

BANCA
NAȚIONALĂ
A ROMÂNIEI

Caiete de studii Nr. 65

Anca Gălățescu

De la căutările online la prognoze
macroeconomice:
Ce spun datele *Google Trends*
despre starea actuală a economiei
via metode de *machine learning*

CAIETE DE STUDII

Nr. 65

August 2025

NOTĂ

Opiniile prezentate în această lucrare aparțin în întregime autorului și nu implică sau angajează în vreun fel Banca Națională a României.

Toate drepturile rezervate. Reproducerea informațiilor este permisă numai în scopuri educative și necomerciale și numai cu indicarea sursei.

ISSN 1584-0883 (versiune online)

De la căutările online la prognoze
macroeconomice:
Ce spun datele *Google Trends*
despre starea actuală a economiei
via metode de *machine learning*

Anca Gălățescu

Cuprins

Rezumat	7
1. Introducere	9
2. Datele	11
3. Modelele	13
4. Rezultatele	15
5. Concluzii	18
Bibliografie	19
Anexă	21

Rezumat

Această lucrare analizează potențialul beneficiu adus de setul de date *Google Trends* pentru prognoza în timp real a PIB al României în contextul utilizării unor tehnici de reducere a dimensionalității care păstrează o cantitate suficientă de informație pentru prognoză (*shrinkage methods*). Analiza arată că utilizarea metodelor de reducere a dimensionalității setului de date și includerea setului de date *Google Trends* pentru prognozele pe termen scurt ale PIB al României, poate conduce la performanțe predictive superioare situației în care sunt utilizate doar seturile tradiționale de date și, respectiv, doar setul de date *Google Trends*. În linie cu rezultatele din literatură, în special cele cu referire la cazul statelor dezvoltate, aceste date apar a fi utile în principal pentru proiecțiile aferente perioadei curente (*nowcasts*), când observațiile indicatorilor la frecvență mai redusă nu sunt încă disponibile.

Cuvinte-cheie: Nowcasting, Google data, Shrinkage

Coduri de clasificare JEL: C55, C81, E01

1. Introducere

Având în vedere prezența în creștere a internetului în ultimii ani în viața cotidiană, devine naturală întrebarea dacă se pot extrage informații referitoare la activitatea economică din căutările personale și profesionale. Investigarea proprietăților predictive ale căutărilor Google a început mai recent să capete atenție în literatură. În plus, pe măsură de acest tip de date de mari dimensiuni a început să fie colectat în mod sistematic, s-au dezvoltat și tehnicile de analizare a acestora. Căutările Google pot oferi o privire mai în profunzime cu privire la planurile consumatorilor și, mai extins, ale agenților economici, în principal referitoare la cheltuieli. În acest context, căutările Google au un rol complementar rezultatelor din sondajele elaborate de DG ECFIN care au o structură predefinită a chestionarelor.

Lucrările care au deschis calea cercetării în domeniul utilizării căutărilor Google în scopul obținerii proiecțiilor altor variabile economice aparțin lui Choi și Varian (2009 și 2012). Autorii arată că folosirea datelor Google contribuie la îmbunătățirea calității proiecțiilor indemnizațiilor de șomaj, respectiv vânzărilor cu amănuntul, aferente perioadei curente/prezente (*nowcasting*).

Ulterior, datele privind căutările Google au început să fie folosite tot mai mult pentru elaborarea de prognoze sectoriale, încorporarea acestora indicând în general sporirea acurateței predicțiilor. D'Amuri și Marcucci (2017) folosesc serii de căutări Google pentru locuri de muncă în scopul obținerii proiecțiilor ratei șomajului din SUA pe baza unor modele liniare și neliniare, detectând îmbunătățiri ale capacității predictive pentru orizonturile de proiecție de una până la trei luni. Nymand-Andersen și Pantelidis (2018) explorează potențialul datelor referitoare la căutările utilizatorilor pe internet ca indicatori timpurii ai evoluțiilor macroeconomice. În particular, autorii testează capacitatea predictivă a căutărilor privind „auto și vehicule” pentru prognoza vânzărilor de mașini noi în zona euro, relevând de asemenea performanțe superioare când aceste date sunt incluse în analiză.

Folosirea surselor non-standard de date s-a dovedit utilă pentru evaluarea situației conjuncturale a economiei și în cazul unor perioade în care comportamentul consumatorilor se schimbă brusc, așa cum a fost cazul la declanșarea pandemiei COVID-19. Austin *et al.* (2021) subliniază utilitatea datelor referitoare la căutările Google pentru aprecierea stării afacerilor în diferite orașe pe parcursul pandemiei. În plus, autorii propun o metodologie de transformare a datelor *Google Trends* care să se suprapună cât mai bine cu cea a definițiilor din conturile naționale, cu scopul de a extinde instrumentarul metodologic pentru institutele de statistică interesate de obținerea cât mai timpurie a unor indicatori economici.

În ceea ce privește în mod particular proiecțiile pentru activitatea economică de ansamblu pe baza căutărilor utilizatorilor în motorul Google, Goetz și Knetsch (2019) folosesc aceste date pentru prognoza PIB al Germaniei cu ajutorul modelelor *Bridge Equation* și detectează îmbunătățiri ale performanței predictive când aceste date sunt alăturate surselor mai tradiționale de date, fiind o posibilă alternativă la datele din sondaje.

Kohns și Bhattacharjee (2023) utilizează datele *Google Trends* și modele bayesiene pentru prognoza pe termen scurt a PIB din SUA. Rezultatele indică îmbunătățirea acurateței prognozelor în cazul în care este luat în considerare și setul de date GT și că acesta poate fi util în detectarea punctelor de turnură în activitatea economică.

Bantis *et al.* (2023) analizează în ce măsură seturile de date de mari dimensiuni, cum sunt cele furnizate de Google, conțin informație predictivă superioară seturilor mai tradiționale de date disponibile cu frecvență ridicată, pentru cazul Braziliei și al SUA pe baza unui model cu factori comuni dinamici. Rezultatele arată că prognozele bazate pe setul de date care conține indicatori tradiționali și datele Google performează cel mai bine comparativ cu etalonul, un model de autoregresie vectorială de ordinul 1. Autorii remarcă, de asemenea, că preselecția variabilelor folosite în prognoză conduce la performanțe predictive superioare comparativ cu situația utilizării întregului set de date.

Ferrara și Simoni (2023) investighează în ce context sunt utile datele *Google Trends* pentru prognoza PIB din zona euro pe baza unei tehnici de reducere a dimensionalității setului de variabile. Îmbunătățiri ale calității prognozelor folosind acest set de date tind să se remarce pentru prognozele *nowcast*.

Tot cu referire la prognozele *nowcast*, Dauphin *et al.* (2022) aplică modele cu factori comuni dinamici și, respectiv, anumite specificații de modele de *machine learning* pentru anumite state din Europa cu scopul elaborării pentru perioada curentă. Autorii constată că specificațiile de modele de *machine learning* performează semnificativ mai bine comparativ cu modelele cu factori comuni dinamici, fiind utile mai ales în perioadele de turnură economică, așa cum a fost pandemia COVID-19.

