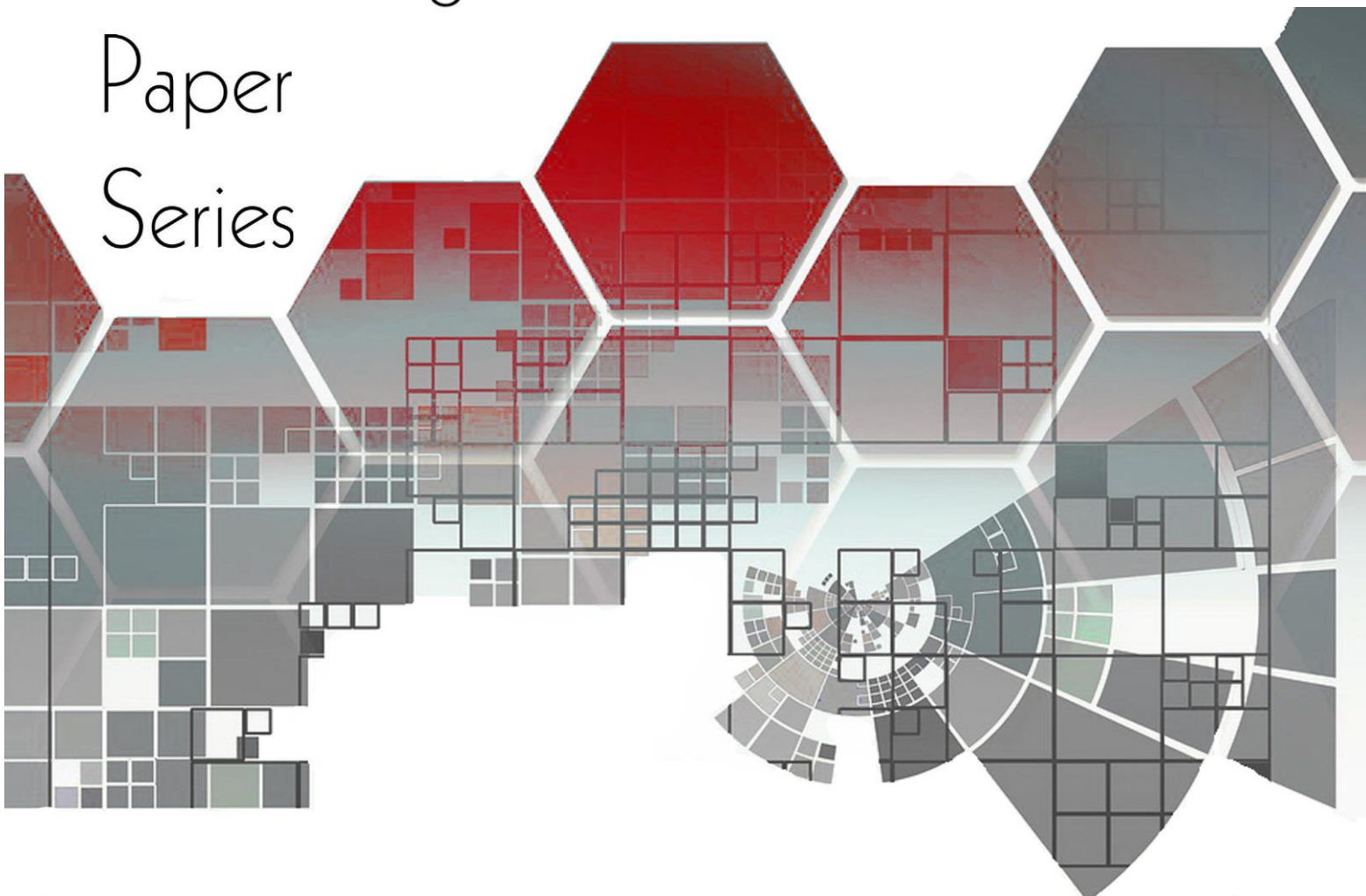


MF Working

Paper

Series



Karolina Konopczak

Kwantyfikacja zmian luki VAT: podejście ekonometryczne

Kwantyfikacja zmian luki VAT: podejście ekonometryczne

Karolina Konopczak¹

Abstrakt

W niniejszym artykule zaproponowano alternatywną względem stosowanych w literaturze metodykę szacowania zmian luki VAT, pozwalającą na ich bieżący (kwartalny) monitoring. Łączy ona dwie tradycje modelowania dochodów podatkowych – metodę rozbieżności oraz metodę szacowania elastyczności dochodów podatkowych względem bazy podatkowej. Kształtowanie się luki VAT utożsamiane jest ze zmianami reaktywności rzeczywistych dochodów podatkowych względem teoretycznych (czyli – ekwiwalentnie – rzeczywistych dochodów skorygowanych o zmiany stawek podatkowych względem bazy podatkowej). Metoda została wykorzystana do oszacowania zmian luki VAT w Polsce w latach 2016-18. Wyniki wskazują na spadek luki w tym okresie o ok. 21 mld zł. Tym samym spadek luki odpowiadał za niemal połowę wzrostu dochodów z VAT w tym okresie. W ujęciu procentowym luka zmniejszyła się z poziomu 20,8% w 2016 r. (Komisja Europejska, 2018) do poziomu poniżej 10% na koniec 2018 r.

Kody JEL: C22, H24, H26

Słowa kluczowe: luka VAT, elastyczność dochodów podatkowych, analiza kointegracji, model ARDL

¹ Ministerstwo Finansów oraz Szkoła Główna Handlowa w Warszawie. Korespondencja: karolina.konopczak@mf.gov.pl.

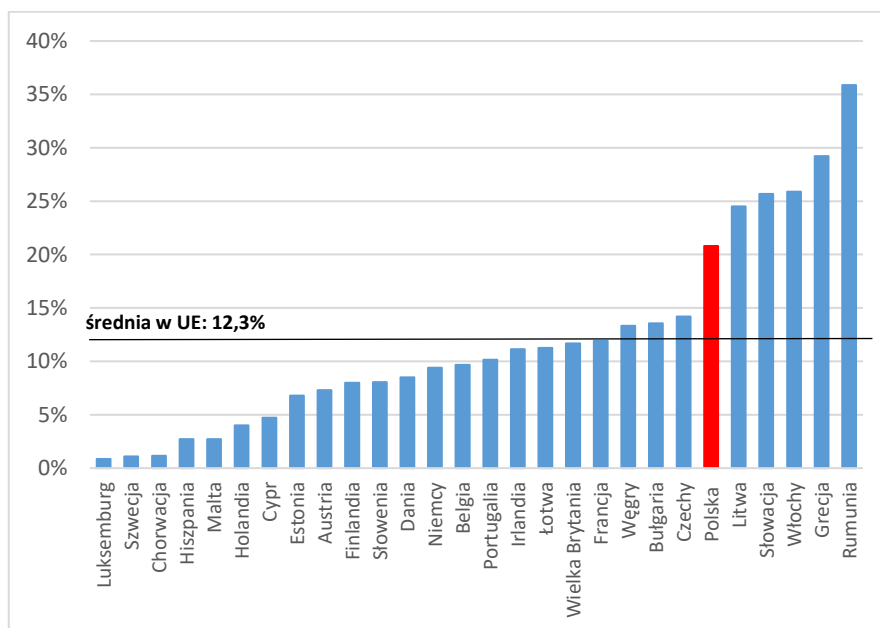
Poglądy wyrażone w niniejszym artykule powinny być traktowane jako poglądy autora(ów), które nie muszą koniecznie wyrażać poglądów Ministerstwa Finansów.

