

Sezónne očisťovanie v európskom kontexte: Poučenie z analýz vývoja nezamestnanosti v časoch pandémie

Matúš Baliak

Matěj Bělín

Štefan Domonkos

Diskusná štúdia

november 2021

PodĎakovanie

Autori ďakujú Lucii Fašungovej a Jánovi Komadelovi (obaja ISP) za cenné rady a technickú podporu pri príprave tejto analýzy. Zároveň ďakujú trom recenzentom výstupu za jeho posúdenie.

Zhrnutie

Sezónne očisťovanie (SA) predstavuje výzvu už viac ako storočie, keď sa objavili prvé seriózne pokusy o úpravu časových radov o sezónny komponent. Dnes sa o sezónne očistené časové rady zaujíma tak odborná, ako aj laická verejnosť, avšak dôvod prečo by sme mali vyžadovať takto upravené dáta často chýba. Napriek výraznému metodologickému posunu za posledné dekády, proces sezónneho očisťovania vzbudzuje naďalej pozornosť, pretože v odbornej literatúre pretrváva polemika nad tým, akým spôsobom proces očisťovania vykonať a či je vôbec nutný. Úprava ekonomických veličín je spojená so stratou informácií a zjednodušením reality, pretože pozorovaný časový rad rozkladáme na nesledované komponenty.

Inštitút sociálnej politiky (ďalej aj „ISP“) preto pripravil diskusnú štúdiu venujúcu sa sezónnemu očisťovaniu, ktorá poskytuje rámcový prehľad súčasného stavu v danej oblasti s dôrazom na postupy odporúčané Eurostatom. Súčasťou našej štúdie je aj **popis praktického postupu sezónneho očistenia miery nezamestnanosti pomocou softvéru JDemetra+** vyvinutého Eurostatom v spolupráci vybranými národnými štatistickými úradmi v EÚ. Miera nezamestnanosti je jednou z najsledovanejších ekonomických indikátorov, u ktorej verejnosť javí vysoký záujem aj o sezónne očistené dáta. **Spôsob sezónneho očisťovania však často ostáva nevysvetlený.** Naša stať **zvyšuje transparentnosť procesu sezónneho očisťovania a znižuje informačnú asymetriu** medzi tvorcom sezónne očistených časových radov nezamestnanosti (ISP) a spotrebiteľmi dát.

Popri prezentácii postupu sezónneho očisťovania sa štúdia venuje aj testovaniu kvality rôznych postupov. Získané výsledky z uvedených testov naznačujú, že obe metódy, X-13ARIMA-SEATS a TRAMO-SEATS, zabezpečujú **adekvátne sezónne očistenie mesačnej miery nezamestnanosti** vypočítanej z celkového počtu UoZ a predstavujú vhodný zdroj informácií pre tvorcov politiky. Výber metódy sa odvíja od štruktúry dát a v prípade mesačnej miery nezamestnanosti mierne adekvátnejšie výsledky zabezpečuje X-13ARIMA-SEATS. V tejto súvislosti naša práca zároveň odporúča využívať viaceré techniky sezónneho očisťovania s cieľom zamedziť vyvodzovaniu hlavných záverov z možných štatistických reliktov.

Pre ilustráciu štatistických reliktov ukazujeme na mesačných dátach nezamestnanosti v SR, že aj štandardne používané techniky sezónneho očisťovania **môžu porušiť nesezónne charakteristiky časových radov.** Tieto porušenia následne vedú k odlišným odhadom dopadu pandémie COVID-19 na slovenský trh práce, ktoré sa môžu rozchádzať o zhruba 30 %, nakoľko nami navrhovaná metóda vedie k odhadu nárastu nezamestnanosti v následku epidémie COVID-19 o asi 2,95 percentuálneho bodu (pb), ale dáta očistené inou zvažovanou metódou vedú k odhadu nárastu nezamestnanosti iba o 2,11 pb.

1 Úvod

Úloha sezónneho očisťovania je pomerne jednoznačná, a to odstrániť pravidelne sa opakujúcu zložku, tzv. sezónny komponent zo zvoleného časového radu. Jednoznačnú definíciu sezónnosti by sme však hľadali márne. Z tohto dôvodu existuje mnoho prístupov ako sa s daným fenoménom vysporiadať. Kým metodologicky sa prístupy odlišujú, jedno majú spoločné: nedokážeme určiť, ktorý z nich je vhodnejší na vykonanie sezónneho očistenia. Napriek tomu, že vychádzajú z rozličných prístupov a aj dochádzajú k mierne rozdielnym záverom, kvôli rôznorodosti dôvodov pre sezónne očisťovanie a neexistencii presnej definície sezónnosti **nie je možné jednoznačne identifikovať najvhodnejší prístup.**

Dagum (1978) označuje za hlavné dôvody sezónnosti klimatické a inštitucionálne faktory, ktoré sú z pohľadu ekonomického systému z krátkodobého hľadiska nemenné a exogénne. S týmto súvisí aj motivácia určiť nesezónny komponent, ktorý na rozdiel od ťažko ovplyvniteľnej sezónnosti, môže byť podľa autorky čiastočne kontrolovaný, a preto považovala sezónne očisťovanie za vhodný vstup pre politické rozhodovanie. Bell a Hillmer (1984) uvádzali za najväčšie výhody sezónneho očisťovania pomoc s krátkodobým prognózovaním, s identifikáciou politických rozhodnutí na časový rad, prípadne hľadaním vzťahov medzi inými časovými radmi. Táto autorská dvojica zdôrazňovala, že sezónnym očisťovaním je zároveň možné získať porovnateľné časové rady nielen v čase, ale aj v priestore a určiť fázu krátkodobého ekonomického cyklu, čím čiastočne nadväzujú na Daguma (1978). Okrem diskusie o dôvodoch prečo vykonávať sezónne očisťovanie, naďalej pretrváva aj polemika o niektorých úzkych miestach úpravy časových radov (pozri napr. Huček a Doliak 2014). Ide napríklad o v čase sa meniacu sezónnosť, nesezónnu nestacionaritu v podobe trendu a cyklu, nemožnosť explicitne matematicky špecifikovať trend, cyklus, sezónny komponent a v neposlednom rade aj o úpravu extrémnych hodnôt. S cieľom poskytnúť odpoveď na tieto výzvy vytvoril Eurostat v spolupráci s národnými inštitúciami niekoľkých európskych štátov **softvér, ktorý prepája dnes najvyužívanejšie metódy X-13ARIMA-SEATS** (ARIMA – Autoregressive integrated moving average, SEATS – Signal Extraction in ARIMA Time Series) a **TRAMO-SEATS** (TRAMO - Time Series Regression with ARIMA Noise, Missing Observations, and Outliers). K **bezplatnému a verejne dostupnému softvérovému nástroju JDemetra+** je k dispozícii aj dokumentácia, určujúca minimálne požiadavky na vykonanie sezónneho očistenia pre krajiny Európy, aby získané údaje boli vzájomne porovnateľné.

Prvé pokusy o “sezónne očisťovanie“ pozostávali z jednoduchého zjemnenia priebehu kriviek ekonomických veličín v čase, bez akéhokoľvek matematického zdôvodnenia. Zvýšený záujem odbornej komunity a aj politických aktérov identifikovať cyklický vývoj hospodárstiev v časových radoch, súčasne s rozmachom výpočtovej techniky, umožnili rozvoj sofistikovanejších postupov, ktoré tvoria základ aj dnešných metód. Klasická deľba časového radu na štyri komponenty podľa Persons (1919), t. j. na sekulárny trend, cyklus, sezónny komponent a reziduálnu variáciu, sa stala súčasťou štandardného uvažovania o časových radoch práve v tomto období.¹ Úvodné pokusy o odhad nepozorovaných komponentov využívali kĺzavé priemery, ktoré boli odpočítané od časového radu, čím dokázali odstrániť trend/cyklus. Zvyšok časového radu pozostával zo sezónneho komponentu a reziduálnej variácie. Metóda s použitím kĺzavých priemerov čelila kritike spojenej s výrazným vplyvom extrémnych pozorovaní a nemožnosťou rozlíšiť medzi cyklom a sezónnosťou. Zároveň prvotné pokusy využívali fixné sezónne faktory pre celé sledované obdobie, čo však aj samotný Persons (1919) spochybňoval. Reakcia na niektoré z výhrad bol prechod na využívanie mediánových hodnôt, ktoré dokázali minimalizovať efekty extrémnych hodnôt (tzv. outlierov), avšak s ostatnými výzvami sa nedokázali vysporiadať.

Postupne narastal vplyv metód využívajúcich **vážené kĺzavé priemery**, ktoré tvoria základ dnešných X-11, X-12ARIMA a **X-13ARIMA-SEATS** metód. Iný prístup predstavuje rozklad časového radu na základe ekonometrických modelov, ktorý v počiatočných fázach nedokázal explicitne matematicky špecifikovať prvky pozorovaného časového radu a ich zápis závisel od svojvoľne zvolených predpokladov vykonávateľa sezónneho očisťovania, s čím súvisela nespoľahlivosť výsledkov. Ak by sme vedeli, čo odhadujeme, tak možno nebude polemika o tom ako to odhadujeme (Bell a Hillmer, 1984). **Výsledkom zdokonaľovania tohto prístupu**

¹ Dnešné metódy však nerozlišujú medzi trendom a cyklom, ktoré sú v procese sezónneho očisťovania odhadnuté spoločne.

je **TRAMO-SEATS**, ktorého prvá časť pozostáva z modelovania časového radu prostredníctvom regresného ARIMA modelu. Cez program SEATS sú v ďalšom kroku odhadnuté jednotlivé komponenty časového radu. Následne je možné časový rad očistiť o jeho sezónnu zložku. Eurostatom podporovaný softvérový nástroj **JDemetra+** umožňuje aplikáciu oboch vyššie uvedených prístupov vrátane manuálneho nastavenia vybraných parametrov dôležitých pre proces sezónneho očisťovania.

