GOSPODARKA NARODOWA

3
(271)
Rok LXXXIV/XXV
maj–czerwiec
2014
s. 63–80

Jacek BIAŁEK*

Pomiar obciążenia wskaźnika cen towarów i usług konsumpcyjnych

Streszczenie: Wskaźnik cen towarów i usług konsumpcyjnych (CPI, *Consumer Price Index*) stosowany jest jako podstawowa miara inflacji. W praktyce do pomiaru CPI używany jest indeks cen Laspeyresa z wagami z okresu bazowego. Tak liczony indeks nie uwzględnia zmian w strukturze konsumpcji, które są spowodowane zmianami cen w badanym przedziale czasowym. Oznacza to, że indeks liczony formułą Laspeyresa może być obciążony z tytułu substytucji dóbr (*commodity substytution bias*). Celem pracy jest omówienie potencjalnych źródeł obciążenia indeksu CPI oraz zastosowanie dwóch podejść do jego oszacowania. Pierwsze z nich wiąże się z pojęciem tzw. indeksów superlatywnych. Drugie związane jest z uogólnionym indeksem Fishera i pozwala na bardziej szczegółową dekompozycję obciążenia indeksu CPI. W artykule, poza wynikami badania empirycznego dla danych dla Polski, zaprezentowano również badanie symulacyjne uwzględniające uogólniony indeks Fishera. Analiza empiryczna dla okresu I 2010 – I 2013 prowadzi do wniosku o niewielkim, choć niestabilnym w czasie obciążeniu CPI wynikającym z tytułu substytucji dóbr. Badanie symulacyjne z kolei pozwala stwierdzić, że imputacja cen dóbr nowych i znikających ma diametralny wpływ na wielkość obciążenia CPI z tytułu ich substytucji. Ponadto można zaobserwować, że znak i wielkość korelacji pomiędzy cenami a ilościami dóbr z analizowanego koszyka CPI z reguły nie mają wpływu na wielkość obciążenia CPI w szerokim sensie.

Słowa kluczowe: CPI, indeks superlatywny, indeks Laspeyresa, indeks Fishera

Kody JEL: E17, E21, E30

Artykuł nadesłany 22 listopada 2013 r., zaakceptowany 14 maja 2014 r.

* Uniwersytet Łódzki, Katedra Metod Statystycznych; email: jbialek@uni.lodz.pl

Wprowadzenie

Wskaźnik cen towarów i usług konsumpcyjnych (CPI, *Consumer Price Index*) stanowi przybliżenie zmiany kosztów konsumpcji gospodarstw domowych, która zapewnia utrzymanie użyteczności na stałym poziomie (COLI, *Cost of Living Index* [Diewert 2004]). Indeks CPI jest najczęściej używaną miarą inflacji [White 1999]. Znaczenie wskaźnika CPI w polityce pieniężnej jest szczególnie istotne w przypadku krajów, w których bank centralny posługuje się strategią bezpośredniego celu inflacyjnego. W Polsce obowiązuje ona od 1999 r. i od początku wskaźnikiem referencyjnym jest indeks cen towarów i usług konsumpcyjnych. Metodyka obliczania CPI w Polsce oparta jest na cenowym indeksie Laspeyresa [Białek 2012] z zastosowaniem wag (zmienianych każdego roku), które obliczane są jako udział wydatków na poszczególne dobra w sumarycznych wydatkach wszystkich gospodarstw domowych w poprzednim roku. Wybór tego rodzaju wag jest determinowany po części łatwością obliczeń i dostępnością danych. Powodem, dla którego indeks CPI powinien jak najdokładniej przybliżać rzeczywistą inflację, jest wykorzystanie wskaźnika do indeksacji wartości nominalnych w gospodarce, która z kolei ma znaczenie w decyzjach cenowych przedsiębiorstw [Hałka i Leszczyńska 2011]. W praktyce do pomiaru CPI używany jest indeks cen Laspeyresa z wagami z okresu bazowego. Tak liczony indeks nie uwzględnia zmian w strukturze konsumpcji, które są spowodowane zmianami cen w badanym przedziale czasowym. Oznacza to, że indeks liczony formułą Laspeyresa może być obciążony z tytułu substytucji dóbr. W pracy omówione zostaną dwa różne podejścia do szacowania obciążenia indeksu CPI, przy czym pierwsze jest związane właśnie z efektem substytucji dóbr. W podejściu tym przyjmuje się, że dobrym oszacowaniem tego rodzaju obciążenia indeksu CPI jest różnica między wskazaniem indeksu cen Laspeyresa a pewnym indeksem superlatywnym, z reguły indeksem Fishera [White 1999]. W niniejszej pracy przedstawiony i omówiony zostanie także tzw. uogólniony indeks Fishera [de Haan 2002]. Teoretycznie wydaje się on lepszą alternatywą dla indeksu Laspeyresa, ponieważ uwzględnia zarówno nowe dobra w koszyku, jak i wycofywane [Von der Lippe 2007]. Ciekawsza jest jednak dekompozycja obciążenia indeksu CPI, rozumianego jako różnica pomiędzy wskazaniami obu indeksów. Dekompozycja ta pozwala ustalić siłę i kierunek efektu znikających (*disappearing goods bias*) i pojawiających się nowych dóbr w koszyku (*new goods bias*) i na niej skoncentrowano uwagę w części pracy. W eksperymencie symulacyjnym dokonano oceny wartości obciążenia indeksu CPI w zależności od korelacji pomiędzy cenami a ilościami dóbr w koszyku.