Criza pandemică a accentuat în particular nevoia de a obține în timp foarte scurt informații despre starea economiei și sectoarele afectate, în special când încă nu sunt disponibile datele la frecvența lunară sau trimestrială. Scopul acestei analize este de a explora potențialele avantaje ale progreselor în domeniul *machine learning*¹ utilizate în contextul în care au devenit disponibile seturi de date de mari dimensiuni (*big data*) și la frecvențe mai ridicate. Concret, este analizat potențialul beneficiu adus de setul de date *Google Trends* furnizat de motorul de căutare Google pentru prognoza PIB al României în contextul utilizării unor tehnici de reducere a dimensionalității (*shrinkage methods*) care păstrează o cantitate suficientă de informație pentru prognoză.

Prima contribuție a acestei lucrări este încorporarea în prognoze a unor informații timpurii oferite de căutările termenilor pe motorul Google care au potențialul de a surprinde modificări bruște în comportamentul agenților economici și al consumatorilor, oferind suport decidenților de politică.

¹ *Machine learning* este definit în sens larg ca o combinație de algoritmi și metode statistice utilizate pentru a descoperi („învăța”) tipare din seturile de date (care pot fi numerice, imagini, cuvinte etc.) și pentru a elabora prognoze.

A doua contribuție se referă la implementarea tehnicilor de reducere a dimensionalității (*shrinkage methods*) seturilor mari de date care păstrează suficientă informație relevantă pentru obținerea prognozelor variabilelor macroeconomice.

Având în vedere revizuirile la care sunt supuse majoritatea datelor statistice, a treia contribuție a acestei lucrări se referă la folosirea bazei de date în timp real, ceea ce dă posibilitatea evaluării prognozelor realizate cu setul de date disponibil analiștilor la momentul realizării fiecărei prognoze.

Analiza arată că utilizarea metodelor de reducere a dimensionalității setului de date și includerea setului de date *Google Trends* pentru prognozele pe termen scurt ale PIB al României pot conduce la performanțe predictive superioare situației în care sunt utilizate doar seturile tradiționale de date și, respectiv, doar setul de date *Google Trends*. În linie cu rezultatele din literatură, în special cele cu referire la cazul statelor dezvoltate, aceste date par a fi utile în principal pentru proiecțiile aferente perioadei curente (*nowcasts*), când observațiile indicatorilor la frecvență mai redusă nu sunt încă disponibile.

Lucrarea este structurată după cum urmează. Secțiunea 2 prezintă mai multe informații referitoare la seturile de date folosite în analiză și specificul datelor *Google Trends*. Secțiunea 3 detaliază modelele de *machine learning* aplicate. Secțiunea 4 descrie modul de elaborare a proiecțiilor și rezultatele obținute. Secțiunea 5 prezintă concluziile.

2. Datele

Pentru a analiza impactul încorporării datelor *Google Trends* (GT) sunt folosite trei seturi de date: (i) setul care cuprinde indicatori *hard* și *soft*, (ii) setul de date *Google Trends* și (iii) setul de date extins, care pune laolaltă primele două seturi.

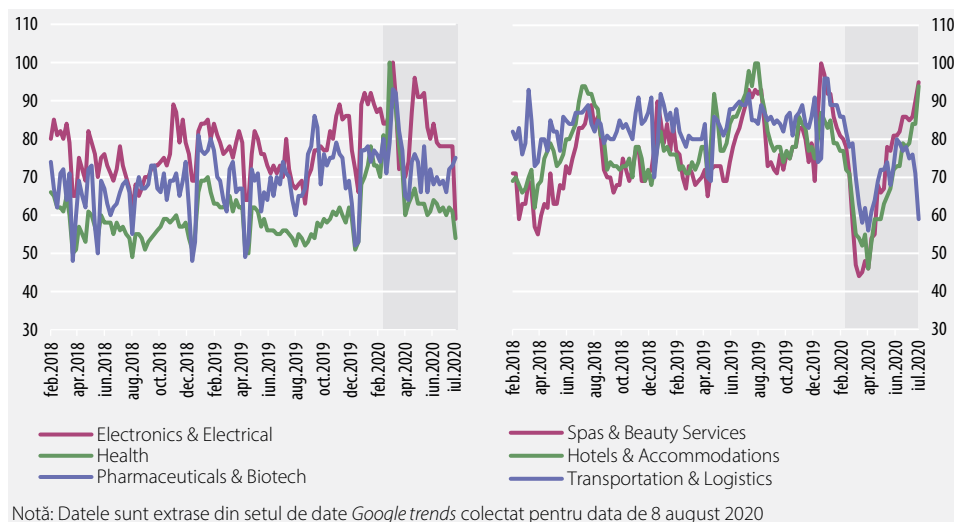
Setul de date standard cuprinde 99 de indicatori folosiți în mod obișnuit în analize și se referă la producția industrială, servicii, prețuri de producție și consum, variabile financiare, internaționale și date din sondajele elaborate de DG ECFIN. Setul de date *Google Trends* pentru România încorporează 293 de serii pe categorii de căutări, furnizate de Google. Al treilea set de date cuprinde totalul celorlalte două, respectiv 392 de indicatori.

Datele *Google Trends*² reprezintă căutările unor cuvinte-cheie grupate pe categorii și măsoară căutările pentru o anumită categorie relativ la totalul căutărilor, pentru anumite zone geografice. În acest fel pot fi identificate modificările relative ale

² Mai multe detalii privind modul de construire a datelor *Google Trends* se găsesc la: <https://newsinitiative.withgoogle.com/resources/trainings/google-trends-understanding-the-data/>

căutărilor pentru o anumită categorie și nu volumul absolut al căutărilor. Căutărilor pe categorii sunt bazate pe eșantioane aleatorii din toate căutărilor Google din fiecare zi și nu sunt revizuite, dar pot apărea modificări ale seriilor istorice atunci când apar actualizări din partea Google. GT sunt agregate de Google la frecvență săptămânală, fiind disponibile pe eșantioane mobile (*rolling sample*) de 5 ani începând cu jumătatea lunii iunie 2020.

Graficul 1. Evoluția seriilor de timp referitoare la o selecție de căutări Google



Pentru perioada de la începutul crizei pandemice, datele privind anumite categorii de căutări Google au suferit modificări semnificative, reflectând interesul populației și putând oferi informații cu privire la evoluția unor preocupări ale acestora în funcție de noul context generat de condițiile de mobilitate. În Graficul 1 sunt prezentate evoluțiile câtorva serii de date de căutări Google care ar putea furniza indicații foarte timpurii privind schimbarea de foarte mare amploare a interesului consumatorilor și implicit asupra evoluției conjuncturale a activității economice pe anumite sectoare de activitate. Din aceste date se remarcă sporirea semnificativă la începutul perioadei de închidere a economiilor, a preocupărilor pentru aspecte precum produse electronice, sănătate sau produse farmaceutice și biotehnologie care au crescut brusc la începutul lunii martie 2020, în unele cazuri reprezentând un nivel maxim al interesului pentru respectivele categorii (Grafic 1, dreapta). În același timp, căutărilor pentru categorii ca servicii de frumusețe, hoteluri și cazare sau transport și logistică s-au prăbușit la niveluri minime ale eșantionului de date în luna martie 2020, revenind ulterior rapid, inclusiv în anticiparea unor relaxări relative a restricțiilor de mobilitate (Grafic 1, stânga).

