



Analytický komentár

„Gúglime“ vývoj trhu práce.

Internet je moderným fenoménom a jeho využitie má do budúcnosti potenciál rásť. Platí to aj o vyhľadávaní informácií pomocou internetových vyhľadávačov. Je preto na mieste uvažovať nad využitím tohto zdroja aj pri krátkodobých prognózach vývoja na trhu práce. Cieľom analytického komentára je sledovať početnosť internetového vyhľadávania výrazov, ktoré by mohli súvisieť s hľadaním práce a otestovať ich potenciál pre využitie na krátkodobé prognózovanie zamestnanosti a nezamestnanosti. Predikčnú schopnosť týchto vyhľadávanií vyhodnocujeme pomocou vyčerpávajúceho súboru jednorovnicových regresných modelov. Ako najlepšie ukazovatele sa javia výrazy „úrad práce“, resp. „ponuka práce“, a to pri prognózovaní nezamestnanosti. Ich odchýlka je o 10, resp. 29 % nižšia v porovnaní s naivným benchmarkovým modelom. Zároveň ponúkame porovnávacie cvičenie, v ktorom testujeme presnosť existujúcej metódy na krátkodobý odhad v NBS (nowcasting zamestnanosti) v porovnaní s jej verziou rozšírenou o vybrané Google indikátory.

Úvod

Služba Trendy Google poskytuje takmer v reálnom čase počty vyhľadávanií pre daný výraz alebo kombináciu výrazov. Zameriavame sa na výrazy, ktoré pravdepodobne zadávajú do vyhľadávača osoby po strate zamestnania, resp. po nadobudnutí postavenia nezamestnaného, a teda testujeme hypotézu, že nárast počtu vyhľadávanií signalizuje nárast nezamestnanosti, prípadne pokles zamestnanosti. Ide o výrazy „ponuka práce“ (PP), „úrad práce“ (UP), „hl'adám prácu“ (HP) a „práca“ (PR)¹. Testovali sme aj výraz „profesia.sk“, ale jeho vývoj sa ukazuje ako pomerne nestabilný, s výrazným nárastom v období krízy a následným utlmením (ľudia pravdepodobne prestávajú tento výraz „gúglit“, radšej ho priamo zadajú). Na druhej strane výraz „profesia“ je pomerne obsírný a v regresných rovniciach nezlepšuje prognózu vysvetľovaných premenných. Preto sa týmto dvom ukazovateľom detailnejšie nevenujeme.

Služba Trendy poskytuje dáta v týždennej frekvencii. Previedli sme ich spriemerovaním na mesačnú frekvenciu, sezónne očistili procedúrou Tramo/Seats a nakoniec spriemerovali na kvartálne údaje. Naším cieľom bolo totiž **prognózovať kvartálne „headline“ ukazovatele zamestnanosti** (v metodike ESA 2010, počet zamestnaných osôb v ekonomike, EMP) a **nezamestnanosti** (metodika VZPS, počet nezamestnaných osôb v ekonomike, UNEM), takisto sezónne očistené. Ide o hlavné prognózované premenné trhu práce v rámci pravidelných predikcií NBS.

Na Slovensku sa pravdepodobne prvýkrát touto témou zaoberala Pécsyová (2011). Jednoduchý Okunov zákon rozšírila o podobné indikátory ako v našom prípade za účelom prognózy mesačnej miery nezamestnanosti v metodike UPSVaR. Hlavným zistením bolo, že pridanie indikátorov do modelu znižuje chybu in-sample prognózy o približne 30 %. **Pridanou hodnotou nášho prístupu je niekoľko nových prvkov:** analyzujeme podstatne širšiu až **vyčerpávajúcu množinu** modelových špecifikácií, okrem in-sample prístupu analyzujeme aj dôležitejšiu **out-of-sample** prognózu a zameriavame sa na **štvrt'ročnú** nezamestnanosť VZPS a zamestnanosť ESA, pri ktorých z dôvodu výrazného oneskorenia existuje najväčší potenciál využitia Google indikátorov. Rozdielom v našom prístupe je tiež **zameranie na ARMA modely ako porovnávací benchmark**, pretože sa chceme

¹ Zohľadňujeme vyhľadávanie s diakritikou aj bez nej a rovnako aj skloňovanie tam, kde je to relevantné. To znamená, že napríklad pri výraze „ponuka práce“ zadávame do nástroja Trendy Google výraz [ponuka prace + ponuka práce]. Pri výraze „práca“ používame ešte širší okruh vyhľadávaných slov: [práca + práce + práci + práci + pracou + praca + prace + praci + pracu + pracou].

Rýchle/analytické komentáre nie sú oficiálnym stanoviskom Národnej banky Slovenska. Prezentujú názory analytikov Úseku pre menu, štatistiku a výskum (ÚMS). Šírenie je povolené bez predchádzajúceho súhlasu, avšak s uvedením zdroja „analytici ÚMS“, resp. „analytici Menového úseku“. Ak nie je uvedené inak, časové rady sú sezónne očistené použitím vlastných sezónnych modelov.