Przyczyny obciążenia indeksu CPI

Większość badaczy zajmujących się problematyką pomiaru inflacji wymienia pięć potencjalnych źródeł obciążenia indeksu CPI [White 1999; Hałka i Leszczyńska 2011], ewentualnie redukuje się ich liczbę do czterech (patrz. raport komisji Boskina [Boskin 1996], badającej obciążenie indeksu CPI w USA):

1. **Obciążenie wynikające z substytucji dóbr** (*commodity substitution bias*). Obciążenie to wynika ze zmian relatywnych cen poszczególnych dóbr wchodzących w skład koszyka CPI. Efekt substytucji polega na tym, że konsumenci reagują na zmiany cen przez zamianę tych dóbr lub usług konsumpcyjnych, które są relatywnie droższe, na dobra relatywnie tańsze [Hałka i Leszczyńska 2011]. Obciążenie substytucyjne estymatora CPI wynika więc z przyjętych stałych wag w formule indeksu Laspeyresa.
2. **Obciążenie wynikające z substytucji rynku zbytu** (*outlet substitution bias*). Obciążenie to wynika z migracji konsumentów w kierunku atrakcyjniejszych, często właśnie się pojawiających rynków dla zakupów. Takim nowym rynkiem może być np. hurtownia internetowa czy punkt sprzedaży wysyłkowej. Formuła Laspeyresa z wagami z okresu bazowego nie jest w stanie nadążać za tego typu zmianami preferencji konsumentów i nowymi kanałami dystrybucji.
3. **Obciążenie wynikające z pojawiania się nowych dóbr** (*new goods bias*). Źródłem tego rodzaju obciążenia są nowe dobra, z jakich w okresie objętym badaniem inflacji zaczęli korzystać konsumenci. Najczęściej są to produkty dotąd innowacyjne, powstałe na skutek wprowadzenia nowej technologii wyrobu, które weszły właśnie do powszechnego użycia. Z oczywistych przyczyn produkty te mogą w ogóle nie być uwzględnione w koszyku dóbr służących oszacowaniu CPI (przykładowo w Polsce opłaty za telefon komórkowy zaczęto uwzględniać dopiero w 2006 r.). Może też się zdarzyć, że spadek cen takich produktów znajduje odzwierciedlenie w CPI dużo później, niż faktycznie nastąpił. Co ciekawe, ocenia się, że ich liczba wynosi od kilku do nawet kilkuset tysięcy w skali roku [Diewert 1996].
4. **Obciążenie wynikające ze zmian jakości produktów** (*quality adjustment bias*). Jest to ten rodzaj obciążenia szacunków CPI, który wynika ze zmieniającej się (np. wraz z rosnącymi oczekiwaniami klientów-konsumentów) jakości oferowanych przez rynek dóbr. Abraham [1995] podaje tu przykład samochodów, których jakość, komfort jazdy, bezpieczeństwo, cena, a także chęć ich posiadania są dziś zupełnie inne niż tych z lat 70. Szacowanie więc inflacji dla długich odcinków czasu musi uwzględniać fakt, że udział tych dóbr w koszyku jest zupełnie inny na początku i na końcu badanego przedziału czasowego. Z definicji wskaźnik CPI powinien mierzyć zmiany cen towarów i usług przy założeniu, że ich cechy nie uległy zmianie w stosunku do okresu bazowego. W rzeczywistości jednak produkty z koszyka dóbr ulegają zmianom – są ulepszane, modyfikowane, a często po prostu wycofywane [Hałka i Leszczyńska 2011].
5. **Obciążenie wynikające z metody kalkulacji** (*formula bias*). Obciążenie to, nazywane także elementarnym obciążeniem indeksu (*elementary index bias* [White 1999]) może powstać jako efekt zastosowania danej metody obliczeń na najniższym poziomie agregacji. W przypadku gdy do kalkulacji wskaźnika ceny danego produktu używana jest średnia arytmetyczna ze wszystkich wskaźników cen danego dobra w kolejnych punktach notowań, wskaźnik cen będzie przeszacowany [Hałka i Leszczyńska 2011]. Jeśli natomiast

