: hlrtierney@yahoo.com; phone: (843) 953-7070; fax: (843) 953-0754, 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, Martin Ridout, Stefan Voß, and two anonymous referees.

1. Introduction

A burgeoning field of research involves using search engine data from the websites: Google Analytics, Google Insights, or Google Trends to predict economic activities.[1] The lure of using Google search volume data 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 raw data. Google is transparent in the fact that it *only* provides data that has been normalized and scaled after it has been amassed from users around the world, all of which further compounds the problems when the data is aggregated from a daily to a weekly level. The main reason that Google will not release the raw search data is due to privacy concerns such as the controversies of revealing a user's identity through search history data (Barbaro and Zeller 2006). Google further normalizes and then scales their raw data in order to present the data in a more user-friendly and understandable format. This necessary process 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 straight forward, which has important policy implications.

The effect of normalization, scaling, and aggregation in regard to website traffic and search engine queries has not been directly examined, and this paper investigates the potential pitfalls of using such data. Geographic Information Systems (GIS) studies have used scaled variables and time series econometrics has used aggregated variables. However, those studies 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). Marvasti (2010) has noted that

[1] The web address for Google Analytics is: http://www.google.com/analytics/; the web address for Google Insights is: http://www.google.com/insights/search/; and the web address for Google Trends is: http://www.google.com/trends.

data regarding information technology (IT) is particularly sensitive to data aggregation to the point where the link between raw data and the aggregate data is lost especially when range reduction occurs.[2] Through the normalization, scaling, and aggregation process, Google data becomes a truncated variable, which also involves range reduction.

Internet activity has been used to predict economic activity and even flu epidemics (Ginsberg, Mohebbi, Patel, Brammer, Smolinski, and Brilliant 2009). In regard to economic activity, Azar (2009) finds a negative relationship between shocks to oil prices and patterns of Google searches for electric cars in a Bayesian Vector Autoregression (BVAR) model. 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.

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. This is important to investigate because regardless of the type of data, *all* Google data goes through the same amassing, normalizing, and scaling processes and this study, by investigating the behavior of Google data, benefits not only the field of Economics, but also other fields of research that intend to use Google data. The second purpose is to understand the relationship between

---

[2] Marvasti (2010) analyzes continuous and discrete data that relates to the Operating Systems metrics such as the central processing unit (CPU), memory, and bandwidth usage; usage metrics such as number of hits, sessions, connections; and environmental metrics such as power supply temperature and processor temperature.

website traffic of a given website and the volumes of keyword searches on Google, 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 generated 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 by Google, 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.[3]

Following the reasoning of Michener and Tighe (1992), a count model, specifically the Poisson quasi-maximum likelihood (QML) regression is used to model the regressand, the website traffic of the Charleston Area Convention and Visitors Bureau (CACVB) website. Website traffic data is a non-negative count variable that does not have an upper bound. To keep the analysis consistent, the Poisson QML regression is used to model both daily and weekly data. The Poisson QML Regression Model relaxes the constraint of the conditional mean being equal to the condition variance. This model is needed since both frequencies of data have both over- and under-dispersion present.

Furthermore, count data typically experiences heteroskedasticity, which the Poisson QML model automatically takes into account (Cameron and Trivedi 1998, Wooldridge 2002). In addition, the use of the Poisson QML regression eliminates the interpretation problem of the regression coefficients being obtained from data that has been normalized and then scaled which trans forms the data into an index. The nature of Poisson regressions permits the interpretation of regression coefficients as elasticities or semi-elasticities, which is not automatically the case with all regression models (Wooldridge 2002).

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

The most important findings of this paper pertain to describing and noting the potential limitations of Google data, a new source of data that has been normalized and scaled. Researchers can neither reverse-engineer nor gain access to the original data. Similar to time series data, 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 the statistical significance of the estimated coefficients. This paper also finds that depending on the frequency, the data can change from having under-dispersion to extreme over-dispersion, which indicates that the normalization, scaling, and aggregation of Google data are affecting its statistical characteristics and the modeling of the data. In addition, this paper also finds that search volumes for certain keyword search queries have a larger impact on website traffic than does the ranking of a website specifically with respect to the CACVB website.

The structure of this paper is of the following format: Section 2 presents the theoretical model. A brief discussion of the data and the empirical results are presented in Section 3. The conclusion is presented in Section 4.

2. Theoretical Model

The data used as the regressands are various forms of website traffic, i.e. the number of web hits on the CACVB website, $\left\{ y_t \in \mathbb{Z}^+ : y_t \geq 0 \right\}$ with $t = \left\{ 1, \cdots, T \right\}$, which is a count variable that takes on only non-negative integer values with no upper bound.

Modeling a discrete variable poses certain challenges. For instance, it is unclear on how to obtain the conditional mean, $E\left(y | \mathbf{x} \right)$, where $\mathbf{x}$ is the regressor matrix, from $E\left(\log\left(1 + y\right) | \mathbf{x} \right)$ if one were to take the log transformation of y in order to use the ordinary least squares (OLS) method. A nonlinear least squares model is inefficient due to the assumption of homoskedasticity and typically, count data is heteroskedastic in nature (Wooldridge 2002).

So, given the nature of the data, this implies that since y_t is a Poisson distributed variable, it should be modeled using the Poisson Count Model as a starting point. That way the predicted values of y_t are ensured to be non-negative. The model is as follows:

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

where $\mathbf{x}_t = \left(1, \quad x_{1t}, \quad x_{2t}, \quad \cdots, \quad x_{kt}\right)$ is the regressor matrix with the total number of regressors being k and where the error term is $\varepsilon_t \sim P\left(0, \sigma_P^2\right)$ (Cameron and Trivedi 1998, Wooldridge 1997). The use of the Poisson Count Model permits one to combine the discrete nature of the regressand while having a condition mean linear in parameters and while permitting heteroskedasticity (Wooldridge 1997).

The potential problem with the Poisson Count Model is that it has the very strict assumption of the conditional mean being equal to the conditional variance, which is as follows:

$$m\left(\mathbf{x}_t, \beta\right) = \exp\left(\mathbf{x}_t'\beta\right) = E\left(y_t \middle| \mathbf{x}_t, \beta\right) = \operatorname{var}\left(y_t \middle| \mathbf{x}_t, \beta\right) \tag{2}$$

(Cameron and Trivedi 1998, Wooldridge 1997).

As recommended by McCullagh and Nelder (1989), the equivalency of the conditional mean to the conditional variance needs to be checked. This can be done using the Pearson Chi-Squared test statistic divided by the degrees of freedom, $df = \left(T - q\right)$ where $q = k + 1$ and is of the following form:

$$\lambda_P = \left(\chi^2 / df\right) = \sum_{t=1}^{T}\left(\frac{\left(y_t - \hat{\mu}_t\right)^2}{\hat{\mu}_t}\right) \Bigg/ df = \sum_{t=1}^{T}\left(\frac{\left(\hat{\varepsilon}_t\right)^2}{\hat{\mu}_t}\right) \Bigg/ df \tag{3}$$

with $\hat{\mu}_t$ and $\hat{\varepsilon}_t$ being the estimated conditional mean and residual for the t^{th} observation, respectively. The reason for using λ_P as a dispersion measure is that $\hat{\lambda}_P$ is a consistent measure of λ_P. The deviance-based measures of dispersion that utilizes the likelihood function do not possess such characteristics.

If λ_P is greater than unity, this indicates that over-dispersion is present and if it is less than unity, then under-dispersion is present. An acceptable range of

λ_p should be no more than 4; otherwise, the difference between the actual and fitted values are deemed too great and generally indicates that the Poisson QML Regression Model is not a good fit for the data (McCullagh and Nelder 1989).

Tables 2A, 2B, 3, and 4 in Appendix show that there are both over- and under-dispersion present in the regressions involving the daily and weekly data, as well as values of $\hat{\lambda}_p > 4$ in the regressions involving the weekly data, which is further discussed in more detail in Section 3. This necessitates that the Poisson Count Model be abandoned in favor of a model that relaxes the assumption of the conditional variance and conditional mean being equal. The two alternatives are the Poisson QML Model and the Negative Binomial QML Model.

The Negative Binomial QML Model is investigated and then rejected as a potential model because it does not produce robust estimators and adds computational complication with a high potential for inconsistencies since both the conditional mean and the fixed variance parameter have to be estimated in a maximum likelihood framework (Wooldridge, 1997). The main reason that the Negative Binomial QML Model is not utilized is that it is only able to incorporate over-dispersion, and thus, is not used to model neither the daily nor the weekly data, since there is under-dispersion in a number of the samples (Cameron and Trivedi 1986).[4]

As a result, the Poisson QML Model is used for modeling both the daily and weekly frequency of Google data due to the following benefits.[5] One of the main benefits is that if the appropriate conditional mean is used, the parameters are consistent even if the underlying distribution is incorrectly specified. Another benefit is that the Poisson QML Model is also relatively efficient with robust

---

[4] The two-step Negative Binomial QML Model, where the estimated fixed variance parameter, $\hat{\eta}^2$ is obtained from the estimated slope coefficient of regressing $\left(u_t^2 - 1\right)$ onto the forecasted values of Equation (4), is also rejected as a potential model since it also is only able to incorporate over-dispersion, (Wooldridge 1997).

[5] For this paper, the quadratic hill climbing method is implemented to estimate the Poisson QML regressions.

standard errors, which take into account heteroskedasticity and is of the following form:

$$y_t = \exp\left(\mathbf{x}_t'\beta + \varepsilon_t\right) = \exp\left(\beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + u_t\right) \tag{4}$$

with the error term being $u_t \sim N\left(0,\sigma^2\right)$ (Cameron and Trivedi 1998, Wooldridge 1997). Since the Poisson QML estimators have asymptotically normal error terms, this transforms y_t into a compound Poisson variable (Cameron and Trivedi 1986, Wooldridge 2002).

The flexibility of the Poisson QML Model permits both over- and under-dispersion to be taken into account though the use of the coefficient of variation, i.e. the variance factor, σ^2, which means that the conditional variance is permitted to differ from the conditional mean:

$$\operatorname{var}\left(y|\mathbf{x}\right) = \sigma^2 m\left(\mathbf{x},\beta\right). \tag{5}$$

When σ^2 is greater than 1, the variance will be greater than the conditional mean, and vice versa. Equation (5) also permits the case when σ^2 is equal to one. For this paper, the coefficient of variation, σ^2, and the weighted standard errors are calculated using the standardized Pearson errors terms (McCullagh and Nelder 1989).

The interpretation of the regression parameters is not as straight forward as a linear regression model since the Poisson QML Model involves the exponential mean function. The regression coefficients automatically lend themselves to interpretation as some form of elasticity, i.e. as a percentage change. This is particularly useful for Google data which has been transformed into a truncated index.