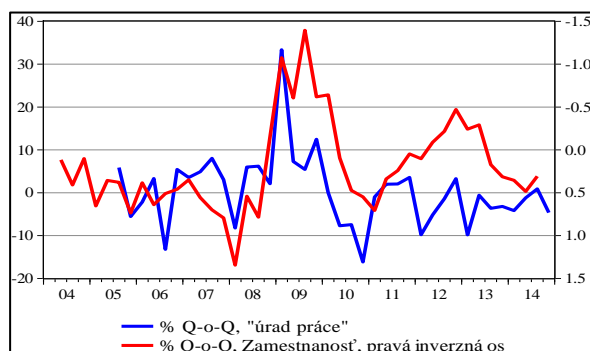
vyhnúť chybe špecifikácie z dôvodu nesprávne zvolenej alebo príliš zjednodušenej štruktúry vysvetľujúcich premenných. **ARMA modely sú vhodnou alternatívou zložitejších štruktúrnych modelov** v prostredí neistoty ohľadne ich správnej špecifikácie (napr. Brooks (2014)). Použitím ARMA špecifikácie a dodatočného Google indikátora v prognostických rovniciach izolujeme nami skúmaný problém, a to **zistiť, či indikátor prispieva k schopnosti vysvetliť vývoj zamestnanosti alebo nezamestnanosti**. Významným príkladom použitia tejto metodiky je aj štúdia autorov z Banca d'Italia D'Amuri, Marucci (2012). Na záver sa tiež venujeme otázke, či by Google indikátory prispeli k predikčnej schopnosti metódy využívané v NBS (tzv. nowcasting zamestnanosti). **Pridávame najlepšie Google indikátory do existujúceho nowcastingu a porovnávame predikčnú schopnosť** takto rozšírenej metódy s pôvodnou verziou bez nich.

Grafický pohľad na Google indikátory

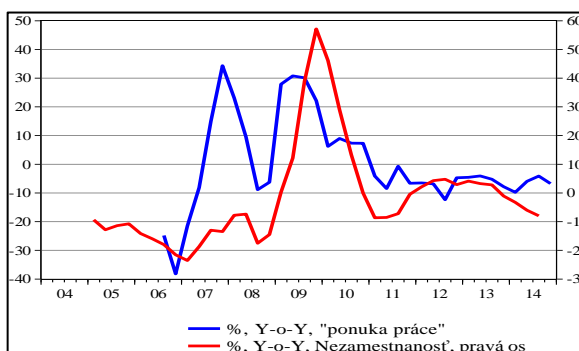
Okrem regresných modelov a vyhodnotenia presnosti prognóz (v ďalších častiach) je na úvod potrebné preskúmať **grafický súlad vývoja indikátorov s vysvetľovanými premennými**. V tejto časti poskytujeme formou grafov 1 až 4 pohľad na indikátory, ktoré graficky najviac súvisia s vývojom zamestnanosti a nezamestnanosti (UP, PP). Zároveň sú tieto indikátory najúspešnejšie z hľadiska modelového vyhodnotenia v ďalších častiach. Vyberáme vždy takú transformáciu, ktorá vizuálne naznačuje najsilnejšiu súvislosť. Z dôvodu prehľadnosti neuvádzame tie indikátory, ktoré z pohľadu vývoja (ne)zamestnanosti neposkytujú žiadnu zaujímavú grafickú informáciu. Zároveň uvádzame aj graf obsahujúci indikátory z konjunkturálnych prieskumov² (očakávania zamestnanosti a nezamestnanosti), ktoré už analytici NBS vo svojich prognózach aktívne využívajú.

Grafy v tejto časti poukazujú na **potenciálnu predikčnú schopnosť Google indikátorov, a to predovšetkým v období globálnej krízy** (vrcholiacej hlavne v roku 2009). Zvýšená početnosť vyhľadávaní vybraných výrazov korešponduje s prudkým nárastom nezamestnanosti. Podobný vývoj počas krízy indikovali aj tradičnejšie indikátory z konjunkturálnych prieskumov, a to očakávania zamestnávateľov (ohľadne zamestnanosti) a očakávania spotrebiteľov (ohľadne nezamestnanosti). Ďalšou zaujímavosťou je **vrchol ukazovateľa PP už v roku 2007**. Zahraničný výskum ukazuje³, že v obdobiach ekonomického boomu si veľká časť zamestnancov hľadá prácu s lepšími podmienkami, čo im daná situácia umožňuje. Je pravdepodobné, že PP zachytáva aj tento fenomén, čo je rizikom tohto konkrétneho indikátora. Riziká rôzneho typu sú však spojené s každým indikátorom, vrátane tých, ktoré nepochádzajú z vyhľadávačov. Preto je potrebné v ďalších častiach podrobnejšie preveriť predikčnú schopnosť Google indikátorov.

Graf 1 Zamestnanosť a „úrad práce“



Graf 2 Nezamestnanosť a „ponuka práce“

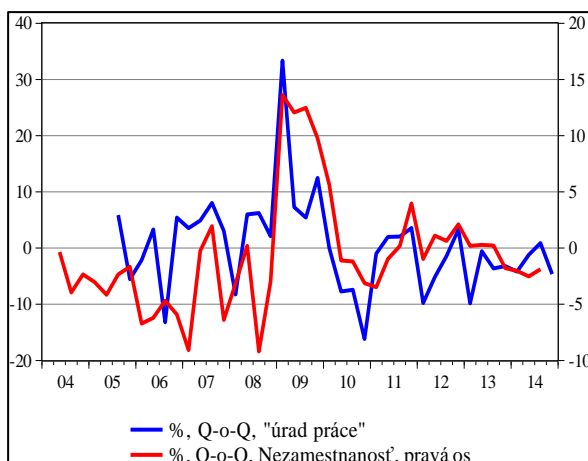


Zdroj: ŠÚ SR, Európska komisia, výpočty autorov. Zamestnanosť v metodike ESA 2010 a nezamestnanosť podľa VZPS.

