Data-Driven Macroeconomic Prediction

CMPUT 654: Modelling Human Strategic Behaviour

Wu & Brynjolfsson (2015)

Wu & Brynjolfsson (2015)

Why:

- Data-driven predictions of behaviour
- Behavioural model (albeit super-simple)
- Macroeconomic domain

Predicting Housing Sales

- **Housing sales** are a major driver of the economy
- Official data on housing sales are reported with large **lags** (quarterly if you're lucky)
- Most work on predicting macroeconomic variable such as housing sales focuses on more powerful models
- This work: Focus on richer and more timely **data** instead

Dataset: Google Trends

- Google Trends reports the **volume** of search queries
  - Individually and aggregated into predefined categories
  - This work: "real estate agencies" and "real estate listings"
- **Relative** volume, not absolute
  - **Percent** of query volume in specified country, state/province, region, city, time period

Behavioural Model

- Housing is a high-commitment, expensive purchase
- People will do a lot of relevant Google searching before undertaking it

Model: Contemporaneous Sales

Seasonal linear autoregression models:

1. **Baseline:**

$$HomeSales_{i,t} = \alpha + \beta_1 HomeSales_{i,t-1} + \beta_2 HPI_{i,t-1} + \beta_3 Population_{i,t} + \sum S_i + \sum R_j + \sum T_t + \epsilon_{i,t}$$

2. **With search data:**

$$HomeSales_{i,t} = \alpha + \beta_1 HomeSales_{i,t-1} + \beta_2 HPI_{i,t-1} + \beta_3 Population_{i,t} + \beta_4 SearchFreq_{i,t} + \beta_5 SearchFreq_{i,t-1} + \sum S_i + \sum R_j + \sum T_t + \epsilon_{i,t}$$

Model:
Sales Prediction

3. **With search data:**

$$HomeSales_{i,t+1} = \alpha + \beta_1 HomeSales_{i,t-1} + \beta_2 HPI_{i,t-1} + \beta_3 Population_{i,t} + \beta_4 SearchFreq_{i,t} + \beta_5 SearchFreq_{i,t-1} + \beta_5 SearchFreq_{i,t-2} + \sum S_i + \sum R_j + \sum T_t + \epsilon_{i,t}$$