najpierw liczona jest średnia cena dobra dla danego okresu ze wszystkich punktów notowań, a następnie jest ona odnoszona do średniej ceny tego dobra w poprzednim okresie, wskaźnik cen nie powinien wykazywać tego rodzaju obciążenia [Ducharme 2000]. Z tą drugą sytuacją mamy do czynienia w Polsce. Obciążenie wskaźnika CPI było niejednokrotnie badane w wielu krajach, jednak z reguły bazowano na danych sprzed kilku lat. Przykładowo, wspomniana komisja Boskina [1996], analizując dane z USA za okres 1995–1996 ustaliła poziom całkowitego obciążenia CPI na 1,1 punktu procentowego[1], w Wielkiej Brytanii w okresie badawczym 1996–2004 nie stwierdzono obciążenia [Cunningham 1996], a w Kanadzie łączne przeszacowanie CPI nie przekroczyło 0,7 punktu procentowego[2] [Crawford 1998]. Te wyniki badań wskazują na dodatnie obciążenie wskaźnika CPI, co sugeruje, że w omawianych krajach przeszacowano rzeczywisty poziom jego wartości. Można znaleźć jednak badania, w których uzyskiwano ujemne obciążenie wskaźnika CPI [Nordhaus 1998]. Hałka i Leszczyńska [2011] zbadały np. obciążenie CPI wynikające z substytucji dóbr dla Polski za lata 2005–2009. Również okazało się, że jego średnia wartość jest ujemna. W dalszej części pracy przeprowadzone zostanie m.in. badanie symulacyjne, którego celem będzie określenie ogólnych warunków wpływających na znak obciążenia wskaźnika CPI.

Indeksy superlatywne w ocenie obciążenia indeksu CPI

Oznaczmy przez $P = [p_1, p_2, ..., p_N]^T$ wektor cen składników koszyka dóbr oraz $Q = [q_1, q_2, ..., q_N]^T$ wektor konsumowanych ilości tych dóbr. Niech $E(P, \bar{u}) = \min_Q \{P^T Q | U(Q) \geq \bar{u})$ będzie funkcją wydatków konsumenta dla pewnej funkcji użyteczności $U(Q)$ i ustalonego referencyjnego poziomu jej wartości $\bar{u}$ – a zatem $E(P, \bar{u})$ to minimalny nakład konsumenta, aby osiągnąć referencyjny pułap wartości jego funkcji użyteczności. Przy wprowadzonych oznaczeniach indeks Konüsa [White 1999] kosztów utrzymania ma postać:

$$I_K = \frac{E(P^t, \bar{u})}{E(P^s, \bar{u})}, \tag{1}$$

gdzie t oznacza okres badany, s oznacza okres bazowy, natomiast $P^\tau = [p_1^\tau, p_2^\tau, ..., p_N^\tau]^T$ jest wektorem cen dóbr w momencie τ. Indeks kosztów utrzymania I_K charakteryzuje się tym, iż uwzględnia zmiany ilości w koszyku dóbr determinowane

1 W tym obciążenie wynikające z substytucji dóbr oszacowano na poziomie 0,4 punktu procentowego, wynikające z substytucji rynku zbytu – na poziomie 0,1 punktu procentowego, wynikające ze zmiany jakości lub nowych dóbr – na poziomie 0,6 punktu procentowego.

2 W tym obciążenie wynikające z substytucji dóbr wyniosło 0,1 punktu procentowego, wynikające z nowych dóbr i zmiany jakości produktów oszacowano na 0,5 punktu procentowego (dane z okresu 1962–1994).

zmianami cen. Tymczasem indeks CPI, zgodny z formułą Laspeyresa, analizuje zmiany cen w koszyku dóbr, ale przy ustalonych ilościach z okresu bazowego, tzn. $Q^s = [q_1^s, q_2^s, ..., q_N^s]^T = Q^t$. Indeks CPI przyjmuje zatem postać

$$I_{La} = \frac{\sum_{i=1}^{N} q_i^s p_i^t}{\sum_{i=1}^{N} q_i^s p_i^s}. \tag{2}$$

Przy założeniu, że wektor Q^t realizuje zadanie minimalizacyjne związane z funkcją wydatków, można pokazać [Diewert 1993], że

$$I_K = \frac{E(P^t, U(Q^s))}{E(P^s, U(Q^s))} \leq I_{La}. \tag{3}$$

A zatem różnica $I_{La} - I_K$ określa obciążenie indeksu CPI z tytułu substytucji dóbr. Wielu ekonomistów za najlepszą aproksymację indeksu I_K uznaje oszacowanie przy pomocy indeksu superlatywnego[3] [White 1999]. Aby wyjaśnić, czym jest indeks superlatywny, należy najpierw zdefiniować pojęcie indeksu *dokładnego*[4] względem jednorodnej w stopniu pierwszym[5] funkcji agregującej[6] $f : R_+^N -> R$, której odpowiada funkcja kosztów jednostkowych $c(P)$. Jest to indeks, który spełnia

