Google it!

Forecasting the US unemployment rate with a Google job search index

Francesco D'Amuri   Juri Marcucci

Bank of Italy

2^{nd} International Conference in Memory of Carlo Giannini

Time Series Econometrics and Macroeconomic Forecasting in a Policy Environment

Rome - January 19, 2010

Outline

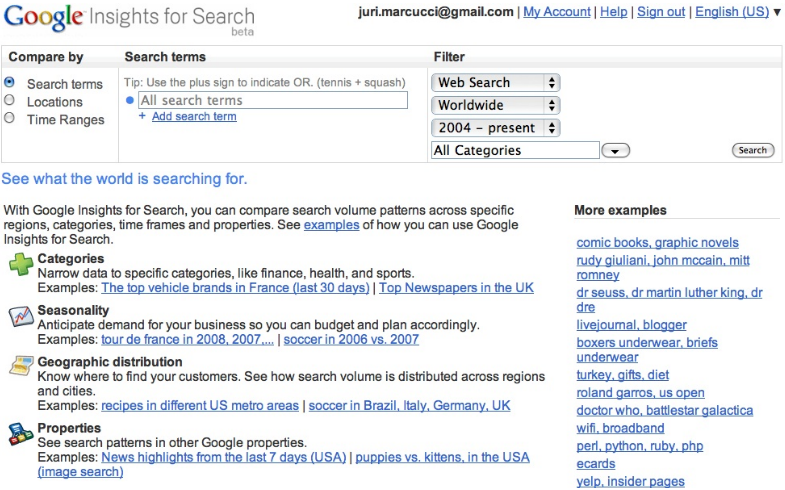
- Introduction and Motivations
- Data and Leading indicators for the US unemployment rate
  - Initial jobless claims (traditional!)
  - Google job web search index (New!)
- Forecasting models
- Out-of-sample evaulation
  - Tests Equal forecast accuracy and forecast encompassing
  - Reality Check test for superior predictive ability
- Some Robustness
  - Results from aggregation of States' forecasts
  - Comparison with Survey of Professional Forecasters
- Discussion and Conclusion

Introduction

- Having *reliable* and *updated* **unemployment forecasts** has become increasingly important, in particular during economic downturns
- The literature on US unemployment forecasting has primarily dealt either with simple **linear** models or with **non-linear** models
  - For example Montgomery, Zarnowitz, Tsay and Tiao (JASA, 1998), Proietti (CSDA, 2003) or Rothman (RESTAT, 1998)
- These linear models have been augmented with some **leading indicators**: in particular the Initial jobless Claim (IC) seem to be the best indicator for the US unemployment, so far...

Motivation

Google 'job' web search weekly index from **Google Insights**



Our Contribution

- In this paper we suggest an ***alternative leading indicator*** to forecast the US unemployment rate
  ⇒ a **Google** job web search index
- To the best of our knowledge, this indicator has only been used by:
  - Askitas & Zimmermann (2009) to forecast German unemployment
  - D'Amuri (2009) to forecast Italian unemployment
  - Suhoy (2009) to forecast unemployment in Israel
  - Choi and Varian (2009) to predict the US initial claims
- Running an extensive out-of-sample forecasting *horse-race*, we compare the predictive power of linear forecasting models using the Google Index (GI) with those using the Initial Claims or combinations of both.
- Our interest is on *short-term forecasting*, i.e. in forecasting the US monthly unemployment rate from 1- to 3-months ahead

Our results

- Our results show that the **Google index really helps** in predicting the monthly US unemployment rate, even after controlling for the effects of data-snooping.
  - Linear models with GI ***outperform*** all the other models using IC as a leading indicator, both in terms of **equal forecast accuracy** and **superior predictive ability**
- Moreover, linear *models augmented with the GI* ***outperform*** also at the **state level** (to predict the unemployment rate in each state) and in comparison to the *Survey of Professional Forecasters*.
- Our preferred models with *GI* also ***outperform*** standard *non-linear* models

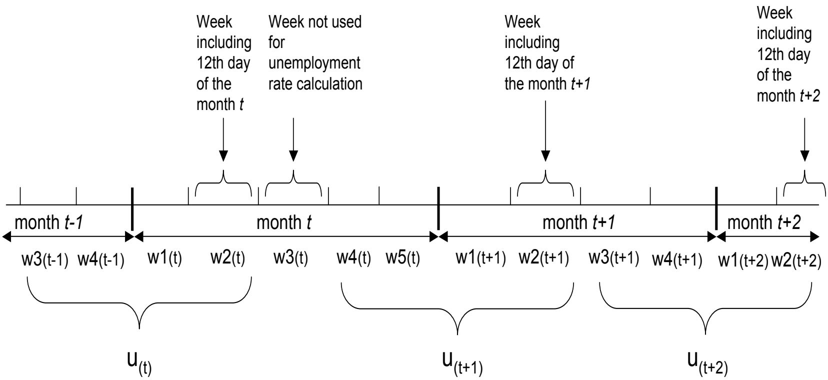
Data

1) Unemployment rate (US and state level)