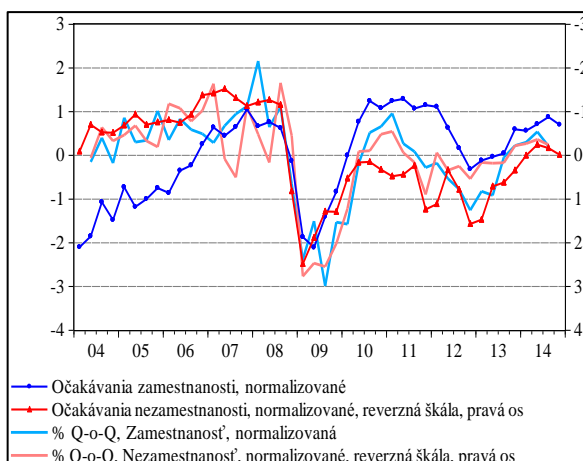
² Zdroj: Európska komisia.

³ Príkladom je štúdia Menzio & Shi (2009).

Graf 3 Nezamestnanosť a „úrad práce“



Graf 4 Očakávania z konjunkturálnych prieskumov



Zdroj: ŠÚ SR, Európska komisia, výpočty autorov. Zamestnanosť v metodike ESA 2010 a nezamestnanosť podľa VZPS.

Metóda testovania predikcií v porovnaní s ARMA modelmi

Za účelom robustného vyhodnotenia predikčnej schopnosti Google indikátorov sme definovali všeobecné **lineárne regresné rovnice**, kde za premennú *ind* sme dosádzali jednu zo štyroch vyššie uvedených vysvetľujúcich premenných (ďalej označované ako „indikátory“):

$$\text{vysvetľovaná premenná} = f(\text{ind}_t, \text{ind}_{t-1}, \text{ind}_{t-2}, \text{členy typu ARMA})$$

Rovnica bola odhadnutá **vo viacerých variantoch pre medziročnú ako aj medzikvartálnu dynamiku zamestnanosti aj nezamestnanosti**. Bližšie sú varianty prezentované v tabuľke P1 v Prílohe. Popri vysvetľujúcich rovniciach boli odhadnuté aj tzv. benchmark modely, čiže jednoduché ARMA procesy zamestnanosti, resp. nezamestnanosti bez ďalších vysvetľujúcich premenných slúžiac len na porovnanie presnosti prognóz z indikátorov⁴.

Každú z týchto rovníc sme odhadli za obdobie približne od rokov 2005-2006⁵ po 3. kvartál 2014. **Presnosť prognózy sme hodnotili tromi spôsobmi**. Prvý z nich berie odhadnutú rovnicu ako danú a vytvára jedнокrokové prognózy za obdobie **2008Q1 až 2014Q3**. Na záver sa vyráta stredná absolútna percentuálna odchýlka (**MAPE**) medzi prognózovanou (ne)zamestnanosťou a skutočnou hodnotou. Druhý spôsob vyhodnotenia poskytne MAPE len za obdobie **2013Q1 až 2014Q3** za účelom vyhodnotenia na aktuálnejšom období. Posledný spôsob odhaduje rovnicu nanovo v krokoch: najprv rovnicu odhadneme za obdobie do 2012Q4 a vytvoríme jedнокrokovú prognózu na 2013Q1; potom odhadujeme do 2013Q1 a prognózujeme na obdobie 2013Q2 atď. až po 2014Q3 (tzv. **rekurzívna regresia s out-of-sample prognózou**). Z výsledných prognóz opäť vypočítame MAPE.

Každá z uvedených metód má **svoje výhody a nedostatky**, súvisiace hlavne s krátkosťou dostupného časového radu. Prvá z nich zachytáva aj výrazný negatívny posun počas krízy, avšak nakoľko ide o in-sample prognózu, je o niečo menej realistická v porovnaní s prognózovaním

⁴ Pre benchmarkové procesy bola zvolená optimálna štruktúra na základe minimálneho Akaikeho informačného kritéria s maximálnym oneskorením 4 kvartály a absencie autokorelácie. Navyše bol ako benchmark otestovaný aj jednoduchý AR(1) model. Celkovo sme teda mali k dispozícii výsledky z 8 benchmarkových rovníc (4 pre zamestnanosť a 4 pre nezamestnanosť) a 384 rovníc s indikátormi (192 + 192).