Pentru a evita potențiala volatilitate excesivă aferentă observațiilor săptămânale ale GT, în cadrul estimărilor din prezenta analiză acestea sunt agregate la frecvență lunară corespunzătoare setului standard de variabile explicative. Datele la frecvență lunară sunt apoi agregate la frecvență trimestrială corespunzătoare variabilei pentru care se dorește obținerea prognozei, respectiv PIB. Lista detaliată a indicatorilor și transformărilor aferente este prezentată în Anexa 1.

Evaluarea acurateții modelelor de prognoză este de tip *real-time rolling window*, fiind realizată pentru perioada T4 2020 – T4 2022, fiind impusă de momentul de la care este disponibil setul de date GT. Performanța predictivă este analizată în timp real, ceea ce înseamnă că fiecare prognoză este bazată pe setul de date existent la momentul la care a fost elaborată fiecare prognoză. Compararea prognozelor se realizează față de setul de date „Provizoriu 2” publicat de INS și, respectiv, cu setul de date disponibil la data de 17 aprilie 2023. Dat fiind tiparul în care sunt disponibile datele GT, precum și având în vedere presupunerea că acestea devin tot mai relevante, estimarea coeficienților se bazează pe eșantioane mobile, începând cu cel corespunzător perioadei T2 2015 – T1 2020, ultimul fiind cel corespunzător perioadei T1 2018 – T4 2022. Modalitatea de obținere a prognozelor în timp real permite reestimarea coeficienților și/sau reselectarea variabilelor la fiecare moment la care se realizează prognozele, astfel încât să poată fi surprinse, într-o oarecare măsură, eventualele modificări structurale dintre variabila explicată și ceilalți indicatori.

Pentru fiecare trimestru este elaborată o secvență de 14 prognoze bilunare, începând de la jumătatea trimestrului anterior și până la sfârșitul primei luni a trimestrului următor, înainte de publicarea datelor semnal de către INS. Momentele de realizare a fiecărei prognoze țin cont de publicarea datelor referitoare la producția industrială și comerț, care are loc până la mijlocul fiecărei luni, respectiv de publicarea datelor financiare și din sondaje, care are loc până la sfârșitul fiecărei luni.

3. Modelele

Având în vedere particularitatea setului de date ales, de a avea comparativ mult mai puține observații în timp (corespunzătoare a 5 ani de date *Google Trends* disponibile) față de numărul foarte mare de indicatori explicativi, este necesară folosirea unor tehnici care să reducă dimensionalitatea setului de date, dar care să păstreze o cantitate suficientă de informații relevante pentru prognoză.

Metodele mai tradiționale de abordare a dimensionalității setului de date au în vedere o extragere a componentelor comune ale acestora. Mai recent, metodele de reducere a dimensionalității (*shrinkage methods*) folosesc ideea diminuării sau chiar anulării impactului anumitor variabile, chiar dacă estimatorii sunt *biased*, astfel încât să implice erori de prognoză mai reduse.

În această lucrare, sunt aplicate trei tipuri de modele din categoria *shrinkage methods*, respectiv regresii Ridge, LASSO și *Elastic Net*. Regresia de tip Ridge propusă de Hoerl și Kennard (1970) are particularitatea de a restricționa coeficienții variabilelor cu contribuții minore la model să se apropie de zero, fără a fi însă exact zero. Estimatorii obținuți pe baza modelelor simple OLS pot fi instabili în cazul în care numărul variabilelor explicative este mai mare decât numărul de observații disponibile. În acest

caz, unul dintre remedii este ca funcția de cost care trebuie minimizată să includă un termen care penalizează coeficienții „mari” și este de forma:

$$S(\beta) = (y - X\beta)'(y - X\beta) + \lambda\beta'\beta$$

Parametrul de regularizare λ controlează gradul în care coeficienții sunt constrânși a se apropia de zero.

Literatura recentă în domeniul proiecțiilor pe termen scurt și *nowcasting* a arătat că folosirea ca atare a unor seturi de date de mari dimensiuni nu conduce neapărat la cele mai bune performanțe predictive. În acest context, folosirea tuturor categoriilor de căutări nu implică neapărat obținerea celor mai bune rezultate. Pentru a evita această situație, în estimări este implementată o preselecție a acestora, fiind utilizate doar acele categorii care prezintă cele mai ridicate corelații cu variabila care se dorește a fi prognozată. Concret, sunt utilizate doar acele serii de date pentru care corelația cu variabila dependentă este de cel puțin 80 la sută, rezultând astfel un set mai restrâns de indicatori care se folosesc efectiv în prognoză. În utilizarea practică a acestor modele se poate avea în vedere utilizarea unui coeficient mai redus de corelație, care să ofere un echilibru între numărul de variabile explicative și performanța predictivă. Această abordare de preselecție a variabilelor este folosită în general în cazul în care setul de date explicative conține multe variabile, de exemplu, pentru modelele cu factori comuni dinamici de Bai și Ng (2008) și în cazul regresiei de tip Ridge de Ferrara și Simoni (2023).

Pentru fiecare eșantion folosit la evaluarea acurateții prognozelor, având în vedere numărul foarte redus de observații de timp disponibile (care face foarte dificilă aplicarea metodelor utilizate în mod obișnuit de *cross-validation*) parametrul de regularizare λ este estimat pe baza tehnicii *generalized cross-validation* (folosită de exemplu de Carrasco și Rossi, 2016). Aceasta presupune selectarea parametrului λ care minimizează expresia:

$$S_T = \frac{T^{-1}(\sum_{t=1}^T Y_t - X_t' \beta_{Ridge})^2}{(1 - T^{-1} \text{tr}(R_t(\lambda)))^2}, \text{ unde}$$

$$R_t(\lambda) = X(T^{-1} \sum_{t=1}^T X_t X_t' + \lambda I)^{-1} T^{-1} \sum_{t=1}^T X_t'.$$

Concret, parametrul de regularizare este selectat din valorile din interiorul șirului de la 0,1 la 1 cu un pas de 0,1 care minimizează valoarea indicatorului S_T .

Regresia LASSO (*least absolute shrinkage and selection operator*) a fost propusă de Tibshirani (1996) și reprezintă o alternativă a regresiei Ridge în care coeficienții variabilelor mai puțin importante pentru model sunt restricționați să aibă valoarea zero. Această abordare permite practic și o selecție a variabilelor care sunt folosite în model. Funcția de cost care trebuie minimizată este de forma:

$$S(\beta) = (y - X\beta)'(y - X\beta) + \lambda|\beta|$$

Acest tip de penalizare determină coeficienții pe baza unui model neliniar și, spre deosebire de regresia Ridge, aceștia nu au o formă analitică.