Podziękowania za dyskusję oraz komentarze dla uczestników seminarium nt. luki VAT w Ministerstwie Finansów, w szczególności dla Doroty Bach, Joanny Bęzy-Bojanowskiej, Sławomira Dudka, Tomasza Mazura, Dawida Pachuckiego oraz Tomasza Szałwińskiego.

Wprowadzenie

Luka VAT stanowi różnicę między rzeczywistymi dochodami z podatku VAT a dochodami teoretycznymi wynikającymi z obowiązujących przepisów prawa. Tym samym stanowi miarę niewywiązywania się z obowiązków podatkowych (ang. *compliance gap*), będącego efektem zarówno nieumyślnych błędów w rozliczeniach, problemów z płynnością, jak i umyślnych działań ukierunkowanych na zmniejszanie ciężaru podatkowego, tj. unikania opodatkowania (ang. *tax avoidance*), uchylania się od opodatkowania (ang. *tax evasion*) lub oszustw podatkowych (ang. *tax fraud*) (Komisja Europejska, 2016). Zgodnie z szacunkami Komisji Europejskiej (2018) w 2016 r. Polska należała do krajów Unii Europejskiej o najwyższej względnej luce VAT (patrz wykres 1). Udział luki w teoretycznych dochodach z podatku VAT kształtował się w Polsce na poziomie ok. 20,8% przy średniej unijnej na poziomie ok. 12,3%. Wyższe szacunki luki VAT otrzymano jedynie dla niektórych krajów Europy Południowej (Włochy, Grecja) i Środkowo-Wschodniej (Rumunia, Słowacja i Litwa).

Wykres 1. Szacunki luki VAT (w procentach teoretycznych dochodów) w 2016 r.



Źródło: Komisja Europejska (2018).

Komisja Europejska opiera swoje szacunki luki VAT na metodzie rozbieżności, w ramach której luka liczona jest jako różnica między teoretycznym podatkiem należnym a rzeczywistymi dochodami z tytułu podatku VAT (patrz Komisja Europejska, 2013). Niezbędne do obliczenia teoretycznych dochodów podatkowych dane dostępne są jednak ze znacznym opóźnieniem. Głównym ich źródłem są bowiem tablice podaży i wykorzystania publikowane z dwuletnim opóźnieniem. Tym samym, metoda ta nie może służyć do bieżącego śledzenia zmian luki VAT, co stanowi istotne jej ograniczenie przy prowadzeniu polityki gospodarczej.

W związku z tym w niniejszym artykule zaprezentowano alternatywną metodykę, pozwalającą na kwartalny monitoring luki VAT. Z racji tego, że bazuje na niepełnym zbiorze danych, ma ona – w przeciwieństwie do algebraicznej metody stosowanej przez Komisję Europejską – charakter ekonometryczny i łączy dwie tradycje modelowania dochodów podatkowych: metodę rozbieżności oraz metodę szacowania elastyczności dochodów względem bazy podatkowej. Kształtowanie się luki VAT utożsamiane jest ze zmianami elastyczności (i tym samym reaktywności) rzeczywistych dochodów względem indykatora dochodów teoretycznych, czyli – ekwiwalentnie – rzeczywistych dochodów skorygowanych o zmiany stawek podatkowych względem bazy podatkowej. Metoda ma charakter agnostyczny w tym sensie, że szacunki zmian luki VAT są wyłącznie pochodną trajektorii analizowanych szeregów czasowych. Tym samym metoda abstrahuje od potencjalnych przyczyn kształtowania się luki VAT, zarówno cyklicznych, jak i strukturalnych. Metoda została wykorzystana do oszacowania zmian luki VAT w Polsce w latach 2016-18.

Artykuł ma następującą strukturę. Rozdział 1 zawiera przegląd metod stosowanych w literaturze do szacowania luki podatkowej. Kolejne rozdziały zawierają opis metodyki opracowanej na potrzeby kwartalnego monitoringu luki VAT (rozdział 2) oraz wykorzystanych do jej implementacji metod ekonometrycznych (rozdział 3). W rozdziale 4 przedstawiono wyniki badania, w tym kwantyfikację zmian luki VAT w Polsce w latach 2016-18. Artykuł kończy podsumowanie.

1. Przegląd literatury

W literaturze wyróżnić można dwie podstawowe grupy metod stosowanych do szacowania luki podatkowej w oparciu o dane makroekonomiczne – metody szacowania rozbieżności między dochodami rzeczywistymi i teoretycznymi oraz metody wskaźnikowe.

Do określenia wielkości luki VAT Komisja Europejska stosuje metodę rozbieżności (ang. *discrepancy method*), w ramach której luka (ang. *VAT gap*) liczona jest jako różnica między teoretycznym podatkiem należnym (ang. *VAT Total Tax Liability*, VTTL) a rzeczywistymi dochodami z tytułu podatku VAT (patrz Komisja Europejska, 2013). W podejściu tym teoretyczne należności podatkowe szacowane są metodą odgórną² (ang. *top-down*) w oparciu o dane z systemu rachunków narodowych. Teoretyczny VAT od konsumpcji finalnej liczony jest przez pomnożenie dla poszczególnych kategorii towarów i usług (według klasyfikacji CPA/COICOP) wielkości ich spożycia prywatnego i publicznego przez efektywne stawki podatku VAT dla tych kategorii. W przypadku zużycia pośredniego odpowiednie agregaty dla poszczególnych sektorów (według klasyfikacji NACE) i kategorii towarów i usług (według klasyfikacji CPA) mnożone są przez efektywne stawki podatku VAT, a następnie powstałe iloczyny korygowane są o udział produkcji zwolnionej podmiotowo lub przedmiotowo z VAT w ogólnej produkcji sektorów. W przypadku nakładów brutto na środki trwałe analogiczne szacunki – jednak bez dezagregacji na poszczególne kategorie dóbr i usług – dokonywane są w podziale na sektory instytucjonalne (przedsiębiorstwa niefinansowe, instytucje finansowe, instytucje rządowe i samorządowe, gospodarstwa domowe, instytucje niekomercyjne działające na rzecz gospodarstw domowych). Formułę, zgodnie z którą Komisja Europejska szacuje VTTL na okres t można zatem zapisać jako:

² Istnieje jeszcze metoda oddolna (bottom-up), w ramach której uogólniane są wyniki kontroli podatkowych przeprowadzane dla reprezentatywnej grupy podatników.