⁵ Každý indikátor má začiatok časového radu v inom štvrtroku z dôvodu rôznej dostupnosti stabilných údajov. Údaje z Trendov Google vo všeobecnosti začínajú od roku 2004, ale v tomto roku sú výsledné hodnoty spravidla nulové alebo veľmi nízke. Procedúra testovania prebehla v prostredí softvéru EViews.

v reálnom čase. Druhá metóda sa sústreďí len na najaktuálnejšie obdobie, ktoré je relevantné z hľadiska aktuálneho prognózovania, ale opäť ide o in-sample prognózu. **Teoreticky najvierohodnejším vyhodnotením by mal byť tretí spôsob. Ten sa približuje vyhodnocovaniu v reálnom čase.** Nevýhodou v tomto prípade je však to, že spoľahlivosť modelov sa najmä kvôli krátkosti časového radu stále len formuje a preto napríklad model z roku 2012 môže byť výrazne odlišný od toho, ktorý je výsledným pri odhade v roku 2014. Z dôvodu týchto kladov a záporov pracujeme ešte aj so sumárnym ukazovateľom presnosti prognózy - aritmetickým priemerom MAPE pre každú rovnicu, vrátane benchmarkov.

Výsledky vyhodnotenia v tabuľkovej forme

Úplné vyhodnotenie všetkých indikátorov na základe štyroch popisovaných hodnotiacich metód (vrátane priemeru MAPE) uvádzame v prílohe vo forme tabuliek P2 a P3⁶. V tejto časti sa však z dôvodu prehľadnosti sústreďíme len na rekurzívnu out-of-sample metódu vyhodnocovania, ktorú považujeme z praktického aj teoretického hľadiska za najdôležitejšiu. Hlavné výsledky sú uvedené v tabuľkách 1 a 2.

Pri vyhodnotení prognóz zamestnanosti najlepšie obstoja indikátory HP a PP. Je však zaujímavé, že ani jeden z modelov nedokázal prekonať presnosť benchmark modelu, čo považujeme za nedostatok indikátorov Google vo vzťahu k zamestnanosti v ekonomike. Na základe výsledkov v tabuľkách ARMA **modely na prognózu zamestnanosti nedokážeme zlepšiť pridaním Google indikátorov, čo nepodporuje ich ďalšie využívanie pri prognózovaní.** Odchýlka je oproti benchmarkovému modelu takmer dvojnásobná.

Vyhodnotenie prognóz nezamestnanosti dopadlo podstatne lepšie. Okrem indikátora HP všetky ostatné zlepšujú prognózu. Odchýlka sa pri využití indikátora PP (s najnižším MAPE) znižuje oproti benchmarku o 29 %.⁷ Napriek odlišnostiam v metodike sme teda dosiahli podobný výsledok ako Pécsyová (2011).

Na ďalšie zhodnotenie spoľahlivosti uvedených indikátorov sme využili aj posúdenie štatistickej významnosti jednotlivých indikátorov v regresných modeloch uvedené nižšie.

Tabuľka 1 – vyhodnotenie out-of-sample prognózy zamestnanosti

Indikátor	Špecifikácia (závislá premenná zamestnanosť)	MAPE, rekurzívne, 2013-2014
benchmark	QoQ c ARMA(2,3)	0.06
HP	QoQ c HP HP(-1) HP(-2) ARMA(2,3)	0.10
PP	QoQ c PP ARMA(2,3)	0.09
UP	QoQ c UP ARMA(2,3)	0.11
PR	QoQ c qoq(PR(-1)) ARMA(2,3)	0.11

Tabuľka 2 – vyhodnotenie out-of-sample prognózy nezamestnanosti

Indikátor	Špecifikácia (závislá premenná nezamestnanosť)	MAPE, rekurzívne, 2013-2014
benchmark	QoQ c AR(1)	0.72
HP	QoQ c HP(-1)	0.74
PP	QoQ c PP ARMA(4,3)	0.51
UP	QoQ c qoq(UP) ARMA(4,3)	0.65
PR	QoQ c PR ARMA(4,3)	0.69

Zdroj: Výpočty autorov.

Výsledné odhady najpresnejších regresných modelov

V tejto časti prezentujeme odhadnuté parametre optimálnych regresíí identifikovaných v predchádzajúcej časti. Nakoľko sme **ne našli žiadny model pre zamestnanosť**, ktorý by dosahoval lepšie výsledky „out-of-sample“ v porovnaní s benchmarkom, **d'alej sa už sústreďíme len na analýzu výsledkov modelov pre nezamestnanosť**, ktoré zlepšili out-of-sample prognózu oproti benchmarku.

⁶ Výsledky úplného vyhodnotenia považujeme za doplňujúce, závery sa nelíšia od zredukovanej podoby v tabuľkách 1 a 2.

⁷ Druhý najlepší indikátor (UP) dosiahol o 10 % nižšiu odchýlku (MAPE) ako benchmarkový AR(1) model.