Parametrul de regularizare λ este determinat pe baza unui criteriu informațional adaptat pentru eșantioane reduse de serii de timp, așa cum este propus de Li (1987). Parametrul este selectat dintre valorile din interiorul șirului de la 0,1 la 0,9 cu un pas de 0,1 care minimizează valoarea criteriului informațional Akaike (corectat, de asemenea, pentru situația eșantioanelor scurte de date). După etapa de selectare a variabilelor care urmează a fi utilizate efectiv în estimare, coeficienții sunt estimați pe baza unei regresii liniare simple, conform abordării propuse de Belloni și Chernozhukov (2013).

Regresia *Elastic Net* combină restricțiile aplicate coeficienților în primele două cazuri menționate mai sus și a fost propusă de Zhou și Hastie (2005). Funcția de cost care trebuie minimizată este de forma:

$$S(\beta) = (y - X\beta)'(y - X\beta) + \lambda_1\beta'\beta + \lambda_2|\beta|$$

Parametrii de regularizare λ_1 și λ_2 sunt determinați pe baza aceluiași procedeu propus de Li (1987), ca și în cazul regresiei LASSO.

4. Rezultatele

Din perspectiva setului de date care oferă o acuratețe mai ridicată a prognozelor elaborate în timp real comparativ cu setul de date „Provizoriu 2” publicat de INS, în Tabelul 1 este prezentată evoluția relativă a indicatorului *root mean square forecast error* (RMSFE) în funcție de fiecare moment al realizării prognozelor. Evoluția indicatorului RMSFE pentru diferitele specificații de modele care folosesc setul extins de date (format din setul folosit în mod obișnuit în analize și cel de date *Google Trends*) comparativ cu cele care folosesc setul standard de date indică anumite câștiguri privind acuratețea prognozelor, în special cu referire la perioada curentă. Sporirea acurateței prognozelor la aproape toate orizonturile de proiecție este însă mult mai evidentă în cazul utilizării setului complet de date (format din setul folosit în mod obișnuit în analize și cel de date *Google Trends*) comparativ cu cel compus doar din datele Google. Acest mesaj se desprinde în cazul fiecărei specificații de model avute în vedere, cât și pentru proiecțiile combinate cu ponderi egale între cele trei specificații de modele. Rezultatele sugerează, în linie cu cele obținute în literatură, că datele Google sunt utile mai ales ca un complement al datelor care se folosesc în mod tradițional în analizele băncilor centrale.

Tabelul 1. Indicatori privind acuratețea relativă a prognozelor utilizând cele trei variante de seturi de date, comparație cu setul de date „Provizoriu 2”

RMSFE of extended dataset relative to standard dataset	Ridge	LASSO	E_net	Combined - equal weights	RMSFE of extended dataset relative to Google Trends dataset	Ridge	LASSO	E_net	Combined - equal weights
Timing of forecast					Timing of forecast				
Forecast Qt-1, mid-M1	1,04	1,01	1,01	1,09	Forecast Qt-1, mid-M1	0,73	0,86	0,91	0,83
Forecast Qt-1, end-M1	1,13	1,01	1,16	1,13	Forecast Qt-1, end-M1	0,82	0,95	0,88	0,90
Forecast Qt-1, mid-M2	0,94	1,05	0,90	0,98	Forecast Qt-1, mid-M2	0,99	1,01	0,96	0,99
Forecast Qt-1, end-M2	0,94	1,04	1,11	1,08	Forecast Qt-1, end-M2	0,98	1,06	1,24	1,11
Forecast Qt-1, mid-M3	0,96	1,01	0,98	1,01	Forecast Qt-1, mid-M3	0,81	0,94	0,94	0,89
Forecast Qt-1, end-M3	0,98	1,02	0,96	1,02	Forecast Qt-1, end-M3	0,71	0,78	0,60	0,70
Nowcast Qt, mid-M1	1,01	1,03	1,02	1,05	Nowcast Qt, mid-M1	0,94	0,90	1,16	1,08
Nowcast Qt, end-M1	0,76	0,99	1,04	0,97	Nowcast Qt, end-M1	0,26	0,70	0,43	0,42
Nowcast Qt, mid-M2	1,14	1,00	1,15	1,11	Nowcast Qt, mid-M2	0,44	0,63	0,80	0,62
Nowcast Qt, end-M2	0,95	1,05	0,95	0,99	Nowcast Qt, end-M2	0,35	0,61	0,61	0,50
Nowcast Qt, mid-M3	0,94	1,03	1,02	1,01	Nowcast Qt, mid-M3	0,29	0,54	0,46	0,43
Nowcast Qt, end-M3	1,03	1,09	0,92	1,02	Nowcast Qt, end-M3	0,30	0,51	0,52	0,42
Backcast Qt+1, mid-M1	1,03	1,09	1,00	1,07	Backcast Qt+1, mid-M1	0,31	0,45	0,46	0,40
Backcast Qt+1, end-M1	0,88	0,99	1,07	1,02	Backcast Qt+1, end-M1	0,99	1,06	0,99	1,15

Notă: Valorile subunitare indică o performanță predictivă superioară comparativ cu etalonul.

Tabelul 2. Indicatori privind acuratețea relativă a prognozelor pe baza fiecărei specificații de model și seturi de date, comparație cu setul de date „Provizoriu 2”

	Standard dataset		GT dataset		Extended dataset	
Timing of forecast	RMSFE of LASSO vs. Ridge	RMSFE of E_net vs. Ridge	RMSFE of LASSO vs. Ridge	RMSFE of E_net vs. Ridge	RMSFE of LASSO vs. Ridge	RMSFE of E_net vs. Ridge
Forecast Qt-1, mid-M1	1,13	1,14	0,93	0,90	1,10	1,11
Forecast Qt-1, end-M1	1,24	1,07	0,95	1,02	1,11	1,10
Forecast Qt-1, mid-M2	0,98	1,09	1,06	1,06	1,09	1,04
Forecast Qt-1, end-M2	0,94	1,03	0,96	0,95	1,04	1,20
Forecast Qt-1, mid-M3	1,14	1,18	1,03	1,03	1,20	1,20
Forecast Qt-1, end-M3	1,03	1,05	0,98	1,22	1,07	1,03
Nowcast Qt, mid-M1	1,09	1,12	1,15	0,93	1,10	1,13
Nowcast Qt, end-M1	0,84	0,90	0,41	0,74	1,10	1,23
Nowcast Qt, mid-M2	1,00	0,99	0,61	0,55	0,88	1,00
Nowcast Qt, end-M2	0,73	0,96	0,46	0,55	0,80	0,96
Nowcast Qt, mid-M3	0,78	0,94	0,46	0,65	0,86	1,02
Nowcast Qt, end-M3	0,85	1,09	0,53	0,56	0,89	0,98
Backcast Qt+1, mid-M1	0,90	1,08	0,65	0,71	0,95	1,05
Backcast Qt+1, end-M1	0,97	0,91	1,02	1,10	1,09	1,10

Notă: Valorile subunitare indică o performanță predictivă superioară comparativ cu etalonul.