Druhá kapitola tejto štúdie bude venovaná stručnému popisu jednotnej metodiky Eurostatu a odporúčaným krokom v procese sezónneho očisťovania. Následne v tretej časti špecifikujeme postup, akým sme vykonali sezónne očistenie mesačných údajov miery nezamesnanosti vypočítanej z celkového počtu UoZ na Slovensku, ktorú zverejňuje Ústredie práce, sociálnych vecí a rodiny (ÚPSVaR). Miera nezamestnanosti, vrátane jej sezónne očistenej verzie, je jednou z najsledovanejších ekonomických ukazovateľov. Primerane detailný popis spôsobu sezónneho očisťovania je preto dôležitý z pohľadu transparentnosti zverejňovaných dát.²

2 Sezónne očisťovanie podľa odporúčaní Eurostatu

2.1 Metodika Eurostatu

Dokument **ESS guidelines on seasonal adjustment** (Eurostat, 2015) predstavuje návrh **jednotnej metodiky procesu sezónneho očisťovania od Eurostatu**, ktorej cieľom je harmonizácia použitých metód v rámci členských štátov EÚ.³ Implementácia odporúčaných štandardov vytvára vhodné podmienky na porovnateľnosť výsledkov medzi krajinami v čase a súčasne zabezpečuje zvýšenú robustnosť európskych agregátnych štatistík (Eurostat, 2015). Zjednodušene môžeme rozdeliť odporúčanú metodiku Eurostatu do niekoľkých parciálnych častí: príprava dátovej základne spolu s voľbou parametrov modelu, výber metódy sezónneho očistenia, kontrola výsledkov - úprava a revízia, komunikácia a zverejňovanie sezónne očistených údajov. Dokument vymedzuje spomenuté kroky a pri každom z nich sú navrhnuté tri alternatívy (A, B, C). Prvé dve sú akceptovateľné, tretej by sa analytici mali vyhnúť.

Zámerom odporúčanej metodiky Eurostatu je získať údaje, ktoré nebudú obsahovať pravidelne sa opakujúce fluktuácie, avšak ostatné nepravidelné a neočakávané pohyby v časovom rade naďalej zostanú viditeľné (Eurostat, 2015). Prostredníctvom *JDemetra+* dokážu analytici implementovať väčšinu odporúčaní bez závažných komplikácií. *JDemetra+* predstavuje spojenie dvoch v súčasnosti najintenzívnejšie využívaných metód, X-13ARIMA-SEATS a TRAMO-SEATS. Ich bližšie fungovanie si uvedieme v Kapitole 3, venovanej rozkladu konkrétneho časového radu.

V nasledujúcich odsekoch sa zameriame na sériu odporúčaní prameniáciach z metodiky Eurostatu.⁴

Box 1 Komponenty časového radu

Úroveň mnohých ekonomických veličín sa vyvíja v závislosti od obdobia v roku. Tieto **sezónne efekty** spôsobujú, že sa v určitom období pozorovaného časového radu vyskytujú **podobné výkyvy v ročných (prípadne kratších) intervaloch**. Aby sme získali **transparentný prehľad** o vývoji ekonomických veličín potrebujeme **odstrániť tento každoročne sa opakujúci sezónny vplyv**. Proces sezónneho očisťovania zahŕňa dekompozíciu sledovaného časového radu Z_t , na nesledované sezónne a nesezónne komponenty S_t a N_t . Tieto prvky môžu tvoriť pozorovaný časový rad buď ako súčet prvkov, teda aditívna dekompozícia (1) alebo ako súčin, tzv. multiplikatívna dekompozícia (2). Ak zlogaritmujeme multiplikatívny zápis, dostaneme aditívny. Ak časový rad obsahuje nulu, prípadne záporné hodnoty je využitá aditívna forma rozkladu.

$$(1) Z_t = TC_t + S_t + I_t$$

$$(2) Z_t = TC_t * S_t * I_t$$

² Pre implementáciu obdobných metód sezónneho očisťovania na iné časové rady pozri napr. Huček a Doliak 2014 a Nikolova a Elliott 2019.

³ Pre viac informácií pozri Eurostat (2015).

⁴ Praktický manuál využitia *JDemetra+* pripravilo aj UNECE pre viac informácií pozri UNECE (2020).

Nesezónny komponent N_t môžeme ďalej rozdeliť na trend/cyklus TC_t a reziduálnu variáciu I_t , ktorá sa v priebehu času vyvíja náhodne (white noise / biely šum) a predstavuje tú časť časového radu, ktorá nie je systematická ani predpovedateľná. Kým komponent TC_t predstavuje dlhodobý vývoj časového radu a zahŕňa aj cyklus, tzn. odchýlky od dlhodobého trendu, sezónny komponent S_t sa vyskytuje v pravidelných ročných intervaloch.

2.1.1 Počiatkové kroky v procese sezónneho očistenia

Skôr ako začneme s importovaním neočistených údajov do softvéru *JDemetra+*, je nevyhnutné mať dáta v požadovanej štruktúre. Softvér sa dokáže vysporiadať s dátami vo väčšine bežne využívaných formátov. Prvý stĺpec však musí byť vo formáte dátum. Dĺžka časového radu s mesačnou frekvenciou musí dosahovať minimálne tri roky, aby bolo možné vykonať sezónne očistenie. Pre dostatočne kvalitné sezónne očistenie sú však odporúčané dlhšie, aspoň päť až sedemročné časové rady (Eurostat, 2015).

Úvodným krokom pri procese sezónneho očistenia je príprava údajov, tzv. **pre-treatment**. Kvalita sezónneho očistenia sa odvíja do veľkej miery práve od prípravy časového radu vo fáze pre-treatmentu, pretože v tomto kroku sa odhadujú jednotlivé komponenty. Preto je dôležité, aby užívatelia venovali tejto fáze primeranú pozornosť. Pre-treatment možno rozdeliť do niekoľkých krokov:

- i. voľba medzi prednastaveným pre-treatmentom alebo špecificky upraveným pre-treatmentom
- ii. grafická analýza neočistených dát
- iii. úprava časového radu o kalendárne efekty
- iv. detekcia a korekcia časového radu o odľahlé hodnoty (tzv. outliers)
- v. výber parametrov modelu a forma dekompozície

Box 2 Úloha predspracovania (pre-treatmentu) v procese sezónneho očistenia

Primárnym cieľom pre-treatmentu je zabezpečiť spoľahlivý odhad sezónneho komponentu. Postupnou aplikáciou jednotlivých krokov získame časový rad, upravený o nelinearity, ktoré by mohli negatívnym spôsobom skresliť veľkosť nepozorovaných komponentov. **Outliery** predstavujú príklad nelineárneho vývoja časového radu, preto je nutné ešte pred dekompozíciou korigovať časový rad o tieto extrémne hodnoty. Odporúčaným nástrojom na úpravu časového radu sú **RegArima modely**, ktoré upravujú hodnoty **extrémnych pozorovaní** (Eurostat, 2015). Zároveň dokážu identifikovať a upraviť časový rad o tzv. **kalendárne efekty**, ktoré plynú z rôznej štruktúry dní v sledovanom období (rôzny počet pracovných dní v mesiaci, prípadne v týždni, priestupný rok, bridging day⁵, efekt školských prázdnin a ďalšie). Kalendárne efekty, ktoré sa opakujú v ročnom intervale sú považované za súčasť sezónneho komponentu. V tomto kroku sa však snažíme odstrániť nesezónne efekty kalendárnych zmien, ktoré neboli identifikované ako sezónne. Najvhodnejšou metódou je **vytvorenie vlastného detailného národného kalendára**, ktorý je súčasťou neautomatického pre-treatmentu. To znamená špecifikovať vlastný RegArima model, ktorý bude obsahovať aj informácie týkajúce sa sviatkov, prázdnin a počtu pracovných dní v mesiacoch. Prostredníctvom tohto druhu úpravy získame údaje, ktoré sú vzájomne porovnateľné medzi mesiacmi. V prípade existencie prekážok pri definovaní samostatného detailného RegArima modelu je vhodné použiť automatický pre-treatment, ktorý je prednastavený v softvéri *JDemetra+* (Eurostat, 2015).

Špecificky upravený pre-treatment predstavuje *Alternatívu A*. To v praxi znamená vykonať štatistickú kontrolu výsledkov minimálne raz ročne, súčasne aplikovať národný kalendár a voľbu parametrov modelu na základe výsledkov vykonaných testov. Výber v prvom kroku zásadne ovplyvňuje postupnosť v nasledujúcich krokoch pre-treatmentu. V mnohých prípadoch je prijateľnou alternatívou aj voľba automatického pre-treatmentu, čo tvorí *Alternatívu B*.

⁵ Deň medzi sviatkom a víkendom - piatkom, ktoré sú síce vnímané ako čisto pracovné dni, ale zároveň sú často využívané ako možnosť na predĺžený víkend, čím dochádza k negatívnym výkyvom napr. v produkcii.

Grafická analýza neočistených dát je počiatočným krokom pri príprave dát, ktorá poskytuje užitočné informácie o ich štruktúre (výskyt chýbajúcich pozorovaní, časová konzistencia pozoraní v podobe trendu, cyklu, prípadne zlomov v časovom rade). Zároveň uľahčuje voľbu metódy dekompozície a parametrov modelu. *Alternatívou A* je v tomto kroku detailná grafická analýza všetkých neočistených časových radov, ktoré chceme sezónne očisťovať (Eurostat, 2015).

Súčasť pre-treatmentu tvorí aj **úprava časového radu o kalendárny efekt**. Existencia kalendárneho efektu plynie z rozličnej štruktúry dní v sledovanom období (napr. v mesiaci, kvartáli, na ktoré môže pripadať odlišný počet víkendov a pracovných dní). Získané údaje sú vzájomne porovnateľné medzi mesiacmi a ich hodnota je upravená o efekt kalendárnych zmien. Treba zdôrazniť, že úprava o kalendárny efekt by sa mala vykonať výhradne pre premenné, u ktorých existuje podozrenie, že štruktúra mesiaca dokáže ovplyvniť sledovanú premennú. *Alternatívou A* je vytvorenie vlastného kalendára, odražajúceho štruktúru dní v jednotlivých mesiacoch (Eurostat, 2015). *Alternatívou B* je aplikácia prednastaveného kalendára v softvéri *JDemetra+*, v ktorom špecifikujeme národné sviatky.

Ďalším krokom v procese pre-treatmentu je **detekcia a korekcia extrémnych hodnôt, tzv. outlierov**. Outliery predstavujú také hodnoty časového radu, ktoré sa výraznejšie odchyľujú od očakávaného vývoja a označujeme ich ako nelinearity. Ich výskyt dokáže negatívne ovplyvniť modelovanie a následnú dekompozíciu časového radu, a teda aj kvalitu procesu sezónneho očisťovania načo upozorňuje aj metodika Eurostatu. *Alternatívou A* je úprava extrémnych hodnôt súvisiacich s chybami v meraní v neupravených dátach ešte pred samotným vykonaním pre-treatmentu. Pri ostatných typoch outlierov sa odporúča detekcia a nahradenie pri procese modelovania. Metodika upozorňuje, že počet identifikovaných outlierov sa odvíja od hodnoty testovacej štatistiky a dĺžky časového radu, z ktorého sa extrémne hodnoty identifikujú. Oba parametre je vhodné prispôbiť. Súčasťou *Alternatívy A* je aj manuálna špecifikácia vopred známych extrémnych hodnôt, o ktorých máme dostatok informácií. *Alternatíva B* je automatická detekcia a úprava extrémnych hodnôt.