- Monthly unemployment rate (u_t) seasonally adjusted from BLS
  - Current Unemployment Statistics (national level)
    - Sample: **1948.1-2009.6** (738 obs.)
  - Local Area Unemployment Statistics (state level)
    - Sample: **1976.1-2009.6** (402 obs.)
- u_t for month t refers to:
  - people who **don't have a job**, but are **available for work**, in the week including the 12^{th} of month t...
  - ...and who **have looked for a job** in the previous 4 weeks (*reference week* included)

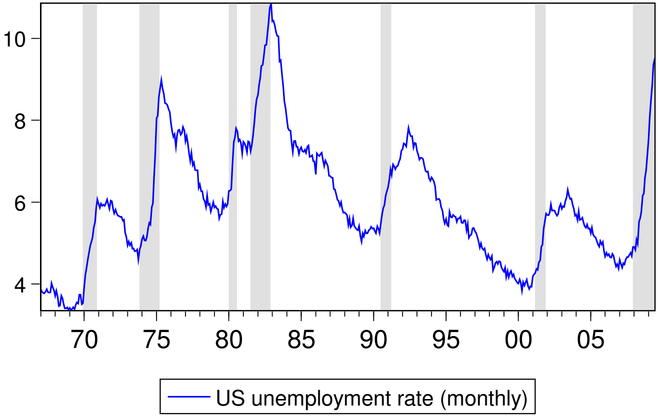
Data (Cont.)

1) Unemployment rate: Exact timing of US monthly unemployment rate



Data (Cont.)

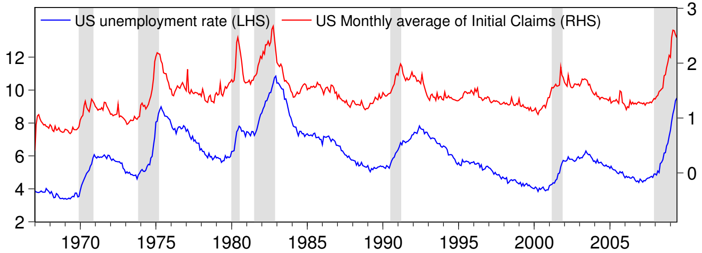
1) Unemployment rate (US) - Sample: 1967.1-2009.6 (NBER recessions - shaded areas)



Data (Cont.)

2) Initial Jobless Claims (US and state level)

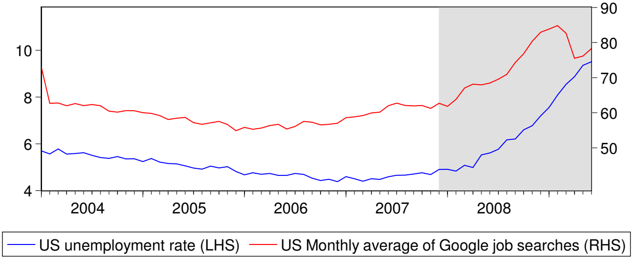
- Weekly **Initial Jobless Claims** (IC) seasonally adjusted from the US Department of Labor
  ⇒ well-known Leading Indicator (Montgomery et al., 1998)
  - National level
    - Sample: **1967.1-2009.6** (510 obs.)
  - State level (SA w/ Tramo-Seats)
    - Sample: **1987.1-2009.6** (271 obs.)



Data (Cont.)

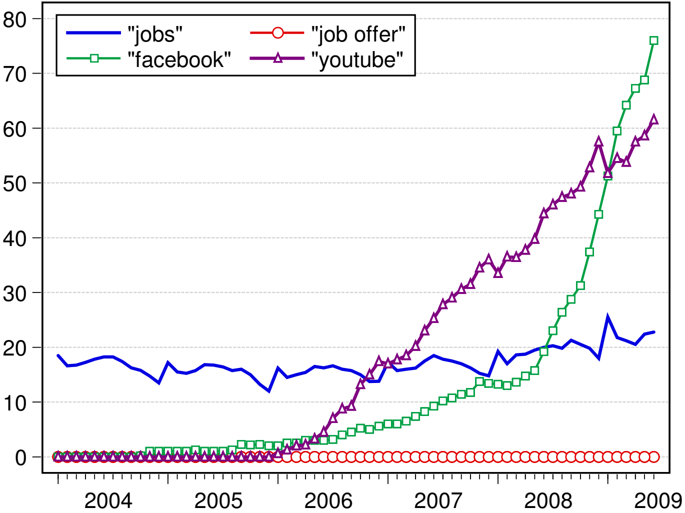
3) Google 'job' web search index from Google Insights (US and state level)

- **Weekly Google Index** (GI) seasonally adjusted from Google Insights (available almost in real time)
  ⇒ suggested Leading Indicator
  (Incidence of "*jobs*" queries on total web queries in relevant week)
  - National level
    - Sample: **2004.1-2009.6** (66 obs.)
  - State level
    - Sample: **2004.1-2009.6** (66 obs.)