$$VTTL_t = \sum_{i \in COICOP/CPA} C_{it} * rate_{it} + \sum_{j \in NACE} \left[\left(\sum_{i \in CPA} IC_{jit} * rate_{jit} \right) * exempt_ratio_{jt} \right] \quad (1)$$

$$+ \sum_{k \in SECTOR} GF_{kt} * rate_{kt} * exempt_ratio_{kt} + net_adjustments_t$$

gdzie:

- C_{it} – wartość netto spożycia finalnego (ang. *consumption*, C) i -ej kategorii klasyfikacji COICOP/CPA;
- IC_{jit} – wartość netto spożycia pośredniego (ang. *intermediate consumption*, IC) i -ej kategorii klasyfikacji CPA j -go sektora klasyfikacji NACE;
- GF_{kt} – wartość netto inwestycji (ang. *gross fixed capital formation*, GF_{kt}) k -go sektora instytucjonalnego ($k=1 \dots 5$);
- $rate_{ijt}$ – efektywna stawka podatku VAT i -ej kategorii dóbr i usług i j -go sektora klasyfikacji NACE;
- $exempt_ratio_{jt}$ – udział produkcji zwolnionej z VAT (ang. *exempt from VAT*) w ogólnej produkcji j -go sektora;
- $net_adjustments_t$ – dyskrecjonalne korekty.

Obliczona w ten sposób wartość teoretycznych należności ($VTTL$) porównywana jest z rzeczywistymi dochodami z podatku VAT (ang. *VAT revenue*). Różnica między tymi wielkościami stanowi oszacowanie luki VAT:

$$VAT_gap_t = VTTL_t - VAT_revenue_t \quad (2)$$

Luka przedstawiana jest najczęściej w ujęciu procentowym jako ubytek teoretycznych dochodów:

$$VAT_gap_t = \frac{VTTL_t - VAT_revenue_t}{VTTL_t} * 100\% \quad (3)$$

Metodyka stosowana przez Komisję Europejską – dzięki oparciu na zdezagregowanych danych – pozwala na uzyskanie możliwie dokładnego szacunku VTTL. Niezbędne do obliczeń dane dostępne są jednak ze znacznym opóźnieniem (głównym źródłem danych są tablice podaży i wykorzystania publikowane z opóźnieniem dwuletnim). Tym samym, metoda ta nie może służyć do bieżącego śledzenia zmian luki VAT. Ponadto, pewne wątpliwości może budzić algebraiczny charakter metody w obliczu szacunkowego charakteru pewnych parametrów, służących do obliczania VTTL (np. wielkości zużycia pośredniego i inwestycji związanych ze sprzedażą zwolnioną przedmiotowo lub podmiotowo z VAT).

Na bieżący monitoring luki podatkowej pozwalają metody wskaźnikowe, m.in. metoda analizy popytu na pieniądź³ (ang. *currency demand approach*, CDA), metoda analizy zużycia energii elektrycznej⁴ (ang. *electricity consumption method*, ECM) czy też model MIMIC⁵ (ang. *multiple indicators / multiple causes*). Porównują one kształtowanie się zmiennych uważanych za indykatory działalności gospodarczej (np. ilość pieniądza w obiegu, zużycie energii) z wielkościami dochodów lub sprzedaży deklarowanymi w celach podatkowych. Część zmienności indykatorów niewyjaśniona oficjalną działalnością gospodarczą uważana jest za aproksymację szarej strefy. W modelu MIMIC szara strefa modelowana jest jako zmienna ukryta, która ma obserwowalne przyczyny (np. wysokość i liczba stawek podatkowych, regulacje rynku pracy, jednostkowe koszty pracy) i obserwowalne skutki (np. popyt na pieniądź, podaż pracy). Luka podatkowa liczona jest jako iloczyn oszacowanej wielkości szarej strefy i efektywnej stawki podatkowej (np. Tax Justice Network, 2011). Wadą metod wskaźnikowych jest jednak to, iż otrzymywane na ich podstawie szacunki wykazują dużą wrażliwość na dobór indykatorów oraz na zmiany strukturalne w gospodarce (choćby zmiany

³ Metoda została zaproponowana przez Cagana (1958) i rozszerzona przez Gutmanna (1977) i Tanziego (1980, 1983).

⁴ Metoda została zaproponowana przez Kaufmann and Kaliberda (1996).

⁵ Model MIMIC opiera się na metodyce modelowania zmiennych ukrytych (Zellner, 1970). Wykorzystanie jej do szacowania wielkości szarej strefy zaproponowali Frey i Weck-Hannemann (1983, 1984). Aplikacje metody dla różnych gospodarek znajdują się m.in. w pracach: Giles (1999), Giles i in. (2002), Giles i Tedds (2002), Dell'Anno i Schneider (2003), Schneider (2005), Bajada i Schneider (2005).

energochłonności produkcji czy preferencji dotyczących płatności gotówkowych), które interpretowane są jako zmiany wielkości szarej strefy (patrz m.in. Breusch 2005; Schneider i Bühn, 2013; Gemmell i Hasseldine, 2015). Z tych względów metody wskaźnikowe zostały uznane przez grupę ekspertów m.in. OECD, IMF i ILO⁶ za całkowicie nieadekwatne w zastosowaniach praktycznych⁷. Dodatkowo, co istotne z punktu widzenia monitoringu luki VAT, metody te nie pozwalają na jednoznaczną dekompozycję luki ze względu na typ podatku.

2. Metodyka badania

W świetle wad wykorzystywanych w literaturze metod – znacznych opóźnień publikacji niezbędnych danych w przypadku analizy rozbieżności oraz niskiej wiarygodności szacunków i braku możliwości dekompozycji luki ze względu na typ podatku w przypadku metod wskaźnikowych – na potrzeby niniejszego badania opracowana została alternatywna metodyka szacowania zmian luki VAT. Łączy ona dwie tradycje ekonometrycznego modelowania dochodów podatkowych – metodę szacowania rozbieżności między dochodami rzeczywistymi a teoretycznymi oraz metodę szacowania elastyczności dochodów względem bazy podatkowej. Podobnie jak w metodzie rozbieżności, porównywane są rzeczywiste i teoretyczne dochody z VAT, jednak analizie podlega nie różnica w ich poziomach, ale w ich trajektorii, czego odzwierciedleniem jest elastyczność jednej zmiennej względem drugiej. Z kolei estymacja elastyczności dochodów rzeczywistych względem teoretycznych jest *de facto* równoznaczna z szacowaniem elastyczności dochodów skorygowanych o zmiany stawek podatkowych względem bazy podatkowej.

⁶ Efektem ich pracy jest wydana w 2002 r. publikacja „*Measuring the Non-Observed Economy - A Handbook*”.