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

$$\frac{\partial E\left[y_t|\mathbf{x}_t\right]}{\partial x_{jt}} = \exp\left(\beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt}\right) \times \beta_j = E\left[y_t|\mathbf{x}_t\right] \times \beta_j \tag{6}$$

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

Another interpretation of the Poisson QML Model parameters involves taking the log of Equation (4):

$$\log(y_t) = \log\left(\exp(\mathbf{x}_t'\beta + \varepsilon_t)\right) = \mathbf{x}_t'\beta + \varepsilon_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_k x_{kt} + \varepsilon_t . \qquad (7)$$

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 $E\left[y|\mathbf{x}\right]$ by $100\beta_j$ (Wooldridge 2002). Following Equations (4) and (7), the regressand will not be transformed into a log variable while the regressor x_j is transformed into a log variable.[6] Thus, the coefficient β_j, obtained from Equation (7), can be interpreted as an elasticity, meaning that β_j is the percentage change in $E\left[y|\mathbf{x}\right]$ due to a percentage change in x_j (Michener and Tighe 1992, Cameron and Trivedi 1998).

When it come to measuring the goodness-of-fit of the Poisson QML Model, the general R^2 term used in OLS Model cannot be used due to the conditional mean being nonlinear and the model containing heteroskedasticity. A *pseudo*-R^2 term based upon the deviance is used as the goodness-of-fit measure. The sum of the deviance refers to the sum of the differences between the unrestricted model, i.e. the fitted model and a model with only a constant term, which is the residual deviance test:

$$G^2 = \sum_{t=1}^{T} d_t = 2\sum_{t=1}^{T}\left[y_t \log(y_t/\hat{\mu}_t) - (y_t - \hat{\mu}_t)\right] \sim \chi^2_{T-q} \qquad (8)$$

where $\sum \hat{u}_t = \sum(y_t - \hat{\mu}_t) = 0$ when an intercept term is included in the regression model. Hence, Equation (8) is reduced to

$$G^2 = \sum_{t=1}^{T} d_t = 2\sum_{t=1}^{T}\left[y_t \log(y_t/\hat{\mu}_t)\right] \qquad (9)$$

Based upon Equation (9), the *pseudo*- R^2 term is formed through the use of a likelihood ratio index:

[6] The regressand is not transformed into a log variable because otherwise when one takes the log of Equation (4) as is done in Equation (7), then one would actually be taking the log of [log (y_t)], which would further complicate policy implementation.

$$pseudo-R^2 = 1 - \frac{l(\hat{\mu}_t, y_t)}{l(\bar{y}_t, y_t)} \qquad (10)$$

where $l(\hat{\mu}_t, y_t)$ is the log-likelihood function of the fitted model and $l(\bar{y}_t, y_t)$ is the log-likelihood function of a model with only a constant term where $\bar{y}_t$ is the average of y_t (Cameron and Windmeijer 1996).

Regarding the out-of-sample performance, the mean absolute percentage error (MAPE) will also be used to measure the forecasting abilities of the regression models. The MAPE is used since it presents the out-of-sample fit in terms of a percentage as opposed to levels. This makes it easier to compare the forecasting results of the models using daily and weekly data, which have been transformed through the use of normalization and scaling methods.

3. Empirical Results

This section is divided into three sub-sections. Subsection 3.1 contains more detail about the data used as the regressands - the five different sources of website traffic, i.e. web hits in level terms (not transformed into log variables) and the regressors, which are the log transformation of the seven keyword search volumes and five keyword rankings that relate to specific search queries.[7] Concerning the univariate Poisson QMLE regression results, which are presented in Subsection 3.2, each of the five different sources of web traffic is modeled using each of the seven keyword search volumes and corresponding keyword ranking as regressors.[8] Subsection 3.2 also concerns the bivariate Poisson QMLE regressions, which model each of the regressands, the five different sources of web traffic using the regressors of keyword search volumes and their respective rank for five different queries. Both daily and weekly data are used for the univariate and bivariate Poisson QMLE regressions. Lastly, in Subsection 3.3, the effects of normalization, scaling, and aggregation are discussed.

[7] The level variables for the search volume of keyword queries have also been used in model estimation, but the findings do not vary significantly.
[8] E-Views 7.1 has been used to calculate all the empirical results of this paper.

3.1 Discussion of Data

When a potential traveler searches for a destination, he or she will most likely type in a query in search engines, look through the returned results, and pick a webpage to investigate. Thus, search volumes for certain 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).

The two regressors in this study are the ranking of one site for certain search queries and the search volumes for those queries. The regressand is online traffic for that specific website. The data used in the Poisson regressions comes from two different Google sources, which are Google Analytics and Google Insights. 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 the Google Analytics account of the CACVB website. Google Analytics uses a short Javascript on every page of a website to capture visitors' visitation behavior. Five regressands are specifically analyzed and listed in Tables 1A and 1B. The first regressand, entitled *all visits*, encompasses all the website traffic to the website. A sub-category of only first visits to the website (as indentified by new Internet Protocol (IP) addresses) forms the second regressand of *new visits*, and website traffic that is from only the search engine, which is referred to as *search traffic*, is the third regressand. The fourth and fifth regressands are also sub-groups of website traffic - one from outside the local Charleston area and is entitled *nonlocal visits* and the other from the local Charleston area entitled *local visits*, respectively.

A program has been designed to download daily search engine results as it pertains to five different keyword search queries in a general overall type of search category, which are *charleston sc, travel charleston, charleston hotels, charleston restaurants*, and *charleston tourism* and two different keyword search queries in the sub-search category of 'travel,' which are *charleston hotels*, and *charleston sc*, bringing the total number of regressors to seven with respect to the search volume regressors. In addition, the rankings of the CACVB site for five search queries were obtained through a custom-built program, and these form the five rank variable regressors. These queries are the popular searches used by visitors to Charleston, SC, according to Google Keyword Tool.[9]

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, daily search volumes for specific searches for up to five queries can be obtained but only for a limited time period (three months), which might be due to the normalization and scaling processes and the need to protect the privacy of Google users. The normalizing and scaling parameters used for daily data change quarterly. Thus, the daily data is examined only one quarter at a time for the sake of consistency and this prohibits the examination of daily data over a longer sample period of more than three months.

Both normalizing and scaling are used in processing amassed raw search engine volume data in Google Insights. Regarding normalization, according to Google Insights (2009a), the raw data is first combined by regions and is then normalized by dividing the number of searches for a given query by the total website traffic in a given region. Thus, even if two regions have the same percentage of search volumes for a given search query, this does not automatically mean that the absolute search volumes are the same. This prevents keywords from 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 is from 0 to 100 with each data point being divided by the highest point or 100 (Google Insights 2009b). Aside from possible distortions caused by the normalization and scaling processes, the aggregation of data may have an effect on the empirical findings as well (Rossana and Seater 1995, Pyle 1999, Marvasti 2010). Here aggregation refers to the transformation of data from a lower frequency to a higher frequency, such as the transformation of daily data into weekly data. This will be discussed in more detail in Subsections 3.2 and 3.3.

---

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

For the regressors, the search volumes for specific search queries are in two categories: an overall type of searches is denoted as 'all' and a subcategory of queries in just travel categories as specified by Google Insights is denoted as 'travel' (Table 1B). The normalized daily data has the same scale for three months, which limits the analysis of the daily datasets to only one quarter at a time (Askitas and Zimmermann 2009, Google Trends 2009). It is not advisable to form a longer sample using daily data by combining the different quarters due to the variability of normalization and scaling parameters from quarter to quarter. 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 website traffic data of the CACVB website are the sum 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*; the queries in the 'travel' search category is as follows: *charleston hotels*, and *charleston sc*. Google Insights did not provide the other three search traffic in the travel category due to their small volumes. The log of the scaled search volumes is used since the coefficients of the regressors can be interpreted as elasticities, which is analogous to Equation (6). 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 elasticities as defined in Equation (7).

Regarding the weekly regressors, the entire period from January 2008 to March 2009 is examined simultaneously because the normalization and scaling parameters are the same for the search volume variables. For the rank variables, the weekly data are the average of a seven-day week. Poisson regressions are used for both daily and weekly data. The number of observations for each Poisson regression is listed in the Appendix in Table 1A. There are some missing observations due to temporary lack of connection to the Internet or the blockage of the custom-built program by the Google Server possibly due to spamming suspicion.

3.2 Interpretation of the Univariate and Bivariate Poisson QMLE Results

The reason for estimating and analyzing univariate Poisson QMLE regressions is to mimic the behavior of a 'typical' consumer looking for a particular type of product on Google such as hotels or restaurants in Charleston, SC. Those queries are related to Charleston tourism and thus affect website traffic to the CACVB website. For the bivariate Poisson QMLE regressions, the rankings of the specific search queries are also included along with the search query volumes in order to see how the search volumes relate to the five different sources of website traffic to the CACVB website.[10]

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. The general relationship between the estimated coefficients of the log of the search volumes and the regressand is mostly a positive one (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 types of website traffic to the CACVB website. This is also true for the estimated coefficients involving the log of the search query 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

---

[10] All of the search queries are not used simultaneously in a Poisson QMLE regression since it does not follow the pattern of a 'typical' consumer looking for a specific product such as hotels, restaurants, or tourist attractions to visit using this specific type of dataset, but in a different field of research, compound queries would be more common.

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, 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 (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, i.e. the existence of a multicollinearity problem. This produced singular matrices in the Poisson QMLE results, which 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.

When it comes to the matter of dispersion, the univariate regressions show the presence of both over- and under-dispersion, which the flexibility of the Poisson QML Regression Models are able to handle. The values of $\hat{\lambda}_p$, which are found in Tables 2A, 2B, 3, and 4, are the estimated measure of the Pearson Chi-Squared test statistic divided by the degrees of freedom as is given in Equation (3).

The univariate regressions that involve the regressands of *all visits*, *new visits*, and *local visits* tend to have $\hat{\lambda}_p$ values greater than unity but less than 4, which indicate that there is over-dispersion present at an acceptable level. Regarding the univariate regressions with the regressand of *search traffic*, they generally show under-dispersion while the regression with the regressand of *nonlocal visits* displays more of a mix of both over- and under-dispersion. Concerning the weekly data, all of the Poisson QML regressions display over-dispersion except for the regressions involving the regressand, *local visits*.

Behavior of some of the daily and weekly data is displayed in Graphs 1A to 1C, 2A to 2C, 3A to 3C, and 4A to 4C. The graphs are organized according to their pseudo-R^2 levels. Graphs that are denoted by the letter "A" and "B," are of univariate regressions using daily data with the lowest pseudo-R^2 terms and highest pseudo-R^2 terms, respectively.