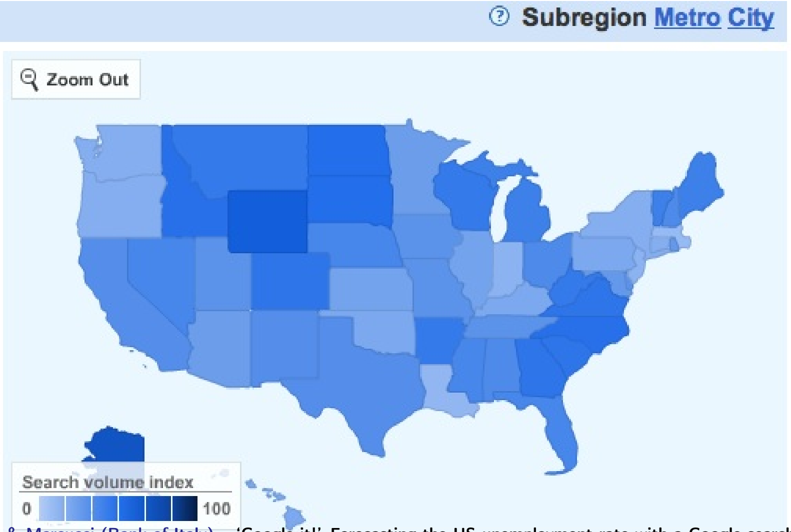
Data (Cont.)

3) Incidence of keyword "jobs" vs other popular keywords



Data (Cont.)

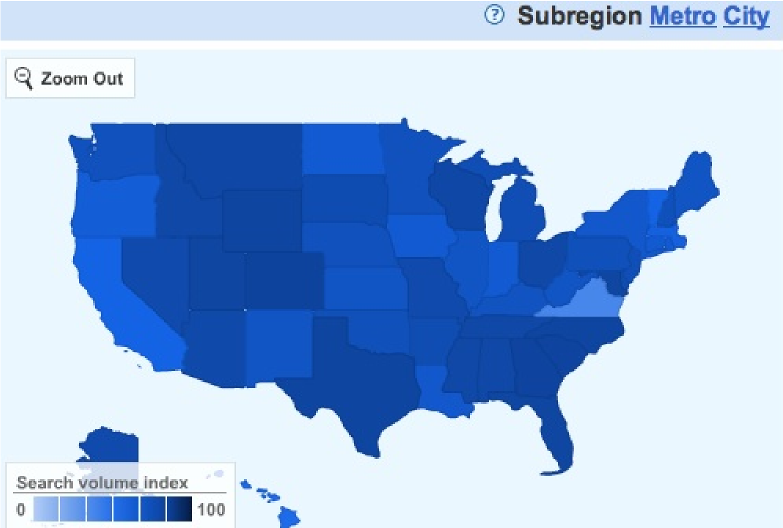
3) Google 'job' web search index from Google Insights
Pre-crisis period: Jan.-Apr. 2007



Data (Cont.)

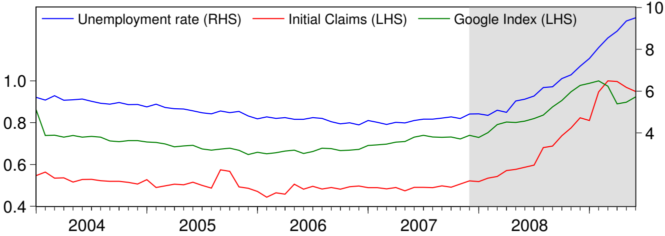
3) Google 'job' web search index from Google Insights

During the crisis: Jan.-Apr. 2009



Data (Cont.)

Unemployment rate (US), Initial Claims and Google index - Sample: 2004.1-2009.6



Data (Cont.)

ADF-GLS Unit Root tests by Elliott et al. (1996). Short and Full Sample

Sample: 1967:1-2009:6			Sample: 2004:1-2009:6		
Variable	Test	Test stat.	Variable	Test	
u_t	$DF-GLS^{\mu}$	-1.054	u_t	$DF-GLS^{\mu}$	-2.881***
	$DF-GLS^{\tau}$	-2.282		$DF-GLS^{\tau}$	-2.902*
$\log(u_t)$	$DF-GLS^{\mu}$	-0.901	$\log(u_t)$	$DF-GLS^{\mu}$	-2.792***
	$DF-GLS^{\tau}$	-2.190		$DF-GLS^{\tau}$	-2.797
u_t^{logit}	$DF-GLS^{\mu}$	-0.912	u_t^{logit}	$DF-GLS^{\mu}$	-2.801***
	$DF-GLS^{\tau}$	-2.203		$DF-GLS^{\tau}$	-2.804
u_t^{HPlog}	$DF-GLS^{\mu}$	-3.752***	u_t^{HPlog}	$DF-GLS^{\mu}$	-2.659***
	$DF-GLS^{\tau}$	-4.414***		$DF-GLS^{\tau}$	-2.523
u_t^{LLD}	$DF-GLS^{\mu}$	-1.344	u_t^{LLD}	$DF-GLS^{\mu}$	-2.823***
	$DF-GLS^{\tau}$	-2.190		$DF-GLS^{\tau}$	-2.797

