THE EFFECT OF GOOGLE TRENDS DATA ON RESIDENTIAL ELECTRICITY DEMAND

Sungjun Park, Hanyang University, Phone +82 2 2220 4479, E-mail: sjpark13@hanyang.ac.kr
Young Yoon, Hanyang University, Phone +82 2 2220 4479, E-mail: youngyoon@hanyang.ac.kr
Jinsoo Kim, Hanyang University, Phone + 82 2 2220 2241, E-mail: jinsookim@hanyang.ac.kr

Overview

Real-time search becomes possible with the increasing penetration of computers and smartphones, and it is becoming popular to obtain information through Internet search. However, traditional economic models still rely on statistics gathered by government data, annual or quarterly reports and financial statements. This limits the usefulness of forecasting, especially novel forecasts, because it allows you to construct an economic model using data from one to three month ago (Wu and Brynjolfsson, 2009). However, with related technologies like Google Trends, we can get an information with no time delay.

Google Trends provides the search frequency for keywords based on worldwide Google searches in real time. We can use information that has not been announced yet with Google Trends, predicting the present is available. In fact, Google Trends has an advantageous aspect to “predict the present”, rather than “predict the future” (Choi and Varian, 2012). Therefore, if the search frequency has more influence than the existing explanatory variables, the researcher should consider constructing a model based on the Internet-based search frequency.

In this study, we analyse the correlation between Google Trends data and US residential electricity demand using a panel analysis. HDD (heating degree days) and CDD (cooling degree days) are used as explanatory variables when constructing the panel analysis to assess the effect of Google Trends. We consider three Google Trends keywords those are “renewable,” “temperature,” and “weather forecast.” In case of "renewable" keywords, there is no research that analyses renewable energy-related variables in a short term because the correlation between US renewable energy and US residential electricity demand cannot be derived from existing data. In addition, due to the problem of personal information and the cost of measurement, it is difficult to identify electricity demand unlike other major sectors such as commercial and industrial sectors (Swan and Ugursal, 2009). Therefore, we use Google Trends to construct models that include interest in renewable energy and other variables and analyse the correlation with residential electricity demand.

Methods

We analyse the correlation between Google Trends data and residential electricity demand through a panel analysis. Temperature has been considered as the most important variable to account for electricity demand and temperature sensitivity to electricity demand is an increasing trend, becoming an important factor in estimating demand (Besse and Fouquau, 2008; Hekkenberg *at al*., 2009). Therefore, HDD and CDD were used instead of temperature variables.

$$E_{i,t} = \alpha + \beta_1 P_{i,t} + \beta_2 HDD_{i,t} + \beta_3 CDD_{i,t} + \beta_4 RE_{i,t} + \beta_5 WF_{i,t} + \beta_6 T_{i,t} + \mu_i + e_{i,t}$$

where $E_{i,t}$ denotes residential electricity demand of each US states, $P_{i,t}$ is the electricity price, $HDD_{i,t}$ and $CDD_{i,t}$ are HDD and CDD of each i states and t time period. $RE_{i,t}$, $WF_{i,t}$, and $T_{i,t}$ are the Google Trends keywords “renewable,” “weather forecast,” and “temperature” by state, and μ and e are the error terms.

Results

After conducting the Hausman test, we selected the fixed effect model to estimate panel regression model. Table 1 shows fixed effects results. Both HDD and CDD variables are statistically significant at 1% significance level. These variables have positive coefficients, and residential electricity demand is increased by 1.792 and 2.523 million kWh, respectively, as HDD and CDD increase one level. The “renewable” variable of the Google keyword has a negative correlation and has a much greater effect than HDD and CDD. The “temperature” keyword variable is statistically significant at 5% significance level and also has a negative correlation. In the case of the “weather forecast” variable, the results cannot reject at the 10% significance level. Thus, the "weather forecast" variable has no explain power for the residential electricity demand and it is preferable to exclude it from the influence factors.

Table 1. Fixed effect results

Residential electricity demand	Coef.	t-statistic	$P > \|t\|$
HDD	**1.792**	**17.27**	**0.000****
CDD	**2.523**	**15.12**	**0.000****
Google Trends Keywords: “renewable”	**-16.746**	**-13.61**	**0.000****
Google Trends Keyword: “temperature”	**2.396**	**1.95**	**0.052***
Google Trends Keyword: “weather forecast”	**-0.398**	**-0.38**	**0.704**
constant	**2136.750**	**23.58**	**0.000****

Note: ** (*) denote statistical significance at the 1% (5%) level.

Conclusions

Proposed model use HDD, CDD, and Google Trends panel data to quantify the interest of each residential. As a result, the correlation of the “renewable” variable is statistically significant and the “temperature” variable is also correlated with residential electricity demand. The "renewable" keyword has a large negative correlation with residential electricity demand, which can be inferred from increased interest in renewable energy. While residential electricity demand patterns are influenced by many variables, this study suggests that interest in renewable energy should also be a key factor affecting demand. Further analysis of electricity demand with more Google Trends keywords is left for future research.

References

Wu, L. and E. Brynjolfsson (2009): “The future of prediction: how Google searches foreshadow housing prices and quantities,” ICIS 2009 Proceedings, 147.

Choi, H. and HAL. Varian (2012): “Predicting the Present with Google Trends,” Economic Record, 88, 2-9.

Swan, L.G. and V.I. Ugursal (2009): “Modeling of end-use energy consumption in the residential sector: A review of modeling techniques,” Renewable and Sustainable Energy Reviews, 13(8), 1819-1835.

Bessec M and J. Fouquau (2008): “The non-linear link between electricity consumption and temperature in Europe: A threshold panel approach,” Energy Economics, 30(5), 2705-2721.

Hekkenberg, M., R. M. J. Benders, H. C. Moll, and AJM Schoot Uiterkamp (2009): “Indications for a changing electricity demand pattern: The temperature dependence of electricity demand in the Netherlands,” Energy Policy, 37(4), 1542-1551.

Acknowledgement

This work was supported by the Human Resources program in Energy Technology (No. 20174010201170) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Trade, Industry and Energy.