Predicting the Present with Search Engine Data

Hal Varian
Google Inc.
hal@google.com

ABSTRACT
Many businesses now have almost real time data available about their operations. This data can be helpful in contemporaneous prediction (“nowcasting”) of various economic indicators. We illustrate how one can use Google search data to nowcast economic metrics of interest, and discuss some of the ramifications for research and policy.

Our approach combines three Bayesian techniques: Kalman filtering, spike-and-slab regression, and model averaging. We use Kalman filtering to whiten the time series in question by removing the trend and seasonal behavior. Spike-and-slab regression is a Bayesian method for variable selection that works even in cases where the number of predictors is far larger than the number of observations. Finally, we use Markov Chain Monte Carlo methods to sample from the posterior distribution for our model; the final forecast is an average over thousands of draws from the posterior.

An advantage of the Bayesian approach is that it allows us to specify informative priors that affect the number and type of predictors in a flexible way.

Categories and Subject Descriptors
J.4 [Economics]: Social and Behavioral Sciences

General Terms
Economics

Keywords
Nowcasting; Macroeconomic indicators; Google Data

Short Bio
Hal R. Varian is the Chief Economist at Google. He started in May 2002 as a consultant and has been involved in many aspects of the company, including auction design, econometric analysis, finance, corporate strategy and public policy.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

KDD’13, August 11–14, 2013, Chicago, Illinois, USA.
ACM 978-1-4503-2174-7/13/08.