The setup of the forecasting horse-race

- **Timing:** $T = R + P$ observations.
  - In the '**full-sample**' (**1967.1-2009.6**) we have $T = 510$
  - In the '**short-sample**' (**2004.1-2009.6**) we have $T = 66$
- The first R are used to estimate the models (**in-sample**) while the last P are used for **out-of-sample** evaluation.
- Want to predict u_t (or transformations) using linear ARMA models w/ and w/o exogenous leading indicators x_t:
  - $x_t = \{IC_t, ..., IC_{t-k}\}$
  - $x_t = \{IC_{wj,t}, ..., IC_{wj,t-k}\}, j = 1, \ldots, 4, k = 1, 2$
  - $x_t = \{G_t, ..., G_{t-k}\}$
  - $x_t = \{G_{wj,t}, ..., G_{wj,t-k}\}, j = 1, \ldots, 4, k = 1, 2$
  - combinations IC and G

The setup of the forecasting horse-race (Cont.)

Forecasting Models: $\phi(L)y_t = \mu + x_t'\beta + \theta(L)\varepsilon_t$

	Full Sample: 1967.1-2007.2								Short Sample: 2004.1-2007.2							
	AR(1)	#	AR(2)	#	ARMA(1,1)	#	ARMA(2,2)	#	AR(1)	#	AR(2)	#	ARMA(1,1)	#	ARMA(2,2)	#
w/o LI																
	u_{t-1}	1	u_{t-k}	1	$u_{t-1}, \varepsilon_{t-1}$	1	$u_{t-k}, \varepsilon_{t-k}$	1	u_{t-1}	1	u_{t-k}	1	$u_{t-1}, \varepsilon_{t-1}$	1	$u_{t-k}, \varepsilon_{t-k}$	1
w/ LI x_t																
(t)																
IC	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1
IC_{wj}	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4
G	-		-		-		-		✓	1	✓	1	✓	1	✓	1
G_{wj}	-		-		-		-		✓	4	✓	4	✓	4	✓	4
IC, G	-		-		-		-		✓	1	✓	1	✓	1	✓	1
IC_{wj}, G_{wj}	-		-		-		-		✓	5	✓	5	✓	5	✓	5
$(t-1)$																
IC	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1
IC_{wj}	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4
G	-		-		-		-		✓	1	✓	1	✓	1	✓	1
G_{wj}	-		-		-		-		✓	4	✓	4	✓	4	✓	4
IC, G	-		-		-		-		✓	1	✓	1	✓	1	✓	1
IC_{wj}, G_{wj}	-		-		-		-		✓	5	✓	5	✓	5	✓	5
$(t-2)$																
IC	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1	✓	1
IC_{wj}	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4	✓	4
G	-		-		-		-		✓	1	✓	1	✓	1	✓	1
G_{wj}	-		-		-		-		✓	4	✓	4	✓	4	✓	4
IC, G	-		-		-		-		✓	1	✓	1	✓	1	✓	1
IC_{wj}, G_{wj}	-		-		-		-		✓	5	✓	5	✓	5	✓	5

$j = 1, \ldots, 4; k = 1, 2$ - w/ or w/o SAR/SMA

The setup of the forecasting horse-race (Cont.)

- Forecasting scheme: we use a **rolling** scheme.
  - 'Short-sample': $T = 66$ with $R = 38$ and $P = 28$.
    - In-sample: 2004.1-2007.2, 2004.2-2007.3, etc.
  - 'Full-sample': $T = 510$ with $R = 482$ and $P = 28$.
    - In-sample: 1967.1-2007.2, 1967.2-2007.3, etc.
- We use **only** the information available **at month** t when we make the prediction.
  - Thus at t we need to forecast future values of our exogenous LI's
  - To predict them, we use different auxiliary ARMA-like models (we present results only for the $AR(1)$ case).

Out-of-sample Results

- For u_t and $u_t - u_{t-1}$ (and all the other transformations) the **best models** out of sample in terms of the lowest MSE are **those including GI** as the leading indicator
- The **best 15 models** at all forecast horizons (1- to 3-months-ahead) **always include GI** as the exogenous variable
- However, the best 3, 5 and 11 models at respectively 1-, 2- and 3-months ahead include **GI *only*** as the LI
- We test for
  - **Equal Forecast Accuracy** (EFA) using the Diebold & Mariano (1995) test
  - **Forecast Encompassing** (FE) using the Harvey, Leybourne & Newbold (1998) (HLN) test

Out-of-sample Results (Cont.)