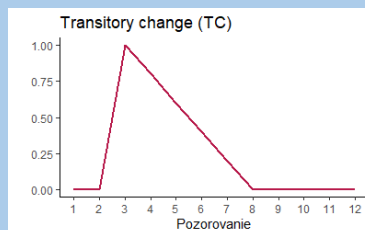
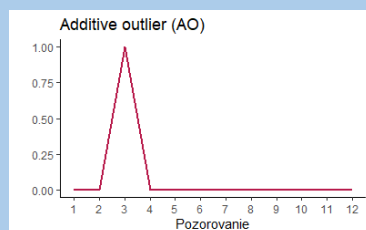
Box 3 Základné druhy odľahlých pozorovaní (outlierov)

AO – jednorazová neočakávaná extrémna hodnota

TC – prudká zmena v úrovni, nasledovaná postupným návratom na počiatočnú hodnotu

LS – permanentná zmena v úrovni

Graficky môžeme znázorniť jednotlivé typy približne nasledovne:



Zdroj: Vlastné spracovanie

Určenie typu outliera na konci a začiatku časového radu môže byť problematické. Pri prvom a poslednom dostupnom pozorovaní nedokážeme rozlíšiť medzi LS a AO. Z tohto plynie problém pri pripisovaní outliera k trendu/cyklu prípadne reziduálnej variácii. Kvalita očistených dát týmto procesom neklesá, pretože očistené údaje obsahujú oba komponenty, avšak nemusia byť pripísané správne. Jednou z vhodných možností je špecifikovať výskyt a typ outlierov (pri disponovaní potrebnými informáciami), čím zabezpečíme, že outlier bude zahrnutý k správnej komponente. Odporúča sa však upozorniť verejnosť na určitú mieru neistoty spojenú so sezónne upravenými dátami a možnosťou dodatočnej korekcie.

Metodika Eurostatu kladie dôraz na **vhodnú špecifikáciu modelu** pomocou ktorého sa identifikujú jednotlivé zložky časového radu. Tento krok pozostáva z výberu viacerých parametrov: stupeň diferenciácie, voľba medzi aditívnym resp. multiplikatívnym rozkladom, zlogaritmovanie respektíve ponechanie pôvodných dát. V *Alternatíve A* dochádza k neautomatickému výberu parametrov modelu pre problematické časové rady. Pri tejto alternatíve zostáva priestor aj na výber z predvolených špecifikácií, pre ktoré je nevyhnutné vykonať kontrolu

vhodnosti. Vo väčšine prípadov postačí diferencovať časový rad iba jedenkrát, voľba medzi aditívnymi a multiplikatívnymi špecifikáciami závisí od toho, či fluktuácie časového radu majú stabilnú variáciu počas celého pozorovaného obdobia. Zmeny vo veľkosti fluktuácií v závislosti na zmenách hodnôt časového radu sú typickejšie pre multiplikatívne modely. *Alternatíva B* sa opiera o automatický proces výberu modelu. Od zvolených nastavení modelu sa odvíja aj odhad vývoja časového radu (prednastavená prognóza na obdobie 12 mesiacov). Prednastavenej prognóze sa nebudeme v tejto práci venovať detailne, nakoľko nejde o nevyhnutnú súčasť hlavnej témy tejto práce, procesu sezónneho očisťovania. Zároveň pri prednastavených prognózach *JDemetra+* vychádza iba z minulého vývoja časového radu a neberie do úvahy exogénne vplyvy (napr. vývoj epidemickej situácie). Takýto zjednodušený pohľad na vývoj na trhu práce môže mať najmä v čase štrukturálnych zlomov obmedzenú výpovednú hodnotu.

Tabuľka 1 Prehľad predvolených špecifikácií modelov v *JDemetra+*

Špecifikácia	Transformácia	Úprava o priestupný rok	Pracovné dni (Working days)	Obchodné dni (Trading days)	Efekt Veľkej noci	Outliery	ARIMA model
TRAMO-SEATS							
RSA0	Nie	Nie	Nie	Nie	Nie	Nie	(0,1,1)(0,1,1)
RSA1	Test	Nie	Nie	Nie	Nie	Test	(0,1,1)(0,1,1)
RSA2	Test	Nie	Test	Nie	Test	Test	(0,1,1)(0,1,1)
RSA3	Test	Nie	Nie	Nie	Nie	Test	AMI
RSA4	Test	Nie	Test	Nie	Test	Test	AMI
RSA5	Test	Nie	Nie	Áno	Test	Test	AMI
RSAfull	Test	Áno	Nie	Test	Test	Test	AMI
X-13ARIMA-SEATS							
X11	Nie	Nie	Nie	Nie	Nie	Nie	(0,1,1)(0,1,1)
RSA1	Test	Nie	Nie	Nie	Nie	Test	(0,1,1)(0,1,1)
RSA2c	Test	Test	Test	Nie	Test	Test	(0,1,1)(0,1,1)
RSA3	Test	Nie	Nie	Nie	Nie	Test	AMI
RSA4c	Test	Test	Test	Nie	Test	Test	AMI
RSA5	Test	Test	Nie	Test	Test	Test	AMI

Zdroj: GRUDKOWSKA, S. (2017a)

Pozn.: Skratka RSA (routine seasonal adjustment) predstavuje predvolené špecifikácie definujúce základné parametre modelov. Pri niektorých špecifikáciách je vykonaný test na zistenie významnosti vybraných faktorov, následne je model rozšírený o dané regresory. AMI predstavuje automatickú identifikáciu modelu. Model (0,1,1)(0,1,1) je tzv. Airline model, ktorý sa osvedčil pri sezónnom očisťovaní mesačných dát. ARIMA model slúži na identifikáciu vlastností časového radu a jeho špecifikácia sa kompaktné uvádza do dvoch zátvoriek, kde prvá zátvorka obsahuje špecifikáciu ne-sezónnej zložky a druhá zátvorka obsahuje špecifikáciu sezónnej zložky. Tri údaje v zátvorke určujú jednotlivé časti ARIMA modelu (AR – počet autoregresných parametrov, I – stupeň diferenciacie, MA – počet parametrov kľzavých priemerov).

2.1.2 Vykonanie sezónneho očistenia

Po príprave dát v prvom kroku a odhade komponentov časového radu, vykonávateľ pristúpi k sérii rozhodnutí týkajúcich sa samotného sezónneho očistenia, ktoré metodika rozdeľuje nasledovne:

- i. voľba metódy sezónneho očistenia
- ii. výber softvéru
- iii. benchmarking
- iv. voľba medzi priamou a nepriamou metódou sezónneho očisťovania

Alternatívou A v prvom bode je využitie jednej z **dvoch odporúčaných metód**: X-13ARIMA-SEATS respektíve TRAMO-SEATS. Prvý zo spôsobov rozkladá časový rad na nesledované komponenty využitím vážených kľzavých priemerov (lineárnych filtrov) a následne je sezónny komponent odstránený. Druhú z metód využíva

odlišný prístup, ktorý rozkladá časový rad na nepozorované komponenty na základe odhadnutého ARIMA modelu (model-based approach), následne dochádza k sezónnemu očisteniu. Metodika uvádza, že pri výbere je potrebné zväžiť výsledky diagnostických testov a reagovať nato pri výbere metódy.

Ďalším dôležitým krokom je **výber vhodného softvéru**. Keďže metodika Eurostatu poskytuje súbor odporúčaní šitých na mieru *JDemetre+*, tak práve tento voľne dostupný a oficiálne schválený softvér predstavuje *Alternatívu A*. V rozhraní *JDemetra+* dokáže vykonávateľ sezónneho očisťovania implementovať väčšinu odporúčaní bez závažných problémov.

Benchmarking v procese sezónneho očisťovania znamená zachovanie sumy sezónne očistených ku sezónne neočisteným dátam (prípadne ich priemerom) za sledované obdobie. Manuál Eurostatu takéto obmedzenie neodporúča, čo tvorí *Alternatívu A*. Existujú však aj prípady (národné účty – nízka frekvencia zverejňovania), v ktorých je potrebná konzistencia údajov a benchmarking predstavuje vhodný nástroj, ako dosiahnuť požadovaný stav.

V nasledujúcej fáze musí vykonávateľ urobiť výber medzi **priamou alebo nepriamou metódou sezónneho očisťovania**. Nepriama metóda aplikuje sezónne očistenie na dis-agregované dáta (napríklad nezamestnanosť v jednotlivých okresoch) a vytvára celkový očistený časový rad (nezamestnanosť na úrovni štátu) ich spojením. Priama metóda aplikuje sezónne očistenie na agregovaný časový rad. Voľba medzi oboma metódami sa vo veľkej miere odvíja od druhu a štruktúry údajov a sezónnosti v jednotlivých kategóriách. Ak komponenty vykazujú podobnú sezónnosť *Alternatívou A* je priama metóda. Zároveň však obsahuje odporúčanie, že v prípade, ak subkategórie vykazujú rozličný stupeň sezónnosti, tak je vhodné využiť nepriamu metódu sezónneho očisťovania.

V súvislosti s nepriamou metódou očisťovania je namieste spomenúť očisťovanie ukazovateľov vypočítaných z viacerých (potenciálne sezónne oscilujúcich) premenných ako je napríklad miera nezamestnanosti. Miera nezamestnanosti je podiel počtu nezamestnaných a ekonomicky aktívneho obyvateľstva. Mohli by sme teda uvažovať o „priamom“ prístupe očistenia celého zlomku, alebo o „nepriamom“ prístupe očistenia čitateľa aj menovateľa zvlášť. Podobne ako dis-agregovaných dátach je ale „nepriama“ metóda menej vhodná. V prvom rade sa nepresnosti v sezónnom očisťovaní čitateľa a menovateľa môžu navzájom umocniť, takže kvalita očistenia jednotlivých komponentov môže byť podstatne nižšia ako kvalita očistenia zlomku priamo. Navyše nakoľko ekonomicky aktívne obyvateľstvo nie je merané mesačne a len ťažko je možné si pri ňom predstaviť sezónne efekty, očisťovanie čitateľa a menovateľa zvlášť nie je opodstatnené.

Box 4 Stručná charakteristika X-13ARIMA-SEATS a TRAMO-SEATS

Obe metódy môžeme rozdeliť do dvoch základných krokov, pre-treatment-modeling a dekompozíciu časového radu. V prvej fáze oba prístupy používajú podobný postup, v rámci ktorého sa aplikuje modelovanie časového radu cez lineárne regresné modely s ARIMA chybami. V druhej fáze, dekompozícii časového radu a odstránení sezónneho komponentu, sa však prístupy X-13ARIMA-SEATS a TRAMO-SEATS odlišujú. Zatiaľ čo pri metóde TRAMO-SEATS sa časový rad rozkladá využitím ARIMA modelu, pri X-13ARIMA-SEATS sa okrem použitia ARIMA modelu aplikuje aj rozklad cez lineárne filtre (napr. filter X-11).