⁷ „(...) there are serious problems with macro-model methods that cast doubt on their utility for any purpose in which precision is important.” OECD 2002, s. 187)

W ramach proponowanej metody dochody teoretyczne konstruowane są w oparciu o trzy agregaty, w przypadku których dostępne są bieżące (z opóźnieniem jednego kwartału) dane o częstotliwości kwartalnej: spożycie prywatne, zużycie pośrednie sektora instytucji rządowych i samorządowych oraz nakłady brutto na środki trwałe sektora instytucji rządowych i samorządowych⁸. Na ich podstawie konstruowany jest indyktor teoretycznych należności podatkowych ($VTTL^*$) według wzoru:

$$VTTL_t^* = C_t * rate_{Ct} + IC_t^{GG} * rate_{IC_{GG}t} + GFCF_t^{GG} * rate_{GFCF_{GG}t} \quad (4)$$

gdzie:

- C_t – wartość netto spożycia finalnego;
- IC_t^{GG} – wartość netto spożycia pośredniego sektora instytucji rządowych i samorządowych (ang. *general government, GG*);
- $GFCF_t^{GG}$ – nakłady brutto na środki trwałe sektora instytucji rządowych i samorządowych;
- $rate_{it}$ – efektywna stawka podatku VAT dla i-tej kategorii bazy podatkowej ($i \in \{C, IC_{GG}, GFCF_{GG}\}$).

Z racji tego, iż sektor instytucji rządowych i samorządowych jest podmiotowo zwolniony z VAT, we wzorze (4) nie występuje parametr $exempt_rate_{GGt}$.

Przy szacowaniu dochodów teoretycznych według wzoru (4) pomijane są zużycie pośrednie i nakłady brutto na środki trwałe związane ze sprzedażą zwolnioną z VAT sektorów innych niż sektor instytucji rządowych i samorządowych. Uproszczenie w stosunku do wzoru (1) stanowi również fakt, iż agregaty konsumpcji i inwestycji nie podlegają rozbiciu na poszczególne kategorie towarów i usług. Do liczenia teoretycznych należności wykorzystywane są zagregowane dane i odpowiadające im stawki efektywne (z okresu, dla którego dostępne są zdezagregowane dane umożliwiające ich policzenie). Ze względu na

⁸ Źródła i szczegółowy opis danych zawarto w publikacji Mazura i in. (2019), s. 11-13.

powyższe uproszczenia – podyktowane brakiem odpowiednich danych – nie można dochodów teoretycznych skonstruowanych na podstawie wzoru (4) wykorzystać bezpośrednio do liczenia luki podatkowej według wzoru (2), szczególnie że były one w okresie analizy często niższe od dochodów rzeczywistych. W związku z tym proponowana metoda analizy luki VAT opiera się na estymacji zmian w czasie elastyczności rzeczywistych dochodów z VAT względem dochodów teoretycznych. Co więcej, zakładając wyjściowy poziom luki (np. na podstawie szacunków Komisji Europejskiej za ostatni dostępny okres), możliwe staje się również szacowanie jej poziomu.

3. Zagadnienia ekonometryczne

3.1. Szacowanie długookresowej elastyczności

Możliwość kwartalnego rozliczania podatku VAT oraz zwroty nadwyżki podatku naliczonego nad należnym powodują, iż w zależności między dochodami rzeczywistymi a teoretycznymi występują przesunięcia w czasie. Z tego względu właściwą metodą modelowania jest w tym przypadku analiza kointegracji, pozwalająca na pełną charakterystykę mechanizmu transmisji w systemie zmiennych, w szczególności oszacowanie elastyczności długookresowej, odzwierciedlającej zmiany dochodów podatkowych w reakcji na zmianę bazy podatkowej po dokonaniu wszystkich rozliczeń i zwrotów. Ze względu na jednoznaczny kierunek przyczynowości (dochody podatkowe dostosowują się do bazy podatkowej) i krótkość dostępnych szeregów (próba obejmuje lata 1999-2018) analiza ma charakter jednowymiarowy i wykorzystuje metodę analizy kointegracji w ramach modelu ARDL zaproponowaną przez Pesarana i Shina (1999). Metodyka ta stanowi alternatywne podejście do analizy związków długookresowych między zmiennymi (obok podejścia Engle'a-Grangera, 1987 i Johansena, 1988) i pozwala na łączną estymację parametrów krótko- i długookresowych w ramach modelu jednowymiarowego.

Pesaran i Shin (1999) udowodnili, że estymator długookresowych parametrów w ramach modelu ARDL jest zgodny i asymptotycznie normalny. W przeciwieństwie do metody Engle'a-Grangera, możliwe jest w jego ramach standardowe wnioskowanie statystyczne. Jako analiza jednowymiarowa, metoda Pesarana i Shina jest – podobnie jak metoda Engle'a-Grangera – dedykowana dla systemów dwóch zmiennych, w których kierunek przyczynowości nie budzi kontrowersji (jako że podział zmiennych na endo- i egzogeniczne dokonywany jest *a priori*). Tym niemniej, metoda ta spotkała się w ostatnich latach z dużym zainteresowaniem badaczy empirycznych dzięki szerokim możliwościom ujmowania i testowania w jej ramach różnych przejawów nieliniowości w relacjach długookresowych (asymetrii, progowości, reżimowości itp.).

Punktem wyjścia analizy jest model ARDL(1,p,q) na poziomach zmiennych:

$$y_{1t} = \alpha_0 + \sum_{i=1}^p \alpha_i y_{1t-i} + \sum_{i=0}^q \beta_i y_{2t-i} + \vartheta_t \quad (5)$$

który szacowany jest po reparametryzacji do jednej z postaci modelu ECM (*unrestricted error correction model*):

$$\Delta y_{1t} = \alpha_0 + \gamma y_{1t-1} + \beta y_{2t-1} + \sum_{i=1}^{p-1} \alpha_i \Delta y_{1t-i} + \sum_{i=0}^{q-1} \beta_i \Delta y_{2t-i} + \vartheta_t \quad (6)$$

gdzie $\gamma = -(1 - \sum_{i=1}^p \alpha_i)$, $\beta = \sum_{i=0}^q \beta_i$.

Po przekształceniu oszacowanego modelu do klasycznej postaci modelu ECM (*restricted ECM*):

$$\Delta \hat{y}_{1t} = \hat{\alpha}_0 + \hat{\gamma} \left(y_{1t-1} + \frac{\hat{\beta}}{\hat{\gamma}} y_{2t-1} \right) + \sum_{i=1}^{p-1} \hat{\alpha}_i \Delta y_{1t-i} + \sum_{i=0}^{q-1} \hat{\beta}_i \Delta y_{2t-i} \quad (7)$$

otrzymujemy oszacowania długookresowej elastyczności $\hat{\delta} = -\frac{\hat{\beta}}{\hat{\gamma}}$ oraz tempa dostosowań do długookresowej równowagi ($\hat{\gamma}$). Istnienie relacji kointegrującej testowane jest w ramach procedury zaproponowanej przez Pesaran i in. (2001) i polega na weryfikacji hipotezy zerowej, zgodnie z którą $\gamma = \beta = 0$.