- DM test and HLN test always reject the null at 10% at 1-month horizon and mostly reject at 2-month horizon.
- This means that our best model with GI outperforms all those models that use the *longest* available time series of u_t and IC, even though our best model is estimated over a rolling sample of 38 obs.
- Our best models with GI outperforms also those models not using GI over the short sample.

Out-of-sample Results (Cont.)

Best Forecasting Models: 1-month ahead

		1-step ahead			
n.	**Model**	**MSE**	**Rank**	**DM**	**HLN**
Panel A1: Best models					
261	$ARX(1) - G_t$	0.0166	1	-	-
398	$ARMAX(1,1) - G_t - SA$	0.0167	2	0.060	2.145**
327	$ARX(2) - G_t$	0.0172	3	0.448	1.063
491	$ARMAX(2,2) - IC_{t-1} - G_{t-1}$	0.0177	4	0.328	1.912*
305	$ARX(1) - G_{t-2}$	0.0179	5	0.616	1.289
464	$ARMAX(2,2) - G_t - SA$	0.0179	6	0.312	1.370
371	$ARX(2) - G_{t-2}$	0.0181	7	0.614	1.642
283	$ARX(1) - G_{t-1}$	0.0182	8	1.516	1.701*
463	$ARMAX(2,2) - G_{w4,t} - SA$	0.0184	9	0.442	2.116**
277	$ARX(1) - IC_t - G_t - SA$	0.0186	10	0.852	1.326
271	$ARX(1) - IC_t - G_t$	0.0186	11	0.709	1.605
266	$ARX(1) - G_t - SA$	0.0188	12	0.998	1.122
337	$ARX(2) - IC_t - G_t$	0.0191	13	0.799	1.875*
343	$ARX(2) - IC_t - G_t - SA$	0.0192	14	0.870	1.550
270	$ARX(1) - IC_{w4,t} - G_{w4,t}$	0.0192	15	0.778	1.807*
Panel B1: Best models without Google					
122	$ARMAX(2,2) - IC_{w4,t-2}$	0.0234	73	2.491**	3.074***
133	$ARMA(1,1)$	0.0301	197	2.152**	2.485**
Panel C1: Non-linear models					
521	$SETAR(2)$	0.0332	258	2.434**	2.925***
522	$LSTAR(2)$	0.0368	362	2.497**	3.015***
523	$AAR(2)$	0.0342	276	2.337**	2.903***

Out-of-sample Results (Cont.)

Best Forecasting Models: 2-month ahead

n.	Model	MSE	Rank	DM	HLN
		2-step ahead			
Panel A2: Best models					
261	$ARX(1) - G_t$	0.0157	1	-	-
464	$ARMAX(2,2) - G_t - SA$	0.0163	2	0.136	1.291
398	$ARMAX(1,1) - G_t - SA$	0.0166	3	0.177	1.219
327	$ARX(2) - G_t$	0.0172	4	0.633	0.864
266	$ARX(1) - G_t - SA$	0.0175	5	0.700	0.869
277	$ARX(1) - IC_t - G_t - SA$	0.0186	6	0.952	1.142
332	$ARX(2) - G_t - SA$	0.0194	7	0.955	1.192
343	$ARX(2) - IC_t - G_t - SA$	0.0206	8	1.150	1.285
283	$ARX(1) - G_{t-1}$	0.0208	9	1.514	1.543
420	$ARMAX(1,1) - G_{t-1} - SA$	0.0217	10	0.981	1.373
288	$ARX(1) - G_{t-1} - SA$	0.0220	11	1.402	1.551
305	$ARX(1) - G_{t-2}$	0.0220	12	1.551	1.718*
349	$ARX(2) - G_{t-1}$	0.0222	13	1.915*	2.024**
293	$ARX(1) - IC_{t-1} - G_{t-1}$	0.0233	14	1.989**	1.994**
299	$ARX(1) - IC_{t-1} - G_{t-1} - SA$	0.0234	15	1.392	1.468
Panel B2: Best models without Google					
122	$ARMAX(2,2) - IC_{w4,t-2}$	0.0514	180	1.814*	1.618
234	$ARMAX(2,2) - IC_{w3,t} - SA$	0.0565	191	1.389	1.131
Panel C2: Non-linear models					
521	$SETAR(2)$	0.0388	97	1.053	1.720*
522	$LSTAR(2)$	0.0447	140	1.190	1.779*
523	$AAR(2)$	0.0436	134	1.183	1.721*

Out-of-sample Results (Cont.)