Tabuľka 3 – regresný model „ponuka práce“ a nezamestnanosť

Vysvetľujúci indikátor: PP		
Závislá premenná: QoQ, Nezamestnanosť		
Dĺžka vzorky: 2005Q3 2014Q3		
R squared: 0.778		
Durbin-Watson štatistika: 2.22		
P-hodnota pre F test celkovej významnosti parametrov: 0.000		
Premenná	Koeficient	P-hodnota
C	-19.42	0.00
PP	0.43	0.00
AR(1)	0.70	0.00
AR(2)	-0.63	0.05
AR(3)	0.30	0.29
AR(4)	-0.42	0.02
MA(1)	-0.22	0.41
MA(2)	0.52	0.06
MA(3)	0.59	0.01

Tabuľka 4 – regresný model „úrad práce“ a nezamestnanosť⁸

Vysvetľujúci indikátor: UP		
Závislá premenná: QoQ, Nezamestnanosť		
Dĺžka vzorky: 2005Q3 2014Q3		
R squared: 0.755		
Durbin-Watson štatistika: 1.98		
P-hodnota pre F test celkovej významnosti parametrov: 0.000		
Premenná	Koeficient	P-hodnota
C	-0.28	0.60
qoq(UP)	0.34	0.00
AR(1)	0.73	0.00
AR(2)	-0.08	0.72
AR(3)	0.04	0.86
AR(4)	-0.02	0.91
MA(1)	0.02	0.94
MA(2)	-0.47	0.02
MA(3)	0.51	0.03

Zdroj: Výpočty autorov. Výrazy „ponuka práce“ a „úrad práce“ zahŕňajú aj varianty bez diakritiky.

Z hľadiska prognózy nezamestnanosti ukazovateľa PP, UP a PR naznačovali zlepšenie oproti benchmarku. Pri bližšom pohľade na najpresnejší model pre indikátor PR je však koeficient ukazovateľa PR nevýznamný. Naopak, v **najpresnejších modeloch s indikátormi PP a UP sú Google indikátory štatisticky významné**, čo znázorňujú tabuľky 3 a 4. Aj počas rekurzívneho odhadovania v období 2013Q1 až 2014Q3 si tieto indikátory zachovávajú koeficient s podobnou veľkosťou (pri štandardnej odchýlke približne do 0,02) a štatistickou významnosťou. Priaznivé výsledky z tabuľky 2 pre indikátory PP a UP sú teda podporené aj pri bližšom preskúmaní koeficientov príslušných rovníc.

Praktická aplikácia Google indikátorov v nowcastingu zamestnanosti NBS

V predošlých častiach sme ukázali, že výrazy **UP a PP dokážu prekonať naivný model**. Na základe tejto skutočnosti je potrebné overiť, **či by zohľadnenie týchto indikátorov prospelo aj aktuálne využívanej metóde krátkodobej prognózy (tzv. nowcastingu) zamestnanosti**. Stručný popis aktuálnej metódy nowcastingu zamestnanosti poskytujeme v prílohe. Za účelom porovnania výsledkov nowcastingu bez Google indikátorov a s nimi sme najprv spustili bežnú procedúru nowcastingu („aktuálny prístup“) a vyhodnotili sme ju rekurzívne za obdobie od začiatku roku 2013, resp. 2012. Následne sme ten istý postup zopakovali, ale s tým, že **do nowcastingu boli vložené dve dodatočné rovnice pre ukazovatele UP a PP** v špecifikácii uvedenej v tabuľkách 3 a 4⁹. Porovnanie ukazuje, že aj keď pridanie Google indikátorov presnosť prognózy výrazne nezhoršuje, **nedokážeme s ich využitím dosiahnuť presnejšie prognózy vývoja na trhu práce**. Možno teda konštatovať, že **Google indikátory majú v priemere marginálne nižšiu predikčnú schopnosť ako indikátory doposiaľ využívané na prognózovanie v NBS** (očakávania z konjunkturálnych prieskumov, nezamestnanosť ÚPSVaR a mesačné štatistiky zamestnanosti za vybrané odvetvia).

⁸ V tabuľke 4 uvádzame model s indikátorom UP len v čase t , napriek tomu, že v tabuľke P3 je najpresnejším modelom pre UP model s indikátorom v čase t a $t-1$. Koeficient pre indikátor v čase $t-1$ je však bez výnimky v období 2013Q1 až 2014Q3 štatisticky nevýznamný.

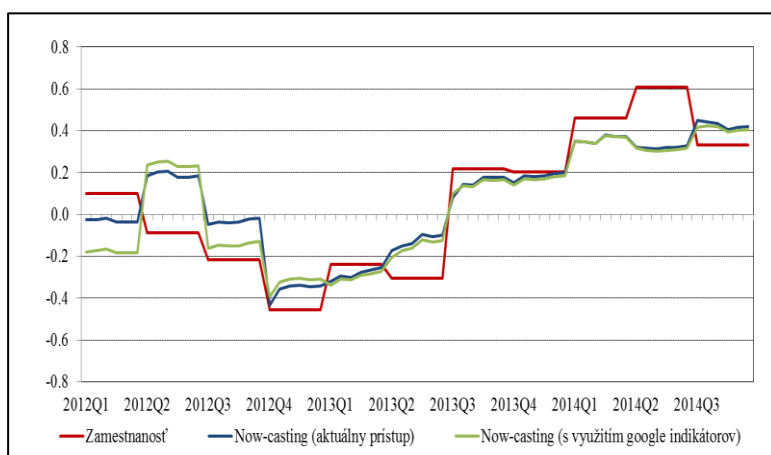
⁹ Tieto rovnice poskytujú odhad nezamestnanosti, ktorý sa následne prepočíta na zamestnanosť s využitím predpokladov popísaných v prílohe.