For example, Graphs 1A to 1C show all the relationship between *all visits* and *charleston hotels (all)*. Graph 1A shows the regressing with the lowest pseudo-R^2 of 0.30, which is produced by the daily data during the time period of January 2008 to March 2008 (Q1). The highest pseudo-R^2 of 0.45 is produced by the daily data during the time period of October 2008 to December 2008 (Q4), and this regression is shown in Graph 1B. Weekly data fares better with a pseudo-R^2 of 078 as is shown in Graph 1C. In terms of explaining the goodness-of-fit of *all visits* at the daily level, the variable *charleston hotels (all)* is not all encompassing.

A similar case can be made for Graphs 2A through 2C, which shows the relationship between *new visits* and *charleston sc (travel).* It is no surprise that the daily data of Quarter 2 (April 2008 to June 2008) (Graph 2B) fits better when compared to Graph 2A because the pseudo-R^2 terms are 0.52 and 0.16 respectively. The best fit is of the weekly data as is shown in Graph 2C with a pseudo-R^2 term of 0.63. In terms of the estimated coefficient of variation $\hat{\lambda}_p$, the weekly data has a measure of 4.88 while the daily data of Quarter 1 and Quarter 4 have $\hat{\lambda}_p$ measures of 1.15 and 0.42 (Table 2A). This seems to signify a difference of behavior between the daily and weekly data, which is further discussed in Sub-section 3.3.

The univariate and bivariate regressions using weekly data have $\hat{\lambda}_p$ values that are much larger than 4 and are also much larger than their daily data counterpart with a few exceptions involving the regressions with *local visits* as a regressand. This shows an underlying change in the statistical characteristics of the weekly data. It should be noted that the average $\hat{\lambda}_p$ values produced by daily data are much smaller than the $\hat{\lambda}_p$ values produced weekly data. This supports Marvasti's (2010) finding that higher levels of aggregations produce distortions.

In terms of the MAPE, the univariate regressions using daily data tend to have values that are around 11% on average while the univariate regressions using weekly data tend to have values that are around 15% on average. This indicates that the univariate regressions utilizing daily data perform better than the regressions involving weekly data.

3.3 Effects of Normalization, Scaling, and Aggregation

As has been discussed in the previous subsection especially as it pertains to $\hat{\lambda}_p$, there appears to be some distortion between the daily and weekly Google data. This is the most dramatic finding of this paper which could have important implications in terms of policy implementation. One would expect some loss of information due to aggregation, but not a complete transformation of the data so much so that the underlying characteristics are lost. This does appear to be the case with respect to IT data as has been found by Marvasti (2010) and with respect to Google data as this paper states.

In order to understand the relationship between daily and weekly data, the average of the regression estimates for the five quarters involving daily data is compared to the regression estimates of the weekly data (Table 4).[11] The regression results quarter-by-quarter are presented in Tables 2A, 2B, and 3. Comparing the results of Tables 2A, 2B, and 3 against Table 4, one can see that the estimated results are all within a given range, meaning that there are no outliers in all quarters when using daily data. So instead of just focusing on one specific quarter, one can use the average to summarize the behavior of the daily data.

However, depending on the level of aggregated data used, not only does the coefficient of variation as captured by the numerator of $\hat{\lambda}_p$ in Equation (3) changes dramatically, but the estimated regression coefficients can also dramatically change even with respect to sign. This indicates a fundamental change in the statistical characteristics of the underlying data.

In some instances, depending on the frequency of data, there is a reversal of the sign of an estimated coefficient and even statistical significance. This 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 search query 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 pseudo-R^2 term the better. For the seven Poisson QMLE regressions that involve *local visits*, the regressions that use weekly data produce pseudo-R^2 terms that are lower than

[11] 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.

when daily data is used.[12] This is not the case for the remaining four regressands. In those cases, the pseudo-R^2 terms are generally higher on average when weekly data is used as opposed to daily data (Table 4).

In addition, the interpretation of the estimated coefficients could be drastically different 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. This indicates 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 (Table 4). The pseudo-R^2 term is approximately 50% higher (Table 4). Furthermore, at the univariate level, there appears to be under-dispersion, with a $\hat{\lambda}_p$ value of 0.85; while at the weekly level there appears to be over-dispersion with a $\hat{\lambda}_p$ value of 4.88. The latter is a little outside the acceptable range for the Poisson QML Regression Model. This indicates that there might be a need for a completely different statistical model depending on the frequency of Google data used.

Another example involves the regressand, *search traffic*, and the regressor *charleston sc (all)* in the general overall search category (Table 4). When daily data is used, the average estimated regression coefficient is 0.984 with a pseudo-R^2 term of 0.30; when weekly data is used, the statistically significant estimated regression coefficient is 1.450 with a pseudo-R^2 term of 0.39. Graphs 3A to 3C show all the relationship between *search traffic* and *charleston sc (all)* for the daily data for Quarters 3 and 4 and for the weekly data set.[13]

Hence, it appears that in terms of explaining the variability of the regressand, *search traffic*, both daily and weekly data capture approximately the same level but the values for $\hat{\lambda}_p$ are very different. All of the daily data sets show under-dispersion

---

[12] This could also indicate that when it comes to researching items related to Charleston tourism, the relationship between local website traffic is not that substantial as it relates to the CACVB website.

[13] Graphs 4A to 4C depict daily data that has under-dispersion while the weekly data appears to have over-dispersion.

while the weekly dataset shows over-dispersion. When daily data is used, the average $\hat{\lambda}_p$ is 0.63, which states there is under-dispersion and when weekly data is used, the $\hat{\lambda}_p$ is 6.82. This also states that the Poisson QML Regression Model is not a good fit for regressions involving weekly data. Thus, based on $\hat{\lambda}_p$, there appears to be a drastic divergence between the Poisson QMLE regression model using daily versus weekly data. Daily data is more efficient due to having a smaller variance when compared to the model using weekly data.

When using weekly data, the regressions produce larger estimated coefficients on average, which naturally produce larger pseudo-R^2 terms. Thus, the pseudo-R^2 terms might not be necessarily the best measure for goodness-of-fit. This leaves the coefficient of variation as a measure of goodness-of-fit of the regression models for consideration, which is represented by $\hat{\lambda}_p$. This is not a prudent conclusion to draw due to the lack of transparency in the normalized and scaled search query volumes.

Thus, in the examination of the relationship between various types of website traffic and the regressor of the log of keyword search volumes, different models and interpretations emerge depending on the frequency of data used.

4. Conclusion

Data from Google Insights or Google Trends can better help businesses to monitor changing consumer interests by providing instantaneous access to the most current search volume data available at any given time. It can also greatly benefit research in microeconomics. This research provides more evidence for the link between one business's revenue-generating activities with Google search volume data. Thus, those are a consumer-driven data source and help explain or predict economic activity without the time lag required for economic time series such as unemployment statistics. Despite the alluring benefits, the potential problems of using Google data as it relates to interpreting economic models for policy implementation should not be ignored, which is universal to any field of research using Google data.

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 and data presentation issues, which is problematic to the researcher wishing to deal with the original level data. For researchers and especially for the purpose of policy implementation, it is difficult to see the direct effect of aggregation when normalized and scaled data is used. As is demonstrated in this paper, the empirical Poisson QMLE results can be drastically different when using daily and weekly data as well as the dispersion levels as measured by the coefficient of variation. These differences could be due to the effects of normalization, scaling, and/or aggregation, but the lack of transparency makes it difficult to identify.

Furthermore, the use of data from Google Insights or 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. In this paper, the problem of interpreting the regression coefficients is averted through the use 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. This will involve a change in the data collection and processing 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 subsequently benefit the consumers versus the protection of consumer 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.

Cameron, A.C. and Windmeijer, F. (1996), "R-Squared Measures for Count Data Regression Models with Applications to Health-Care Utilization," *Journal of Business and Economic Statistics*, 14:2, 209-220.

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).

McCullagh, P., Nelder, J. A. (1989), "*Generalized Linear Models*," Second Edition, New York: Chapman and Hall.

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.

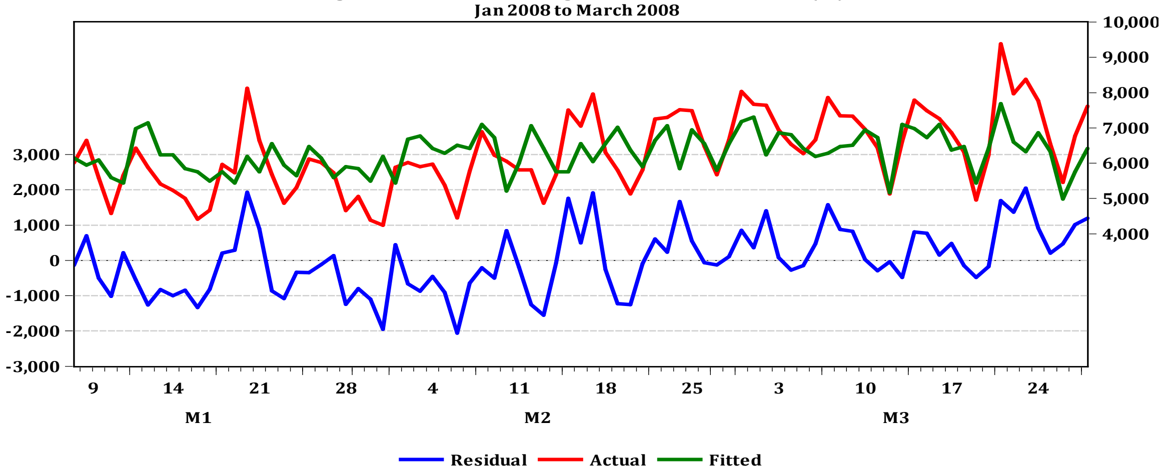
Marvasti, M.A. (2010), "Quantifying Information Loss through Data Aggregation," VMware Technical White Paper, 1-14.