Best Forecasting Models: 3-month ahead

		3-step ahead			
n.	**Model**	**MSE**	**Rank**	**DM**	**HLN**
Panel A3: Best models					
398	$ARMAX(1,1) - G_t - SA$	0.0350	1	-	-
327	$ARX(2) - G_t$	0.0372	2	0.230	0.793
332	$ARX(2) - G_t - SA$	0.0379	3	0.244	0.671
261	$ARX(1) - G_t$	0.0382	4	0.308	0.852
464	$ARMAX(2,2) - G_t - SA$	0.0382	5	0.295	0.579
266	$ARX(1) - G_t - SA$	0.0383	6	0.299	0.777
349	$ARX(2) - G_{t-1}$	0.0488	7	1.164	1.300
354	$ARX(2) - G_{t-1} - SA$	0.0495	8	1.115	1.440
393	$ARMAX(1,1) - G_t$	0.0508	9	0.722	1.060
288	$ARX(1) - G_{t-1} - SA$	0.0510	10	1.142	1.501
283	$ARX(1) - G_{t-1}$	0.0513	11	1.217	1.383
343	$ARX(2) - IC_t - G_t - SA$	0.0528	12	0.659	0.811
277	$ARX(1) - IC_t - G_t - SA$	0.0531	13	0.681	0.852
365	$ARX(2) - IC_{t-1} - G_{t-1} - SA$	0.0548	14	1.275	1.658*
265	$ARX(1) - G_{w4,t} - SA$	0.0555	15	0.938	1.219
Panel B3: Best models without Google					
122	$ARMAX(2,2) - IC_{w4,t-2}$	0.1406	191	1.309	1.249
215	$ARMAX(1,1) - IC_{w4,t-1} - SA$	0.1294	173	1.748*	1.752*
Panel C3: Non-linear models					
521	$SETAR(2)$	0.0589	24	0.758	1.447
522	$LSTAR(2)$	0.0620	30	0.790	1.411
523	$AAR(2)$	0.0652	35	0.814	1.389

Out-of-sample test of Superior Predictive Ability

White's (2000) Reality Check (RC) test

- The RC is a test for **superior unconditional predictive ability** that also accounts for the ***dependence*** among forecasting models (*data-snooping*).
- The **null** hypothesis is that ***all the competing models are no better than the benchmark*** model, i.e. $H_0 : \max_{1\leq k\leq L} E(f_k) \leq 0$, where $f_k = e^2_{0,t} - e^2_{k,t}$
- The ***alternative*** is that H_0 is false, that is, ***there exists a best model which is superior to the benchmark***.
- White's (2000) RC test statistic for H_0 is formed as $\bar{V} = \max_{1\leq k\leq L} \sqrt{P}\bar{f}_k$, where $\bar{f}_k = P^{-1/2}\sum_{t=R+1}^{T} \hat{f}_{k,t}$
- Using the stationary bootstrap of Politis and Romano (1994), the empirical distribution of $\bar{V}^* = \max_{1\leq k\leq L} \sqrt{P}(\bar{f}^*_k - \bar{f}_k)$ is used to compute the RC p-value

Out-of-sample test of Superior Predictive Ability (Cont.)

Reality Check p-values (in **bold** p-values≥**5%** ⇒ fail to reject H_0)

	B=2000	B=5000		B=2000	B=5000
	u_t			u_t^{LLD}	
1-step	Benchmark=403		1-step	Benchmark=327	
q=0.50	**0.073**	**0.070**	q=0.50	**0.076**	**0.076**
q=0.10	**0.053**	**0.057**	q=0.10	**0.053**	**0.060**
2-step	Benchmark=332		2-step	Benchmark=327	
q=0.50	0.037	0.039	q=0.50	0.043	0.040
q=0.10	**0.053**	**0.052**	q=0.10	**0.061**	**0.057**
3-step	Benchmark=332		3-step	Benchmark=266	
q=0.50	0.037	0.045	q=0.50	0.029	0.025
q=0.10	0.046	**0.052**	q=0.10	**0.050**	**0.052**
	$\log(u_t)$			$u_t - u_{t-1}$	
1-step	Benchmark=327		1-step	Benchmark=261	
q=0.50	**0.099**	**0.100**	q=0.50	**0.107**	**0.098**
q=0.10	**0.050**	0.045	q=0.10	**0.055**	**0.057**
2-step	Benchmark=327		2-step	Benchmark=261	
q=0.50	**0.080**	**0.080**	q=0.50	**0.098**	**0.097**
q=0.10	**0.058**	**0.058**	q=0.10	**0.053**	0.045
3-step	Benchmark=266		3-step	Benchmark=398	
q=0.50	**0.114**	**0.114**	q=0.50	**0.073**	**0.073**
q=0.10	**0.058**	**0.066**	q=0.10	0.048	0.048
	u_t^{logit}			u_t^{HPlog}	
1-step	Benchmark=327		1-step	Benchmark=327	
q=0.50	**0.083**	**0.083**	q=0.50	**0.073**	**0.083**
q=0.10	**0.073**	**0.068**	q=0.10	**0.057**	**0.060**
2-step	Benchmark=327		2-step	Benchmark=327	
q=0.50	0.027	0.033	q=0.50	**0.065**	**0.062**
q=0.10	**0.054**	**0.056**	q=0.10	**0.057**	**0.057**
3-step	Benchmark=266		3-step	Benchmark=266	
q=0.50	0.028	0.027	q=0.50	0.041	0.038
q=0.10	**0.052**	**0.054**	q=0.10	**0.061**	**0.052**