Metóda sezónneho očisťovania TRAMO-SEATS je výsledkom práce ekonómov Národnej banky Španielska a predstavuje prepojenie dvoch programov. Kým v prvej fáze (pre-treatment) TRAMO modeluje časový rad využitím regresného modelu a zároveň prognózuje vývoj časového radu, tak SEATS v druhej časti odhadne nesledované komponenty na základe modelu z predchádzajúceho kroku a časový rad je následne sezónne očistený (Caporello a Maravall, 2004).

Metóda X-13ARIMA-SEATS je založená na predchádzajúcich verziách X-11, X-11-ARIMA. V prvom kroku aplikuje ARIMA model na úpravu a prognózu časového radu. Následne v druhej fáze sa využije na dekompozíciu buď metóda X-11 (rozklad cez lineárne filtre) alebo odhad nesledovaných komponentov prostredníctvom SEATS. V ďalšom kroku je časový rad zbavený sezónneho komponentu.

TRAMO-SEATS predpokladá aditívny model dekompozície časového radu, prípadne poskytuje možnosť aplikovať zlogaritmovaný aditívny model. X-13ARIMA-SEATS umožňuje okrem týchto dvoch aditívnych modelov taktiež využiť aj multiplikatívny rozklad.

Ako už bolo niekoľkokrát uvedené v predchádzajúcich častiach, časový rad môže trpieť rôznymi nedokonalosťami. Na modelovanie časového radu a jeho úpravu X-13ARIMA-SEATS a TRAMO-SEATS používajú podobný postup, lineárny regresný model s ARIMA chybami (RegARIMA, resp. TRAMO). Ako je uvedené v **Tabuľke 1**, predvolené špecifikácie ARIMA modelov používajú dve základné nastavenia. Prvou možnosťou je tzv. Airline model $(0,1,1)(0,1,1)$, ktorý sa osvedčil pri sezónnom očisťovaní mesačných dát. Druhou alternatívou je AMI, čo znamená, že je zvolený najvhodnejší ARIMA model automaticky (najmä na základe testov jednotkového koreňa a autokorelácie rezíduí). ARIMA model sa využíva aj na odhad časového radu v horizonte 12 mesiacov (prípadne iný horizont), čo limituje vplyv dodatočných korekcií pri prvom a poslednom pozorovaní, pri využití symetrických lineárnych filtrov.

Pri prvom odhade modelu je vhodné využiť RSAfull alebo RSA5 pri TRAMO-SEATS ako benchmark, kým u X-13ARIMA-SEATS sú to verzie RSA5 a RSA4c, ktoré aplikujú testy na identifikovanie vlastností časového radu. Získané výsledky z prvého kroku umožňujú vhodnejšie špecifikovať parametre sezónneho očisťovania.

2.1.3 Revízia a zmena modelu

Sezónne očisťovanie je opakujúci sa proces, ktorý sa vykonáva zakaždým, keď sú zverejnené nové údaje, čo vytvára priestor na revíziu, tak modelu, ako aj očistených dát. Revízia môže nastať, buď ako výsledok úpravy neočistených dát, alebo v dôsledku získania dodatočných pozorovaní. Pri zmene vstupných údajov dochádza aj k odhadu nového modelu, ktorý poskytuje lepší obraz vývoja časového radu. V oboch prípadoch je výsledkom aj zmena v sezónne očistených dátach, a preto by mali mať analytické jednotky stanovený postup, akým bude dochádzať k takejto revízii. Metodika Eurostatu definuje v tejto fáze dôležitú povinnosť a to:

i. výber medzi pravidelnou aktualizáciou (current adjustment) a priebežnou aktualizáciou (concurrent adjustment)

Obe uvedené stratégie predstavujú hraničné prístupy, akými vykonávať sezónne očisťovanie. Pravidelná aktualizácia minimalizuje revíziu modelu. To v praxi znamená, že pôvodne odhadnutý model sa používa na očistenie všetkých nasledujúcich pozorovaní. K revízii modelu by malo podľa odporúčaní dochádzať raz ročne (Eurostat, 2015). Pri výbere tejto možnosti získame najstabilnejší vývoj očistených radov. Priebežná aktualizácia predstavuje opačný extrém, keď sú parametre modelu odhadnuté nanovo, ak sú dodané dodatočné pozorovania. Prostredníctvom priebežnej aktualizácie získavame aktuálnejší pohľad na vývoj časového radu, pretože model zohľadňuje všetky dostupné informácie. Cenou pri výbere priebežnej aktualizácie je nestabilita sezónne očistených údajov, ktoré sa dodaním dodatočného pozorovania revidujú. *Alternatívou A* je v prípade stabilného sezónneho komponentu využitie pravidelnej aktualizácie. Ak dochádza k úprave pôvodných dát, metodika odporúča výber priebežnej, prípadne čiastočnej priebežnej (partial concurrent) úpravy, ktorá predstavuje medzistupeň medzi oboma hraničnými prístupmi. Bližšie vysvetlenie čiastočnej priebežnej úpravy nájdeme v Huček a Doliak (2014).

2.1.4 Kontrola výsledkov sezónneho očisťovania

Po špecifikácii modelu a získaní sezónne očistených údajov nastáva záverečná fáza, a to analyzovanie výsledkov. Tu metodika Eurostatu definuje niekoľko požiadaviek na správne vyhodnotenie kvality procesu:

- i. použitie detailnej validácie výsledkov, ktoré zahŕňajú širokú paletu štatistických testov (parametrické aj neparametrické kritéria), následne vykonať úpravu v špecifikácii modelu, ak je to nevyhnutné,
- ii. porovnať vhodnosť zvoleného prístupu sezónneho očisťovania prostredníctvom skupiny testov (M-štatistiky, výskyt (reziduálnej) sezónnosti a ďalšie)

Box 5 Diagnostické nástroje

Sekcia výstupu z *JDemetra+* venovaná kontrole vhodnosti použitého prístupu (Diagnostics) je rozdelená do piatich častí, ktoré sú rovnaké pre X-13ARIMA-SEATS a TRAMO-SEATS. Nájdeme v nej nasledovné

kategórie testov⁶: Seasonality tests, Spectral analysis, Sliding spans, Revisions history, Model stability. Každá kategória je rozdelená do niekoľkých indikátorov kvality, ktorých farba a popis dokumentujú kvalitu procesu.

Úvodná strana časti diagnostických nástrojov (**Diagnostics**) zachytáva všeobecné informácie o kvalite procesu sezónneho ošišovania. Poskytuje stručný prehľad o kvalite procesu sezónneho ošišovania, ktorému sa detailnejšie venujú jednotlivé subkategórie.

V časti zameranej na zistenie prítomnosti sezónnosti (**Seasonality tests**) je vykonaných niekoľko testov na kontrolu výskytu sezónnosti tak v pôvodnom časovom rade, ako aj časových radoch odvodených z pôvodného (napr. linearizovaný časový rad⁷, zisťovanie prítomnosti sezónnosti v rezíduách). V každej časti je deskriptívne opísaných sedem testov. V prípade rozporuplných výsledkov Grudawska (2017b) odporúča venovať väčšiu váhu parametrickým testom (testy 1, 6 a 7).

Spektrálna analýza poskytuje grafickú kontrolu rezíduí, nepravidelného komponentu a sezónne upravených dát na výskyt sezónnych prípadne obchodných (trading) efektov. Ak grafy obsahujú vrcholy na vybraných časových horizontoch, je nutné upraviť špecifikáciu modelu. Za kritickú hranicu sa považuje hodnota presahujúca hodnotu 10.

Analýza kĺzavých časových rozmedzí (**Sliding spans analysis**) je vhodná pre časové rady, ktoré trpia meniacou sa sezónnosťou alebo obsahujú mnoho extrémnych hodnôt (outlierov). Z časového radu sa vytvoria štyri osemročné časové intervaly, ktoré sa čiastočne prekrývajú. Tento druh analýzy sa vykoná pre sezónny komponent, sezónne očistené dáta a aj pre efekt obchodných dní. Kritická hranica je stanovená na úrovni 3 %.

História revízií (**Revision history**) poskytuje informácie o efekte rozšírenia časového radu o nové pozorovanie. Poskytuje grafickú analýzu vývoja sezónne očistených dát a trendu, súčasne poskytuje aj ich absolútnu (relatívnu) zmenu. Postup v tomto bode kontroly je nasledovný: najprv je odhadnutý model pre celý časový rad, následne sa časový horizont, z ktorého sa odhaduje model skracuje a dekompozícia sa upravuje. Pre každé pozorovanie je vykonaný odhad niekoľkokrát pre rôzne časové obdobia. Výsledky sa považujú za nestabilné v prípade, ak získaný rozdiel medzi odhadom v prvom období a poslednom období (obyčajne obdobie štyroch rokov) presahuje dvojnásobok RMSE (root mean squared error). Cieľom tohto testu je potvrdiť, či rozšírenie časového radu o nové hodnoty spôsobuje výrazné zmeny vo výsledkoch.

Poslednou časťou kontroly výsledkov je test stability modelu (**Model stability**), ktorý analyzuje stabilitu ARIMA parametrov pre kĺzavé časové obdobia v horizonte ôsmich rokov. V prípade, ak sledujeme nestabilný odhad parametrov v čase, indikuje to možnú revíziu údajov pri dodaní ďalších pozorovaní.

JDemetra+ poskytuje užívateľsky priateľské rozhranie a výsledky sú jednoducho čitateľné. Výsledky väčšiny testov dokážeme interpretovať na základe farby a popisu, ktoré prislúchajú k danému indikátoru. *JDemetra+* rozlišuje tri farby: zelenú, žltú a červenú. Kým zelená predstavuje želaný stav, kedy neboli nájdené žiadne problémy. Žltá znamená, že výsledok je neistý. V takom prípade je vhodné uvažovať nad úpravou procesu sezónneho očistenia. Červená predstavuje prípad, kedy je nutné prijať nápravu, pretože niekde v procese nastal problém.

Stručne sme opísali všetky nevyhnutné kroky v procese sezónneho ošišovania. Poslednou časťou dokumentu bude navrhovaná špecifikácia pri úprave mesačných údajov miery nezamestnanosti z celkového počtu UoZ, ktoré zverejňuje Ústredie práce, sociálnych vecí a rodiny. Prostredníctvom sezónneho očistenia by sme mali byť schopní lepšie porozumieť dynamike vývoja nezamestnanosti po hektickom období spojenom s vypuknutím pandémie Covid-19.

