DEUTSCHE
BUNDESBANK
EUROSYSTEM

Discussion "Forecasting U.S. Birth Rates using Google Trends"

Workshop Harnessing Big Data & Machine Learning techniques for Central Banks
26 – 27 March 2018

Simon Oehler, Directorate General Statistics

Agenda
Discussion "Forecasting U.S. Birth Rates using Google Trends"

- **Summary**

- **Data**

- **Methodology / Modelling**

- **Results**

- **Comments**

Summary
Discussion "Forecasting U.S. Birth Rates using Google Trends"

- Authors **study birth rates**, or fertility, **in the U.S.** as a major determinant of (future) population size in a time series setting.

- Generally population dynamics realize over the long run (trend), however, there seems to be a kind of **short term cyclicality,** which this paper foucusses on.

- In several **single equation models fertility is forecasted** on different **horizons up to 24 month**

- Models including leading indicators (GDP, UR, EPU, Google) are **benchmarked** against a simple autoregressive model

- Response:
  - Birth rates in the U.S.

- Independent variables:
  - Macro indicators: GDP, unemployment
  - News/Search indicators: EPU, Google Trends (Maternity, pregnancy, ovulation)

- Two Samples:
  - Long sample 1990M1 – 2008M12 (for GDP, UR, EPU)
  - Short sample 2004M1 – 2013M12 (for Google)

- **Single equation models**:

  - Benchmark: AR(p) process, lag length selection done by using BIC criterion

  - Models with additional RHS variables are introduced one at a time & estimated

  - H-step (6, 12, 18, 24) ahead forecasts are computed and compared to the benchmark

  - Exercise is done on federal & State level

Results
Discussion "Forecasting U.S. Birth Rates using Google Trends"

- Short sample:

  - Especially model with Google Indicator 1 (GI1) outperforms benchmark over all forecast horizons

  - Similar result for EPU Index

  - Models including Google perform well also on state-level

- EPU performs well in the long sample

- Data:

  - SA is done on Google Indixes, accounting for yearly seasonality? (Prob. Due to Christmas related searches?)
    - Possibly still other seasonal patterns?
    - Total # of searches is unknown, probably increases over time & changes due to specific events: How to cure?

  - Short sample ranges from 2004 – 2013
    - Extend up to 2018?

Comments
Discussion "Forecasting U.S. Birth Rates using Google Trends"

- Data:

  - Google Indexes are seasonally adjusted, accounting for yearly seasonality (Prob. due to Christmas related searches?)
    - Possibly still other seasonal patterns?
    - Total # of searches is unknown, probably increases over time & changes due to specific events: How to cure?

  - Short sample ranges from 2004 – 2013
    - Extend up to 2018?

  - Are birth rate & RHS variables stationary in levels?

Comments

Discussion "Forecasting U.S. Birth Rates using Google Trends"

GI1 – 'Maternity'
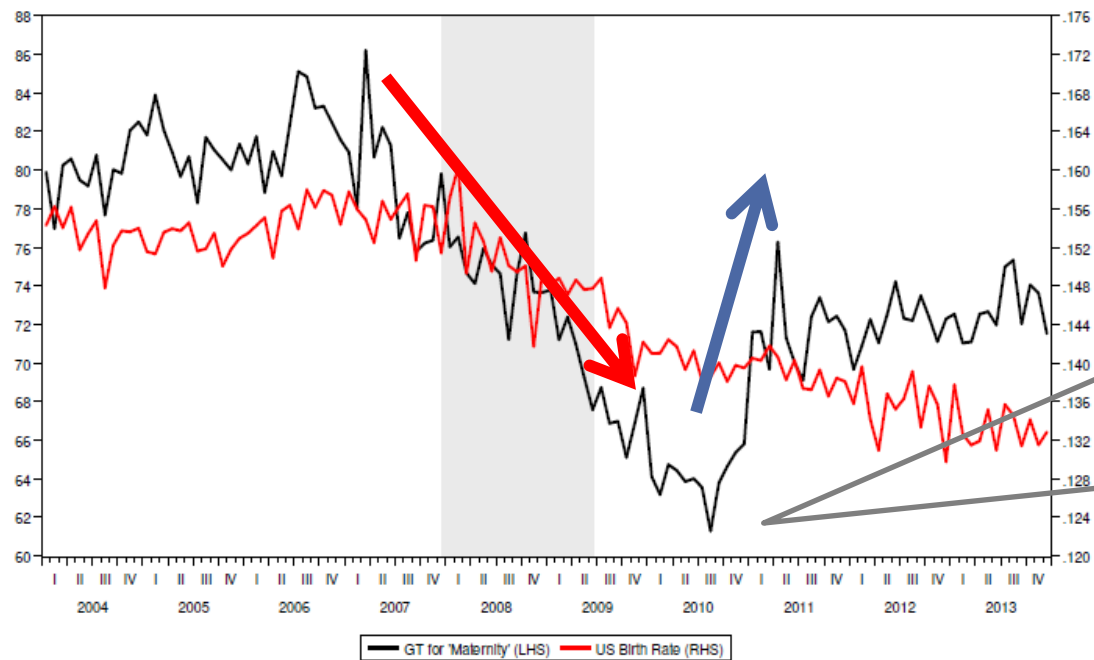
December 2009: President Obama signs federal teen pregnancy initiative into law. Budget: 115 $ for several, private & public initiatives.

Comments
Discussion "Forecasting U.S. Birth Rates using Google Trends"

- Possible further steps:

    - Regress Google Indexes on event dummies? (Also accounting for Financial crisis)

    - Alternatively, refine search terms in order to exclude specific „noisy" searches

    - Include also government expenditures, particularly social welfare etc.

    - Data selection: Extend # of your search queries e.g. by using Google Correlate

    - Model selection: subsequently use PCA or LASSO / RIDGE regressions