Tabuľka 5 – Porovnanie nowcastingu NBS s využitím Google indikátorov a bez nich

Poradie nowcastu	2013Q1-2014Q3		Poradie nowcastu	2012Q1-2014Q3	
	MAE, aktuálny prístup	MAE, s Google indikátormi		MAE, aktuálny prístup	MAE, s Google indikátormi
1	0.128	0.118	1	0.136	0.144
2	0.115	0.118	2	0.137	0.148
3	0.119	0.121	3	0.140	0.151
4	0.106	0.110	4	0.131	0.143
5	0.103	0.108	5	0.130	0.142
6	0.101	0.106	6	0.130	0.142
Priemer	0.112	0.113	Priemer	0.134	0.145

Zdroj: Výpočty autorov. Poradie nowcastu je v rámci jedného štvrtroka. Stredná absolútna odchýlka (MAE) je vypočítaná na základe absolútného rozdielu medzi prognózovanou a skutočnou medzikvartálnou dynamikou zamestnanosti.

Graf 5 – Nowcasting NBS s využitím Google indikátorov a bez nich (% , QoQ)



Zdroj: ŠÚ SR, výpočty autorov.

Záver

Z identifikovaných potenciálne užitočných vyhľadávaných výrazov z Trendov Google sa javia ako **najúžitočnejšie z hľadiska prognózy indikátory UP a PP, a to pre prognózu nezamestnanosti**. Oproti naivnému ARMA benchmarku zaznamenali o 10, resp. 29 % nižšiu chybu.

Skutočnosť, že **nezamestnanosť je prognózovaná lepšie ako zamestnanosť** môže byť čiastočne vysvetlená tým, že indikátory Google teoreticky skôr vystihujú výrazy vyhľadávané nezamestnanými. Výrazný rozdiel môže vzniknúť napríklad pri príchode uchádzačov o zamestnanie na trh práce z predošlej neaktivity (napríklad z rodičovskej dovolenky, po vzdelávaní, období na PN, v domácnosti). V tomto prípade rastie nezamestnanosť, ale zamestnanosť sa okamžite zvýšiť nemusí.

Zároveň treba zdôrazniť, že **s Google indikátormi sa spája vyššia volatilita** v porovnaní s tradičnejšími ukazovateľmi vývoja na trhu práce. **Indikátory Google môžu byť citlivé na zmenu internetových zvyklostí populácie** (v ktorom prehliadači/vyhľadávači vyhľadávame, aký výraz zadáme, či sa učíme danú tému hľadať na konkrétnej stránke, koľko z nás má okamžitý prístup na internet atď.). Tieto dodatočné faktory môžu zahmlievať pohľad na skutočný vývoj (ne)zamestnanosti.

Aj to môže byť jeden z dôvodov, prečo **zohľadnenie Google indikátorov nedokáže zlepšiť presnosť už aktuálne využívanej metódy krátkodobej prognózy vývoja na trhu práce** (tzv. nowcasting zamestnanosti), ako ukázalo porovnávacie cvičenie.

Aktuálnu iniciatívu testovania rozmerných databáz vrátane Trendov Google na získavanie cenných informácií možno zaradiť pod termín **Big Data**. Intenzívne sa ňou zaoberá aj ECB¹⁰. Aj v NBS je pravdepodobné ďalšie rozvíjanie tejto tematiky za účelom prognózovania a porozumenia vývoju rôznych makroekonomických ukazovateľov.

Alexander Karšay, Stanislav Tvrz (analytici@nbs.sk)

Použitá literatúra:

1. Brooks, C. (2014): „Introductory Econometrics for Finance.“ 3rd edition. Cambridge.
2. D’Amuri, F., Marcucci, J. (2012): „The predictive power of Google searches in forecasting unemployment.“ Banca d’Italia Working paper Nr. 891.
3. Menzio, G., Shi, S. (2009). „Efficient search on the job and the business cycle.“ NBER working paper Nr. w14905.
4. Pécsyová, M. (2011): „„Vygooglime“ si budúcnosť.“ Komentár IFP 2011/29.

¹⁰ https://www.ecb.europa.eu/pub/conferences/html/20140407_workshop_on_using_big_data.en.html

Príloha

Stručný popis metodiky nowcastingu zamestnanosti

NBS za účelom krátkodobých prognóz zamestnanosti (najmä na aktuálne prebiehajúci štvrťrok) pravidelne využíva tzv. nowcasting zamestnanosti. Pozostáva zo siedmich¹¹ rovníc odhadovaných metódou najmenších štvorcov. Tieto rovnice sú vážené na základe ich historickej predikčnej schopnosti (strednej absolútnej odchýlky, tzv. Mean Absolute Error) tak, že čím menšia je historická chyba, tým väčšiu váhu priradíme danej rovnici. Do jednotlivých rovníc vstupuje vždy ako vysvetľovaná premenná zmena zamestnanosti (alebo nezamestnanosti – podľa kontextu) a ako vysvetľujúca premenná jeden z mesačných ukazovateľov trhu práce (očakávania zamestnanosti alebo nezamestnanosti, nezamestnanosť v metodike ÚPSVaR, zamestnanosť za vybrané odvetvia súkromného sektora).

