Supporting Information

Yang et al. 10.1073/pnas.1515373112

SI Materials and Methods

Details of our methodology are presented as follows. First, the predictive distribution in the formulation of the ARGO model and the corresponding assumptions are described; second, the statistical strategy to determine the hyperparameters of the ARGO model is explained; third, the results of two sensitivity analysis aimed at testing the robustness of the ARGO methodology—(*i*) with respect to subsequent revisions of CDC's ILI activity reports and (*ii*) with respect to observed variation of the input variables coming from Google Trends data—are presented; fourth, the exact search query terms identified by Google Correlate with different data access dates are presented; and fifth, a heat map showing the coefficients for the time series and Google search terms dynamically trained by ARGO is included.

The R package that implements the ARGO method is available at the authors' websites (www.people.fas.harvard.edu/~skou/publication.htm).

SI Predictive Distribution in the Formulation of ARGO Model

To improve normality for both the input variables and the dependent variables, the CDC-reported ILI activity level was logit-transformed, and the linearly normalized volume of Google search queries were log-transformed. To avoid taking the log of 0, we add a small number $\delta = 0.5$ before the log transformation. These transformations led to two sets of variables, the intrinsic (influenza epidemics activity) time series of interest $\{y_t\}$ and the (Google search) variable vector $\boldsymbol{X}_t$ at time t (that depends only on y_t). Our formal mathematical assumptions are

(assumption 1) $y_t = \mu_y + \sum_{j=1}^{N} \alpha_j y_{t-j} + \epsilon_t, \epsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$

(assumption 2) $\boldsymbol{X}_t | y_t \sim \mathcal{N}_K(\boldsymbol{\mu}_x + y_t \boldsymbol{\beta}, \boldsymbol{Q})$

(assumption 3) conditional on $y_t, \boldsymbol{X}_t$ is independent of $\{y_l, \boldsymbol{X}_l : l \neq t\}$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_K)^\top, \boldsymbol{\mu}_x = (\mu_{x_1}, \mu_{x_2}, \ldots, \mu_{x_K})^\top$, and $\boldsymbol{Q}$ is the covariance matrix. The predictive distribution $f(y_{t+1} | y_{1:t}, \boldsymbol{X}_{1:(t+1)})$ is given by

$$f(y_{t+1} | y_{1:t}, \boldsymbol{X}_{1:(t+1)}) \sim \mathcal{N}\left(\left(\frac{1}{\sigma^2} + \boldsymbol{\beta}^\top \boldsymbol{Q}^{-1} \boldsymbol{\beta} \right)^{-1} \left(\frac{\mu_y + \boldsymbol{\alpha}^\top y_{(t-N+1):t}}{\sigma^2} + \boldsymbol{\beta}^\top \boldsymbol{Q}^{-1} (\boldsymbol{X}_{t+1} - \boldsymbol{\mu}_x) \right), \left(\frac{1}{\sigma^2} + \boldsymbol{\beta}^\top \boldsymbol{Q}^{-1} \boldsymbol{\beta} \right)^{-1} \right), \quad \textbf{[S1]}$$

which is a normal distribution, whose mean is a linear combination of $y_{(t-N):(t-1)}$ and $\boldsymbol{X}_t$, and whose variance is a constant.

SI Determination of the Hyperparameters for ARGO

The optimized parameters of the ARGO model, $\mu_y, \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)$, are obtained by

$$\underset{\mu_y, \boldsymbol{\alpha}, \boldsymbol{\beta}}{\arg\min} \sum_t \left(y_t - \mu_y - \sum_{j=1}^{52} \alpha_j y_{t-j} - \sum_{i=1}^{100} \beta_i X_{i,t} \right)^2 + \lambda_\alpha \|\boldsymbol{\alpha}\|_1 + \eta_\alpha \|\boldsymbol{\alpha}\|_2^2 + \lambda_\beta \|\boldsymbol{\beta}\|_1 + \eta_\beta \|\boldsymbol{\beta}\|_2^2. \quad \textbf{[S2]}$$