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

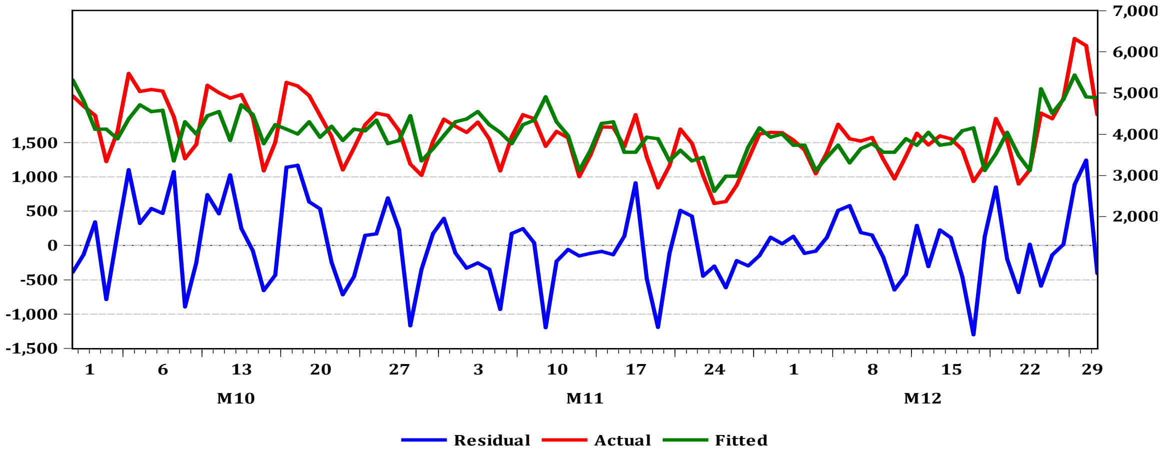
Wooldridge, J.M. (1997), "Quasi-Likelihood Methods for Count Data," *Handbook of Applied Econometrics, Volume II: Microeconomics*, M.H. Pesaran and P. Schmidt, eds., Oxford: Blackwell, 352-406.

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

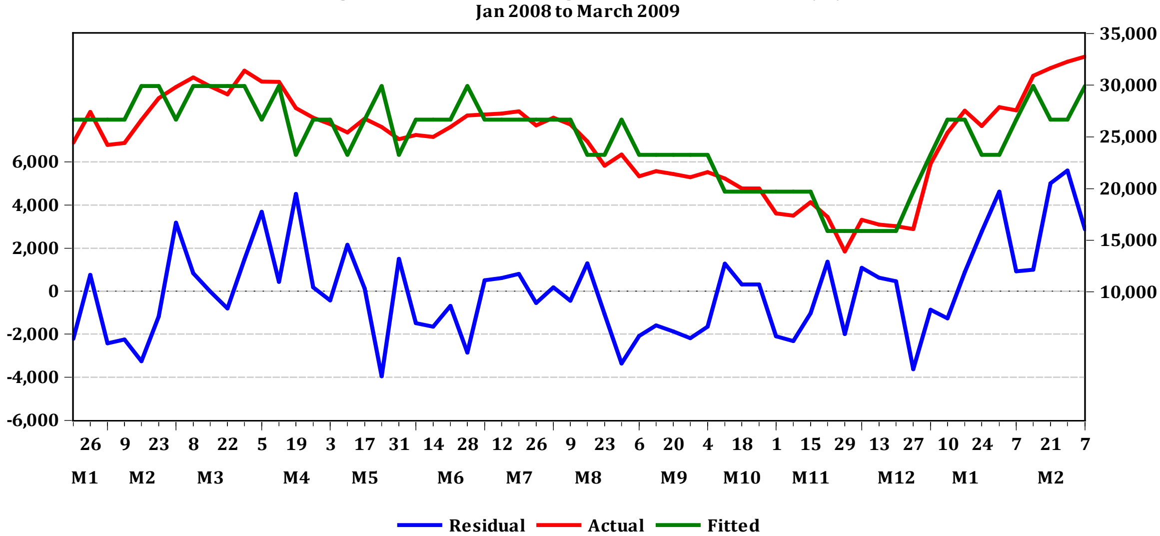
Graph 1A--Quarter 1
Regressand-All Visits and Regressor-"charleston hotels" (all)
Jan 2008 to March 2008



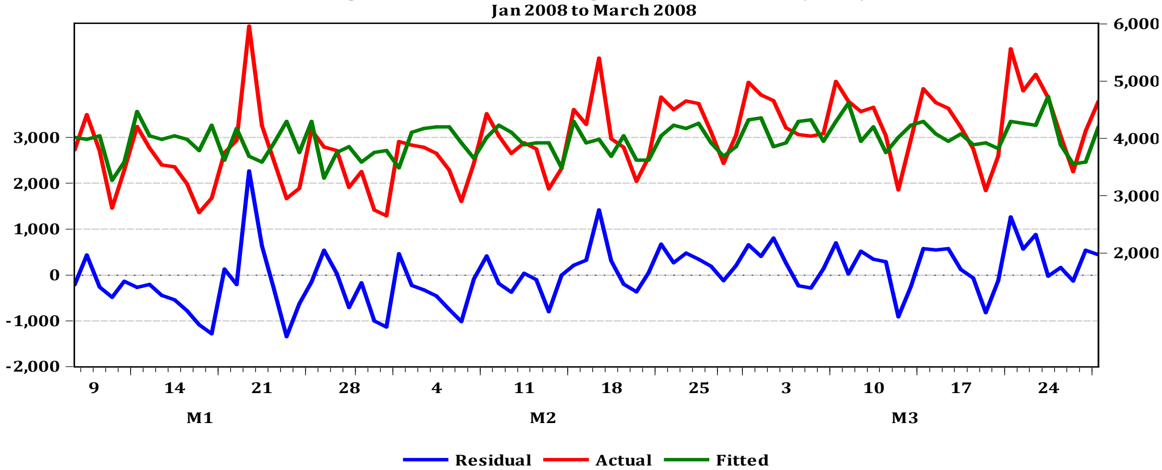
Graph 1B--Quarter 4
Regressand-All Visits and Regressor-"charleston hotels" (all)
Oct 2008 to Dec 2008



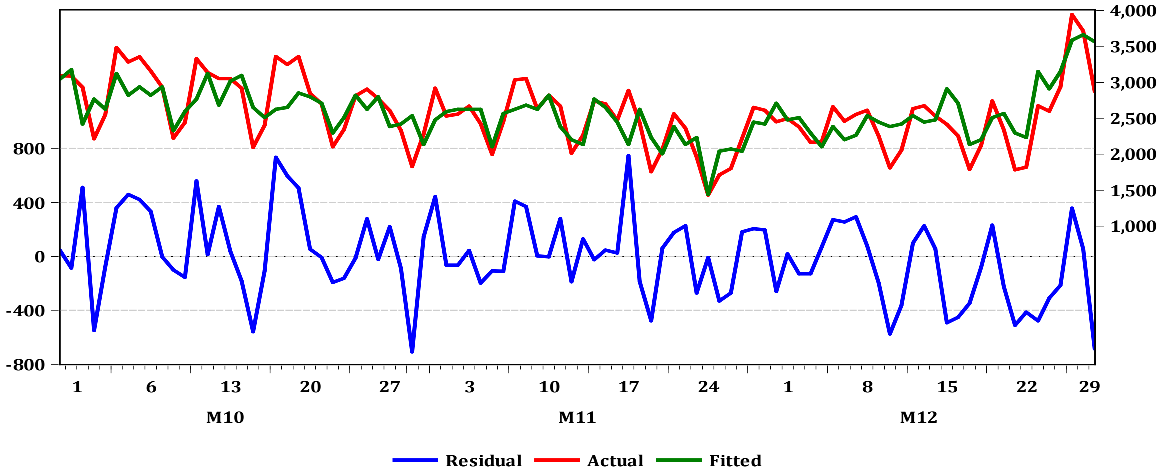
Graph 1C--Weekly Data
Regressand-All Visits and Regressor-"charleston hotels" (all)
Jan 2008 to March 2009



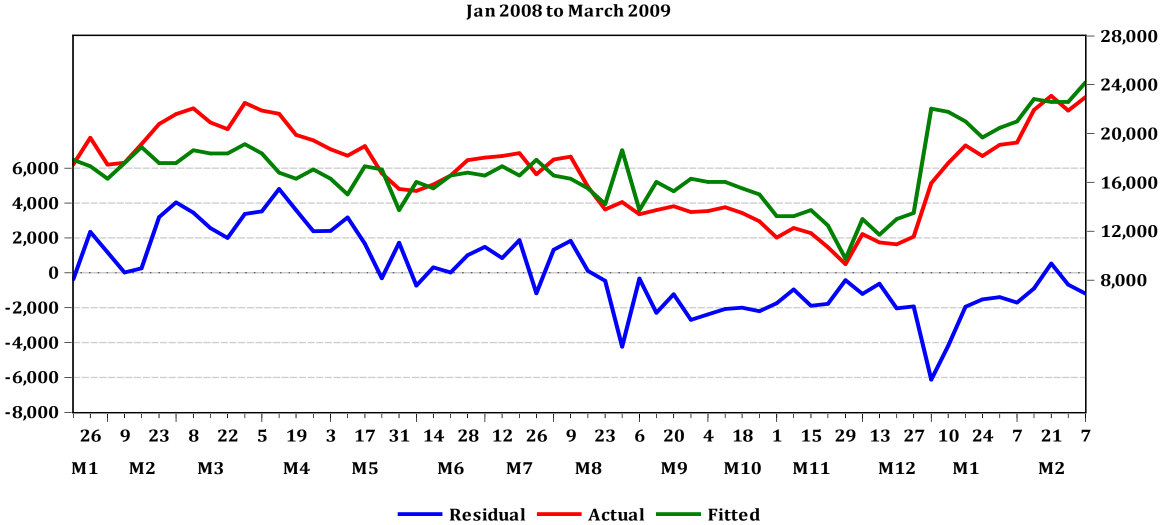
Graph 2A--Quarter 1
Regressand-New Visits and Regressor-"charleston sc" (travel)
Jan 2008 to March 2008



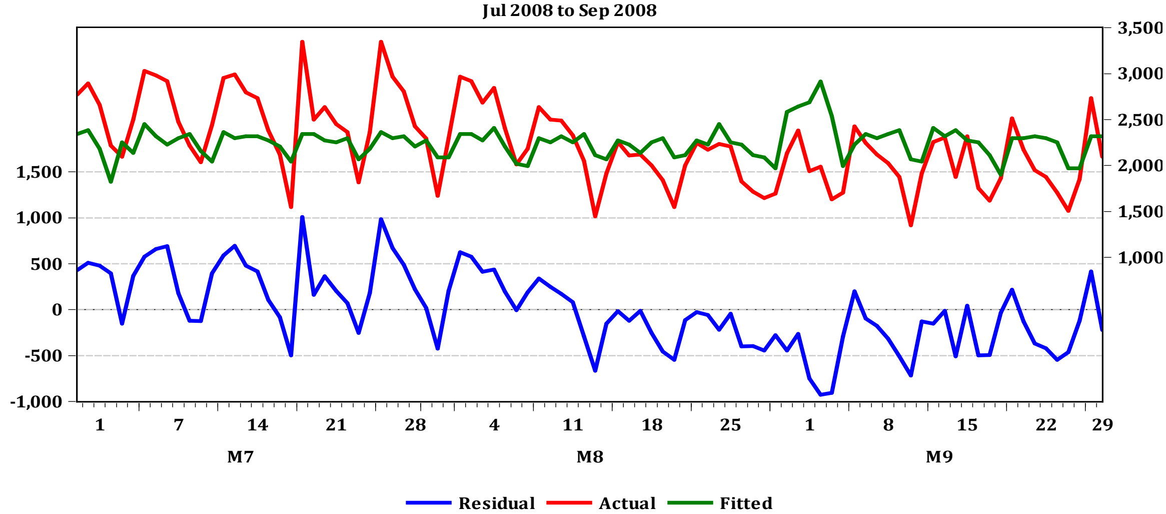
Graph 2B--Quarter 4
Regressand-All Visits and Regressor-"charleston sc" (travel)
Oct 2008 to Dec 2008