$$I = \frac{c(P^t)}{c(P^s)}. \tag{4}$$

Przyjęta funkcja f ma odzwierciedlać preferencje konsumentów. Można pokazać, że indeks Fishera (zdefiniowany poniżej) jest *dokładny* względem następującej funkcji agregującej [Diewert 1976]

$$f(x) = (x^T Ax)^{0,5}, \tag{5}$$

gdzie A jest dowolną macierzą symetryczną. Jeżeli nałożymy pewne dodatkowe warunki[7] na funkcję f, wymagając m.in., aby pochodne cząstkowe I i II rzędu funkcji f istniały i były ciągłe, to o indeksie spełniającym (4) powiemy właśnie,

---

3 Najczęściej w tym celu stosuje się indeks Fishera.

4 W literaturze anglojęzycznej pisze się o takich indeksach jako *„exact for a linearly homogeneous aggregator function"* [Von der Lippe 2007]. Autor artykułu nie spotkał się nigdy z polskim tłumaczeniem tego określenia.

5 Jednorodność stopnia pierwszego (liniowa homogeniczność) oznacza, iż zachodzi $f(\lambda x) = \lambda f(x)$ dla $\lambda \geq 0$.

6 W naszym przypadku jest to funkcja użyteczności.

7 Najważniejszym wymogiem jest tzw. giętkość (termin ten wprowadził Diewert [1974]). Powiemy, że funkcja f jest giętka (*flexible*), jeżeli można za jej pomocą dokonać aproksymacji II rzędu arbitralnie przyjętej funkcji $\hat{f}$ o ciągłych pochodnych cząstkowych I i II rzędu.

że jest superlatywny[8]. Okazuje się [Diewert 1976], że funkcja określona w (5) spełnia te dodatkowe warunki, co oznacza, że indeks Fishera (I_F) jest indeksem superlatywnym. Indeks ten definiujemy jako średnią geometryczną z indeksów wyznaczonych formułami Laspeyresa (I_{La}) i Paaschego (I_{Pa}), czyli

$$I_F = \sqrt{I_{La} I_{Pa}}\,. \tag{6}$$

W pracach Afriata [1972] oraz Samuelsona i Swamy'ego [1974] możemy znaleźć przykłady innych indeksów superlatywnych.

Indeksami superlatywnymi są również indeks Walsha

$$I_W = \frac{\sum_{i=1}^{N} p_i^t \cdot \sqrt{q_i^s q_i^t}}{\sum_{i=1}^{N} p_i^s \cdot \sqrt{q_i^s q_i^t}} \tag{7}$$

oraz indeks Törnqvista

$$I_T = \prod_{i=1}^{N} \left(\frac{p_i^t}{p_i^s}\right)^{\overline{w}_i}, \tag{8}$$

gdzie

$$w_i^s = \frac{p_i^s q_i^s}{\sum_{k=1}^{N} p_k^s q_k^s}, \quad w_i^t = \frac{p_i^t q_i^t}{\sum_{k=1}^{N} p_k^t q_k^t}, \quad \overline{w}_i = \frac{1}{2}(w_i^s + w_i^t). \tag{9}$$

Ponieważ, jak wspomniano, indeksy superlatywne uchodzą za bardzo dobre przybliżenie indeksu kosztów utrzymania I_K, zatem zgodnie z relacją (3), obciążenie indeksu CPI z tytułu substytucji dóbr może być szacowane jako [Hałka i Leszczyńska 2011]:

$$B_{csub} \approx I_{La} - I_{\sup}, \tag{10}$$

gdzie I_{sup} oznacza dowolny indeks superlatywny cen.

Zauważmy, że indeks Laspeyresa, bazujący na strukturze konsumpcji z okresu bazowego, może przeszacowywać poziom inflacji. Z drugiej strony indeks Paaschego może jej niedoszacowywać (wagi wyliczane są już po zmianie cen). Stosowanie uśrednionej struktury konsumpcji z okresu bazowego i badanego to cecha indeksów superlatywnych.

8 Generalnie indeksy superlatywne mają tą zaletę, że aproksymują indeks COLI, abstrahując od postaci funkcji użyteczności. Nie jest więc wymagana estymacja parametrów funkcji $U(Q)$.