3.2. Szacowanie zmian długookresowej elastyczności w czasie

W ramach proponowanego podejścia, trajektoria luki VAT utożsamiana jest ze zmianami reaktywności rzeczywistych dochodów podatkowych względem teoretycznych (czyli – ekwiwalentnie – rzeczywistych dochodów skorygowanych o zmiany stawek podatkowych względem bazy podatkowej). Celem analizy jest zatem oszacowanie szeregu zmiennej w czasie długookresowej elastyczności rzeczywistych dochodów z VAT względem teoretycznych. W tym celu opracowano nowe narzędzie, stanowiące uogólnienie estymacji *rolling window*, oparte na algorytmach uśredniania wyników estymacji zaproponowanych przez Pesarana i Timmermanna (2007). Oryginalne algorytmy Pesarana i Timmermanna przewidziane są dla celów prognostycznych i pozwalają na wybór szerokości okna estymacji lub też uśrednienie wyników uzyskanych dla wszystkich możliwych szerokości okien estymacji tak, aby uzyskać prognozy obarczone możliwie małym błędem. Na potrzeby niniejszej analizy wykorzystano tę ideę, jednak nie w odniesieniu do prognoz tylko do szacowania parametrów modelu. Celem algorytmów było uśrednienie szacunków długookresowych elastyczności z kointegracyjnego modelu ARDL dla wszystkich możliwych szerokości okien estymacji w taki sposób, aby możliwie wiernie odzwierciedlić zmiany w czasie procesu generującego dane. Podejście to nie wymaga zatem – w przeciwieństwie do modeli zmiennych parametrów⁹ – przyjmowania arbitralnych założeń dotyczących procesu, jakiemu podlegają zmiany parametrów ani określania wartości początkowych.

Do celów prognostycznych Pesaran i Timmermann zaproponowali, aby wagi służące do uśredniania wyników estymacji oprzeć na błędach prognoz. Dla celów niniejszej analizy wagi skonstruowano w oparciu o jakość dopasowania modelu do danych. Dla kolejnych obserwacji ($t = 1 \dots T$) szacowano modele we

⁹ W ramach analizy kointegracji zostały opracowane modele zmiennych parametrów, w których szereg elastyczności aproksymowany jest przy pomocy funkcji wielomianowych (Bierens i Martins, 2010) lub trygonometrycznych (Park i Hahn, 1999). Z kolei w przypadku wykorzystania do modelowania zmiennej w czasie kointegracji modelu przestrzeni stanów (Koop i in., 2011) zakłada się, iż parametry długookresowe podlegają procesowi autoregresji.

wszystkich możliwych szerokościach okien¹⁰ ($m = 1 \dots M$), przyjmując pewną minimalną szerokość (zgodnie z zaleceniami Pesarana i Timmermanna trzykrotność liczby szacowanych parametrów), a następnie dla każdej szerokości wyznaczano jakość dopasowania modelu do danego punktu w czasie, wykorzystując w tym celu sumę kwadratów błędów predykcji (*SSE*, ang. *sum of squared errors*)¹¹. Na tej podstawie każdej szerokości okna (m) przypisano wagę (w_{mt}) proporcjonalną do jakości dopasowania (definiowanej jako odwrotność sumy kwadratów reszt, $\frac{1}{SSE_{mt}}$) modelu szacowanego w tym oknie do obserwacji empirycznej w punkcie t :

$$w_{mt} = \frac{\frac{1}{SSE_{mt}}}{\sum_{m=1}^M \left(\frac{1}{SSE_{mt}} \right)} \quad (8)$$

Następnie, na podstawie skonstruowanych wag dla każdego punktu w czasie liczona była ważona średnia ($\hat{\delta}_t$) z elastyczności oszacowanych w kolejnych oknach ($\hat{\delta}_t^m$):

$$\hat{\delta}_t = \sum_{m=1}^M w_{mt} \hat{\delta}_t^m \quad (9)$$

Wariancja uśrednionych oszacowań ($var(\hat{\delta}_t)$) liczona była według wzoru (patrz Buckland i in., 1997):

$$var(\hat{\delta}_t) = \left[\sum_{m=1}^M w_{mt} \sqrt{var(\hat{\delta}_t^m) + (\hat{\delta}_t^m - \hat{\delta}_t)^2} \right]^2 \quad (10)$$

¹⁰ Okna mają charakter obustronny (zarówno prawo-, jak i lewostronny), tj. do estymacji procesu generującego dane dla danego punktu w czasie wykorzystywane są zarówno informacje przeszłe, jak i przyszłe.

¹¹ Dopasowanie modelu do tylko jednej obserwacji generuje dużą zmienność przebiegu szeregu elastyczności. W związku z tym – celem wygładzenia szeregu – dopasowanie było liczone dla danego punktu w czasie oraz dla obserwacji sąsiadujących.

gdzie $var(\widehat{\delta}_t^m)$ oznacza wariancję oszacowania parametru w m -tym oknie. Wariancja uśrednionych oszacowań jest zatem funkcją zarówno błędów szacunku parametru w poszczególnych oknach, jak i rozrzutu oszacowanych parametrów wokół wartości średniej.

Przedziały ufności dla uśrednionych parametrów konstruowane były według standardowego wzoru:

$$\widehat{\delta}_t \pm z_\alpha \sqrt{var(\widehat{\delta}_t)} \quad (11)$$

gdzie z_α oznacza $(1 - \alpha)$ kwantyl rozkładu normalnego standaryzowanego.

3.3. Kwantyfikacja zmian w procesie generującym dane

Kwantyfikacja zmian w procesie generującym dane (tj. zmian w zakresie reaktywności dochodów podatkowych względem bazy), utożsamianych ze zmianami luki podatkowej, wymaga określenia okresu referencyjnego. Najbardziej prawdopodobna data zmiany strukturalnej w procesie generującym dane określona została poprzez minimalizację wartości kryteriów informacyjnych w ramach następującej specyfikacji modelu:

$$\Delta y_{1t} = \alpha_0 + \gamma(y_{1t-1} + \delta y_{2t-1} + \delta^* \mathbb{I}_{\{t \geq k\}} y_{2t-1}) + \sum_{i=1}^{p-1} \alpha_i \Delta y_{1t-i} + \sum_{i=0}^{q-1} \beta_i \Delta y_{2t-i} + \sum_{i=0}^{q-1} \beta_i^* \mathbb{I}_{\{t \geq k\}} \Delta y_{2t-i} + \vartheta_t \quad (12)$$

gdzie k oznacza podejrzewaną datę zmiany strukturalnej, zaś $\mathbb{I}_{\{t \geq k\}}$ jest funkcją wskaźnikową przyjmującą wartość 1, gdy spełniony jest warunek w nawiasie klamrowym (a więc począwszy od daty podejrzewanego załamania), i wartość 0 dla pozostałych obserwacji.

Zmiany w zakresie luki VAT liczone były jako różnica między prognozami dochodów z VAT otrzymanymi w oparciu o zmienny w czasie proces generujący dane oraz przy założeniu, że proces generujący dane pozostawał na poziomie z okresu referencyjnego (scenariusz *no policy change*).