Results: Nowcasting vs. Baseline

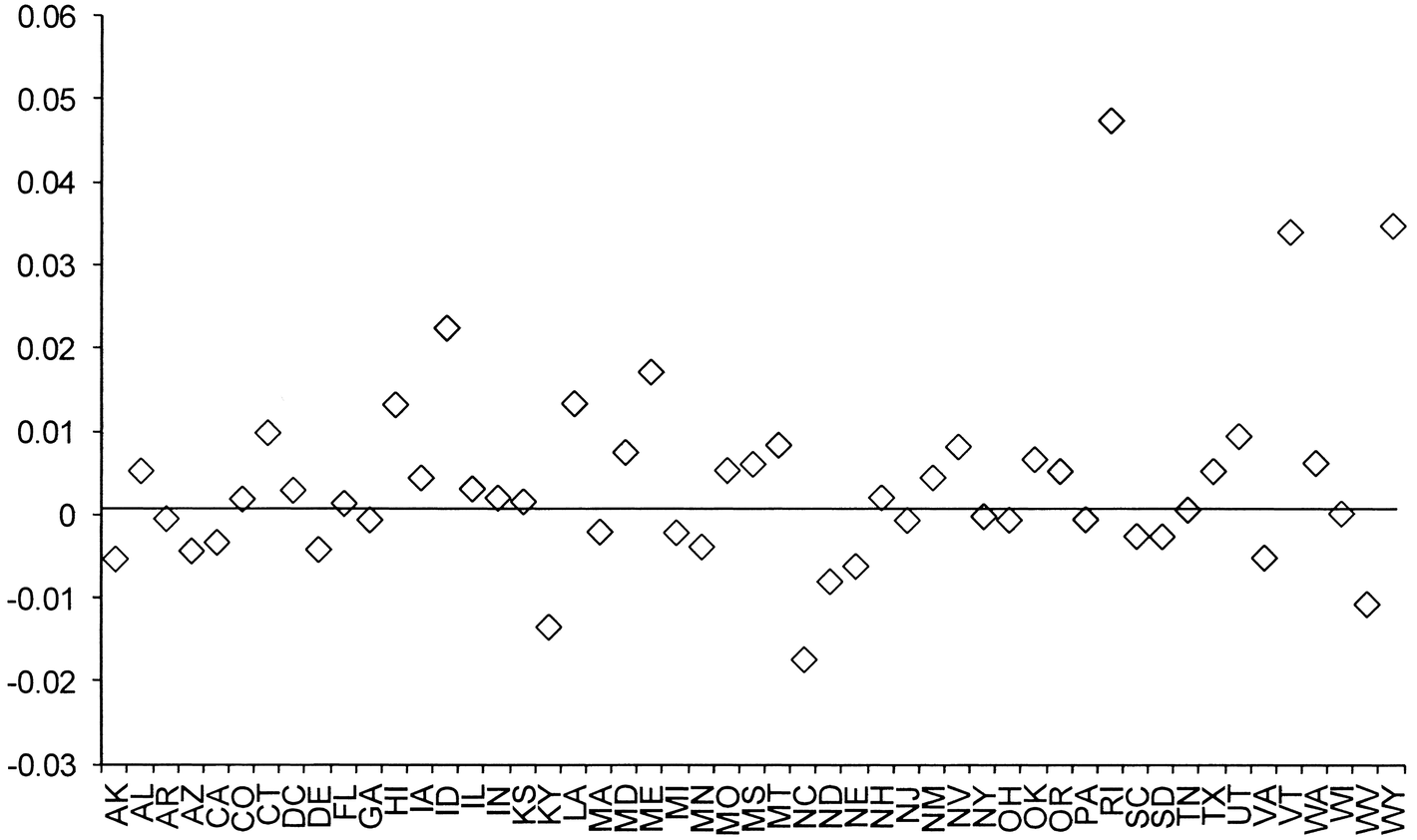

Fig. 3.3 The Y-axis indicates the average difference in MAE between the baseline model (equation [1]) and the model that uses search indices (equation [2])

Results: Prediction vs. Baseline

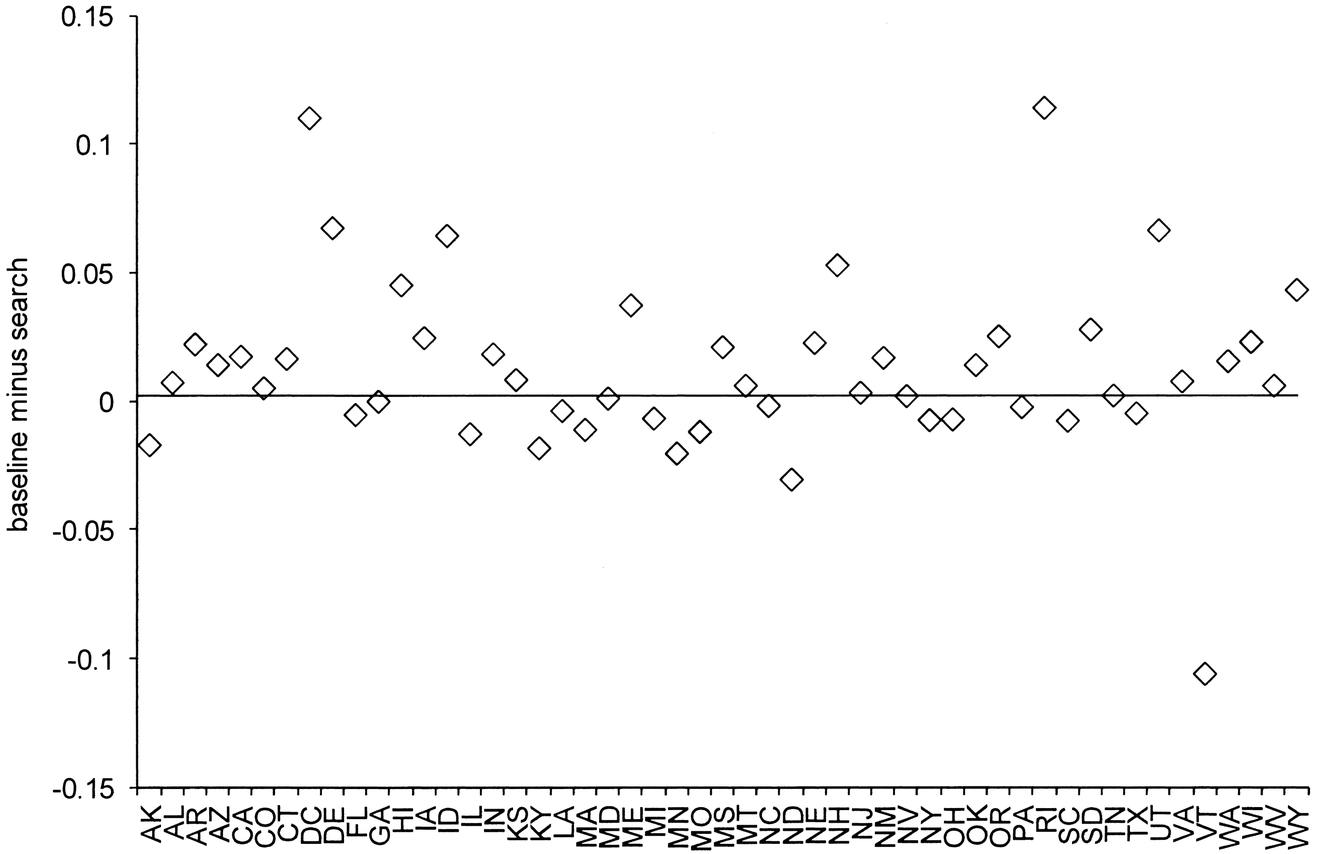

Fig. 3.4 MAE differences between the baseline model and predictions using search indices

Results: Prediction vs. National Ass'n of Realtors

Table 3.2 **Comparing with predictions from the National Association of Realtors for home sales in the United States**

MAE for sales$_{t+1}$	Obs.	Mean	Std. err.	Min.	Max.
Search	10	0.084	0.031	0.012	0.156
NAR	10	0.110	0.026	0.050	0.169
Diff.		23.6%		$p < 0.01$	