Uogólniony indeks Fishera

Klasyczne indeksy cen bazują na ustalonym zbiorze dóbr, stałym w czasie ze względu na strukturę. Oznaczmy przez D^s i D^t zbiór tych dóbr, które należą do rozważanego koszyka dóbr odpowiednio w chwili s oraz t. Wówczas indeks Paaschego można wyrazić jako:

$$I_{Pa} = \frac{\sum_{t \in D^t} q_i^t p_i^t}{\sum_{t \in D^t} q_i^t p_i^s}. \tag{11}$$

Oznaczmy przez D_{new}^t dobra, które pojawiły się w koszyku po raz pierwszy w chwili t, oraz przez D_{dis}^s dobra, które funkcjonowały jedynie w chwili s, by zniknąć w chwili t. Oczywiście zachodzi $D_{dis}^s \subseteq D^s$ oraz $D_{new}^t \subseteq D^t$. Przy powyższych oznaczeniach uogólniony indeks Fishera[9] definiujemy następująco [de Haan 2002]:

$$I_F^G = \left[\frac{\sum_{i \in D^s} q_i^s p_i^t}{\sum_{i \in D^s} q_i^s p_i^s} \frac{\sum_{i \in D^t} q_i^t p_i^t}{\sum_{i \in D^t} q_i^t p_i^s} \right]^{\frac{1}{2}}, \tag{12}$$

gdzie $p_i^t = \hat{p}_i^t$ dla $i \in D_{dis}^s$ oraz $p_i^s = \hat{p}_i^s$ dla $i \in D_{new}^t$ są imputowanymi cenami. Oznaczmy przez $D^{st} = D^s \cap D^t$ zbiór tych dóbr, które są dostępne w rozważanym koszyku w momentach s i t.

Klasyczny indeks Fishera zdefiniowany na zbiorze D^{st} ma postać

$$I_F = \left[\frac{\sum_{i \in D^{st}} q_i^s p_i^t}{\sum_{i \in D^{st}} q_i^s p_i^t} \frac{\sum_{i \in D^{st}} q_i^t p_i^t}{\sum_{i \in D^{st}} q_i^t p_i^s} \right]^{\frac{1}{2}}. \tag{13}$$

Wobec (12) i (13) uogólniony indeks Fishera można zapisać jako [de Haan 2002]:

$$I_F^G = I_F \left[\frac{\sum_{i \in D^t} q_i^t p_i^t / \sum_{i \in D^{st}} q_i^t p_i^t}{\sum_{i \in D^s} q_i^s p_i^s / \sum_{i \in D^{st}} q_i^s p_i^s} \right]^{\frac{1}{2}} \left[\frac{\sum_{i \in D^s} q_i^s p_i^t / \sum_{i \in D^{st}} q_i^s p_i^t}{\sum_{i \in D^t} q_i^t p_i^s / \sum_{i \in D^{st}} q_i^t p_i^s} \right]^{\frac{1}{2}}. \tag{14}$$

Drugi i trzeci czynnik stojący po prawej stronie formuły (14) zawiera ceny i ilości dóbr nowych i znikających. Oznaczmy dodatkowo przez

---

9 Wybór do rozważań formuły Fishera nie jest przypadkowy. Diewert [1992] pokazał, że indeks Fishera spełnia 20 najważniejszych aksjomatów (tzw. testów) wymaganych od poprawnie zdefiniowanych indeksów statystycznych. Pełną listę tych testów znajdzie czytelnik w [Balk 1995; Von der Lippe 2007].

$$\mu = \left[\frac{1 + \sum_{D^t_{new}} q_i^t p_i^t / \sum_{i \in D^{st}} q_i^t p_i^t}{1 + \sum_{i \in D^t_{new}} q_i^t \hat{p}_i^s / \sum_{i \in D^{st}} q_i^t p_i^s} \right]^{\frac{1}{2}}, \tag{15}$$

$$\lambda = \left[\frac{1 + \sum_{i \in D^s_{dis}} q_i^s \hat{p}_i^t / \sum_{i \in D^{st}} q_i^s p_i^t}{1 + \sum_{i \in D^s_{dis}} q_i^s p_i^s / \sum_{i \in D^{st}} q_i^s p_i^s} \right]^{\frac{1}{2}}. \tag{16}$$

Można pokazać [Von der Lippe 2007], że zachodzi wówczas

$$I_F^G = I_F \cdot \mu \cdot \lambda . \tag{17}$$

Parametry μ i λ pokazują odpowiednio, jaki jest wpływ efektu nowych i znikających dóbr na wartość uogólnionego indeksu Fishera.

Dekompozycja (17) tego indeksu pozwala lepiej zrozumieć naturę obciążenia indeksu CPI z tytułu substytucji dóbr. Zanim to jednak uczynimy, zauważmy, iż można oczekiwać, że nowe dobra, wprowadzone do koszyka dóbr dopiero w chwili t, charakteryzowały się prawdopodobnie wysoką ceną w okresie bazowym s. Analogicznie, w przypadku dóbr znikających, imputowane ceny ich substytutów z okresu badanego powinny być relatywnie wyższe niż funkcjonujące jeszcze w okresie bazowym ceny tych już w zasadzie „przestarzałych" dóbr. Z powyższego powodu raczej spodziewamy się wartości $\mu < 1$ oraz $\lambda > 1$. Dokonajmy jeszcze rozróżnienia obliczanych formuł Laspeyresa, w zależności od danych, na jakiej bazuje formuła:

$$I_{La} = \frac{\sum_{i \in D^s} q_i^s p_i^t}{\sum_{i \in D^s} q_i^s p_i^s}, \quad I_{La}^{st} = \frac{\sum_{i \in D^{st}} q_i^s p_i^t}{\sum_{i \in D^{st}} q_i^s p_i^s}, \quad \hat{I}_{La} = \frac{\sum_{i \in \hat{D}^s} q_i^s p_i^t}{\sum_{i \in \hat{D}^s} q_i^s p_i^s}, \tag{18}$$

gdzie $\hat{I}_{La}$ jest estymatorem indeksu Laspeyresa cen wyznaczonym przez daną agencję statystyczną (GUS) na podstawie pobranej próby $\hat{D}^s$ ze zbiorowości D^s. Można pokazać, że różnica pomiędzy oczekiwaną wartością indeksu $\hat{I}_{La}$ a uogólnionym indeksem Fishera (czyli poziom obciążenia estymatora CPI, jakim jest formuła Laspeyresa) można wyrazić następująco [de Haan 2002]:

$$E\hat{I}_{La} - I_F^G = \left[E\hat{I}_{La} - I_{La} \right] + \lambda^2 \left[I_{La}^{st} - I_F \right] + \left[\frac{\lambda}{\mu} - 1 \right] I_F^G . \tag{19}$$

Formuła (19) pokazuje, z jakich składowych składa się to obciążenie.

Pierwszy z czynników występujących po prawej stronie formuły (19) reprezentuje obciążenie wynikające w głównej mierze z niedoskonałości metody imputacji cen (w okresie badanym) znikających dóbr. Znak tego obciążenia jest nieznany *a priori*, przy czym będzie dodatni, jeśli GUS przeszacuje imputowane ceny. Drugi z czynników wyraża obciążenie, które w literaturze przedmiotu nazwano obciążeniem CPI w wąskim sensie [de Haan 2002]. Oczekuje się dodatniego znaku tego obciążenia, przy czym jego wielkość jest proporcjonalna do wartości λ^2. Trzeci czynnik cytowany autor nazwał obciążeniem dynamicznym, i jego wartość zależy zarówno od wielkości μ, jak i λ, a zatem czynnik ten uwzględnia efekt zarówno dóbr nowo pojawiających się, jak i znikających (również oczekujemy jego dodatniej wartości, o ile tylko spełnione będą warunki: $\mu < 1$ oraz $\lambda > 1$). Naturalnym problemem wydaje się być ustalenie, która spośród tych trzech składowych obciążenia CPI (i w jakich sytuacjach) ma relatywnie największe znaczenie. W przeprowadzonym badaniu symulacyjnym pokazano, iż największą rolę odgrywa obciążenie dynamiczne.

Badanie empiryczne

Przeprowadzone badanie empiryczne miało na celu ustalenie obciążenia (w wąskim sensie) CPI z tytułu substytucji dóbr w Polsce. Na podstawie danych dostępnych w GUS[10] za okres I 2010 – I 2013 wyznaczono zgodnie z formułą (10) następujące warianty tego obciążenia:

$$BIAS_1 = I_{La} - I_F,\ BIAS_2 = I_{La} - I_T \text{ oraz } BIAS_3 = I_{La} - I_J,$$

gdzie I_J jest nieważonym, cenowym indeksem Jevonsa [Silver i Herai 2007] często stosowanym w ocenie obciążenia CPI [Hałka i Leszczyńska 2011]. Otrzymane rezultaty prezentuje tabela 1.

Tabela 1. Wartości wybranych indeksów cen oraz obciążenie CPI dla Polski w okresie I 2010–I 2013

Formuła	I 2010 – I 2011	I 2011 – I 2012	I 2012 – I 2013
I_{LA}	1,036244	1,040849	1,016906
I_{PA}	1,035568	1,040583	1,016496
I_F	1,035906	1,040716	1,016701
I_T	1,035910	1,040713	1,016702
I_J	1,024879	1,032853	1,010544
$BIAS_1$	0,000338	0,000133	0,000205
$BIAS_2$	0,000333	0,000135	0,000203
$BIAS_3$	0,011365	0,007996	0,006362