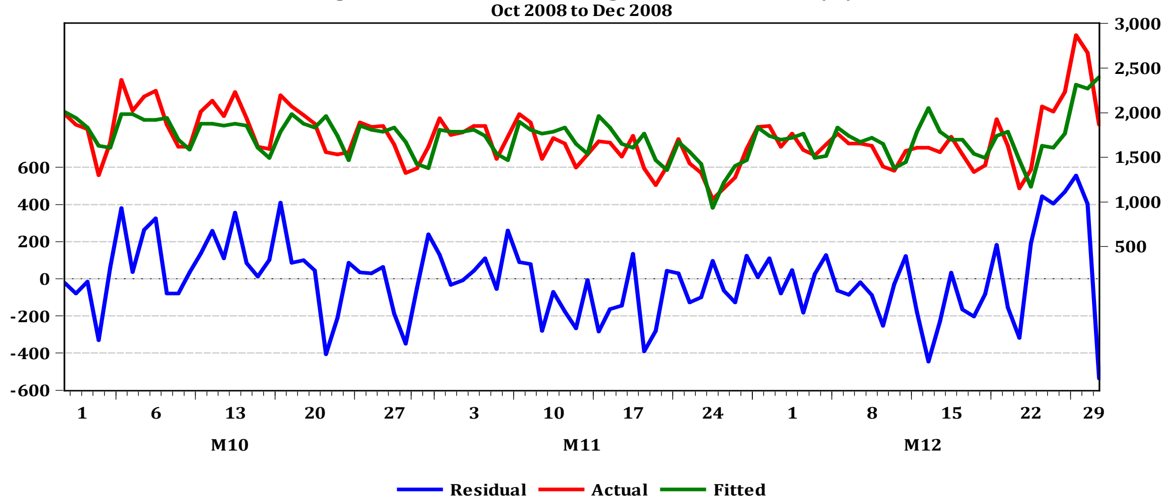
Graph 2C--Weekly Data
Regressand-All Visits and Regressor-"charleston sc" (travel)
Jan 2008 to March 2009



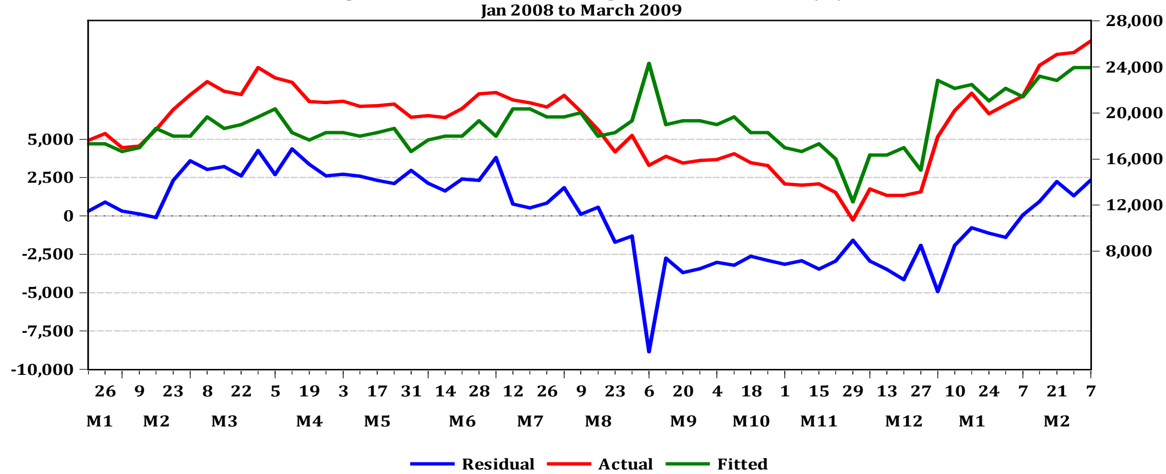
Graph 3A--Quarter 3
Regressand-Search Traffic and Regressor-"charleston sc" (all)
Jul 2008 to Sep 2008



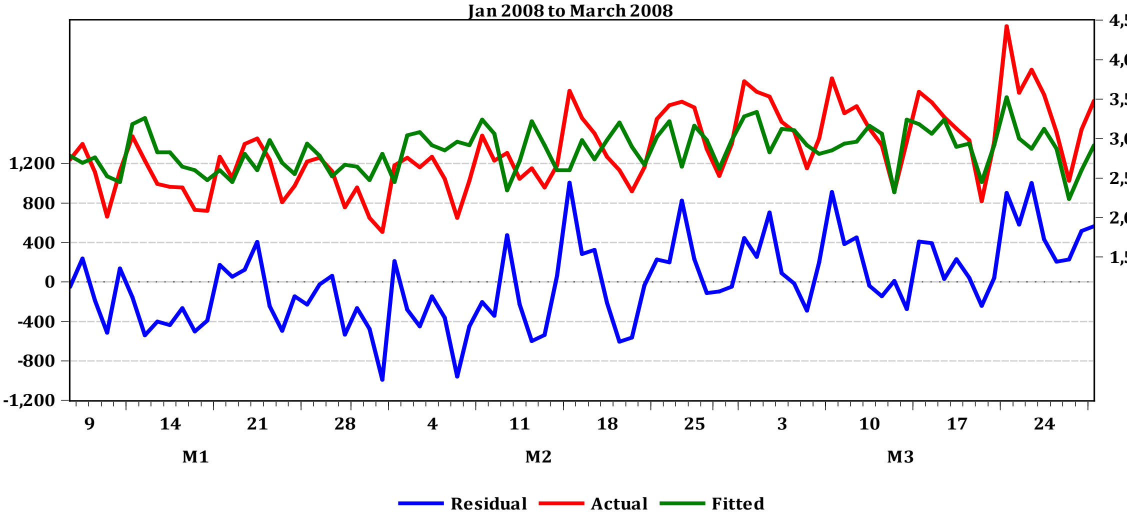
Graph 3B--Quarter 4
Regressand-Search Traffic and Regressor-"charleston sc" (all)
Oct 2008 to Dec 2008



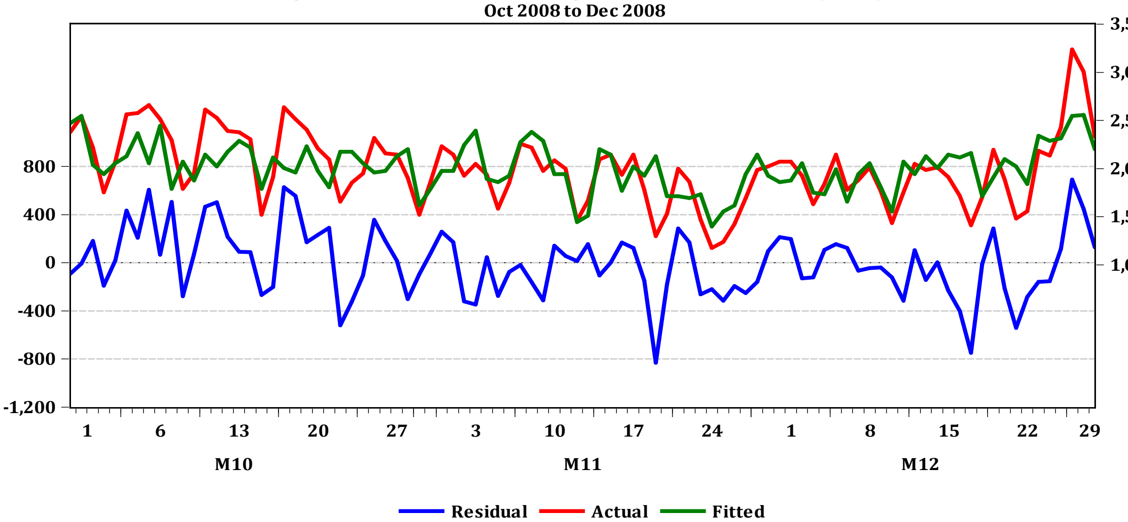
Graph 3C--Weekly Data
Regressand-Search Traffic and Regressor-"charleston sc" (all)
Jan 2008 to March 2009



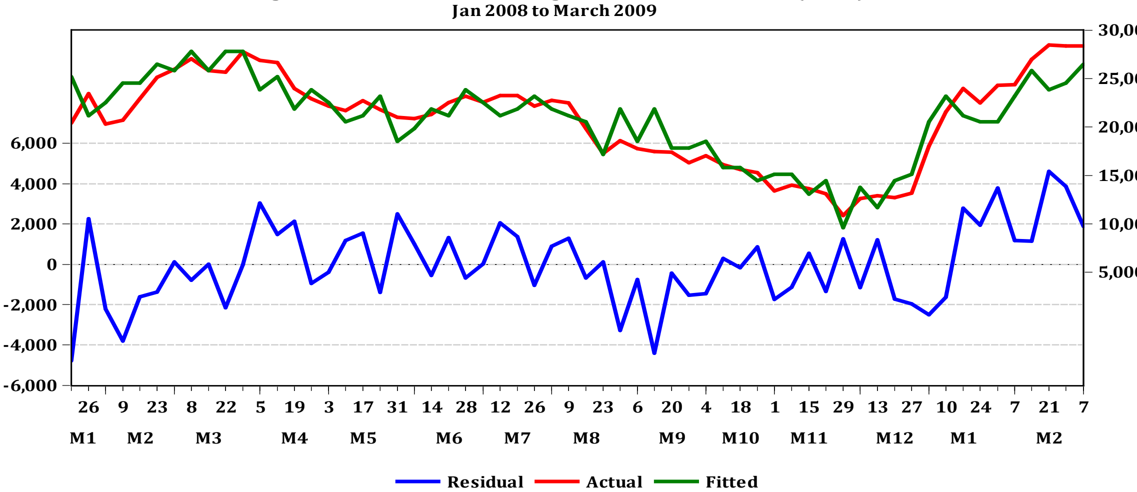
Graph 4A--Quarter 1
Regressand-Nonlocal Visits and Regressor-"charleston hotels" (travel)
Jan 2008 to March 2008



Graph 4B--Quarter 4
Regressand-Nonlocal Visits and Regressor-"charleston hotels" (travel)
Oct 2008 to Dec 2008