The training period consists of a 2-y (104-wk) rolling window that immediately precedes the desired date of estimation. The hyperparameters are $\lambda_\alpha, \lambda_\beta, \eta_\alpha$, and η_β. We tested the performance of ARGO with the following specifications of hyperparameters: (specification 1) restrict $\eta_\alpha = \eta_\beta = 0$ and $\lambda_\alpha = \lambda_\beta$, cross-validate on λ_α. This is our proposed ARGO with the same L_1 penalty for Google search terms and autoregressive lags; (specification 2) restrict $\eta_\alpha = \eta_\beta = 0$, cross-validate on $(\lambda_\alpha, \lambda_\beta)$. This is ARGO with separate L_1 penalties for Google search terms and autoregressive lags; (specification 3) restrict $\eta_\alpha = \eta_\beta$ and $\lambda_\alpha = \lambda_\beta = 0$, cross-validate on η_α. This is ARGO with the same L_2 penalty for Google search terms and autoregressive lags; (specification 4) restrict $\lambda_\alpha = \lambda_\beta = 0$, cross-validate on $(\eta_\alpha, \eta_\beta)$—this is ARGO with separate L_2 penalties for Google search terms and autoregressive lags; and (specification 5) restrict $\lambda_\alpha = \lambda_\beta, \eta_\alpha = \eta_\beta$, cross-validate on $(\lambda_\alpha, \eta_\alpha)$. This is ARGO with the same elastic net (both L_1 and L_2) penalty for Google search terms and autoregressive lags.

Table S5 summarizes the in-sample estimation performance for our proposed ARGO, together with the other specifications of hyperparameters. It is apparent from the table that the L_1 penalty generally outperforms the L_2 penalty. The L_1 penalty tends to shrink the coefficients of unnecessary independent variables to be exactly zero, and thus eliminates redundant information; on the other hand, the L_2 penalty can only shrink the coefficients to be close to zero. As a result, L_2 penalized coefficients are not as sparse as their L_1 counterparts. Furthermore, from Table S5, we see that ARGO with separate L_1 penalties (specification 2) outperforms ARGO with separate L_2 penalties (specification 4), in terms of both RMSE and MAE. Similarly, ARGO with the same L_1 penalty (specification 1) outperforms ARGO with the same L_2 penalty (specification 3), in terms of both RMSE and MAE.

The elastic net model, which combines L_1 penalty and L_2 penalty, does not provide any error reduction. In the cross-validation process of setting $(\lambda_\alpha, \eta_\alpha)$ for the elastic net model, 70 wk out of 116 in-sample weeks showed that the smallest cross-validation mean error when restricting $\eta_\alpha = 0$ (i.e., zero L_2 penalty) is within 1 SE of the global smallest cross-validation mean error, suggesting that restricting L_2 penalty term to be zero (i.e., $\eta_\alpha = 0$) will introduce little bias. Therefore, for the simplicity and sparsity of the model, we drop the L_2 penalty terms and use only the L_1 penalty.

Next, we want to decide between the remaining two specifications, ARGO with separate L_1 penalties (specification 2) and ARGO with the same L_1 penalty (specification 1). One might argue that Google search terms and autoregressive lags are different sources of information and thus should have different L_1 penalties. However, empirical evidence in Table S5 shows that, again, giving extra flexibility to $(\lambda_\alpha, \lambda_\beta)$ does not generate improvement compared with fixing $\lambda_\alpha = \lambda_\beta$. In the cross-validation process of setting $(\lambda_\alpha, \lambda_\beta)$ for separate L_1 penalties, 99 wk out of 116 in-sample weeks showed that the smallest cross-validation mean error when restricting $\lambda_\alpha = \lambda_\beta$ (i.e., same L_1 penalty) is within 1 SE of the global smallest cross-validation mean error. This may well be due to the gain from variance reduction when imposing the restriction $\lambda_\alpha = \lambda_\beta$. Based on the same simplicity and sparsity consideration, we finally decided to restrict $\eta_\alpha = \eta_\beta = 0$ and $\lambda_\alpha = \lambda_\beta$ in the setting of hyperparameters for ARGO.

SI Revision of CDC's ILI Activity Reports

Within a flu season, CDC reports are constantly revised to improve their accuracy as new information is incorporated. Thus, CDC's weighted ILI figures displayed in previously published reports may change in subsequent weeks. As a consequence, in a given week, the available CDC ILI information from the most recent weeks may be inaccurate. To test the robustness of ARGO in the presence of these revisions and mimic the real-time tracking in our retrospective predictions, we trained ARGO and all other alternative models based on the following schedule.