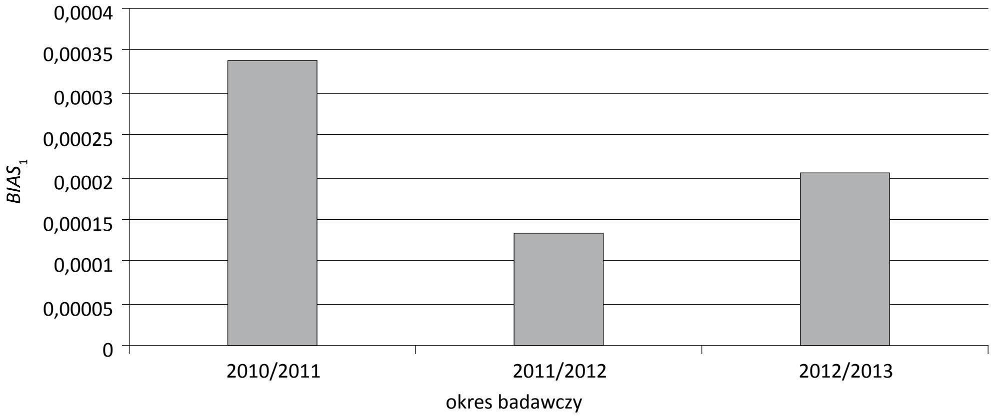
Źródło: Obliczenia własne w programie Excel na podstawie danych z GUS.

---

10 Uwzględniono dane roczne o wskaźnikach cen towarów i usług konsumpcyjnych oraz dane GUS o przyjętym systemie wag za dany rok (dane dostępne na stronie: http://www.stat.gov.pl). Z uwagi na charakter dostępnych danych (wysoki stopień agregacji) ograniczono się tylko do zastosowania niektórych indeksów superlatywnych. Porównano sytuację Polski na koniec stycznia dla lat 2010–2013.

Ponieważ, jak wspomniano, obciążenie CPI z tytułu substytucji dóbr najczęściej wyznaczane jest zgodnie z formułą $BIAS_1$, prezentujemy również wykres słupkowy obrazujący, jak obciążenie to zmieniało się w kolejnych latach objętych badaniem:

Rysunek 1. Wartości obciążenia CPI ($BIAS_1$) w okresie I 2010 – I 2013



Źródło: Opracowanie własne w programie Excel na podstawie danych z GUS.

Badanie symulacyjne

Można zauważyć, że w sytuacji relatywnie licznego zbioru D^{st} w stosunku do zbioru $D^s \cup D^t$ otrzymamy wartości parametrów μ i λ zbliżone do jedności. Innymi słowy, jeżeli zbiór nowych i znikających dóbr jest nieliczny, wówczas uogólniony indeks Fishera dobrze przybliża klasyczny indeks Fishera. W takim przypadku badanie obciążenia z tytułu substytucji dóbr sprowadzić można do pomiaru obciążenia w wąskim sensie, tzn. za pomocą formuły (10), stosując, jako superlatywny, indeks Fishera. Ponieważ interesuje nas wymierna korzyść ze stosowania uogólnionego indeksu Fishera, przyjmiemy w badaniu, iż 25% wszystkich dóbr stanowią dobra nowe i tak samo 25% dóbr stanowią dobra znikające. Rozważać będziemy koszyk $N = 12$ dóbr, przy czym trzy pierwsze będą dobrami znikającymi, a trzy ostatnie dobrami nowymi. Przy wprowadzonych oznaczeniach mamy $D^s = \{1,2,3,4,5,6,7,8,9\}$, $D^t = \{4,5,6,7,8,9,10,11,12\}$, $D^s_{dis} = \{1,2,3\}$, $D^t_{new} = \{10,11,12\}$, $D^{st} = \{4,5,6,7,8,9\}$. Przyjmijmy, że rozkłady cen i ilości dóbr w chwili bazowej s opisują następujące równomierne rozkłady prawdopodobieństwa[11] o parametrach zaprezentowanych w tabeli 2:

11 Zastosowano tu rozkład równomierny $U(m,n)$ [Białek i Depta 2010] o dość znacznej różnicy pomiędzy parametrami rozkładu m i n, aby symulacja objęła stosunkowo szerokie spektrum przypadków. Podobną analizę z zastosowaniem rozkładu równomiernego znajdzie czytelnik w pracy [Białek 2012].

Tabela 2. Parametry m i n rozkładów równomiernych $U(m,n)$ zastosowanych do opisu cen i ilości dóbr w chwili bazowej

Nr dobra	Cena p_i^s		Ilość q_i^s	
	parametr m	parametr n	parametr m	parametr n
1	100	500	3 000	6 000
2	1 000	5 000	20	50
3	3 000	9 000	80 000	90 000
4	1 400	1 700	30 000	80 000
5	1 000	7 000	40 000	90 000
6	300	800	20 000	60 000
7	1 000	6 000	200	600
8	300	900	300	900
9	4 000	8 000	30 000	70 000
10	200	500	300 000	900 000
11	40	80	40 000	80 000
12	3 000	8 000	4 000	1 000

Źródło: Opracowanie własne.

Następnie wprowadzono do badania zmienną określającą tendencję zmian cen w koszyku dóbr jako

$$\Delta p_i \sim U(inf\,1, inf\,2), \text{ dla } i = 1,2,...,12,$$

gdzie parametry *inf1* oraz *inf2* określają tak naprawdę zakres dla stopy inflacji, przy czym oczekiwana wartość tej zmiennej to (*inf1*+*inf2*)/2.