⁶ Využijeme anglické názvy

⁷ Linearizovaný časový rad je upravený pôvodný časový rad o outliery a kalendárny efekt

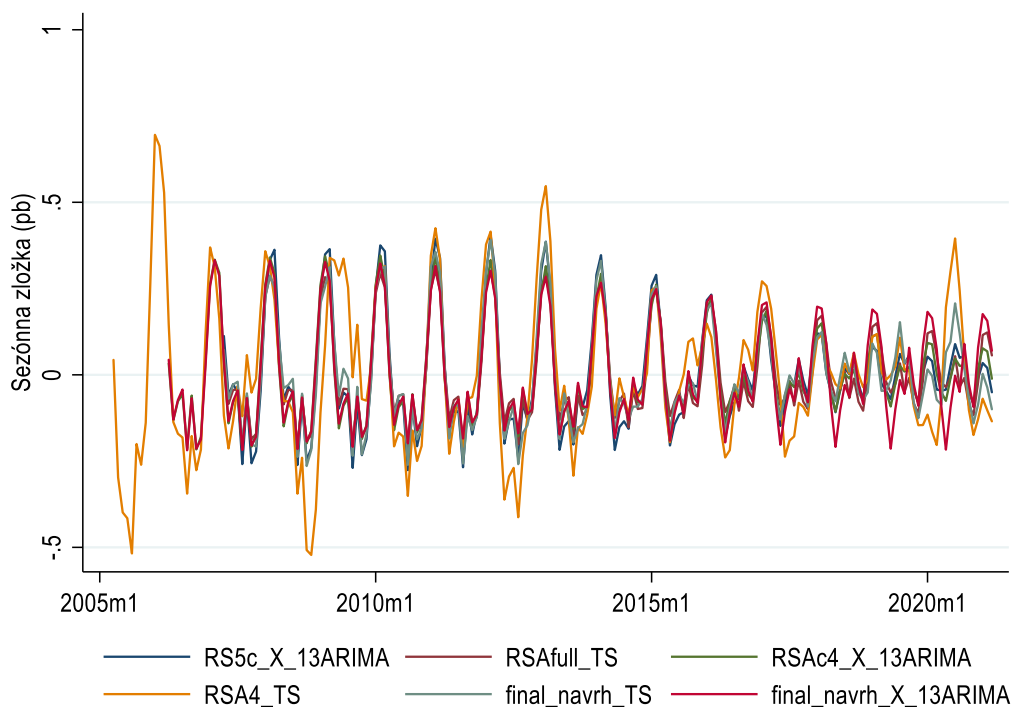
3 JDemetra+ v praxi: Návrh postupu sezónneho očisťovania mesačnej miery nezamestnanosti vypočítanej z celkového počtu UoZ

3.1 Úvodné kroky v procese sezónneho očisťovania

Nasledujúca časť sa zameria na **stanovenie postupu pri očisťovaní mesačného časového radu miery nezamestnanosti pomocou JDemetra+**. Už tým, že sme zvolili softvér *JDemetra+* na odstránenie sezónnych vplyvov sme splnili niekoľko predpokladov nato, aby naše upravené údaje boli konzistentné a porovnateľné v čase, tak ako odporúča metodika Eurostatu. Ešte pred samotným vykonaním sezónneho očistenia v softvéri, je vhodné využiť vstavané nástroje *JDemetra+*, ako grafické zobrazenie dát a identifikáciu outlierov. Nevyhnutnou súčasťou je detekcia výskytu sezónnosti v časovom rade. Spomínané úkony dokážu uľahčiť prácu pri špecifikácii nastavení sezónneho očisťovania.

Sezónne očistenie vykonáme tak spôsobom X-13ARIMA-SEATS, ako aj TRAMO-SEATS a následne identifikujeme najvhodnejší postup, akým sa dopracovať k sezónne očisteným údajom celkovej miery nezamestnanosti. **Tabuľka 1** opisuje detaily jednotlivých prednastavených špecifikácií, bližšie informácie sú dostupné v Grudawska (2017a). Obrázok 1 potom porovnáva sezónne zložky izolované rôznymi technikami. Je vidno, že rôzne techniky vedú k podobným odhadom, hoci RSA4_TS vykazuje výrazne volatilnejšiu sezónnu zložku.

Obrázok 1: Sezónne zložky izolované z časového radu miery nezamestnanosti

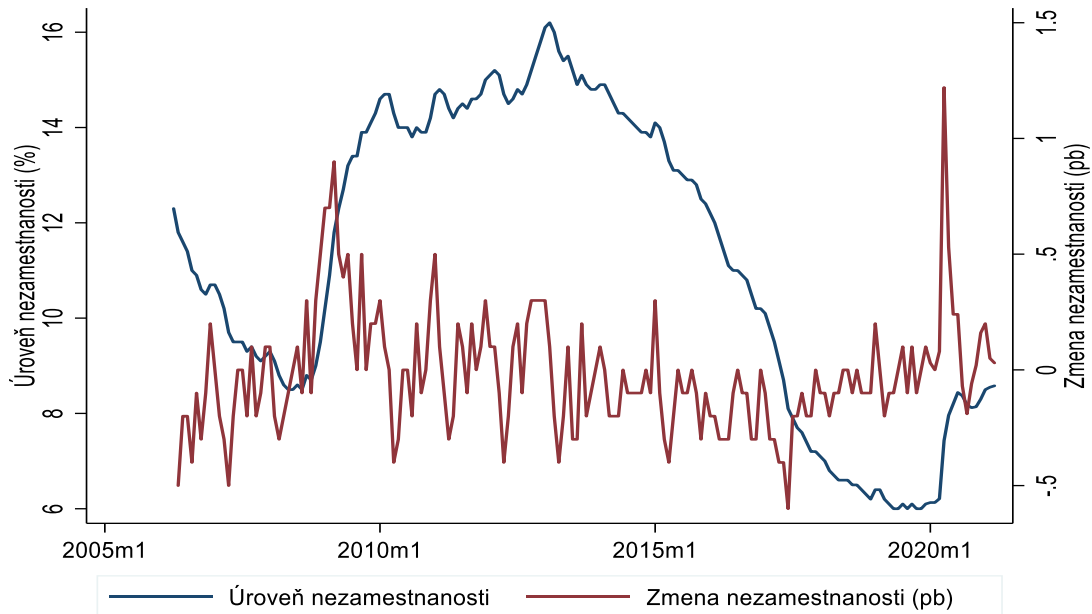


Zdroj údajov: Ústredie práce, sociálnych vecí a rodiny.
Vlastné spracovanie.

Pred prispôbením procesu sezónneho očisťovania na základe odporúčanej metodiky Eurostatu, je ideálne využiť prednastavené špecifikácie v softvéri *JDemetra+*. Pri TRAMO-SEATS je to RSAfull pre časový rad s predpokladanými výraznými kalendárnymi efektmi, zatiaľ čo RSA3 je vhodný pre údaje, ktoré nepodliehajú výrazným kalendárnym efektom. RSA3 je rovnako vhodná základná špecifikácia u X-13ARIMA-SEATS. Naopak, u časových radov, ktoré podliehajú kalendárnym efektom sa v prípade metodického prístupu X-

13ARIMA odporúčajú špecifikácie RSA4c a RSAc. Pri sezónnom očisťovaní miery nezamestnanosti vychádzame z predpokladu, že zvolený časový rad netrpí výraznými kalendárnymi efektami. Možno to zdôvodniť tým, že miera nezamestnanosti nie je ovplyvnená v závislosti od dĺžky mesiaca, priestupného roku, alebo ďalších nepravidelných kalendárných efektov. Z veľkej časti sledujeme skôr sezónne vplyvy, ako príchod stredoškôľakov a vysokoškôľakov na trh práce, dopady prázdninového obdobia prípadne obdobia Vianoc. Všetky uvedené príklady sú v tomto prípade považované za sezónny komponent, preto v našej špecifikácii nebudeme predpokladať existenciu kalendárných efektov.

Obrázok 2 Vývoj miery nezamestnanosti vypočítanej z celkového počtu UoZ od januára 2004 po marec 2021



Zdroj údajov: Ústredie práce, sociálnych vecí a rodiny.
Vlastné spracovanie.

Znázornený časový rad predstavuje celkovo 207 pozorovaní. Pri kontrole prítomnosti outlierov v časovom rade mesačnej nezamestnanosti sme detegovali výskyt jednej extrémnej hodnoty, a to apríl 2020. Prudký nárast v miere nezamestnanosti v tomto mesiaci vieme pripísať razantným proti-epidemickým opatreniam v dôsledku potvrdenia výskytu vírusu Covid-19 na území Slovenska. Disponujeme pomerne dlhým časovým radom s viditeľnými sezónnymi výkyvmi, ktoré sa snažíme odstrániť. Na odhad komponentov a sezónne očistenie časového radu využijeme v oboch prístupoch rôznu dĺžku časového radu, z ktorého sa budú odhadovať nepozorované komponenty. Prostredníctvom tejto úpravy minimalizujeme vplyv úvodných rokov časového radu, ktoré vykazujú vyššiu mieru sezónnosti, čo by mohlo spôsobiť suboptimálne výsledky, a tým aj znehodnotiť kvalitu sezónne očistených údajov najmä u posledných pozorovaní.

3.2 Sezónne očistenie prostredníctvom X-13ARIMA-SEATS

V tejto časti aplikujeme prvú z odporúčaných metód. Cieľom je zdokumentovať skutočný stav miery nezamestnanosti očistený o sezónne vplyvy s prihliadnutím na vypuknutie pandémie Covid-19 a získať predstavu o dynamike jej vývoja v posledných mesiacoch.