În ceea ce privește tipul de modele care oferă prognoze mai precise, în Tabelul 2 sunt prezentate valorile relative ale indicatorului RMSFE, ale prognozelor bazate pe modelele LASSO și, respectiv, *Elastic Net* comparativ cu Ridge, pentru toate cele trei seturi de date avute în vedere. Rezultatele indică faptul că specificațiile LASSO și *Elastic Net* tind să producă *nowcasts* mai precise comparativ cu Ridge în cazul tuturor seturilor de date analizate. Pentru orizontul de prognoză de un trimestru, modelele Ridge oferă însă o acuratețe a proiecțiilor superioară celor pe baza specificațiilor LASSO și *Elastic Net*. În linie și cu alte rezultate din literatură, analiza performanței predictive a unei suite de specificații de modele oferă indicii analiștilor cu privire la tipul de model care tinde să aibă o performanță predictivă superioară în funcție de conjunctura economică și de momentul realizării prognozelor.

Pentru comparația cu setul de date „definitive” disponibil la data de 17 aprilie 2023, acuratețea prognozelor este superioară în cazul utilizării setului extins de date comparativ cu setul de date standard și, în special față de cazul setului GT, pentru marea majoritate a orizonturilor de proiecție. Câștigurile privind acuratețea prognozelor sunt mai substanțiale în cazul *nowcasts* comparativ cu proiecțiile pentru orizontul de un trimestru (Tabelul 3).

Tabelul 3. Indicatori privind acuratețea relativă a prognozelor utilizând cele trei variante de seturi de date, comparație cu setul de date „definitive” disponibil la data de 17 aprilie 2023

RMSFE of extended dataset relative to standard dataset	Ridge	LASSO	E_net	Combined - equal weights	RMSFE of extended dataset relative to Google trends dataset	Ridge	LASSO	E_net	Combined - equal weights
Timing of forecast					Timing of forecast				
Forecast Qt-1, mid-M1	0,57	1,01	1,00	1,05	Forecast Qt-1, mid-M1	0,55	0,75	0,84	0,72
Forecast Qt-1, end-M1	0,87	1,01	1,47	1,25	Forecast Qt-1, end-M1	0,67	0,82	0,75	0,78
Forecast Qt-1, mid-M2	0,66	1,07	0,77	0,85	Forecast Qt-1, mid-M2	0,85	1,02	0,90	0,95
Forecast Qt-1, end-M2	0,66	1,07	1,00	0,92	Forecast Qt-1, end-M2	0,99	1,12	1,45	1,24
Forecast Qt-1, mid-M3	0,74	1,00	0,97	0,92	Forecast Qt-1, mid-M3	1,01	0,81	0,90	0,88
Forecast Qt-1, end-M3	0,71	1,01	1,01	0,91	Forecast Qt-1, end-M3	0,59	0,59	0,48	0,55
Nowcast Qt, mid-M1	0,72	1,03	0,95	0,93	Nowcast Qt, mid-M1	0,70	0,82	0,90	0,96
Nowcast Qt, end-M1	0,53	0,96	0,96	0,83	Nowcast Qt, end-M1	0,15	0,47	0,29	0,25
Nowcast Qt, mid-M2	0,98	0,98	1,13	1,05	Nowcast Qt, mid-M2	0,27	0,41	0,47	0,37
Nowcast Qt, end-M2	0,88	1,16	0,98	0,99	Nowcast Qt, end-M2	0,25	0,42	0,45	0,34
Nowcast Qt, mid-M3	0,91	1,14	1,07	1,07	Nowcast Qt, mid-M3	0,21	0,35	0,31	0,29
Nowcast Qt, end-M3	1,11	1,28	0,94	1,13	Nowcast Qt, end-M3	0,23	0,31	0,35	0,28
Backcast Qt+1, mid-M1	1,12	1,21	0,97	1,17	Backcast Qt+1, mid-M1	0,22	0,28	0,33	0,27
Backcast Qt+1, end-M1	0,66	0,97	1,05	0,94	Backcast Qt+1, end-M1	0,56	1,16	0,86	0,96
Notă: Valorile subunitare indică o performanță predictivă superioară comparativ cu etalonul.									

Referitor la specificațiile care oferă prognoze mai precise, LASSO și *Elastic Net* tind să aibă performanțe predictive superioare specificației Ridge la toate orizonturile de prognoză și pentru toate cele trei seturi de date considerate (Tabelul 4).

Tabelul 4. Indicatori privind acuratețea relativă a prognozelor pe baza fiecărei specificații de model și seturi de date, comparație cu setul de date „definitive” disponibil la data de 17 aprilie 2023

Timing of forecast	Standard dataset		GT dataset		Extended dataset	
	RMSFE of LASSO vs. Ridge	RMSFE of E_net vs. Ridge	RMSFE of LASSO vs. Ridge	RMSFE of E_net vs. Ridge	RMSFE of LASSO vs. Ridge	RMSFE of E_net vs. Ridge
Forecast Qt-1, mid-M1	0,75	0,79	0,98	0,90	1,33	1,38
Forecast Qt-1, end-M1	1,05	0,73	0,99	1,10	1,22	1,23
Forecast Qt-1, mid-M2	0,75	0,95	1,01	1,05	1,22	1,12
Forecast Qt-1, end-M2	0,57	0,83	0,82	0,85	0,93	1,26
Forecast Qt-1, mid-M3	0,69	0,74	1,14	1,09	0,92	0,97
Forecast Qt-1, end-M3	0,62	0,70	0,89	1,22	0,89	1,00
Nowcast Qt, mid-M1	0,78	0,87	0,96	0,89	1,13	1,15
Nowcast Qt, end-M1	0,56	0,71	0,33	0,68	1,02	1,28
Nowcast Qt, mid-M2	0,75	0,86	0,49	0,56	0,75	0,99
Nowcast Qt, end-M2	0,50	0,87	0,38	0,53	0,65	0,97
Nowcast Qt, mid-M3	0,52	0,91	0,39	0,72	0,66	1,07
Nowcast Qt, end-M3	0,53	1,03	0,46	0,57	0,61	0,87
Backcast Qt+1, mid-M1	0,68	1,14	0,58	0,67	0,73	0,99
Backcast Qt+1, end-M1	0,75	0,73	0,54	0,75	1,11	1,16

Notă: Valorile subunitare indică o performanță predictivă superioară comparativ cu etalonul.

5. Concluzii

În această lucrare a fost analizat potențialul beneficiu al utilizării datelor *Google Trends* pentru proiecțiile pe termen scurt ale PIB în cazul României. În acest context, obiectivul a fost analizarea performanței predictive în timp real (*real-time rolling window*) a unor specificații de model care să permită reducerea dimensionalității setului de date care prezintă un număr considerabil mai mare de variabile explicative comparativ cu numărul de observații.

Rezultatele arată că setul de date compus din indicatorii standard și datele *Google Trends* tinde să ofere prognoze mai precise decât cele elaborate doar pe baza setului standard de indicatori și, respectiv, doar pe baza datelor *Google Trends*. Tehnicile de reducere a dimensionalității tind în general să fie mai precise pentru prognozele realizate pentru trimestrul în curs. În ceea ce privește tipurile de modele, specificațiile LASSO și *Elastic Net* tind să ofere prognoze mai precise decât specificația Ridge pentru marea majoritate a orizonturilor de prognoză.

Direcțiile viitoare de analiză au în vedere evaluarea acurateței predictive a acestor tipuri de date și modele cu frecvențe mixte care implică utilizarea reprezentării *state-space*, precum și comparații ale acurateței predictive cu alte entități, cum ar fi zona euro.