Suppose $z_{i,j}$ is the CDC-reported ILI activity level of week i accessed at week j. Since CDC's ILI activity report is often delayed for 1 wk, on week j, the historical ILI activity-level data we have are $\{z_{i,j} : i \le j-1\}$. Due to revisions, ILI activity level of week i accessed at different weeks $z_{i,i+1}, z_{i,i+2}, \ldots$ may be different but will converge to a finalized value $z_{i,\infty}$ eventually. Hence, to avoid using forward-looking information, in week j, we train all models with the ILI activity level accessed at that week, $\{z_{i,j} : i \le j-1\}$. In this sense, any future revision beyond week j will not be incorporated in the training at week j. However, for the accuracy metrics, the estimation target remains the finalized the ILI activity level $(z_{i,\infty}, i = 1,2, \ldots)$.

Table S1 shows the estimation results when using the aforementioned schedule. Note that ARGO still outperforms all other alternative models. Moreover, the absolute values of all four accuracy metrics for ARGO trained this way essentially do not change compared with ARGO trained with finalized ILI activity level as studied in Table 1 of the main text, indicating the robustness of ARGO.

The weekly revisions of CDC's ILI activity reports are available at the CDC website from week 40 of the year to week 20 of the subsequent year for all seasons studied in this article. For example, ILI activity level revisions at week 50 of season 2012–2013 are available at www.cdc.gov/flu/weekly/weeklyarchives2012-2013/data/senAllregt50.htm; ILI activity report revision at week 9 of season 2014–2015 is available at www.cdc.gov/flu/weekly/weeklyarchives2014-2015/data/senAllregt09.html (the webpage has suffix "htm" for seasons before 2014–2015 and suffix "html" for 2014–2015 season). In this retrospective case study, when the revisions of ILI activity level were not available for a particular week during the off-season period, the finalized ILI activity level was used instead.

SI Variations of Google Trends Data

Google Trends historical data constantly change as a consequence of renormalizations and algorithm updates. To study the robustness of ARGO to Google Trends data revisions, we obtained the search frequencies of the search query terms identified by Google Correlate on May 22, 2010 (see Fig. S1 and Table S4) from the Google Trends website (www.google.com/trends) on 25 different days in April 2015. We studied the variability of ARGO's performance when using these 25 different versions of Google Trends data as input variables for the common time period of September 28, 2014 to March 29, 2015. We studied the 2014–2015 flu season only partially (up to March 2015) because this is the longest study period covered by all of the obtained versions of Google Trends data, at the time (May 1, 2015) of the first submission of this article. We want to emphasize that Google Correlate data were only available up to February 2014 when accessed in April 2015.

Despite the inevitable variation to the revision of the low-quality data from Google Trends, ARGO still achieves considerable stability compared with the method of Santillana et al. (16) during this time period. Table S2 suggests that ARGO is threefold more robust than the method of ref. 16. The incorporation of time series information helps ARGO achieve stability. As an extreme example, the AR(3) model focuses entirely on the time series information and is thus independent of Google Trends data revisions. GFT, formulated with the original search variables as inputs, is, by construction, insensitive to the changes in Google Trends data. For this portion of the study, we included the signal from GFT for context only, and we treat it as exogenous in our analysis. Based on the results from previous time periods, it is highly likely that if we had access to Google's internal raw data (i.e., historical search volume for disease-related phrases), we would have achieved the same stability as well. However, even with these low-quality data, ARGO outperforms GFT uniformly on all versions of data in terms of both RMSE and MAE.

Detailed Description of Google Correlate Data. Tables S3 and S4 list the search query phrases identified by Google Correlate as of March 28, 2009 and May 22, 2010, respectively. The March 2009 version included spurious terms such as "college.basketball.standings," "march.vacation," "aloha.ski," "virginia.wrestling," etc. These spurious terms did not appear in the May 2010 version.

Dynamic Coefficients for ARGO. Fig. S1 shows the coefficients for the time series and Google search terms dynamically trained by ARGO via a heat map. The level of ILI activity last week is seen to have a significant effect on the current level of ILI activity, and ILI activity half a year ago and/or 1 y ago could provide further information, as the figure shows. Among Google Correlate query terms, ARGO selected 14 terms out of 100, on average, each week.