Graph 4C--Weekly Data
Regressand-Nonlocal Visits and Regressor-"charleston hotels" (travel)
Jan 2008 to March 2009



Appendix

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	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-Stat	PV	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-Stat	PV	$\hat{\lambda}$	Psu-R^2
Q1	ch hotels	0.573	0.094	6.126	**0.00**	1.56	0.30	0.528	0.096	5.530	**0.00**	1.03	0.26	0.487	0.096	5.068	**0.00**	0.75	0.22
Q2	(all)	0.604	0.090	6.705	**0.00**	1.53	0.32	0.586	0.088	6.648	**0.00**	0.90	0.30	0.536	0.099	5.400	**0.00**	0.77	0.22
Q3		0.828	0.095	8.748	**0.00**	1.41	0.44	0.729	0.098	7.426	**0.00**	0.96	0.35	0.735	0.094	7.805	**0.00**	0.57	0.36
Q4		0.543	0.059	9.172	**0.00**	0.83	0.45	0.523	0.058	9.069	**0.00**	0.51	0.43	0.512	0.056	9.147	**0.00**	0.32	0.41
Q5		0.727	0.103	7.051	**0.00**	1.52	0.38	0.654	0.097	6.711	**0.00**	0.82	0.35	0.692	0.100	6.950	**0.00**	0.60	0.35
W		0.746	0.053	14.183	**0.00**	3.22	0.78	0.825	0.069	12.007	**0.00**	3.66	0.72	0.714	0.061	11.781	**0.00**	3.29	0.71
Q1	ch restaurants	0.107	0.082	1.317	0.19	2.22	0.02	0.065	0.081	0.810	0.42	1.38	0.01	0.037	0.080	0.467	0.64	0.96	0.00
Q2	(all)	-0.114	0.085	-1.339	0.18	2.21	0.02	-0.103	0.084	-1.234	0.22	1.30	0.01	-0.054	0.090	-0.599	0.55	1.00	0.00
Q3		0.089	0.093	0.953	0.34	2.56	0.01	0.075	0.090	0.831	0.41	1.53	0.01	0.125	0.088	1.426	0.15	0.94	0.02
Q4		0.192	0.079	2.432	**0.02**	1.51	0.06	0.165	0.077	2.145	**0.03**	0.92	0.04	0.208	0.075	2.788	**0.01**	0.58	0.07
Q5		0.012	0.080	0.151	0.88	2.54	0.00	0.000	0.074	0.001	1.00	1.32	0.00	0.016	0.077	0.204	0.84	1.00	0.00
W		0.806	0.138	5.820	**0.00**	9.41	0.36	0.921	0.159	5.799	**0.00**	8.48	0.36	0.782	0.143	5.477	**0.00**	7.64	0.33
Q1	ch sc	1.290	0.171	7.544	**0.00**	1.33	0.39	1.239	0.173	7.179	**0.00**	0.86	0.36	1.086	0.180	6.033	**0.00**	0.67	0.28
Q2	(all)	1.473	0.139	10.570	**0.00**	1.02	0.54	1.462	0.133	10.975	**0.00**	0.57	0.54	1.264	0.169	7.496	**0.00**	0.63	0.36
Q3		0.870	0.185	4.700	**0.00**	2.10	0.19	0.932	0.173	5.390	**0.00**	1.16	0.22	0.695	0.182	3.818	**0.00**	0.83	0.12
Q4		1.018	0.084	12.175	**0.00**	0.60	0.59	1.053	0.071	14.754	**0.00**	0.28	0.64	0.900	0.086	10.424	**0.00**	0.28	0.48
Q5		1.214	0.178	6.805	**0.00**	1.56	0.37	1.165	0.161	7.219	**0.00**	0.77	0.38	0.977	0.186	5.248	**0.00**	0.73	0.25
W		1.353	0.239	5.665	**0.00**	9.41	0.36	1.420	0.285	4.976	**0.00**	9.36	0.29	1.450	0.230	6.297	**0.00**	6.82	0.39
Q1	ch tourism	-0.215	0.097	-2.231	**0.03**	2.64	0.08	-0.198	0.092	-2.159	**0.03**	1.50	0.07	-0.220	0.088	-2.510	**0.01**	0.98	0.09
Q2	(all)	-0.139	0.075	-1.851	0.06	2.45	0.04	-0.140	0.073	-1.904	0.06	1.42	0.04	-0.179	0.079	-2.277	**0.02**	1.10	0.05
Q3		-0.187	0.145	-1.283	0.20	2.78	0.02	-0.175	0.140	-1.245	0.21	1.63	0.02	-0.222	0.136	-1.630	0.10	0.99	0.03
Q4		0.550	0.106	5.210	**0.00**	1.33	0.22	0.522	0.102	5.102	**0.00**	0.81	0.21	0.525	0.100	5.273	**0.00**	0.51	0.21
Q5		0.044	0.124	0.355	0.72	2.71	0.00	0.018	0.115	0.158	0.87	1.41	0.00	0.106	0.121	0.877	0.38	1.10	0.01
W		0.457	0.042	10.777	**0.00**	5.47	0.63	0.540	0.045	11.983	**0.00**	5.23	0.61	0.417	0.049	8.448	**0.00**	5.33	0.53
Q1	ch travel	-0.126	0.096	-1.315	0.19	2.22	0.02	-0.128	0.094	-1.356	0.18	1.35	0.02	-0.095	0.093	-1.024	0.31	0.95	0.01
Q2	(all)	0.064	0.083	0.773	0.44	2.23	0.01	0.046	0.081	0.561	0.57	1.31	0.00	0.083	0.087	0.955	0.34	1.01	0.01
Q3		0.112	0.090	1.248	0.21	2.60	0.02	0.059	0.087	0.673	0.50	1.55	0.01	0.142	0.084	1.686	0.09	0.94	0.03
Q4		0.239	0.096	2.477	**0.01**	1.28	0.32	0.216	0.093	2.319	**0.02**	0.76	0.31	0.237	0.094	2.523	**0.01**	0.52	0.28
Q5		0.151	0.084	1.797	0.07	2.33	0.10	0.148	0.078	1.911	0.06	1.22	0.04	0.129	0.082	1.573	0.12	0.96	0.03
W		0.750	0.075	9.986	**0.00**	5.01	0.66	0.843	0.089	9.419	**0.00**	3.84	0.70	0.693	0.084	8.203	**0.00**	5.20	0.54
Q1	ch hotels	0.485	0.089	5.439	**0.00**	1.66	0.26	0.458	0.090	5.117	**0.00**	1.06	0.23	0.427	0.089	4.773	**0.00**	0.76	0.20
Q2	(travel)	0.526	0.077	6.867	**0.00**	1.49	0.33	0.516	0.074	6.940	**0.00**	0.86	0.32	0.516	0.082	6.307	**0.00**	0.70	0.28
Q3		0.677	0.106	6.379	**0.00**	1.79	0.30	0.628	0.104	6.020	**0.00**	1.09	0.27	0.594	0.103	5.784	**0.00**	0.69	0.24
Q4		0.474	0.059	7.973	**0.00**	0.94	0.39	0.456	0.058	7.889	**0.00**	0.57	0.37	0.440	0.057	7.739	**0.00**	0.37	0.34

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	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-Stat	PV	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-Stat	PV	$\hat{\lambda}$	Psu-R^2
Q5	ch hotels	0.546	0.094	5.789	**0.00**	1.77	0.30	0.478	0.090	5.336	**0.00**	0.96	0.25	0.466	0.095	4.910	**0.00**	0.76	0.22
W	(travel)	0.765	0.046	16.692	**0.00**	5.39	0.63	0.875	0.055	15.897	**0.00**	2.36	0.81	0.729	0.056	13.079	**0.00**	2.81	0.74
Q1	ch sc	0.640	0.145	4.431	**0.00**	1.81	0.18	0.600	0.145	4.145	**0.00**	1.15	0.16	0.484	0.147	3.289	**0.00**	0.86	0.11
Q2	(travel)	0.872	0.105	8.273	**0.00**	1.29	0.41	0.859	0.102	8.399	**0.00**	0.74	0.41	0.896	0.111	8.084	**0.00**	0.59	0.39
Q3		0.745	0.149	5.006	**0.00**	2.06	0.21	0.760	0.141	5.382	**0.00**	1.17	0.23	0.655	0.143	4.582	**0.00**	0.78	0.17
Q4		0.707	0.062	11.419	**0.00**	0.66	0.55	0.674	0.062	10.958	**0.00**	0.42	0.52	0.667	0.058	11.431	**0.00**	0.25	0.51
Q5		0.910	0.129	7.076	**0.00**	1.52	0.39	0.850	0.118	7.176	**0.00**	0.78	0.38	0.837	0.127	6.600	**0.00**	0.63	0.33
W		0.969	0.095	10.151	**0.00**	2.46	0.82	1.106	0.110	10.028	**0.00**	4.88	0.63	0.976	0.096	10.156	**0.00**	4.17	0.63
Q1	ch sc	0.010	0.005	1.947	0.05	2.19	0.04	0.008	0.005	1.552	0.12	1.36	0.03	0.006	0.005	1.166	0.24	0.95	0.01
Q2	(rank)	-0.193	0.141	-1.374	0.17	2.42	0.02	-0.179	0.136	-1.322	0.19	1.40	0.02	-0.319	0.155	-2.055	**0.04**	1.08	0.05
W		0.031	0.018	1.684	0.09	7.98	0.07	0.042	0.023	1.813	0.07	9.30	0.08	0.020	0.021	0.970	0.33	7.49	0.02
Q1	ch hotel	0.000	0.001	-0.327	0.74	4.15	0.00	-0.001	0.001	-0.382	0.70	2.50	0.00	-0.001	0.001	-0.547	0.58	1.59	0.01
Q2	(rank)	0.002	0.006	0.352	0.72	7.26	0.01	0.002	0.005	0.392	0.70	3.78	0.01	0.001	0.006	0.228	0.82	3.47	0.00
Q3		0.002	0.001	1.681	0.09	4.37	0.06	0.002	0.001	1.655	0.10	2.50	0.05	0.001	0.001	1.226	0.22	1.62	0.03
Q4		0.000	0.001	0.166	0.87	1.87	0.00	0.000	0.001	0.397	0.69	1.09	0.00	0.000	0.001	0.043	0.97	0.64	0.00
Q5		0.020	0.006	3.216	**0.00**	2.24	0.11	0.018	0.006	3.117	**0.00**	1.17	0.10	0.021	0.006	3.551	**0.00**	0.85	0.13
W		-0.001	0.001	-0.914	0.36	21.97	0.02	-0.002	0.001	-1.230	0.22	18.61	0.03	-0.001	0.001	-0.956	0.34	16.90	0.02
Q1	ch restaurants	0.004	0.002	1.964	**0.05**	2.19	0.04	0.003	0.002	1.562	0.12	1.36	0.03	0.004	0.002	1.926	0.05	0.93	0.04
Q2	(rank)	-0.006	0.002	-2.702	**0.01**	2.09	0.07	-0.006	0.002	-2.681	**0.01**	1.22	0.07	-0.009	0.002	-3.839	**0.00**	0.88	0.12
Q3		0.002	0.003	0.784	0.43	2.63	0.01	0.003	0.003	0.887	0.37	1.54	0.01	0.001	0.003	0.278	0.78	0.97	0.00
Q4		0.006	0.004	1.545	0.12	0.61	0.59	0.007	0.004	1.744	0.08	0.95	0.03	0.004	0.004	0.934	0.35	0.63	0.01
Q5		0.003	0.004	0.727	0.47	2.52	0.01	0.002	0.004	0.634	0.53	1.31	0.00	0.003	0.004	0.855	0.39	0.99	0.01
W		0.004	0.003	1.428	0.15	14.31	0.03	0.002	0.003	0.691	0.49	13.08	0.01	0.003	0.003	1.001	0.32	11.29	0.02
Q1	ch tourism	-0.031	0.008	-3.730	**0.00**	1.91	0.14	-0.027	0.008	-3.240	**0.00**	1.22	0.11	-0.026	0.008	-3.166	**0.00**	0.85	0.10
Q3	(rank)	-0.203	0.045	-4.505	**0.00**	2.10	0.18	-0.171	0.044	-3.834	**0.00**	1.32	0.13	-0.174	0.043	-4.011	**0.00**	0.81	0.14
Q4		-0.092	0.040	-2.316	**0.02**	0.61	0.60	-0.101	0.038	-2.627	**0.01**	0.91	0.06	-0.086	0.038	-2.258	**0.02**	0.61	0.05
W		-0.015	0.027	-0.537	0.59	14.45	0.00	0.001	0.031	0.020	0.98	13.07	0.00	-0.033	0.028	-1.155	0.25	11.10	0.02
Q1	ch travel	-0.026	0.009	-2.763	**0.01**	2.06	0.08	-0.022	0.009	-2.410	**0.02**	1.29	0.06	-0.023	0.009	-2.495	**0.01**	0.89	0.07
Q2	(rank)	-0.580	0.233	-2.490	**0.01**	2.08	0.07	-0.547	0.225	-2.435	**0.01**	1.22	0.07	-0.647	0.251	-2.577	**0.01**	0.94	0.08
Q3		0.193	0.231	0.834	0.40	2.62	0.01	0.240	0.227	1.053	0.29	1.52	0.01	0.127	0.212	0.597	0.55	0.96	0.00
Q4		0.246	0.153	1.610	0.11	1.57	0.03	0.217	0.146	1.484	0.14	0.95	0.02	0.246	0.146	1.681	0.09	0.61	0.03
Q5		-0.197	0.031	-6.404	**0.00**	1.65	0.34	-0.173	0.029	-5.866	**0.00**	0.91	0.29	-0.184	0.030	-6.109	**0.00**	0.67	0.30
W		-0.041	0.034	-1.206	0.23	14.05	0.02	-0.041	0.039	-1.029	0.30	12.76	0.02	-0.054	0.035	-1.531	0.13	10.76	0.04