Bibliografie

- Austin, P.,
Marini, M.,
Sanchez, A.,
Simpson-Bell, C.
Tebrake, J. „Using the Google Places API and Google Trends data to develop high frequency indicators of economic activity“, *IMF Working Paper* No. 21/295, 2021
- Bai, J.,
Ng, S. „Large dimensional factor analysis“, *Foundations and Trends in Econometrics*, 3(2), pp. 89-163, 2008
- Bantis, E.,
Clements, M. P.,
Urquhart, A., „Forecasting GDP growth rates in the United States and Brazil using Google Trends“, *International Journal of Forecasting*, Elsevier, Volume 39(4), pp. 1909-1924, 2023
- Belloni, A.,
Chernozhukov, V. „Least squares after model selection in high-dimensional sparse models“, *Bernoulli*, Volume 19, no. 2, pp. 521-547, 2013
- Choi, H.,
Varian, H. „Predicting initial claims for unemployment benefits“, *Google Technical Report*, 2009
- „Predicting the present with Google Trends“, *Google Technical Report*, 2012
- Carrasco, M.,
Rossi, B. In-sample inference and forecasting in misspecified factor models, *Journal of Business & Economic Statistics*, Volume 34(3), pp. 313-338, 2016
- D'Amuri, F.,
Marcucci, J. „The Predictive power of Google searches in forecasting US unemployment“, *International Journal of Forecasting*, 33, issue 4, pp. 801-816, 2017
- Dauphin, J.-F.,
Dybczak, K.,
Maneely, M.,
Sanjani, M. T.,
Suphaphiphat, N.,
Wang, Y.,
Zhang, H. Nowcasting GDP: A scalable approach using DFM, machine learning and novel data, applied to European economies, *IMF Working Paper* No. 22/52, 2022
- Ferrara, L.,
Simoni, A. „When are Google data useful to nowcast GDP? An approach via preselection and shrinkage“, *Journal of Business and Economic Statistics*, Volume 41, issue 4, pp. 1188-1202, 2023
- Goetz, T. B.,
Knetsch, T. A. „Google data in bridge equation models for German GDP“, *International Journal of Forecasting*, Volume 35, Issue 1, pp. 45-66, 2019
- Hoerl, A. E.,
Kennard, W. R. „Ridge Regression: Biased estimation for nonorthogonal problems“, *Technometrics*, Volume 12(1), pp. 55-67, 1970
- Kohns, D.,
Bhattacharjee, A. „Nowcasting growth using Google Trends data: A Bayesian Structural Time Series model“, *International Journal of Forecasting*, Volume 39, pp. 1384-1412, 2023

- Li, K.-C. „Asymptotic optimality for C_p , CL , cross-validation and generalized cross-validation: Discrete index set, the annals of statistics”, Volume 15, no. 3, pp. 958-975, 1987
- Nymand-Andersen, P., Pantelidis, E. „Google econometrics: nowcasting euro area car sales and big data quality requirements”, *ECB Statistic Paper* no. 30, 2018
- Tibshirani, R. „Regression shrinkage and selection via the lasso”, *Journal of the Royal Statistical Society, Series B (Methodological)*, Volume 58, no. 1, pp. 267-88, 1996
- Zou, H., Hastie, T. „Regularization and variable selection via the elastic net”, *Journal of the Royal Statistical Society, Volume Series B* 101(476), pp. 301-320, 2005

Anexă

Tabel A1. Lista variabilelor și transformărilor aferente

	STANDARD DATASET	Logarithm	Differences	Group			Logarithm	Differences	Group
1	MONEY M1	1	1	Money	51	3-year	0	0	FX
2	MONEY M2	1	1	Money	52	5-year	0	0	FX
3	MONEY M3	1	1	Money	53	M2_EA	1	1	Intl
4	MONEY Loans	1	1	Money	54	IP_EA	1	1	Intl
5	EER	1	1	FX	55	CPI_EA	1	1	Intl
6	EER_CPI	1	1	FX	56	PPI_EA	1	1	Intl
7	EER_PPI	1	1	FX	57	retail_EA	1	1	Intl
8	EURRON	1	1	FX	58	Ret.Tov xm	1	1	Serv
9	HICP total	1	1	HICP	59	Ret.Tov f&b&t	1	1	Serv
10	HICP x e&f	1	1	HICP	60	Ret.Tov n.food	1	1	Serv
11	HICP svc	1	1	HICP	61	Ret.Tov tex	1	1	Serv
12	HICP energy	1	1	HICP	62	Ret.Tov a&v	1	1	Serv
13	HICP goods	1	1	HICP	63	Ret.Tov comp	1	1	Serv
14	HICP food	1	1	HICP	64	Ret.Tov comm	1	1	Serv
15	HICP bv&tb	1	1	HICP	65	Ret. Tov	1	1	Serv
16	IP total	1	1	IP	66	BUIL.RO.TOT.COF.BS.M	0	1	Surv
17	IP xc	1	1	IP	67	BUIL.RO.TOT.1.BS.M	0	1	Surv
18	IP manuf	1	1	IP	68	BUIL.RO.TOT.2.F1S.M	0	1	Surv
19	IP constr	1	1	IP	69	BUIL.RO.TOT.2.F2S.M	0	1	Surv
20	IP xce	1	1	IP	70	BUIL.RO.TOT.2.F3S.M	0	1	Surv
21	IP energy xDE	1	1	IP	71	BUIL.RO.TOT.2.F4S.M	0	1	Surv
22	IP capital	1	1	IP	72	BUIL.RO.TOT.2.F5S.M	0	1	Surv
23	IP D cons	1	1	IP	73	BUIL.RO.TOT.2.F6S.M	0	1	Surv
24	IP MIG energy	1	1	IP	74	BUIL.RO.TOT.2.F7S.M	0	1	Surv
25	IP IM goods	1	1	IP	75	BUIL.RO.TOT.3.BS.M	0	1	Surv
26	IP ND cons	1	1	IP	76	BUIL.RO.TOT.4.BS.M	0	1	Surv
27	IP metals	1	1	IP	77	BUIL.RO.TOT.5.BS.M	0	1	Surv
28	IP chemicals	1	1	IP	78	INDU.RO.TOT.COF.BS.M	0	1	Surv
29	IP electric	1	1	IP	79	INDU.RO.TOT.1.BS.M	0	1	Surv
30	IP machinery	1	1	IP	80	INDU.RO.TOT.2.BS.M	0	1	Surv
31	IP paper	1	1	IP	81	INDU.RO.TOT.3.BS.M	0	1	Surv
32	IP plastic	1	1	IP	82	INDU.RO.TOT.4.BS.M	0	1	Surv
33	PPI tot xc	1	1	PPI	83	INDU.RO.TOT.5.BS.M	0	1	Surv
34	PPI manuf	1	1	PPI	84	INDU.RO.TOT.6.BS.M	0	1	Surv
35	PPI tot xce	1	1	PPI	85	INDU.RO.TOT.7.BS.M	0	1	Surv
36	PPI energy xDE	1	1	PPI	86	RETA.RO.TOT.COF.BS.M	0	1	Surv
37	PPI capital	1	1	PPI	87	RETA.RO.TOT.1.BS.M	0	1	Surv
38	PPI D cons	1	1	PPI	88	RETA.RO.TOT.2.BS.M	0	1	Surv
39	PPI energy	1	1	PPI	89	RETA.RO.TOT.3.BS.M	0	1	Surv
40	PPI IM goods	1	1	PPI	90	RETA.RO.TOT.4.BS.M	0	1	Surv
41	PPI ND cons	1	1	PPI	91	RETA.RO.TOT.5.BS.M	0	1	Surv
42	PPI metals	1	1	PPI	92	RETA.RO.TOT.6.BS.M	0	1	Surv
43	PPI chemicals	1	1	PPI	93	SERV.RO.TOT.COF.BS.M	0	1	Surv
44	PPI electric	1	1	PPI	94	SERV.RO.TOT.1.BS.M	0	1	Surv
45	PPI machinery	1	1	PPI	95	SERV.RO.TOT.2.BS.M	0	1	Surv
46	PPI paper	1	1	PPI	96	SERV.RO.TOT.3.BS.M	0	1	Surv
47	PPI plastic	1	1	PPI	97	SERV.RO.TOT.4.BS.M	0	1	Surv
48	IP new cars	1	1	IP	98	SERV.RO.TOT.5.BS.M	0	1	Surv
49	3-mon	0	0	FX	99	SERV.RO.TOT.6.BS.M	0	1	Surv
50	1-year	0	0	FX					