Fig. S1. Dynamic coefficients for ARGO. Red color represents positive coefficients, blue color represents negative coefficients, white color represents zero, and gray color represents missing values. Missing values can be the result of (*i*) query terms not identified by Google Correlate and (*ii*) Google Trends data not available for particular query terms. Black horizontal dashed line separates Google query queries from autoregressive lags. Yellow vertical dashed line separates coefficients trained on Google Correlate data from those trained on Google Trends data, and green vertical dashed line separates query terms identified on March 28, 2009 from those identified on May 22, 2010.

Table S1. Comparison of different models for the estimation of influenza epidemics, with weekly CDC's ILI activity level that excludes forward-looking information from ILI activity report revision

	Whole period: Mar 29, 2009 to Jul 11, 2015	Off-season flu H1N1: Mar 29, 2009 to Dec 27, 2009	Regular flu seasons (week 40 to week 20 next year)				
			2010–2011: Oct 3, 2010 to May 22, 2011	2011–2012: Oct 2, 2011 to May 20, 2012	2012–2013: Sep 30, 2012 to May 19, 2013	2013–2014: Sep 29, 2013 to May 18, 2014	2014–2015: Sep 28, 2014 to May 17, 2015
RMSE							
ARGO	**0.565**	0.630	**0.509**	**0.608**	**0.622**	**0.298**	**0.434**
GFT (Oct 2014)	2.003	0.702	0.971	1.878	4.387	0.885	0.714
Ref. 16	0.897	0.858	0.760	1.179	1.248	0.373	0.691
GFT+AR(3)	0.825	**0.530**	0.616	0.680	1.168	0.981	0.898
AR(3)	0.963	0.805	0.986	1.136	1.087	0.946	0.931
Naive	1.000 (0.385)	1.000 (0.661)	1.000 (0.388)	1.000 (0.263)	1.000 (0.506)	1.000 (0.391)	1.000 (0.456)
MAE							
ARGO	**0.557**	0.595	**0.483**	**0.555**	**0.627**	**0.339**	**0.501**
GFT (Oct 2014)	1.465	0.670	1.093	2.026	5.082	0.747	0.787
Ref. 16	0.865	0.723	0.875	1.283	1.087	0.472	0.847
GFT+AR(3)	0.790	**0.485**	0.672	0.643	1.000	1.036	0.890
AR(3)	0.999	0.808	0.982	1.158	1.094	0.943	0.920
Naive	1.000 (0.252)	1.000 (0.494)	1.000 (0.299)	1.000 (0.218)	1.000 (0.322)	1.000 (0.253)	1.000 (0.289)
MAPE							
ARGO	**0.587**	**0.587**	**0.511**	**0.560**	**0.588**	**0.350**	**0.582**
GFT (Oct 2014)	1.350	0.603	1.163	2.163	4.827	0.688	0.906
Ref. 16	0.970	0.709	1.141	1.363	1.143	0.545	0.937
GFT+AR(3)	0.848	0.599	0.749	0.669	0.819	1.068	0.964
AR(3)	1.067	0.915	1.051	1.169	1.050	0.945	0.935
Naive	1.000 (0.129)	1.000 (0.166)	1.000 (0.126)	1.000 (0.129)	1.000 (0.123)	1.000 (0.108)	1.000 (0.095)
Correlation							
ARGO	**0.985**	0.979	**0.988**	0.911	**0.971**	**0.992**	**0.992**
GFT (Oct 2014)	0.875	**0.989**	0.968	0.833	0.926	0.969	0.986
Ref. 16	0.965	0.956	0.985	**0.937**	0.938	0.987	0.973
GFT+AR(3)	0.971	0.984	0.983	0.853	0.931	0.943	0.960
AR(3)	0.961	0.965	0.955	0.815	0.921	0.920	0.953
Naive	0.956	0.943	0.946	0.828	0.928	0.910	0.945
Correlation of increment							
ARGO	**0.742**	0.751	**0.772**	0.262	**0.633**	0.898	0.892
GFT (Oct 2014)	0.706	**0.863**	0.702	0.484	0.502	0.847	**0.918**
Ref. 16	0.625	0.680	0.719	**0.619**	0.293	**0.917**	0.837
GFT+AR(3)	0.536	0.703	0.703	0.155	0.220	0.514	0.621
AR(3)	0.420	0.562	0.554	0.067	0.106	0.360	0.549
Naive	0.455	0.552	0.556	0.162	0.247	0.345	0.586

The estimation target is the finalized CDC's ILI activity level. RMSE, MAE, and MAPE are relative to the error of the naive method. The absolute error of the naive method is reported in parentheses. Boldface highlights the best performance for each metric in each study period.