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	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-stat	PV	$\hat{\lambda}$	Psu-R^2
Q1	ch hotels	0.615	0.093	6.623	**0.00**	0.70	0.32	0.538	0.101	5.318	**0.00**	1.00	0.24
Q2	(all)	0.554	0.088	6.308	**0.00**	0.70	0.28	0.650	0.096	6.800	**0.00**	0.90	0.32
Q3		0.806	0.091	8.807	**0.00**	0.63	0.42	0.849	0.104	8.176	**0.00**	0.88	0.40
Q4		0.544	0.056	9.647	**0.00**	0.38	0.45	0.542	0.066	8.168	**0.00**	0.52	0.38
Q5		0.773	0.099	7.823	**0.00**	0.66	0.41	0.686	0.113	6.083	**0.00**	0.96	0.31
W		0.918	0.067	13.724	**0.00**	4.17	0.77	-0.047	0.065	-0.727	0.47	0.95	0.01
Q1	ch restaurants	0.149	0.083	1.801	0.07	1.03	0.03	0.073	0.085	0.860	0.39	1.31	0.01
Q2	(all)	-0.127	0.082	-1.559	0.12	0.96	0.02	-0.103	0.091	-1.126	0.26	1.32	0.01
Q3		0.111	0.090	1.227	0.22	1.17	0.02	0.068	0.099	0.691	0.49	1.14	0.00
Q4		0.194	0.077	2.527	**0.01**	0.72	0.06	0.190	0.084	2.247	**0.02**	0.86	0.05
Q5		0.020	0.080	0.246	0.81	1.21	0.00	0.005	0.083	0.062	0.95	1.44	0.00
W		0.950	0.171	5.557	**0.00**	12.09	0.34	0.039	0.112	0.352	0.72	0.96	0.00
Q1	ch sc	1.423	0.164	8.664	**0.00**	0.56	0.44	1.179	0.190	6.200	**0.00**	0.90	0.30
Q2	(all)	1.414	0.132	10.693	**0.00**	0.44	0.52	1.527	0.154	9.923	**0.00**	0.64	0.50
Q3		0.738	0.185	3.985	**0.00**	1.01	0.14	0.991	0.192	5.158	**0.00**	1.18	0.21
Q4		1.030	0.077	13.439	**0.00**	0.25	0.59	1.005	0.098	10.216	**0.00**	0.41	0.49
Q5		1.260	0.174	7.228	**0.00**	0.70	0.38	1.172	0.192	6.100	**0.00**	0.96	0.31
W		1.598	0.295	5.417	**0.00**	12.30	0.33	0.073	0.188	0.386	0.70	0.96	0.00
Q1	ch tourism	-0.216	0.103	-2.086	**0.04**	1.38	0.07	-0.215	0.095	-2.255	**0.02**	1.41	0.08
Q2	(all)	-0.100	0.072	-1.386	0.17	1.08	0.02	-0.175	0.080	-2.194	**0.03**	1.43	0.05
Q3		-0.198	0.141	-1.400	0.16	1.26	0.02	-0.176	0.154	-1.142	0.25	1.62	0.01
Q4		0.536	0.104	5.157	**0.00**	0.65	0.21	0.564	0.112	5.027	**0.00**	0.75	0.20
Q5		0.066	0.127	0.521	0.60	1.34	0.00	0.024	0.126	0.191	0.85	1.49	0.00
W		0.557	0.052	10.785	**0.00**	7.09	0.62	-0.066	0.045	-1.470	0.14	0.96	0.00
Q1	ch travel	-0.116	0.099	-1.176	0.24	1.06	0.02	-0.135	0.099	-1.362	0.17	1.29	0.02
Q2	(all)	0.072	0.079	0.913	0.36	0.97	0.01	0.056	0.088	0.639	0.52	1.32	0.00
Q3		0.096	0.088	1.093	0.27	1.19	0.01	0.128	0.095	1.341	0.18	1.52	0.02
Q4		0.203	0.095	2.132	**0.03**	0.73	0.05	0.275	0.102	2.683	**0.01**	0.83	0.07
Q5		0.150	0.084	1.779	0.08	1.13	0.04	0.152	0.087	1.742	0.08	1.34	0.04
W		0.909	0.094	9.642	**0.00**	6.22	0.66	-0.042	0.083	-0.511	0.61	0.92	0.03
Q1	ch hotels	0.509	0.090	5.648	**0.00**	0.77	0.26	0.465	0.095	4.904	**0.00**	1.02	0.21
Q2	(travel)	0.493	0.074	6.670	**0.00**	0.67	0.30	0.556	0.082	6.776	**0.00**	0.89	0.31
Q3		0.657	0.103	6.368	**0.00**	0.81	0.29	0.695	0.114	6.089	**0.00**	1.08	0.27
Q4		0.475	0.057	8.324	**0.00**	0.41	0.38	0.473	0.066	7.192	**0.00**	0.57	0.33

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	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-stat	PV	$\hat{\lambda}$	Psu-R^2
Q5	ch hotels	0.589	0.091	6.485	**0.00**	0.78	0.33	0.507	0.102	4.973	**0.00**	1.09	0.23
W	(travel)	0.946	0.057	16.526	**0.00**	3.08	0.83	-0.066	0.064	-1.044	0.30	0.94	0.02
Q1	ch sc	0.702	0.145	4.848	**0.00**	0.82	0.20	0.590	0.154	3.834	**0.00**	1.12	0.14
Q2	(travel)	0.814	0.103	7.936	**0.00**	0.59	0.38	0.924	0.113	8.198	**0.00**	0.77	0.40
Q3		0.704	0.146	4.832	**0.00**	0.95	0.19	0.783	0.158	4.943	**0.00**	1.21	0.20
Q4		0.702	0.059	11.902	**0.00**	0.30	0.54	0.712	0.071	10.089	**0.00**	0.42	0.48
Q5		0.917	0.128	7.157	**0.00**	0.72	0.37	0.903	0.137	6.603	**0.00**	0.91	0.34
W		1.162	0.120	9.704	**0.00**	7.08	0.61	-0.026	0.099	-0.259	0.80	0.96	0.00
Q1	ch sc	0.011	0.005	2.077	**0.04**	1.03	0.04	0.009	0.005	1.735	0.08	1.29	0.03
Q2	(rank)	-0.161	0.133	-1.213	0.23	1.06	0.02	-0.224	0.014	-16.087	**0.00**	1.45	0.04
W		0.037	0.023	1.602	0.11	11.05	0.06	0.037	0.023	1.602	0.11	11.05	0.06
Q1	ch hotel	0.000	0.001	0.007	0.99	2.00	0.00	-0.001	0.001	-0.592	0.55	2.36	0.01
Q2	(rank)	0.000	0.006	0.083	0.93	3.85	0.00	0.003	0.006	0.608	0.54	3.64	0.02
Q3		0.002	0.001	1.485	0.14	2.03	0.04	0.002	0.001	1.807	0.07	2.48	0.06
Q4		0.000	0.001	0.282	0.78	0.88	0.00	0.000	0.001	0.097	0.92	1.09	0.00
Q5		0.020	0.006	3.264	**0.00**	1.06	0.11	0.019	0.006	3.044	**0.00**	1.28	0.10
W		-0.002	0.002	-1.000	0.32	27.21	0.02	-0.002	0.002	-1.000	0.32	27.21	0.02
Q1	ch restaurants	0.004	0.002	2.014	**0.04**	1.05	0.04	0.004	0.002	1.817	0.07	1.29	0.04
Q2	(rank)	-0.005	0.002	-2.284	**0.02**	0.94	0.05	-0.007	0.002	-2.984	**0.00**	1.22	0.08
Q3		0.002	0.003	0.698	0.49	1.19	0.01	0.003	0.003	0.831	0.41	1.53	0.01
Q4		0.006	0.004	1.453	0.15	0.76	0.02	0.007	0.004	1.572	0.12	0.89	0.02
Q5		0.003	0.004	0.622	0.53	1.20	0.00	0.003	0.004	0.792	0.43	1.42	0.01
W		0.004	0.003	1.232	0.22	17.90	0.02	0.004	0.003	1.232	0.22	17.90	0.02
Q1	ch tourism	-0.030	0.009	-3.539	**0.00**	0.92	0.13	-0.032	0.009	-3.663	**0.00**	1.12	0.14
Q3	(rank)	-0.211	0.043	-4.899	**0.00**	0.92	0.20	-0.195	0.049	-4.013	**0.00**	1.28	0.14
Q4		-0.095	0.039	-2.443	**0.01**	0.74	0.06	-0.090	0.043	-2.109	**0.04**	0.88	0.04
W		-0.015	0.033	-0.444	0.66	17.95	0.00	-0.015	0.033	-0.444	0.66	17.95	0.00
Q1	ch travel	-0.024	0.010	-2.440	**0.01**	1.00	0.06	-0.028	0.010	-2.878	**0.00**	1.20	0.09
Q2	(rank)	-0.516	0.217	-2.380	**0.02**	0.92	0.06	-0.641	0.255	-2.510	**0.01**	1.23	0.07
Q3		0.274	0.233	1.177	0.24	1.18	0.02	0.123	0.237	0.517	0.60	1.53	0.00
Q4		0.165	0.145	1.137	0.26	0.77	0.01	0.336	0.169	1.995	**0.05**	0.87	0.04
Q5		-0.210	0.030	-7.073	**0.00**	0.72	0.36	-0.186	0.034	-5.553	**0.00**	1.03	0.27
W		-0.052	0.043	-1.232	0.22	17.39	0.03	-0.052	0.043	-1.232	0.22	17.39	0.03