4. Wyniki empiryczne

Szacunki zmiennej w czasie elastyczności rzeczywistych dochodów z podatku VAT względem teoretycznych otrzymano zgodnie z metodyką opisaną w rozdziale 3 jako ważone średnie elastyczności oszacowanych we wszystkich możliwych szerokościach okien estymacji, przyjmując jako wagi jakość dopasowania modelu do kolejnych obserwacji empirycznych. Ze względu na silną sezonowość dochodów z VAT dopasowanie było liczone dla czterech obserwacji wokół danego punktu empirycznego¹², co powodowało uśrednienie szacunków w ujęciu rocznym i tym samym wygładzenie oszacowanego szeregu elastyczności. Przeprowadzono również analizę wrażliwości, biorąc pod uwagę różne kombinacje parametrów algorytmów ważących (dotyczących kryterium doboru wag¹³, minimalnej szerokości okna estymacji¹⁴ czy sposobu doboru liczby opóźnień modelu¹⁵). Jej wyniki wskazują, iż szacunki są w niewielkim stopniu wrażliwe na liczbę obserwacji braną pod uwagę przy konstruowaniu kryterium, przy czym – zgodnie z przewidywaniami – maksymalizacja dopasowania do tylko jednej obserwacji generuje dużą zmienność przebiegu szeregu elastyczności, zaś uwzględnienie dodatkowych obserwacji wokół danego punktu w czasie wygładza szereg (patrz wykres 2).

Uzyskane szacunki (patrz wykres 3) wskazują na zmienną w czasie reaktywność dochodów podatkowych na zmiany bazy podatkowej. Na początku próby elastyczność utrzymywała się przez kilka lat na poziomie zbliżonym do jedności, tj. jednoprocentowemu wzrostowi dochodów teoretycznych w długim okresie (tj. po dokonaniu rozliczeń i zwrotów) towarzyszył około jednoprocentowy wzrost dochodów rzeczywistych, co można utożsamiać z utrzymywaniem się luki podatkowej na stałym poziomie. W latach 2003-2005 odnotowano jej przejściowy spadek do poziomu istotnie poniżej jedności (średnio ok. 0,6). W okresie tym

¹² Brano pod uwagę dany punkt w czasie, dwa punkty poprzedzające oraz jeden po nim następujący.

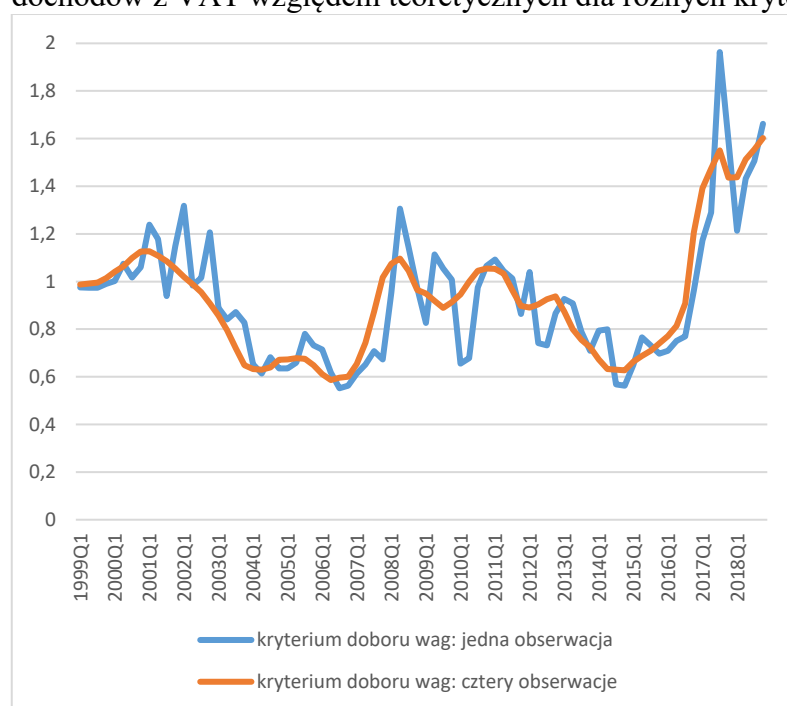
¹³ Brano pod uwagę od jednej do czterech obserwacji.

¹⁴ Brano pod uwagę dwu-, trzy- i czterokrotność liczby szacowanych parametrów.

¹⁵ Stosowano w tym celu kryterium Akaike lub Schwarza.

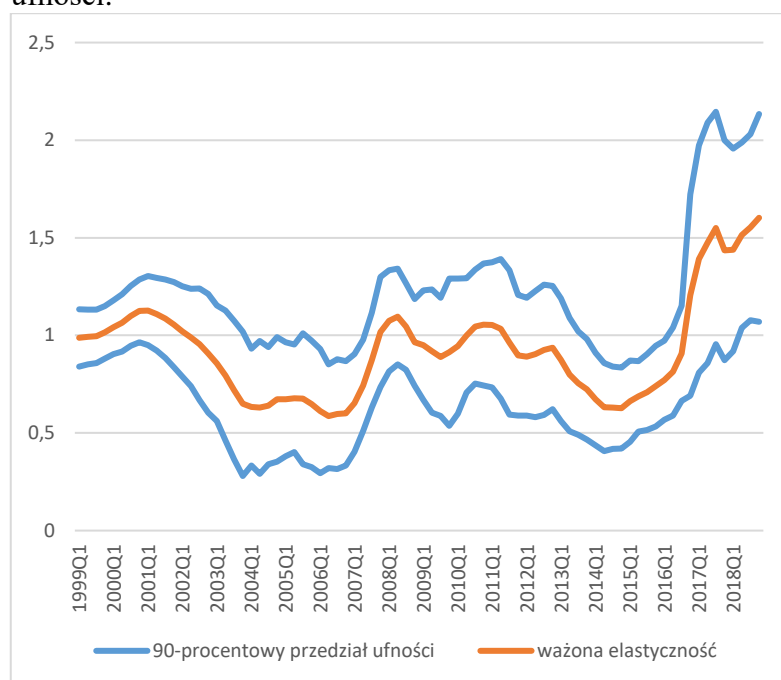
jednoprocentowemu wzrostowi dochodów teoretycznych towarzyszył zatem mniej niż jednoprocentowy wzrost dochodów rzeczywistych, co można interpretować jako rozwieranie się luki podatkowej. Począwszy od 2007 r. elastyczność rosła do poziomu nieco powyżej jedności, który utrzymał się – za wyjątkiem okresu kryzysu finansowego – do roku 2012. Po spadku do poziomu istotnie niższego od jedności elastyczność zaczęła ponownie rosnąć począwszy od 2016, osiągając poziom zbliżony do 1,5 w 2018 r. Pod koniec próby długookresowa elastyczność była po raz pierwszy w analizowanym okresie istotnie wyższa od jedności (przy poziomie ufności 90%). Jednoprocentowemu wzrostowi dochodów teoretycznych towarzyszył zatem w długim okresie więcej niż jednoprocentowy wzrost dochodów rzeczywistych, co wskazuje na domykanie się luki VAT.

Wykres 2. Oszacowany przebieg długookresowej elastyczności rzeczywistych dochodów z VAT względem teoretycznych dla różnych kryteriów doboru wag.



Źródło: Szacunki własne na podstawie danych GUS i MF.