Table S2. Mean and SD of accuracy metrics when using Google Trends data accessed at different dates

	RMSE	MAE	MAPE	Correlation	Correlation of increment
Mean					
ARGO	0.226	0.304	0.079	0.981	0.831
GFT (Oct 2014)	0.262	0.366	0.089	0.985	0.920
Ref. 16	0.306	0.398	0.116	0.973	0.803
GFT+AR(3)	0.303	0.482	0.090	0.948	0.581
AR(3)	0.332	0.492	0.096	0.936	0.492
SD					
ARGO	0.013	0.017	0.005	0.002	0.016
GFT (Oct 2014)	0.000	0.000	0.000	0.000	0.000
Ref. 16	0.029	0.049	0.013	0.005	0.050
GFT+AR(3)	0.000	0.000	0.000	0.000	0.000
AR(3)	0.000	0.000	0.000	0.000	0.000

The common study period is 2014–2015 partial season (September 28, 2014 to March 29, 2015). At the time of first submitting this article, Google Correlate data covered only up to February 2014, which inspired us to study the robustness of ARGO with respect to Google Trends data variability on the 2014–2015 season.

Table S3. All search phrases identified by Google Correlate using data as of March 28, 2009

influenza.type.a	painful.cough	treatment.for.the.flu	weather.march
flu.incubation	fever.flu	basketball.standing	fevers
bronchitis	over.the.counter.flu	flu.test	duration.of.flu
influenza.contagious	pneumonia	tussionex	flu.contagious.period
flu.fever	how.long.is.the.flu	reduce.a.fever	cold.vs.flu
influenza.a	flu.how.long	how.long.is.the.flu.contagious	cure.the.flu
influenza.incubation	treatment.for.flu	treat.flu	walking.pneumonia
flu.contagious	fever.cough	spring.break.family	flu.vs..cold
treating.the.flu	flu.medicine	las.vegas.shows.march	length.of.flu
type.a.influenza	dangerous.fever	how.to.reduce.a.fever	influenza.a.and.b
symptoms.of.the.flu	high.fever	flu.or.cold	flu.and.pregnancy
influenza.symptoms	is.flu.contagious	incubation.period.for.the.flu	sinus.infections
flu.duration	normal.body	harlem.globe	influenza.treatment
flu.report	normal.body.temperature	tussin	jiminy.peak.ski
symptoms.of.flu	how.long.does.the.flu.last.	basketball.standings	baseball.preseason
influenza.incubation.period	symptoms.of.pneumonia	sinus	spring.break.date
how.to.treat.the.flu	signs.of.the.flu	upper.respiratory	indoor.driving
treat.the.flu	flu.vs.cold	get.over.the.flu	z.pack
symptoms.of.bronchitis	low.body	acute.bronchitis	college.spring.break.dates
flu.treatment	cough.fever	body.temperature	aloha.ski
symptoms.of.influenza	vegas.shows.march	college.basketball.standings	concerts.in.march
treating.flu	is.the.flu.contagious	strep	break.a.fever
flu.in.children	type.a.flu	march.weather	influenza.duration
fever.reducer	flu.treatments	getting.over.the.flu	robitussin
cold.or.flu	remedies.for.the.flu	march.vacation	virginia.wrestling

Table S4. All search phrases identified by Google Correlate using data as of May 22, 2010