Navrhovaný postup pri sezónnom očisťovaní prostredníctvom X-13ARIMA-SEATS je nasledovný:

i. V prvom kroku využijeme vstavané nástroje softvéru na detekciu prítomnosti sezónnosti (Seasonality Tests). Skupina štatistických testov potvrdzuje prítomnosť sezónnosti v údajoch.

ii. Následne sa zameriame na **výber modelu a jeho špecifikáciu**. Ako sme zdôvodnili v predchádzajúcej časti, predpokladáme, že časový rad netrpí výraznými kalendárnymi efektmi,⁸ preto zvolíme RSA3 prednastavenú špecifikáciu, ktorú následne upravíme na základe odporúčaní.

iii. Dôležitou súčasťou je **výber obdobia**, z ktorého sa budú odhadovať parametre RegARIMA modelu, tzv. *model span*. Na odhad využijeme posledných 108 pozorovaní (9 ročný interval), na základe ktorého bude sezónne očistený celý časový rad (series span⁹). Rozšírením modelu o skoršie pozorovania, môžeme zapríčiniť skreslenie vo výsledkoch sezónneho očistenia. Možno to zdôvodniť tým, že parametre modelu budú odhadnuté aj z obdobia, ktoré nie je štrukturálne porovnateľné so situáciou dnes a zároveň časový rad trpel odlišnou sezónnosťou. Na základe uvedeného sme sa rozhodili skrátiť obdobie, z ktorého odhadujeme parametre.

iv. Špecifikácia RSA3 vykonáva test na **overenie potreby transformácie** časového radu na logaritmus. V záložke *Transformation* nevykonávame žiadne úpravy.

v. V ďalšej záložke pod názvom *Regression*, označíme druh a časový **výskyt extrémnych hodnôt**, ktoré sme identifikovali ešte v počiatočných krokoch sezónneho očisťovania. Softvér našiel jednu extrémnu hodnotu, a to apríl 2020 (pozri tiež obrázok v Boxe 3). Ide o zlom (tzv. level shift),¹⁰ čo taktiež špecifikujeme manuálne. Tento postup umožní lepšie modelovať časový rad, čím ovplyvníme kvalitu odhadu jeho komponentov. Zároveň však ponecháme naďalej aktivovanú možnosť detegovať aj ďalšie extrémne hodnoty, ktoré sme v prvom kroku nedokázali zachytiť. Do modelu sú pridávané nové koeficienty dovtedy, dokým nie sú odhalené žiadne ďalšie extrémne hodnoty.

vi. Špecifikácia RSA3 v základnom nastavení využíva automatický výber najvhodnejšieho ARIMA modelu, ktorá v niektorých prípadoch nemusí poskytovať optimálne výsledky. V našom prípade sme preto namiesto automatického výberu zvolili model (0,2,1)(0,1,1), ktorý kvalitnejšie odráža vývoj časového radu a poskytuje robustnejšie výsledky. Zvolený model sa použije aj na prednastavenú predpoveď vývoja časového radu na obdobie nasledujúcich 12 mesiacov.¹¹

vii. Dostávame sa k ďalšej záložke, a to **výberu sezónnych filtrov**. RSA3 využíva automatickú detekciu najvhodnejšieho sezónneho filtra. V našom prípade sme však zvolili dlhší sezónny filter 3x9,¹² ktorého diagnostická časť poskytuje vyššiu štatistickú významnosť. Nepredpokladáme špecifický typ rozkladu, preto zvolíme automatickú voľbu medzi rôznymi formami rozkladu časového radu.

viii. Metodika Eurostatu **neodporúča využiť benchmarking** (pozri podkapitolu 2.1.2 Vykonanie sezónneho očistenia), preto ani v našom prípade nebudeme klásť podmienku rovnosti sezónne očistených a sezónne neočistených údajov.

ix. Pri **revízii postupu sezónneho očisťovania** zvolíme pravidelnú aktualizáciu. Odôvodňujeme to tým, že časový rad trpí stabilnou sezónnosťou, a teda pridanie dodatočných pozorovaní neovplyvní kvalitu procesu sezónneho očisťovania. Postačuje, aby k revízii modelu dochádzalo v ročnom intervale.

⁸ Tieto závery potvrdzujú aj výsledky získané RSAfull (vykonáva test na prítomnosť kalendárnych efektov).

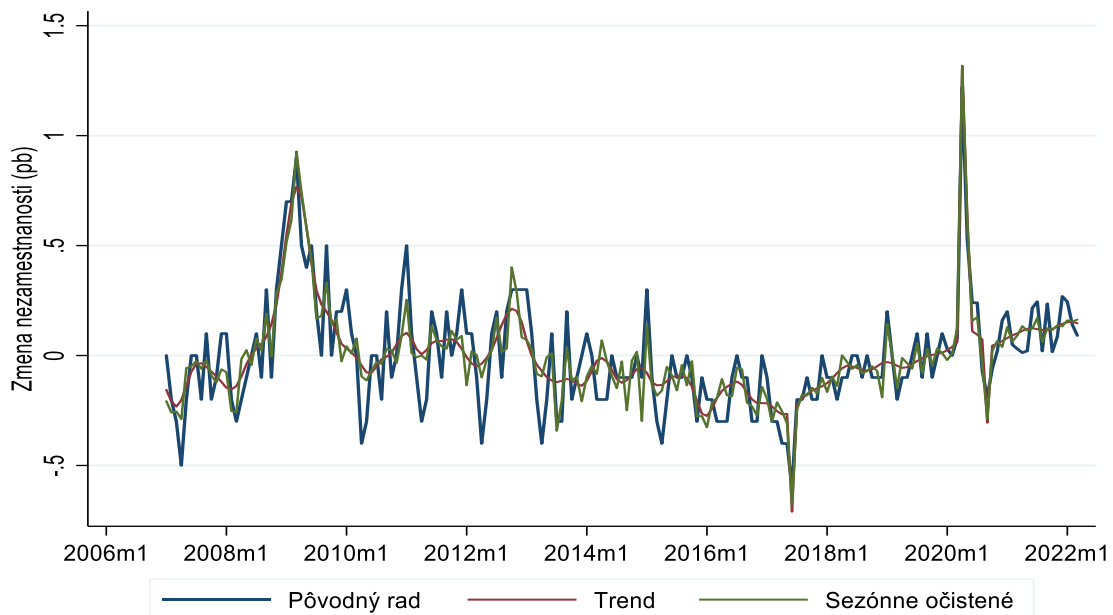
⁹ Dĺžka časového radu, ktorý sezónne očistíme. Požadujeme odstrániť sezónny komponent z celého dostupného časového radu.

¹⁰ Permanentná zmena v úrovni.

¹¹ Horizont prognózy sa dá taktiež upraviť.

¹² Zatiaľ čo krátky filter 3x1 dokáže zachytiť zmenu dynamiky sezónnosti pomerne rýchlo, nevýhodou je pomerne veľká revízia po zahrnutí dodatočných údajov. Čím stabilnejšiu sezónnosť predpokladáme, tým vhodnejšie je využiť dlhší sezónny filter, tak ako aj v našom prípade.

Obrázok 3 Sezónne očistený časový rad prostredníctvom X-13ARIMA-SEATS



Zdroj údajov: Ústredie práce, sociálnych vecí a rodiny.
Vlastné spracovanie.

Obrázok 3 kombinuje tri prvky, ktoré sú znázornené ako medzimesačná zmena: pôvodný časový rad, odhad trendu a zároveň poskytuje pohľad na očistený časový rad prostredníctvom X-13ARIMA-SEATS. Modelovanie časového radu prebehlo prostredníctvom RegARIMA modelu $(0,2,1)(0,1,1)$. Dekompozícia bola vykonaná lineárnym filtrom X11 (3x9). Získaný očistený časový rad poskytuje štatisticky významný odhad komponentov, pretože všetky diagnostické nástroje signalizujú štatistickú významnosť. Nami zvolený proces sezónneho očistenia poskytuje adekvátny odhad sezónne očistených údajov mesačnej miery nezamestnanosti vypočítanej z celkového počtu UoZ. Avšak netvrdíme, že zvolený postup je jediným možným. Aj v prípade využitia prednastavenej základnej špecifikácie RSA3¹³ je celková kvalita procesu sezónneho očistenia dobrá, ale úpravou sme dosiahli kvalitnejší pohľad na vývoj nezamestnanosti.

3.3 Sezónne očistenie prostredníctvom TRAMO-SEATS

Postup pri metóde sezónneho očistenia prostredníctvom TRAMO-SEATS sa vo viacerých bodoch ponáša na predchádzajúci zoznam krokov navrhnutý pre X-13ARIMA-SEATS.

Navrhovaný postup pri sezónnom očisťovaní prostredníctvom TRAMO-SEATS je nasledovný:

- i. V prvom kroku opäť využijeme vstavené nástroje softvéru *JDemetra+* na detekciu **prítomnosti sezónnosti** (Seasonality Tests). Skupina štatistických testov potvrdzuje prítomnosť sezónnosti v údajoch.
- ii. Následne venujeme pozornosť **výberu modelu** a jeho špecifikácii. Rovnako ako sme zdôvodnili v predchádzajúcej časti, predpokladáme, že časový rad netrpí výraznými kalendárnymi efektami,¹⁴ preto zvolíme RSA3 prednastavenú špecifikáciu, ktorej parametre upravíme na základe odporúčaní.
- iii. Dôležitou súčasťou je **výber obdobia**, z ktorého sa bude odhadovať RegARIMA model, tzv. *model span*. Na odhad využijeme posledných 120 pozorovaní (10 ročný interval), na základe ktorého bude sezónne očistení

¹³ Na odhad parametrov modelu sme využili posledných 108 pozorovaní. Zároveň sme označili jednu extrémnu hodnotu (apríl 2020).

¹⁴ Tieto závery potvrdzujú aj výsledky získané RSAfull (vykonáva test na prítomnosť kalendárných efektov).

časový rad (series span¹⁵). Dôvody na úpravu obdobia, z ktorého dochádza k odhadu parametrov modelu sme poskytli v vyššie (pozri podkapitolu 3.2).

iv. Špecifikácia RSA3 vykonáva test na **overenie potreby transformácie** časového radu na logaritmus. V záložke *Transformation* nevykonáme žiadne zmeny.

v. V ďalšej záložke pod názvom *Regression*, identifikujeme druh a časový výskyt extrémnych hodnôt, ktoré sme detegovali ešte v počiatočných krokoch sezónneho očisťovania. Softvér identifikoval jednu extrémnu hodnotu, a to apríl 2020. Jedná sa o Level Shift¹⁶, čo upresníme v tejto záložke.

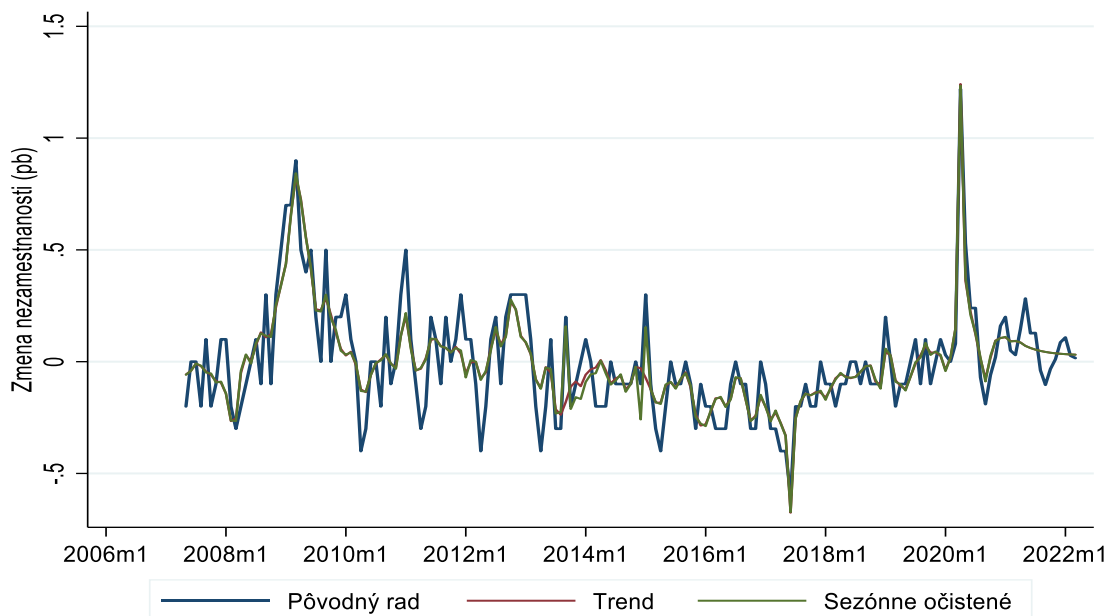
vi. RSA3 v základnej špecifikácii využíva automatickú detekciu ARIMA modelu. Na rozdiel od postupu v prípade X-13ARIMA-SEATS, v tomto prípade nezvolíme špecifický model, ale ponecháme automatický výber. Softvérom vybraný model sa použije aj na prognózu vývoja časového radu na obdobie nasledujúcich 12 mesiacov.¹⁷

vii. Nie je potrebné upravovať ani dĺžku prognózovaného obdobia, ani metódu rozkladu. V záložke venovanej dekompozícii časového radu (SEATS) je vhodné ponechať prednastavenú špecifikáciu.