Wykres 3. Oszacowany przebieg długookresowej elastyczności rzeczywistych dochodów z VAT względem teoretycznych wraz z 90-procentowym przedziałem ufności.



Źródło: Szacunki własne na podstawie danych GUS i MF.

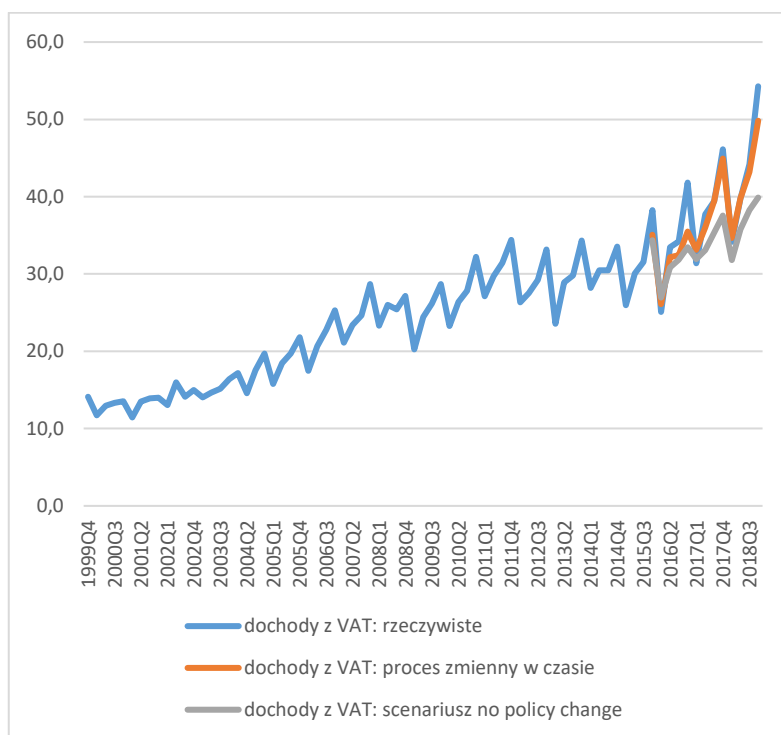
Celem kwantyfikacji zaobserwowanego pod koniec próby domykania się luki VAT określono najbardziej prawdopodobną datę zmiany strukturalnej w procesie generującym dane. W tym celu posłużono się specyfikacją zadaną wzorem (12), przyjmując jako możliwe daty zmian kolejne kwartały lat 2015-17. Kryteria informacyjne jednoznacznie wskazują na IV kw. 2015 r. jako najbardziej prawdopodobną datę zmiany (patrz tabela 1). Efekt domykania się luki policzono, porównując to, jakie byłyby dochody z podatku VAT, gdyby nie nastąpiła zmiana w procesie generującym dane (scenariusz *no policy change*) z dochodami, na które wskazuje zmieniający się proces (patrz wykres 4). Wartości dopasowane otrzymane na podstawie procesu zmiennego w czasie w dużym stopniu odzwierciedlają kształtowanie się rzeczywistych dochodów z VAT. Z kolei proces dopasowany do warunków roku 2015 systematycznie zaniża dochody podatkowe z tytułu VAT w latach 2016-18, co jest odzwierciedleniem wzrostu reaktywności dochodów podatkowych względem bazy w owym okresie w stosunku do roku 2015.

Tabela 1. Kryteria informacyjne dla potencjalnych dat załamania strukturalnego.

Data zmiany strukturalnej	SIC	AIC	HQC
2015Q1	-2.170	-2.502	-2.369
2015Q2	-2.174	-2.507	-2.373
2015Q3	-2.224	-2.587	-2.442
2015Q4	-2.262	-2.595	-2.462
2016Q1	-2.216	-2.549	-2.415
2016Q2	-2.186	-2.518	-2.385
2016Q3	-2.188	-2.520	-2.387
2016Q4	-2.156	-2.489	-2.355
2017Q1	-2.157	-2.490	-2.357
2017Q2	-2.156	-2.488	-2.355
2017Q3	-2.153	-2.486	-2.353
2017Q4	-2.159	-2.491	-2.358

Uwagi: SIC oznacza kryterium informacyjne Schwarza, AIC – kryterium Akaike, zaś HQC – kryterium Hannana-Quinna.

Wykres 4. Dochody podatkowe z tytułu VAT (w mld zł)

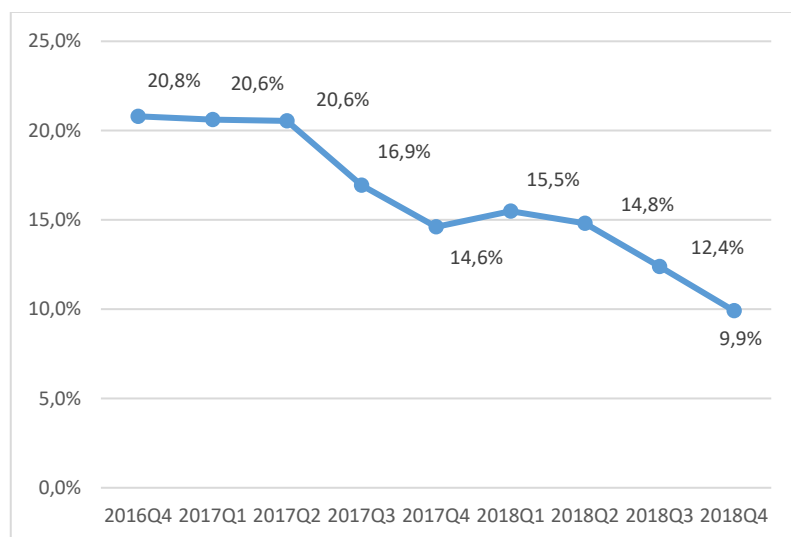


Źródło: Szacunki własne na podstawie danych GUS i MF.

Kwantyfikacji zmian luki VAT dokonano przez odjęcie od wartości dopasowanych otrzymanych na podstawie zmiennego w czasie procesu generującego dane wartości dopasowanych otrzymanych w oparciu o proces sprzed załamania strukturalnego. Szacunki wskazują, iż w związku z domykaniem się luki VAT dochody podatkowe w latach 2016-18 r. były wyższe o ok. 21 mld zł¹⁶. Stanowiło to ok. 45% ogólnego wzrostu dochodów z tytułu podatku VAT w stosunku do roku 2015, zaś pozostała część wynikała ze wzrostu bazy podatkowej. Zgodnie z otrzymanymi szacunkami w 2016 r. luka VAT zmniejszała się o ok. 5 mld, w 2017 ok. 10 mld, zaś w 2018 r. ok. 6 mld zł.

Zakładając skalę luki VAT w IV kw. 2016 r. na poziomie obliczonym przez Komisję Europejską dla roku 2016 (20,8%, patrz wykres 1) oraz wykorzystując oszacowany szereg długookresowej elastyczności rzeczywistych dochodów z VAT względem teoretycznych, możliwe było określenie trajektorii luki w latach 2017-18 (patrz wykres 4). Zgodnie z otrzymanymi szacunkami luka VAT zmniejszyła się w tym okresie o ponad połowę, osiągając poziom poniżej 10% w IV kw. 2018 r.