Prognózu z rovníc v rámci nowcastingu, ktoré poskytnú prognózu nezamestnanosti (ako napríklad nezamestnanosť ÚPSVaR) je potrebné previesť na prognózu zamestnanosti. Na tento účel sa použijú predpoklady z prognóz NBS o budúcom vývoji ekonomicky aktívneho obyvateľstva, pričom platí, že zmenu ekonomicky aktívnych osôb považujeme za danú, zmena nezamestnanosti pochádza z rovnice a zmena zamestnanosti sa dopočíta, keďže ekonomicky aktívne obyvateľstvo je sumou zamestnanosti a nezamestnanosti¹². V rámci nowcastingu máme teda k dispozícii 7 odhadov zamestnanosti, ktoré po prevážení poskytujú finálny odhad zmeny zamestnanosti pre daný štvrťrok.

Spravidla spúšťame nowcasting za jeden štvrťrok 6 krát pri zverejňovaní rôznych mesačných indikátorov za jednotlivé mesiace. Chýbajúce mesiace v danom štvrťroku sa doprognozujú na základe jednoduchého ARMA modelu.

¹¹ Rovníc je sedem ale mesačné ukazovatele sú štyri. Niektoré rovnice používajú ten istý mesačný ukazovateľ ale v inej transformácii (medziročný vs. medzikvartálny rast) alebo s dodatočným dynamickým členom – oneskorenou hodnotou závislej premennej.

¹² Tento prepočet sa používa aj v nowcastingu rozšírenom o Google indikátory, nakoľko tie vo svojich rovniciach vysvetľujú nezamestnanosť.

Tabuľka P1 – Špecifikácia regresných rovníc

	Benchmark	QoQ c ARMA(2,3) YoY c ARMA(2,2) QoQ c AR(1) YoY c AR(1)
Indikátor ako vysvetľujúca premenná (bez ARMA)	Vzťah v čase t	QoQ c ind QoQ c qoq(ind) YoY c ind YoY c yoy(ind)
	Len t-1	QoQ c ind(-1) QoQ c qoq(ind(-1)) YoY c ind(-1) YoY c yoy(ind(-1))
	Vzťah v čase t a t-1	QoQ c ind ind(-1) QoQ c qoq(ind) qoq(ind(-1)) YoY c ind ind(-1) YoY c yoy(ind) yoy(ind(-1))
	Vzťah v čase t a t-2	QoQ c ind ind(-1) ind(-2) QoQ c qoq(ind) qoq(ind(-1)) qoq(ind(-2)) YoY c ind ind(-1) ind(-2) YoY c yoy(ind) yoy(ind(-1)) yoy(ind(-2))
Indikátor ako vysvetľujúca premenná + optimálna ARMA štruktúra	Vzťah v čase t	QoQ c ind ARMA(2,3) QoQ c qoq(ind) ARMA(2,3) YoY c ind ARMA(2,2) YoY c yoy(ind) ARMA(2,2)
	Len t-1	QoQ c ind(-1) ARMA(2,3) QoQ c qoq(ind(-1)) ARMA(2,3) YoY c ind(-1) ARMA(2,2) YoY c yoy(ind(-1)) ARMA(2,2)
	Vzťah v čase t a t-1	QoQ c ind ind(-1) ARMA(2,3) QoQ c qoq(ind) qoq(ind(-1)) ARMA(2,3) YoY c ind ind(-1) ARMA(2,2) YoY c yoy(ind) yoy(ind(-1)) ARMA(2,2)
	Vzťah v čase t a t-2	QoQ c ind ind(-1) ind(-2) ARMA(2,3) QoQ c qoq(ind) qoq(ind(-1)) qoq(ind(-2)) ARMA(2,3) YoY c ind ind(-1) ind(-2) ARMA(2,2) YoY c yoy(ind) yoy(ind(-1)) yoy(ind(-2)) ARMA(2,2)
Indikátor ako vysvetľujúca premenná + AR(1) štruktúra	Vzťah v čase t	QoQ c ind AR(1) QoQ c qoq(ind) AR(1) YoY c ind AR(1) YoY c yoy(ind) AR(1)
	Len t-1	QoQ c ind(-1) AR(1) QoQ c qoq(ind(-1)) AR(1) YoY c ind(-1) AR(1) YoY c yoy(ind(-1)) AR(1)
	Vzťah v čase t a t-1	QoQ c ind ind(-1) AR(1) QoQ c qoq(ind) qoq(ind(-1)) AR(1) YoY c ind ind(-1) AR(1) YoY c yoy(ind) yoy(ind(-1)) AR(1)
	Vzťah v čase t a t-2	QoQ c ind ind(-1) ind(-2) AR(1) QoQ c qoq(ind) qoq(ind(-1)) qoq(ind(-2)) AR(1) YoY c ind ind(-1) ind(-2) AR(1) YoY c yoy(ind) yoy(ind(-1)) yoy(ind(-2)) AR(1)

Skratky QoQ alebo YoY na začiatku špecifikácie predstavujú transformáciu pre závislú premennú v danej rovnici. Za „ind“ sa dosádza Google indikátor; c je konštanta; (-1) resp. (-2) znamená časový posun, čiže indikátor v kvartáli t-1, resp. t-2. Tabuľka znázorňuje optimálnu ARMA štruktúru pre modely zamestnanosti. Modely nezamestnanosti sú špecifikované analogicky (so svojou vlastnou optimálnou ARMA štruktúrou, vid' tab. P3)