– continuare –

	GOOGLE TRENDS DATASET	Logarithm	Differences	Group
100	Celebrities & Entertainment News	1	1	GT
101	Comics & Animation	1	1	GT
102	Entertainment Industry	1	1	GT
103	Clubs & Nightlife	1	1	GT
104	Events & Listings	1	1	GT
105	Fun & Trivia	1	1	GT
106	Humor	1	1	GT
107	Movies	1	1	GT
108	Music & Audio	1	1	GT
109	Offbeat	1	1	GT
110	Online Games	1	1	GT
111	Web Portals	1	1	GT
112	Online Media	1	1	GT
113	Performing Arts	1	1	GT
114	TV & Video	1	1	GT
115	CAD & CAM	1	1	GT
116	Photographic & Digital Arts	1	1	GT
117	Visual Art & Design	1	1	GT
118	Arts & Entertainment	1	1	GT
119	Automotive Industry	1	1	GT
120	Bicycles & Accessories	1	1	GT
121	Boats & Watercraft	1	1	GT
122	Campers & RVs	1	1	GT
123	Classic Vehicles	1	1	GT
124	Commercial Vehicles	1	1	GT
125	Custom & Performance Vehicles	1	1	GT
126	Hybrid & Alternative Vehicles	1	1	GT
127	Microcars & City Cars	1	1	GT
128	Motorcycles	1	1	GT
129	Off-Road Vehicles	1	1	GT
130	Personal Aircraft	1	1	GT
131	Scooters & Mopeds	1	1	GT
132	Trucks & SUVs	1	1	GT
133	Vehicle Brands	1	1	GT
134	Vehicle Codes & Driving Laws	1	1	GT
135	Vehicle Licensing & Registration	1	1	GT
136	Vehicle Maintenance	1	1	GT
137	Vehicle Parts & Accessories	1	1	GT
138	Vehicle Shopping	1	1	GT
139	Vehicle Shows	1	1	GT
140	Autos & Vehicles	1	1	GT
141	Beauty Pageants	1	1	GT
142	Body Art	1	1	GT
143	Cosmetic Procedures	1	1	GT
144	Cosmetology & Beauty Professionals	1	1	GT
145	Face & Body Care	1	1	GT
146	Fashion & Style	1	1	GT
147	Fitness	1	1	GT
148	Hair Care	1	1	GT
149	Spas & Beauty Services	1	1	GT
150	Weight Loss	1	1	GT

		Logarithm	Differences	Group
151	Beauty & Fitness	1	1	GT
152	Biographies & Quotations	1	1	GT
153	Book Retailers	1	1	GT
154	Children's Literature	1	1	GT
155	E-Books	1	1	GT
156	Fan Fiction	1	1	GT
157	Literary Classics	1	1	GT
158	Magazines	1	1	GT
159	Poetry	1	1	GT
160	Writers Resources	1	1	GT
161	Books & Literature	1	1	GT
162	Advertising & Marketing	1	1	GT
163	Aerospace & Defense	1	1	GT
164	Agriculture & Forestry	1	1	GT
165	Business Education	1	1	GT
166	Business Finance	1	1	GT
167	Business News	1	1	GT
168	Business Operations	1	1	GT
169	Business Services	1	1	GT
170	Chemicals Industry	1	1	GT
171	HVAC & Climate Control	1	1	GT
172	Construction & Maintenance	1	1	GT
173	Energy & Utilities	1	1	GT
174	Enterprise Technology	1	1	GT
175	Grocery & Food Retailers	1	1	GT
176	Hospitality Industry	1	1	GT
177	Industrial Materials & Equipment	1	1	GT
178	Manufacturing	1	1	GT
179	Metals & Mining	1	1	GT
180	Pharmaceuticals & Biotech	1	1	GT
181	Journalism & News Industry	1	1	GT
182	Printing & Publishing	1	1	GT
183	Professional & Trade Associations	1	1	GT
184	Retail Trade	1	1	GT
185	Small Business	1	1	GT
186	Textiles & Nonwovens	1	1	GT
187	Transportation & Logistics	1	1	GT
188	Business & Industrial	1	1	GT
189	Computer Hardware	1	1	GT
190	Computer Security	1	1	GT
191	Consumer Electronics	1	1	GT
192	Electronics & Electrical	1	1	GT
193	Networking	1	1	GT
194	Programming	1	1	GT
195	Web Apps & Online Tools	1	1	GT
196	Software	1	1	GT
197	Technology News	1	1	GT
198	Computers & Electronics	1	1	GT
199	Accounting & Auditing	1	1	GT
200	Banking	1	1	GT