viii. Metodika Eurostatu neodporúča využiť benchmarking (pozri podkapitolu 2.1.2 Vykonanie sezónneho očistenia), preto ani v našom prípade nebudeme klásť podmienku rovnosti sezónne očistených a sezónne neočistených údajov.

ix. Pri revízii postupu sezónneho očisťovania zvolíme **pravidelnú aktualizáciu**. Odôvodňujeme to tým, že časový rad v posledných rokoch trpí pomerne stabilnou sezónnosťou,¹⁸ a teda pridanie dodatočných pozorovaní neovplyvní kvalitu procesu sezónneho očisťovania. Postačuje, aby k revízii modelu dochádzalo v ročnom intervale.

Obrázok 4 Sezónne očistený časový rad prostredníctvom TRAMO-SEATS



Zdroj údajov: Ústredie práce, sociálnych vecí a rodiny.
Vlastné spracovanie.

¹⁵ Dĺžka časového radu, ktorý sezónne očistíme. Požadujeme odstrániť sezónny komponent z posledných 168 pozorovaní (14 rokov) dostupného časového radu.

¹⁶ Permanentná zmena v úrovni.

¹⁷ Horizont prognózy sa dá taktiež upraviť.

¹⁸ Výsledky potvrdzuje aj časť v diagnostickej časti, ktorá skúma dopady rozšírenia časového radu o dodatočné pozorovania

Obrázok 4 znázorňuje pôvodný časový rad, trend a sezónne očistené údaje ako medzimesačnú zmenu. Celková kvalita procesu sezónneho očistenia zvolenou metódou je dobrá, čo dokumentujú aj výsledky v diagnostickej časti. Pri automatickom výbere bola za najvhodnejšiu vybraná špecifikácia $(1,1,0)(1,0,0)$. Prognóza časového radu obsahuje vývoj v nasledujúcich 12 mesiacoch. V porovnaní s prognózou prostredníctvom X-13ARIMA-SEATS sledujeme mierne optimistickjší scenár, avšak aj na základe tohto odhadu budeme sledovať postupný nárast v miere nezamestnanosti až k hranici presahujúcej 9 %.

3.4 Porovnanie oboch prístupov

Pri vizuálnej kontrole priebehu kriviek sezónne očistených dát na Obrázku 1,2, respektíve Obrázku 3, sledujeme **podobný vývoj pri oboch metódach**. Zatiaľ čo TRAMO-SEATS počas prvých mesiacov roku 2021 odhaduje vyššiu úroveň sezónne očistenej miery nezamestnanosti, v letných mesiacoch minulého roku to bola X-13ARIMA-SEATS, ktorej výsledky vykazovali vyššiu mieru sezónne očistenej miery nezamestnanosti. Výraznejšiu odchýlku zaznamenávame v horizonte nasledujúcich 12 mesiacoch, ktorá je zapríčinená rôznymi koeficientami ARIMA modelu. Kým odhad sezónne očistenej nezamestnanosti prostredníctvom TRAMO-SEATS môžeme z oboch metód považovať za konzervatívny, miera nezamestnanosti by mala podľa obidvoch modelov rásť, aj keď tempo rastu by sa malo spomaľovať. Ako sme však zdôraznili v podkapitole 2.1.1 *Počiatočné kroky v procese sezónneho očisťovania* prednastavené prognózy v *JDemetra+* sa opierajú výlučne o informácie o minulom vývoji časového radu a neberú do úvahy exogénne vplyvy, ako napr. rýchle zlepšenie/zhoršenie epidemickej situácie.

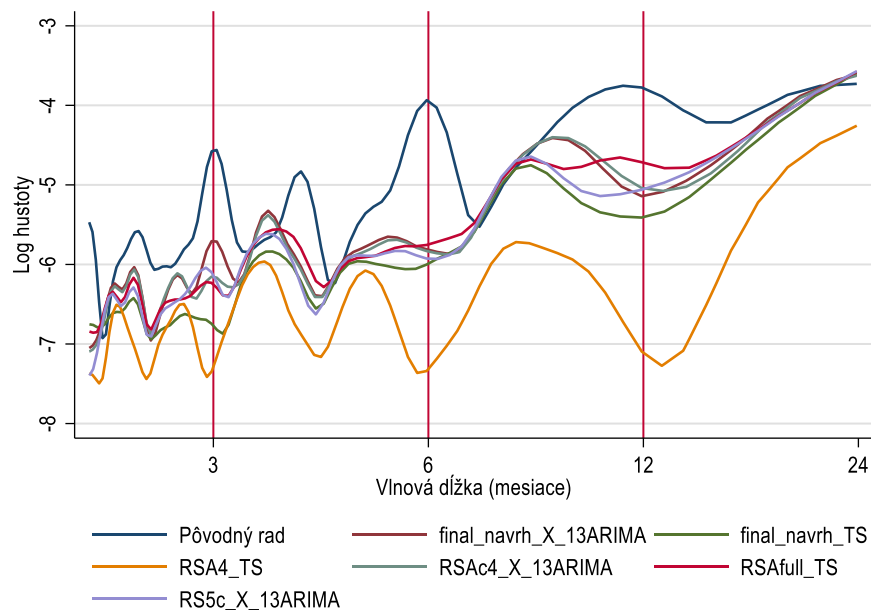
Získané sezónne upravené údaje dokumentujú dopad vypuknutia pandémie a s ním súvisiaci negatívny vplyv na slovenský trh práce. Aj napriek ekonomickému oživeniu počas letných mesiacov, obe metódy zachytili medzimesačný **pokles** nezamestnanosti po sezónnom očistení len za **september 2020**. X-13ARIMA-SEATS zdokumentovala medzimesačný pokles aj za august, avšak za ostatné mesiace sezónne očistená nezamestnanosť na území Slovenska rástla. Postupný nárast v miere nezamestnanosti sledujeme aj v prvých mesiacoch roku 2021. S postupným otváraním ekonomiky však očakávame stabilizáciu situácie na trhu práce. Očakávame, že pozitívny impulz pre ekonomiku v podobe ďalšieho otvárania budeme sledovať aj v klesajúcom počte nezamestnaných.

Celková kvalita procesu sezónneho očistenia využitím TRAMO-SEAT, ako aj X-13ARIMA-SEATS je dobrá, čo potvrdzujú aj dostupné diagnostické nástroje. Získané výsledky využitím oboch metód poskytujú jasnejší pohľad na vývoj miery nezamestnanosti vypočítanej z celkového počtu UoZ na Slovensku. Porovnanie diagnostických nástrojov u oboch metód naznačuje, že **mierne adekvátnejšie je sezónne očistenie prostredníctvom X-13ARIMA-SEATS**, ktorého štatistické vlastnosti sú priaznivejšie ako u TRAMO-SEATS. To však v žiadnom prípade neznamená, že TRAMO-SEATS poskytuje neadekvátny odhad sezónneho komponentu. Výber metódy vo veľkej miere závisí od typu dát, a preto oba prístupy považujeme za vhodný nástroj na detekciu a odstránenie sezónneho komponentu, čo potvrdzujú aj získané výsledky.

3.5 Kotrola prítomnosti reziduálnej sezónnosti

Pre jednoduchú grafickú kontrolu adekvátnosti sezónneho očistenia je možné použiť tzv. spektrálnu analýzu, kde sa časový rad podrobí Fourierovej transformácii, čím dekomponujeme dáta na sínusové a kosínusové krivky o rôznych frekvenciách. Kľúčom je zistiť, či krivky na sezónnej frekvencii nie sú nadmerne zastúpené v očistenom časovom rade alebo či na týchto frekvenciách nebadáť zníženú spektrálnu hustotu (Granger, 1978).

Obrázok 4 Spektrálna hustota údajov o nezamestnanosti pred a po sezónnom očistení



Zdroj údajov: Ústredie práce, sociálnych vecí a rodiny.
Vlastné spracovanie.

Obrázok 4 ukazuje jasnú nadmieru zastúpenia 6-mesačných a 12-mesačných vlnových dĺžok v pôvodných dátach o nezamestnanosti. O niečo slabší signál pozorujeme aj na kvartálnej vlnovej dĺžke, ktorý by tiež mohol indikovať krátkodobú sezónnosť, avšak 3 a 6 mesačné frekvencie môžu byť taktiež harmonické zložky (aliquotné frekvencie) 12-mesačnej sezónnosti. **V ideálnom prípade by sezónne očistené dáta mali potlačiť nadbytočné vrcholy na sezónnych frekvenciách bez vytvorenia „preliačín“ na týchto frekvenciách.** Ako vidno, 3, 6 aj 12-mesačné vrcholy sú skutočne potlačené, aj keď badať aj potlačenie vyšších frekvencií (všetky očistené rady ukazujú menšie zastúpenie kriviek s frekvenciami kratšími ako 3 mesiace), čo naznačuje čiastočné vyhladzovanie aj ne-sezónnych komponentov. Uvažované modely vykazujú podobnú spektrálnu dekompozíciu s výnimkou RSA4_TS, ktorá jasne potláča vysoké frekvencie a zároveň príliš redukuje aj sezónne frekvencie nakoľko vytvára výrazné preliačiny na všetkých uvažovaných sezónnych frekvenciách. Obrázok 4 taktiež dokladá, že **nie je žiadny „ideálny“ prístup k sezónnemu očisteniu**, ktorý by dokonale **eliminoval sezónnu variáciu bez skreslenia ostatných frekvencií.** Možno ale povedať, že model RSAfull naozaj poskytuje dobrý prvý krok pri analýze sezónnosti: aj keď potláča viacero vysoko-frekvenčných komponentov ako nami navrhované postupy, na druhej strane má miernejšiu preliačinu na 12-mesačnej frekvencii, takže RSAfull o niečo lepšie zachováva ne-sezónnu 12-mesačnú dynamiku dát.