Tabuľka P2 – vyhodnotenie prognóz zamestnanosti*

Indikátor	Špecifikácia	MAPE, 2008-2014	MAPE, 2013-2014	MAPE, rekurzívne, 2013-2014	MAPE, priemer
benchmark	QoQ c ARMA(2,3)	0.21	0.08	0.06	0.12
benchmark	YoY c ARMA(2,2)	0.26	0.11	0.13	0.17
benchmark	QoQ c AR(1)	0.32	0.14	0.15	0.20
benchmark	YoY c AR(1)	0.72	0.45	0.47	0.54
HP	QoQ c HP HP(-1) HP(-2) ARMA(2,3)	0.10	0.05	0.10	0.08
HP	QoQ c qoq(HP) qoq(HP(-1)) qoq(HP(-2)) ARMA(2,3)	0.11	0.05	0.17	0.11
HP	QoQ c qoq(HP(-1)) ARMA(2,3)	0.12	0.05	0.17	0.12
HP	QoQ c qoq(HP) ARMA(2,3)	0.18	0.10	0.11	0.13
PP	QoQ c PP PP(-1) ARMA(2,3)	0.17	0.08	0.10	0.12
PP	QoQ c PP ARMA(2,3)	0.20	0.08	0.09	0.12
PP	YoY c yoy(PP) ARMA(2,2)	0.18	0.07	0.14	0.13
PP	QoQ c PP(-1) ARMA(2,3)	0.14	0.09	0.19	0.14
UP	QoQ c UP ARMA(2,3)	0.12	0.05	0.11	0.09
UP	QoQ c qoq(UP) qoq(UP(-1)) qoq(UP(-2)) ARMA(2,3)	0.10	0.05	0.15	0.10
UP	QoQ c UP(-1) ARMA(2,3)	0.15	0.07	0.16	0.12
UP	QoQ c qoq(UP) qoq(UP(-1)) qoq(UP(-2)) AR(1)	0.27	0.12	0.12	0.17
PR	QoQ c qoq(PR) qoq(PR(-1)) qoq(PR(-2)) ARMA(2,3)	0.17	0.08	0.13	0.13
PR	QoQ c qoq(PR(-1)) ARMA(2,3)	0.21	0.08	0.11	0.13
PR	QoQ c qoq(PR) ARMA(2,3)	0.18	0.08	0.14	0.13
PR	QoQ c PR(-1) ARMA(2,3)	0.19	0.07	0.37	0.21

Tabuľka P3 – vyhodnotenie prognóz nezamestnanosti*

Indikátor	Špecifikácia	MAPE, 2008-2014	MAPE, 2013-2014	MAPE, rekurzívne, 2013-2014	MAPE, Priemer
benchmark	QoQ ARMA(4,3)	2.03	0.44	0.74	1.07
benchmark	YoY c MA(4)	3.83	1.84	2.30	2.66
benchmark	QoQ c AR(1)	2.51	0.71	0.72	1.31
benchmark	YoY c AR(1)	5.58	1.96	2.12	3.22
HP	QoQ c HP(-1) ARMA(4,3)	0.85	0.18	1.37	0.80
HP	QoQ c qoq(HP) qoq(HP(-1)) qoq(HP(-2)) ARMA(4,3)	0.98	0.52	1.42	0.97
HP	QoQ c HP(-1) AR(1)	2.40	0.71	0.81	1.31
HP	QoQ c HP(-1)	3.53	0.66	0.74	1.64
PP	QoQ c qoq(PP) qoq(PP(-1)) qoq(PP(-2)) ARMA(4,3)	1.09	0.35	0.99	0.81
PP	QoQ c PP ARMA(4,3)	1.95	0.70	0.51	1.05
PP	QoQ c PP PP(-1) PP(-2) ARMA(4,3)	1.70	0.40	1.17	1.09
PP	QoQ c PP	3.26	0.55	0.57	1.46
UP	QoQ c qoq(UP) qoq(UP(-1)) ARMA(4,3)	1.90	0.51	0.48	0.96
UP	QoQ c qoq(UP) ARMA(4,3)	1.87	0.50	0.65	1.01
UP	QoQ c qoq(UP(-1)) ARMA(4,3)	1.91	0.63	0.93	1.15
UP	QoQ c qoq(UP) qoq(UP(-1)) qoq(UP(-2)) ARMA(4,3)	1.59	0.64	1.43	1.22
PR	QoQ c PR ARMA(4,3)	2.03	0.45	0.69	1.06
PR	QoQ c qoq(PR) qoq(PR(-1)) ARMA(4,3)	1.40	0.88	0.94	1.07
PR	QoQ c qoq(PR(-1)) ARMA(4,3)	1.97	0.52	0.75	1.08
PR	QoQ c PR AR(1)	2.51	0.65	0.70	1.29

* Skratky QoQ alebo YoY na začiatku špecifikácie predstavujú transformáciu pre závislú premennú v danej rovnici; c je konštanta. Zelenou farbou písma je označený najlepší benchmark model. Zelené pozadie majú tie indikátorové modely, ktoré sú lepšie v danej metóde vyhodnotenia ako najlepší benchmark model. Červené hrubé písmo má model, ktorý je v danej metóde vyhodnotenia najlepší spomedzi všetkých indikátorových modelov. Čierne hrubé písmo má model, ktorý je najlepší pre daný výraz a danú metódu.