Further robustness checks out-of-sample

- Also **recursive** scheme with similar results (unreported).
- Different **auxiliary** models to predict the LI's: $AR(2)$, $ARMA(1,1)$, $ARMA(2,2)$ with similar (unreported) results.
- Comparison of our best models (overall and without Google indicator) with the **Survey of Professional Forecasters** for the quarterly unemployment rate
- **State-level** forecasts with different aggregation schemes
- Some **non-linear models** typically adopted in the literature
- We also ran the horse-race for different **transformation** of u_t typically used in the literature, such as
  - $\log(u_t)$
  - $u_t^{LLD} = \log(u_t) - \hat{\alpha} - \hat{\beta}t$
  - $u_t^{logit} = \log[u_t/(1-u_t)]$
  - $u_t^{HPlog} = \log(u_t) - [\log(u_t)]^{HP}$.

A further out-of-sample check: comparison with the SPF

Sample: 2007:Q1-2009:Q2

- We also compared our forecasting models with the Survey of Professional Forecasters (SPF) (mean, median and best)
- At the 'middle' of $Q(J)$ (around Feb, May, Aug and Nov 15) SPF issues forecasts for $Q(J+1)$ to $Q(J+5)$ (true deadline for forecasters is around 10$^{\text{th}}$ of same month)
- We compare SPF^{best}, SPF^{median} and SPF^{mean} with 3 different forecasts of quarterly US unemployment from the following models (for u_t)
  - Best model overall, i.e. model with Google (# 403)
  - Best model overall without Google, i.e. model with Initial Claims (# 128)
  - Best model in the short sample without Google (# 205)

A further check: comparison with the SPF (Cont.)

Sample: 2007:Q1-2009:Q2

- For each model we compute 3 sets of quarterly forecasts
  1. At the end of $Q(J)$, e.g. 2007.3: forecast 1-month ahead
  $$\hat{u}_{t+1|t} \Rightarrow \mathbf{x}^{\mathbf{1st-month}}$$
  is our forecast for $Q(J+1)$ (conservative)
  2. At the end of $Q(J)$, e.g. 2007.3: forecast 2-month ahead
  $$\hat{u}_{t+2|t} \Rightarrow \mathbf{x}^{\mathbf{2nd-month}}$$
  is our forecast for $Q(J+1)$ (conservative)
  3. Around the 10th of the second month of $Q(J)$, e.g. 2007.5: forecast 1- and 2-month ahead
  $$[u_t + \hat{u}_{t+1|t} + \hat{u}_{t+2|t}]/3 \Rightarrow \mathbf{x}^{\mathbf{Comb}}$$
  is our forecast for $Q(J+1)$ (less conservative and similar timing to SPF)

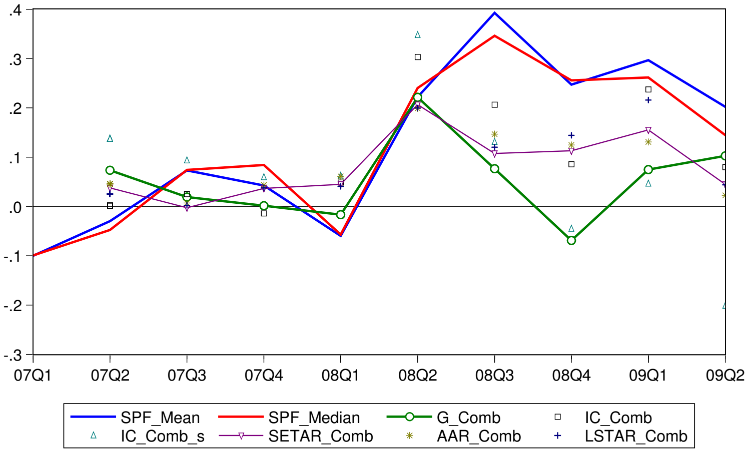
A further check: comparison with the SPF (Cont.)

Sample: 2007:Q1-2009:Q2. Benchmark: G^{Comb}

	MSE	Rank	DM	HLN
SPF^{best}	1.373	21	1.911*	2.177**
SPF^{mean}	0.415	11	1.545	2.784***
SPF^{med}	0.360	7	1.317	2.892***
$G^{1st-month}$	0.530	15	-1.522	2.401**
$G^{2nd-month}$	0.419	12	1.724*	1.925*
G^{Comb}	**0.082**	1	-	-
$IC^{1st-month}$	0.893	17	-0.337	2.621***
$IC^{2nd-month}$	0.361	8	-0.919	1.457
IC^{Comb}	0.208	5	-2.012**	-1.875*
$IC_s^{1st-month}$	0.612	16	0.048	2.386**
$IC_s^{2nd-month}$	0.413	10	1.810*	1.759*
IC_s^{Comb}	0.218	6	1.306	1.239
$SETAR^{1st-month}$	1.123	19	2.881***	2.596***
$SETAR^{2nd-month}$	0.373	9	1.098	2.902***
$SETAR^{Comb}$	*0.098*	2	-1.401	2.587***
$LSTAR^{1st-month}$	1.228	20	2.558**	2.407**
$LSTAR^{2nd-month}$	0.433	14	1.550	2.723***
$LSTAR^{Comb}$	0.127	4	-1.265	2.315**
$AAR^{1st-month}$	1.060	18	2.630***	2.418**
$AAR^{2nd-month}$	0.432	13	1.768*	2.900***
AAR^{Comb}	0.102	3	-1.37	2.662***

A further check: comparison with the SPF (Cont.)