Wykres 4. Oszacowana trajektoria luki VAT w latach 2017-18.



Źródło: Szacunki własne na podstawie danych GUS, MF oraz KE.

¹⁶ Analiza wrażliwości wskazuje na duży stopień odporności wyników na zmiany parametrów algorytmów użytych do estymacji szeregu zmiennych w czasie parametrów. Otrzymane w jej ramach szacunki zmian luki VAT w latach 2016-18 wahają się od 20,2 do 23,5 mld zł.

Podsumowanie

W niniejszym artykule przedstawiono metodykę szacowania zmian luki VAT opracowaną na potrzeby jej bieżącego monitoringu. Szacunki reaktywności dochodów z podatku VAT względem jego bazy jednoznacznie wskazują na jej wzrost w latach 2016-18 do poziomu istotnie powyżej jedności i tym samym na domykanie się luki VAT. Zgodnie z wynikami estymacji, luka VAT zmniejszyła się w owym okresie o ponad połowę z poziomu – zgodnie z szacunkami Komisji Europejskiej (2018) – 20,8% w 2016 r. do poziomu poniżej 10% na koniec 2018 r. Dochody z podatku VAT były 2018 r. wyższe o ok. 46,5 mld zł niż w 2015 r., z czego – zgodnie z wynikami estymacji – ok. 21 mld (45%) wynikało ze wzrostu reaktywności dochodów podatkowych względem bazy, a zatem z domykania się luki VAT.

Literatura

Bajada C., Schneider F. (2005) *“The Shadow Economies of the Asia-Pacific”*, Pacific Economic Review 10(3), s. 379-401.

Bierens H., Martins L. (2010) *“Time-varying cointegration”*, Econometric Theory 26(5), s. 1453-1490.

Breusch T. (2005) *“The Canadian underground economy: An examination of Giles and Tedds”*, Canadian Tax Journal 53(2), s. 367–391.

Buckland S., Burnham K., Augustin N. (1997) *“Model selection: an integral part of inference”*, Biometrics 53(2), s. 603-618.

Cagan P. (1958) *“The demand for currency relative to the total money supply”*, Journal of Political Economy 66(1), s. 34-46.

Del’Anno R., Schneider F. (2003) *“The shadow economy of Italy and other OECD countries: what do we know?”*, Journal of Public Finance and Public Choice 21(2-3), s. 223-245.

Engle R., Granger C.W. (1987) *“Co-integration and error correction representation, estimation and testing”*, Econometrica 55(2), s.251–276.

Frey B., Weck H. (1983) “*Estimating the shadow economy: a ‘naïve’ approach*”, Oxford Economic Papers 35(1), s. 23–44.

Frey B., Weck-Hanneman H. (1984) “*The Hidden Economy as an ‘unobserved’ Variable*”, European Economic Review 26(1-2), s. 33-53.

Gemmell N., Hasseldine J. (2015) “*Taxpayers’ Behavioural Responses and Measures of Tax Compliance ‘Gaps’: A Critique and a New Measure*”, Fiscal Studies 35(3), s. 275-296.

Giles D. (1999) “*Measuring the Hidden Economy: Implications for Econometric Modelling*”, Economic Journal 109(456), s.370–380.

Giles D., Tedds L. (2002) “*Taxes and the Canadian underground economy*”, Canadian Tax Paper 106.

Giles D., Tedds L., Gugsa W. (2002) “*The Canadian underground and measured economies*”, Applied Economics 34(18), s. 2347–2352.

Gutmann P. (1977) “*The Subterranean Economy*”, Financial Analysts Journal 33(6), s.26-27.

Johansen S. (1988) “*Statistical Analysis of Cointegration Vectors*”, Journal of Economic Dynamics and Control 12(2-3), s. 231–254.

Kaufmann D., Kaliberda A. (1996) “*Integrating the unofficial economy into the dynamics of post socialist economies: a framework of analyses and evidence*”, w: B. Kaminski (red.), Economic Transition in Russia and the New States of Eurasia, M.E. Sharpe, London.

Komisja Europejska (2013) “*Study to quantify and analyse the VAT Gap in the EU-27 Member States*”,
https://ec.europa.eu/taxation_customs/sites/taxation/files/docs/body/vat-gap.pdf

Komisja Europejska (2016) “*The Concept Of Tax Gaps: Report on VAT Gap Estimation*”,
https://ec.europa.eu/taxation_customs/sites/taxation/files/resources/documents/common/publications/studies/tgpg_report_en.pdf

Komisja Europejska (2018) “*Study and Reports on the VAT gap in the EU-28 Member States: 2018 Final Report*”,
https://ec.europa.eu/taxation_customs/sites/taxation/files/2018_vat_gap_report_en.pdf

Koop G., León-González R., Strachan R. (2011) “*Bayesian Inference in a Time Varying Cointegration Model*”, *Journal of Econometrics* 165(2), s. 210-220.

Mazur T., Bach D., Jużwik A., Czechowicz I., Bieńkowska J. (2019) „*Raport na temat wielkości luki podatkowej w podatku VAT w Polsce w latach 2004-2017*”, MF Opracowania i Analizy 3, <https://www.gov.pl/web/finanse/no-3-2019>.

OECD (2002) “*Measuring the Non-Observed Economy - A Handbook*”, <https://www.oecd.org/sdd/na/1963116.pdf>

Park J., Hahn S. (1999) “*Cointegrating Regression with Time Varying Coefficients*” *Econometric Theory* 15(5), s. 664–703.

Pesaran H.M., Timmermann A. (2007) “Selection of estimation window in the presence of breaks”, *Journal of Econometrics* 137, s. 495-510.

Pesaran M.H., Shin Y. (1999) “*An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis*”, w: Strom S. (red.) *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium*, Cambridge University Press.

Pesaran M.H., Shin Y., Smith R.J. (2001) “*Bounds testing approaches to the analysis of level relationships*”, *Journal of Applied Econometrics* 16(3), s. 289-326.

Schneider F. (2005) “*Shadow economies around the world: what do we really know?*”, *European Journal of Political Economy* 21(3), s. 598–642.

Schneider F., Bühn A. (2013) “*Estimating the Size of the Shadow Economy: Methods, Problems and Open Questions*”, CESifo Working Paper 4448.

Tanzi V. (1980) “*The underground economy in the United States: estimates and implications*”, Banca Nazionale del Lavoro Working Paper 135.

Tanzi V. (1983) “*The underground economy in the United States: annual estimates, 1930–1980*”, IMF Staff Paper 30.

Tax Justice Network (2011) “*The Cost of Tax Abuse. A Briefing Paper on the Cost of Tax Evasion Worldwide*”, <https://www.taxjustice.net/wp-content/uploads/2014/04/Cost-of-Tax-Abuse-TJN-2011.pdf>.

Zellner A. (1970) “*Estimation of regression relationships containing unobservable independent Variables*”, *International Economic Review* 11(3), s. 441-454.