– continuare –

		Logarithm	Differences	Group			Logarithm	Differences	Group
201	Credit & Lending	1	1	GT	251	Contests, Awards & Prizes	1	1	GT
202	Currencies & Foreign Exchange	1	1	GT	252	Crafts	1	1	GT
203	Financial Planning	1	1	GT	253	Cycling	1	1	GT
204	Grants & Financial Assistance	1	1	GT	254	Outdoors	1	1	GT
205	Insurance	1	1	GT	255	Paintball	1	1	GT
206	Investing	1	1	GT	256	Animal Products & Services	1	1	GT
207	Retirement & Pension	1	1	GT	257	Pets	1	1	GT
208	Finance	1	1	GT	258	Wildlife	1	1	GT
209	Alcoholic Beverages	1	1	GT	259	Pets & Animals	1	1	GT
210	Candy & Sweets	1	1	GT	260	Radio Control & Modeling	1	1	GT
211	Cooking & Recipes	1	1	GT	261	Recreational Aviation	1	1	GT
212	Culinary Training	1	1	GT	262	Running & Walking	1	1	GT
213	Non-Alcoholic Beverages	1	1	GT	263	Special Occasions	1	1	GT
214	Restaurants	1	1	GT	264	Subcultures & Niche Interests	1	1	GT
215	Food & Drink	1	1	GT	265	Water Activities	1	1	GT
216	Arcade & Coin-Op Games	1	1	GT	266	Hobbies & Leisure	1	1	GT
217	Board Games	1	1	GT	267	Bed & Bath	1	1	GT
218	Card Games	1	1	GT	268	Domestic Services	1	1	GT
219	Computer & Video Games	1	1	GT	269	Gardening & Landscaping	1	1	GT
220	Family-Oriented Games & Activities	1	1	GT	270	Home Appliances	1	1	GT
221	Party Games	1	1	GT	271	Home Furnishings	1	1	GT
222	Puzzles & Brainteasers	1	1	GT	272	Home Improvement	1	1	GT
223	Roleplaying Games	1	1	GT	273	Home Storage & Shelving	1	1	GT
224	Table Games	1	1	GT	274	Homemaking & Interior Decor	1	1	GT
225	Games	1	1	GT	275	Kitchen & Dining	1	1	GT
226	Aging & Geriatrics	1	1	GT	276	Laundry	1	1	GT
227	Alternative & Natural Medicine	1	1	GT	277	Nursery & Playroom	1	1	GT
228	Health Conditions	1	1	GT	278	Pest Control	1	1	GT
229	Health Education & Medical Training	1	1	GT	279	Swimming Pools & Spas	1	1	GT
230	Health Foundations & Medical Research	1	1	GT	280	Yard & Patio	1	1	GT
231	Health News	1	1	GT	281	Home & Garden	1	1	GT
232	Medical Devices & Equipment	1	1	GT	282	Communications Equipment	1	1	GT
233	Medical Facilities & Services	1	1	GT	283	Email & Messaging	1	1	GT
234	Medical Literature & Resources	1	1	GT	284	Mobile & Wireless	1	1	GT
235	Men's Health	1	1	GT	285	Search Engines	1	1	GT
236	Mental Health	1	1	GT	286	Service Providers	1	1	GT
237	Nursing	1	1	GT	287	Teleconferencing	1	1	GT
238	Nutrition	1	1	GT	288	Web Services	1	1	GT
239	Oral & Dental Care	1	1	GT	289	Internet & Telecom	1	1	GT
240	Pediatrics	1	1	GT	290	Education	1	1	GT
241	Pharmacy	1	1	GT	291	Jobs	1	1	GT
242	Public Health	1	1	GT	292	Jobs & Education	1	1	GT
243	Reproductive Health	1	1	GT	293	Government	1	1	GT
244	Substance Abuse	1	1	GT	294	Legal	1	1	GT
245	Vision Care	1	1	GT	295	Military	1	1	GT
246	Women's Health	1	1	GT	296	Public Safety	1	1	GT
247	Health	1	1	GT	297	Social Services	1	1	GT
248	Antiques & Collectibles	1	1	GT	298	Law & Government	1	1	GT
249	Bowling	1	1	GT	299	Broadcast & Network News	1	1	GT
250	Clubs & Organizations	1	1	GT	300	Gossip & Tabloid News	1	1	GT

– continuare –

		Logarithm	Differences	Group
301	Local News	1	1	GT
302	Newspapers	1	1	GT
303	Politics	1	1	GT
304	Sports News	1	1	GT
305	Weather	1	1	GT
306	World News	1	1	GT
307	News	1	1	GT
308	Blogging Resources & Services	1	1	GT
309	Dating & Personals	1	1	GT
310	File Sharing & Hosting	1	1	GT
311	Forum & Chat Providers	1	1	GT
312	Online Goodies	1	1	GT
313	Online Journals & Personal Sites	1	1	GT
314	Photo & Video Sharing	1	1	GT
315	Social Networks	1	1	GT
316	Virtual Worlds	1	1	GT
317	Online Communities	1	1	GT
318	Disabled & Special Needs	1	1	GT
319	Family & Relationships	1	1	GT
320	Kids & Teens	1	1	GT
321	Religion & Belief	1	1	GT
322	Seniors & Retirement	1	1	GT
323	Social Issues & Advocacy	1	1	GT
324	Social Sciences	1	1	GT
325	People & Society	1	1	GT
326	Apartments & Residential Rentals	1	1	GT
327	Commercial & Investment Real Estate	1	1	GT
328	Property Development	1	1	GT
329	Property Inspections & Appraisals	1	1	GT
330	Property Management	1	1	GT
331	Real Estate Agencies	1	1	GT
332	Real Estate Listings	1	1	GT
333	Timeshares & Vacation Properties	1	1	GT
334	Real Estate	1	1	GT
335	Directories & Listings	1	1	GT
336	General Reference	1	1	GT
337	Geographic Reference	1	1	GT
338	Humanities	1	1	GT
339	Language Resources	1	1	GT
340	Libraries & Museums	1	1	GT
341	Technical Reference	1	1	GT
342	Reference	1	1	GT
343	Astronomy	1	1	GT
344	Biological Sciences	1	1	GT
345	Chemistry	1	1	GT
346	Computer Science	1	1	GT
347	Earth Sciences	1	1	GT
348	Ecology & Environment	1	1	GT
349	Engineering & Technology	1	1	GT
350	Mathematics	1	1	GT

		Logarithm	Differences	Group
351	Physics	1	1	GT
352	Scientific Equipment	1	1	GT
353	Scientific Institutions	1	1	GT
354	Science	1	1	GT
355	Apparel	1	1	GT
356	Auctions	1	1	GT
357	Classifieds	1	1	GT
358	Consumer Resources	1	1	GT
359	Entertainment Media	1	1	GT
360	Luxury Goods	1	1	GT
361	Mass Merchants & Department Stores	1	1	GT
362	Shopping Portals & Search Engines	1	1	GT
363	Sporting Goods	1	1	GT
364	Swap Meets & Outdoor Markets	1	1	GT
365	Tobacco Products	1	1	GT
366	Toys	1	1	GT
367	Wholesalers & Liquidators	1	1	GT
368	Shopping	1	1	GT
369	College Sports	1	1	GT
370	Combat Sports	1	1	GT
371	Extreme Sports	1	1	GT
372	Fantasy Sports	1	1	GT
373	Individual Sports	1	1	GT
374	Motor Sports	1	1	GT
375	Sports Coaching & Training	1	1	GT
376	Team Sports	1	1	GT
377	Winter Sports	1	1	GT
378	World Sports Competitions	1	1	GT
379	Sports	1	1	GT
380	Air Travel	1	1	GT
381	Bus & Rail	1	1	GT
382	Car Rental & Taxi Services	1	1	GT
383	Carpooling & Ridesharing	1	1	GT
384	Cruises & Charters	1	1	GT
385	Hotels & Accommodations	1	1	GT
386	Luggage & Travel Accessories	1	1	GT
387	Specialty Travel	1	1	GT
388	Tourist Destinations	1	1	GT
389	Travel Agencies & Services	1	1	GT
390	Travel Guides & Travelogues	1	1	GT
391	Travel	1	1	GT
392	All categories	1	1	GT