Table 3--Bivariate Poisson Regressions for Regressand--All Visits													
Time	**Regressors**	**Coef.**	**S.E.**	**Z-Stat**	**PV**	**Regressors**	**Coef.**	**S.E.**	**Z-Stat**	**PV**	$\hat{\lambda}$	**Psu-R^2**
Q1	ch sc (all)	1.265	0.170	7.439	**0.00**	ch (rank)	0.007	0.004	1.790	0.07	1.30	0.41
Q2		1.478	0.146	10.107	**0.00**		-0.199	0.095	-2.100	**0.04**	1.10	0.56
W		0.787	0.239	3.289	**0.00**		0.036	0.016	2.219	**0.03**	6.23	0.27
Q1	ch hotels (all)	0.607	0.142	4.279	**0.00**	ch hotel (rank)	0.000	0.001	0.416	0.68	2.92	0.30
Q2		0.303	0.281	1.079	0.28		0.004	0.006	0.615	0.54	7.31	0.07
Q3		0.779	0.140	5.571	**0.00**		0.001	0.001	1.243	0.21	2.78	0.42
Q4		0.522	0.077	6.758	**0.00**		-0.001	0.001	-0.955	0.34	1.12	0.39
Q5		0.693	0.118	5.887	**0.00**		0.004	0.006	0.613	0.54	1.53	0.39
W		0.745	0.047	15.910	**0.00**		-0.001	0.000	-2.699	**0.01**	3.09	0.86
Q1	ch restaurants (all)	0.088	0.082	1.076	0.28	ch restaurants (rank)	0.004	0.002	1.805	0.07	2.17	0.06
Q2		-0.128	0.083	-1.542	0.12		-0.006	0.002	-2.802	**0.01**	2.07	0.09
Q3		0.090	0.093	0.964	0.34		0.002	0.003	0.802	0.42	2.62	0.02
Q4		0.206	0.078	2.644	**0.01**		0.007	0.004	1.869	0.06	1.47	0.09
Q5		0.004	0.081	0.053	0.96		0.003	0.004	0.709	0.48	2.55	0.01
W		0.790	0.141	5.590	**0.00**		0.002	0.002	0.863	0.39	9.56	0.36
Q1	ch tourism (all)	-0.165	0.087	-1.904	0.06	ch tourism (rank)	-0.142	0.034	-4.136	**0.00**	2.05	0.28
Q3		-0.182	0.134	-1.363	0.17		-0.200	0.045	-4.407	**0.00**	2.24	0.20
Q4		0.512	0.117	4.365	**0.00**		-0.032	0.042	-0.762	0.45	1.34	0.23
W		0.456	0.043	10.660	**0.00**		-0.006	0.016	-0.347	0.73	5.09	0.66
Q1	ch travel (all)	-0.150	0.092	-1.641	0.10	ch travel (rank)	-0.027	0.009	-2.928	**0.00**	1.41	0.11
Q2		0.039	0.081	0.484	0.63		-0.569	0.235	-2.422	**0.02**	2.10	0.07
Q3		0.104	0.091	1.148	0.25		0.162	0.233	0.695	0.49	2.65	0.02
Q4		0.237	0.094	2.523	**0.01**		0.269	0.135	1.988	***0.05***	1.43	0.11
Q5		0.039	0.072	0.539	0.59		-0.186	0.032	-5.896	**0.00**	1.62	0.33
W		0.756	0.072	10.501	**0.00**		-0.045	0.019	-2.376	**0.02**	5.02	0.67

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	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-Stat	PV	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-Stat	PV	$\hat{\lambda}$	Psu-R^2
D	CH hotels	0.655	0.088	7.560	**0.00**	1.37	0.38	0.604	0.087	7.077	**0.00**	0.84	0.34	0.592	0.089	6.874	**0.00**	0.60	0.32
W	(all)	0.746	0.053	14.183	**0.00**	3.22	0.78	0.825	0.069	12.007	**0.00**	3.66	0.72	0.714	0.061	11.781	**0.00**	3.29	0.71
D	CH hotels	0.541	0.085	6.489	**0.00**	1.53	0.31	0.507	0.083	6.260	**0.00**	0.91	0.29	0.489	0.085	5.902	**0.00**	0.66	0.26
W	(travel)	0.765	0.046	16.692	**0.00**	5.39	0.63	0.875	0.055	15.897	**0.00**	2.36	0.81	0.729	0.056	13.079	**0.00**	2.81	0.74
D	CH rests	0.057	0.084	0.703	0.32	2.21	0.02	0.040	0.081	0.510	0.41	1.29	0.01	0.066	0.082	0.857	0.44	0.90	0.02
W	(all)	0.806	0.138	5.820	**0.00**	9.41	0.36	0.921	0.159	5.799	**0.00**	8.48	0.36	0.782	0.143	5.477	**0.00**	7.64	0.33
D	CH sc	1.173	0.151	8.359	**0.00**	1.32	0.42	1.170	0.142	9.104	**0.00**	0.73	0.43	0.984	0.161	6.604	**0.00**	0.63	0.30
W	(all)	1.353	0.239	5.665	**0.00**	9.41	0.36	1.420	0.285	4.976	**0.00**	9.36	0.29	1.450	0.230	6.297	**0.00**	6.82	0.39
D	CH sc	0.775	0.118	7.241	**0.00**	1.47	0.35	0.749	0.114	7.212	**0.00**	0.85	0.34	0.708	0.117	6.798	**0.00**	0.62	0.30
W	(travel)	0.969	0.095	10.151	**0.00**	2.46	0.82	1.106	0.110	10.028	**0.00**	4.88	0.63	0.976	0.096	10.156	**0.00**	4.17	0.63
D	CH tourism	0.011	0.109	0.040	0.20	2.38	0.07	0.005	0.105	-0.010	0.24	1.35	0.07	0.002	0.105	-0.053	0.10	0.93	0.08
W	(all)	0.457	0.042	10.777	**0.00**	5.47	0.63	0.540	0.045	11.983	**0.00**	5.23	0.61	0.417	0.049	8.448	**0.00**	5.33	0.53
D	CH travel	0.088	0.090	0.996	0.19	2.13	0.09	0.068	0.087	0.822	0.27	1.24	0.08	0.099	0.088	1.142	0.17	0.88	0.07
W	(all)	0.750	0.075	9.986	**0.00**	5.01	0.66	0.843	0.089	9.419	**0.00**	3.84	0.70	0.693	0.084	8.203	**0.00**	5.20	0.54

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	$\hat{\lambda}$	Psu-R^2	Coef.	S.E.	Z-stat	PV	$\hat{\lambda}$	Psu-R^2
D	CH hotels	0.659	0.085	7.842	**0.00**	0.61	0.38	0.653	0.096	6.909	**0.00**	0.85	0.33
W	(all)	0.918	0.067	13.724	**0.00**	4.17	0.77	-0.047	0.065	-0.727	0.47	0.95	0.01
D	CH hotels	0.545	0.083	6.699	**0.00**	0.69	0.31	0.539	0.092	5.987	**0.00**	0.93	0.27
W	(travel)	0.946	0.057	16.526	**0.00**	3.08	0.83	-0.066	0.064	-1.044	0.30	0.94	0.02
D	CH rests	0.069	0.082	0.848	0.25	1.02	0.03	0.047	0.088	0.547	**0.42**	1.21	0.01
W	(all)	0.950	0.171	5.557	**0.00**	12.09	0.34	0.039	0.112	0.352	0.72	0.96	0.00
D	CH sc	1.173	0.147	8.802	**0.00**	0.59	0.41	1.175	0.165	7.520	**0.00**	0.82	0.36
W	(all)	1.598	0.295	5.417	**0.00**	12.30	0.33	0.073	0.188	0.386	0.70	0.96	0.00
D	CH sc	0.768	0.116	7.335	**0.00**	0.68	0.34	0.782	0.126	6.734	**0.00**	0.89	0.31
W	(travel)	1.162	0.120	9.704	**0.00**	7.08	0.61	-0.026	0.099	-0.259	0.80	0.96	0.00
D	CH tourism	0.018	0.110	0.161	0.19	1.14	0.06	0.004	0.114	-0.075	0.23	1.34	0.07
W	(all)	0.557	0.052	10.785	**0.00**	7.09	0.62	-0.066	0.045	-1.470	0.14	0.96	0.00
D	CH travel	0.081	0.089	0.948	0.20	1.02	0.02	0.095	0.094	1.008	0.19	1.26	0.03
W	(all)	0.909	0.094	9.642	**0.00**	6.22	0.66	-0.042	0.083	-0.511	0.61	0.92	0.03