3.6 Dopady SA: Kumulatívny efekt epidémie COVID-19

Dopady sezónneho očistenia na analýzu časových radov je možné vidieť na jednoduchom modeli nezamestnanosti v SR po dobu 167 mesiacov od mája 2007 do marca 2021. Použitie kratšieho pozorovaného okienka by ohrozilo asymptotické vlastnosti, ktoré vyžadujú, aby dĺžka časového radu bola čo najvyššia. Tento časový rad modelujeme ako ARMA(2,2) proces, nakoľko táto špecifikácia viedla k stacionárnym rezíduám pri použití relatívne skromného modelu. Formálne:

$$u_t^{SA_m} = +\beta_1 \mathbb{I}[t = \text{apríl 2020}] + \beta_2 \mathbb{I}[t \geq \text{máj 2020}] + v_t,$$

kde perzistencia v miere nezamestnanosti je modelovaná:

$$v_t = \rho_1 v_{t-1} + \rho_2 v_{t-2} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-2} + \varepsilon_t.$$

Závisle premenná $u_t^{SA_m}$ je nezamestnanosť v SR v mesiaci t , upravená pomocou metódy sezónneho očistenia m (vrátane žiadneho sezónneho očistenia) a $\mathbb{I}[\cdot]$ je indikátorová funkcia, ktorá nadobúda hodnotu 1 v prípade, ak je podmienka v argumente splnená a hodnotu 0 vo všetkých ostatných prípadoch. Korelácie medzi rezíduami (v_t) modelujeme pomocou dvoch autoregresných parametrov (ρ_1 a ρ_2 , tzv. AR parametre) a dvoch parametrov pre kľzavé priemery (θ_1 a θ_2 , tzv. MA parametre). Inovácie ε_t sú biely šum. Tento ARMA(2,2) proces nám umožňuje zachytiť vývoj na trhu práce pomocou jednoduchého modelu, ktorý si nevyžaduje veľký počet parametrov. Koeficient β_1 potom zodpovedá nárastu nezamestnanosti nad rámec trendu a sezónnej variácie v nadväznosti na prvý výrazný výskyt vírusu COVID-19 a s ním spojenými proti-epidemickými opatreniami. Koeficient β_2 potom meria následnú zmenu v nezamestnanosti v dôsledku prebiehajúcej epidémie. Kumulatívny efekt sme následne počítali ako $\beta_1 + \beta_2$. Pre výpočet smerových odchýlok sme používali Bollerslev-Wooldridgeovu variančnú maticu, ktorá nevyžaduje normalitu ani homoskedasticitu ε_t . Experimentálne sme uvažovali aj so špecifikáciami s jemnejšie rozčlenenou periódou od mája 2020, ale zvýšenie počtu parametrov viedlo k veľmi nepresným odhadom.

Tabuľka 2 Kumulatívny efekt pandémie na mieru nezamestnanosti v SR (pb.)

Časový rad	Odhad $\beta_1 + \beta_2$	95% Interval spoľahlivosti	
Neupravené dáta	2,66	2,29	3,03
final_navrh_X_13ARIMA	2,95	2,77	3,14
final_navrh_TS	2,39	2,18	2,60
RSA4_TS	2,11	2,04	2,19
RSAc4_X_13ARIMA	2,72	2,53	2,91
RSAfull_TS	2,73	2,47	3,00
RS5c_X_13ARIMA	2,59	2,42	2,76

Tabuľka 2 ukazuje, že hoci všetky uvažované časové rady (pôvodné dáta aj sezónne očistené) dávajú podobné výsledky indikujúce nárast miery nezamestnanosti o vyše 2 pb, je tu jednoznačne badať rozdiely. Pri použití nami navrhovaného postupu X_13ARIMA dostaneme nárast o 2,95 pb avšak pri použití RSA4_TS je odhadnutý nárast takmer o tretinu miernejší (2,11 pb).

4 Záver

Analýza ukazuje, že v závislosti na použítom postupe sezónneho očistenia sa môže odhadnutý kumulatívny efekt epidémie COVID-19 oproti predchádzajúcemu trendu nezamestnanosti pohybovať od približne +2,11 pb ak použijeme RSA4_TS až po +2,95 pb pri použití nami navrhutej špecifikácie X_13ARIMA. Tento odhad je podobný ako keď sa výpočet realizuje na pôvodných dátach (+2,66 pb), čo naznačuje, že X_13ARIMA málo porušuje ne-sezónne vlastnosti dát.

Táto analýza odráža praktické potreby sezónneho očistenia, nakoľko tvorba dlhodobých politík si zväčša vyžaduje abstrakciu od efektu sviatkov, ročných období a ďalších sezónnych vplyvov. Typický rozhodovací problém si v tejto situácii vyžaduje pohľad na súčasný (prípadne historický) trend ekonomickej situácie, aby sa zabránilo extrapoláciám na základe sezónnych fluktuácií. V tejto súvislosti analýza podáva jednoznačné odporúčanie pre procesy tvorby politík, ktoré sa opierajú o sezónne očistené dáta. V rámci možnosti je dôležité verifikovať robustnosť výsledku sezónneho očistenia konkrétneho časového radu voči zmenám metodiky sezónneho očistenia. Analýza ukázala, že aj pri použití štandardných metód môže dôjsť k závažným porušeniam nesezónnych charakteristík dátových radov (viď výsledky RSA4_TS). Aj v prípadoch, keď sezónne očistenie možno považovať za „úspešné“, voľba „najlepšej“ varianty očistenia si typicky vyžaduje istú mieru kompromisu: modely, ktoré vedú k najmenej korelovaným rezíduám či iným klasickým kritériám vhodnosti môžu napríklad viesť k pod-reprezentácii sezónnych vlnových dĺžok. Naopak, modely, ktoré zachovávajú nesezónne komponenty dát môžu dosahovať horšie výsledky v iných ohľadoch. Preto je nutné pozorne zvažovať, či sa

odporúčania nezakladajú na štatistických reliktoch zanechaných konkrétnou voľbou eliminácie sezónnych vplyvov.

Prostredníctvom softvéru JDemetra+ na základe odporúčaní jednotnej metodiky Eurostatu sme odstránili sezónny komponent z dát o mesačnej miere nezamestnanosti. V úvodných častiach sme stručne opísali počítačové pokusy o sezónne očisťovanie, zdôraznili sme dôvody prečo je vhodné vykonávať úpravu časového radu o sezónny komponent a zároveň sme uviedli niektoré problémy, ktoré sú naďalej spojené s týmto procesom. V nasledujúcich častiach sme **definovali komponenty časového radu a zamerali svoju pozornosť na implementáciu metodiky Eurostatu v praxi.** Postupne sme v skratke opísali odporúčané kroky sezónneho očisťovania a v záverečnej časti sme **navrhli postup sezónneho očisťovania** prostredníctvom oboch odporúčaných metód, ktoré sme podrobili aj spektrálnej analýze.

Analýza poskytla základný prehľad zvolených parametrov sezónneho očisťovania v dvoch diskutovaných metódach. Získané výsledky považujeme za vhodný zdroj informácií, ktoré dokážu uľahčiť vykonávanie sezónneho očisťovania v budúcnosti. Cieľom dokumentu bola minimalizácia informačnej asymetrie medzi vykonávateľom sezónneho očisťovania a čitateľom, ktorý vďaka prehľadnej organizácii analýzy dokáže porozumieť prijatým rozhodnutiam v procese sezónneho očistenia. Dokument tak poslúži ako referenčný zdroj pre budúce diskusie týkajúce sa sezónne očistených dát v gescii ISP. Naďalej zdôrazňujeme, že výber metódy sa odvíja od štruktúry dát a v prípade mesačnej miery nezamestnanosti vypočítanej z celkového počtu UoZ mierne adekvátnejšie výsledky zabezpečuje X-13ARIMA-SEATS. V tejto súvislosti naša práca zároveň odporúča využívať viaceré techniky sezónneho očisťovania s cieľom zamedziť vyvodzovaniu hlavných záverov z možných štatistických reliktov.

Získané výsledky poskytujú jasnejší pohľad na dynamiku vývoja miery nezamestnanosti, ale súčasne dokumentujú aj dopad pandémie. Obe diskutované metódy (X-13ARIMA-SEATS a TRAMO-SEATS) využívajú inú formu dekompozície časového radu, preto sa získané výsledky odlišujú. Rozdiely sú badateľné najmä v horizonte 12 mesačnej prognózy, kde obe využívajú iné koeficienty ARIMA modelu, čo ovplyvňuje aj proces dekompozície. Na základe diagnostických vlastností oboch metód **v našom prípade preferujeme X-13ARIMA-SEATS a využitie sezónnych filtrov.** Avšak tento záver neimplikuje akúkoľvek všeobecnú nadradenosť tejto metódy na úkor TRAMO-SEATS. Sezónne očisťovanie je stále v procese vývoja, a preto môžeme očakávať ďalšie sofistikovanejšie prístupy k riešeniu tohto komplexného problému.

Bibliografia

- Bell, W. R., & Hillmer, S. C. (1984). Issues Involved with the Seasonal Adjustment of Economic Time Series. *Journal of Business & Economic Statistics*, 2(4), 291-320.
- Caporello, G., Maravall, A. (2004a), 'Program TSW. Revised Reference Manual', Banco de España.
- Dagum, E. B. (1978). A Comparison and Assessment of Seasonal Adjustment Methods for Employment and Unemployment Statistics. *National Commission on Employment and Unemployment Statistics*, Background Paper No. 5, Washington, D.C.: Government Printing Office, 1-94.
- Eurostat, (2015), 'ESS Guidelines on Seasonal Adjustment', Eurostat Methodological Working Papers.
- Granger, C. W. J. (1978). Seasonality: Causation, Interpretation, and Implications. In A. Zellner (Ed.), *Seasonal Analysis of Economic Time Series*: NBER.
- Grudkowska, S. (2017a). 'JDemetra+ Reference Manual', Eurostat.
- Grudkowska, S. (2017b). 'JDemetra+ User Guide', Eurostat.
- Huček, J., Doliak, M. (2014). Sezónne očisťovanie časových radov (jednoduchý sprievodca postupmi, možnosťami a výsledkami s aplikáciou na očisťovanie HICP). *Biatec*, 7-18.
- Nikolova, A., Elliott, D. (2019). Dealing with High Frequency Time Series: Seasonal Adjustment of Road Traffic Data. Dostupné na: https://ec.europa.eu/eurostat/cros/system/files/s06p03_-_dealing_with_high_frequency_time_series.pdf
- Persons, W. M. (1919). Indices of Business Conditions. *Review of Economics and Statistics*, 1, 5-107.
- UNECE. (2020). Practical Guide to Seasonal Adjustment with JDemetra+, UNECE.