influenza.type.a	get.over.the.flu	type.a.influenza	flu.care
symptoms.of.flu	treating.flu	i.have.the.flu	how.long.contagious
flu.duration	flu.vs..cold	taking.temperature	fight.the.flu
flu.contagious	having.the.flu	flu.versus.cold	reduce.a.fever
flu.fever	treatment.for.flu	bronchitis	cure.the.flu
treat.the.flu	human.temperature	how.long.flu	medicine.for.flu
how.to.treat.the.flu	dangerous.fever	flu.germs	flu.length
signs.of.the.flu	the.flu	cold.vs..flu	cure.flu
over.the.counter.flu	remedies.for.flu	flu.and.cold	exposed.to.flu
how.long.is.the.flu	influenza.a.and.b	thermoscan	low.body
symptoms.of.the.flu	contagious.flu	flu.complications	early.flu.symptoms
flu.recovery	how.long.does.the.flu.last	high.fever	remedies.for.the.flu
cold.or.flu	fever.flu	flu.children	flu.report
flu.medicine	oscillococcinum	the.flu.virus	incubation.period.for.flu
flu.or.cold	flu.remedies	how.to.treat.flu	break.a.fever
normal.body	how.long.is.flu.contagious	pneumonia	flu.contagious.period
is.flu.contagious	flu.treatments	flu.headache	influenza.incubation.period
treat.flu	influenza.symptoms	flu.cough	cold.versus.flu
body.temperature	cold.vs.flu	ear.thermometer	flu.in.children
is.the.flu.contagious	braun.thermoscan	how.to.get.rid.of.the.flu	what.to.do.if.you.have.the.flu
reduce.fever	fever.cough	flu.how.long	medicine.for.the.flu
flu.treatment	signs.of.flu	symptoms.of.bronchitis	flu.and.fever
flu.vs.cold	how.long.does.flu.last	cold.and.flu	flu.lasts
how.long.is.the.flu.contagious	normal.body.temperature	over.the.counter.flu.medicine	incubation.period.for.the.flu
fever.reducer	get.rid.of.the.flu	treating.the.flu	do.i.have.the.flu

Table S5. Comparison of different specifications of hyperparameters for in-sample study period

	Whole in-sample period: Jan 7, 2007 to Mar 29, 2009	2006–2007 partial season: Jan 7, 2007 to May 20, 2007	2007–2008 season: Sep 30, 2007 to May 18, 2008	2008–2009 partial season: Sep 28, 2008 to Mar 29, 2009
RMSE				
ARGO w/ same L_1	0.644	0.697	0.602	0.653
ARGO w/ sep. L_1	0.658	0.672	0.637	0.629
ARGO w/ same L_2	1.165	0.817	1.175	1.243
ARGO w/ sep. L_2	1.010	0.740	0.946	1.173
ARGO w/ ElasticNet	0.669	0.757	0.585	0.766
Naive	1.000 (0.316)	1.000 (0.286)	1.000 (0.473)	1.000 (0.304)
MAE				
ARGO w/ same L_1	0.678	0.651	0.584	0.634
ARGO w/ sep. L_1	0.691	0.671	0.621	0.593
ARGO w/ same L_2	1.223	0.836	1.094	1.469
ARGO w/ sep. L_2	1.149	0.753	0.943	1.401
ARGO w/ ElasticNet	0.738	0.718	0.613	0.780
Naive	1.000 (0.206)	1.000 (0.245)	1.000 (0.335)	1.000 (0.226)
Correlation				
ARGO w/ same L_1	0.987	0.977	0.983	0.977
ARGO w/ sep. L_1	0.986	0.980	0.980	0.976
ARGO w/ same L_2	0.969	0.984	0.976	0.955
ARGO w/ sep. L_2	0.979	0.987	0.983	0.967
ARGO w/ ElasticNet	0.987	0.984	0.986	0.975
Naive	0.965	0.949	0.950	0.935
Correlation of increment				
ARGO w/ same L_1	0.779	0.643	0.857	0.646
ARGO w/ sep. L_1	0.708	0.545	0.758	0.697
ARGO w/ same L_2	0.828	0.793	0.864	0.799
ARGO w/ sep. L_2	0.845	0.795	0.881	0.824
ARGO w/ ElasticNet	0.814	0.835	0.852	0.738
Naive	0.623	0.473	0.756	0.322

"ARGO w/ same L_1" is ARGO with the same L_1 penalty for Google search terms and autoregressive lags (specification 1). "ARGO w/ sep. L_1" is ARGO with separate L_1 penalties for Google search terms and autoregressive lags (specification 2). "ARGO w/ same L_2" is ARGO with the same L_2 penalty for Google search terms and autoregressive lags (specification 3). "ARGO w/ sep. L_2" is ARGO with separate L_2 penalties for Google search terms and autoregressive lags (specification 4). "ARGO w/ ElasticNet" is ARGO with the same elastic net penalty for Google search terms and autoregressive lags (specification 5). The first column is for the entire in-sample study period. The second column is for 2006–2007 partial season; 2006–2007 full season is not available because data before January 2007 are used for training. The third column is for 2007–2008 full season. The fourth column is for 2008–2009 partial season; 2008–2009 full season is not available because our out-of-sample study period starts in April 2009. RMSE and MAE are relative to the error of the naive method. The absolute error of the naive method is reported in parentheses.