Forecast errors of 'best' models - Sample: 2007:Q1-2009:Q2



A further check: aggregation of State-level forecasts

- For each 51 states (including District of Columbia) we ran the same horse-race with the same 520 forecasting models.
- For $u_t - u_{t-1}$ the percentage of best models for each state using the Google indicator as a LI ranges between 75% and 84% for 1-, 2- and 3-month-ahead forecasts.
- For u_t such percentage ranges between 69 and 82%.
- We test whether the **aggregation** of the 51 best state models could improve the forecasting performance over the federal benchmark. We use the following weights:
  - equal weight
  - % or share of labor force w.r.t. US total
  - % of labor force $\times$ share of internet use among labor force
  - % of labor force $\times$ share of internet use among active
  - % of labor force $\times$ share of internet use among unemployed
  - % of unemployed w.r.t. US total $\times$ share of internet use among unemployed

A further check: aggregation of State-level forecasts (Cont.)

	1-Step					2-Step					3-Step				
Variable: $d(u_t)$	MSE	Rk1	Rk2	DM	HLN	MSE	Rk1	Rk2	DM	HLN	MSE	Rk1	Rk2	DM	HLN
Model															
best	0.0166	1	1	-	-	0.0157	1	1	-	-	0.0350	1	4	-	-
simple avg	0.2845	7	525	5.30^a	4.92^a	0.3391	7	524	2.77^a	2.31^b	0.3966	7	510	1.99^b	2.31^b
labor force (LF)	0.0292	2	181	-0.13	2.68^a	0.0310	2	48	-0.30	1.31	0.0411	2	7	-1.17	1.31
IU all × LF	0.0299	5	196	-0.06	2.75^a	0.0314	3	51	-0.28	1.32	0.0413	3	8	-1.16	1.32
IU active × LF	0.0296	3	190	-0.09	2.69^a	0.0318	4	56	-0.26	1.30	0.0423	4	9	-1.14	1.30
IU UN × LF	0.0298	4	194	-0.07	2.71^a	0.0322	5	57	-0.25	1.31	0.0425	5	10	-1.13	1.31
IU UN × UN	0.0917	6	519	2.33^b	3.33^a	0.0690	6	239	0.65	1.66^c	0.0618	6	32	-0.53	1.66^c

Variable: u_t	MSE	Rk1	Rk2	DM	HLN	MSE	Rk1	Rk2	DM	HLN	MSE	Rk1	Rk2	DM	HLN
Model															
best	0.0167	1	1	-	-	0.0169	1	7	-	-	0.0482	6	15	-	-
simple avg	0.3000	7	526	5.29^a	4.70^a	0.3700	7	522	2.48^b	2.15^b	0.4560	7	514	1.83^c	1.73^c
labor force (LF)	0.0280	2	120	0.24	2.95^a	0.0293	2	29	-1.23	0.37	0.0459	3	3	-1.06	0.54
IU all × LF	0.0283	3	131	0.26	2.98^a	0.0294	3	30	-1.24	0.36	0.0454	2	2	-1.07	0.54
IU active × LF	0.0286	4	137	0.29	2.94^a	0.0303	5	33	-1.21	0.38	0.0474	5	5	-1.04	0.55
IU UN × LF	0.0287	5	140	0.30	2.96^a	0.0302	4	32	-1.21	0.38	0.0469	4	4	-1.05	0.56
IU UN × UN	0.0709	6	513	2.06^b	3.31^a	0.0519	6	152	-0.65	1.41	0.0373	1	1	-1.16	0.70

^{a}, ^{b}, and ^{c} significant at 1, 5 & 10%

Conclusion and discussion

- In this paper we have suggested **a new leading indicator** based on **Google job web search index** (GI) to forecast the monthly US unemployment rate
- We have tested the predictive power of different models using the Google index running an out-of-sample horse-race for 1- to 3-month-ahead forecasts
- Our results show that simple time series models augmented with GI **outperform** similar models using IC even when estimated over *longer* samples

Conclusion and discussion (Cont.)

- We assess the out-of-sample predictive ability of our best model (with GI) using DM and HLN test of EFA and FE, finding that **our best model is more accurate**
- We also assess the **superior predictive ability** of our best models with the Reality Check, thus controlling for *data-snooping* biases.
- Our results are robust to **different transformations** of u_t, to **state-level** data and aggregation, and our models also **outperform the SPF**
- Some **caveats** remain: we have a *very short* sample but our results seem very *promising*.