Określono w sposób następujący ceny i ilości w okresie badanym t:

$$p_i^t = p_i^s \cdot (1+\Delta p_i),\ i=3,4,\ldots,12,$$

$$q_i^t = q_i^s (1+\Delta p_i)^c,\ i=1,2,\ldots,12,$$

gdzie parametr c określa reakcję ilości dóbr w koszyku na zmianę ich cen (ujemne wartości tego parametru wskazują na ujemną korelację cen i ilości). Następnie określono wielkości imputowanych[12] cen w okresie t dóbr znikających:

$$\hat{p}_i^t = p_i^s (1+a),\ i=1,2,3,\ \text{a} > 0,$$

oraz ustalono imputowane ceny w okresie s dóbr nowych:

12 Imputację rozumiemy tu jako proces uzupełniania brakujących danych. W przypadku dóbr znikających nie dysponujemy informacjami o ich cenach w okresie t, w przypadku dóbr nowych nie mamy danych o ich poziomie w okresie s. Imputację, wedle potrzeby, przeprowadza urząd statystyczny wg metodologii wypracowanej w danym kraju.

$$\hat{p}_i^s = p_i^t(1+b),\ i=10,11,12,\ \mathrm{b}>0.$$

W ten sposób w eksperymencie wprowadzono możliwość manipulacji wielkością odstępstw cen imputowanych (dóbr nowych i znikających) od cen zarejestrowanych dla jednego z momentów czasowych. Dla każdego zestawu parametrów *inf1, inf2, a, b* i *c* wykonano 100 000 repetycji doświadczenia. Otrzymano wyniki:

Sytuacja 1: *inf1* = –0,02 oraz *inf2* = 0,05

Tabela 3. Średnie wartości badanych formuł dla danych wartości parametrów *a, b* i *c*

Formuła	$c=-5$			
	$a=0{,}03$ $b=0{,}3$	$a=0{,}3$ $b=0{,}03$	$a=0{,}03$ $b=0{,}03$	$a=0{,}3$ $b=0{,}3$
I_{La}	1,0166	1,0578	1,0175	1,0576
I_F	1,0142	1,0150	1,0152	1,0156
I_F^G	0,9799	1,0319	1,0121	1,0004
μ	0,9651	0,9960	0,9959	0,9654
λ	1,0012	1,0208	1,0011	1,0204
Obciążenie CPI	0,0365	0,0256	0,0053	0,0570
Formuła	$c=-1$			
	$a=0{,}03$ $b=0{,}3$	$a=0{,}3$ $b=0{,}03$	$a=0{,}03$ $b=0{,}03$	$a=0{,}3$ $b=0{,}3$
I_{La}	1,0168	1,0581	1,0178	1,0571
I_F	1,046	1,0162	1,0157	1,0150
I_F^G	0,9805	1,0328	1,0129	0,9998
μ	0,9654	0,9961	0,9962	0,9652
λ	1,0011	1,0204	1,0010	1,0205
Obciążenie CPI	0,0364	0,0251	0,0049	0,0572
Formuła	$c=0$			
	$a=0{,}03$ $b=0{,}3$	$a=0{,}3$ $b=0{,}03$	$a=0{,}03$ $b=0{,}03$	$a=0{,}3$ $b=0{,}3$
I_{La}	1,0173	1,0575	1,0163	1,0575
I_F	1,0152	1,0157	1,0138	1,0150
I_F^G	0,9814	1,0326	1,0115	1,0002
μ	0,9657	0,9963	0,9965	0,9654
λ	1,0010	1,0204	1,0012	1,0206
Obciążenie CPI	0,0360	0,0249	0,0048	0,0572

Formuła	$c = 1$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,0167	1,0563	1,0166	1,0579
I_F	1,0144	1,0145	1,0144	1,0156
I_F^G	0,9810	1,0315	1,0120	1,0004
μ	0,9660	0,9965	0,9965	0,9651
λ	1,0011	1,0204	1,0011	1,0206
Obciążenie CPI	0,0356	0,0247	0,0047	0,0575
Formuła	$c = 2$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,0172	1,0579	1,0165	1,0567
I_F	1,150	1,0158	1,0143	1,0147
I_F^G	0,9812	1,0328	1,0119	0,9995
μ	0,9657	0,9964	0,9965	0,9652
λ	1,0011	1,0205	1,0011	1,0205
Obciążenie CPI	0,0359	0,0249	0,0046	0,0571

Źródło: Obliczenia własne w programie Mathematica 8.0.

Sytuacja 2: $inf1 = 0{,}05$ oraz $inf2 = 0{,}2$

Tabela 4. Średnie wartości badanych formuł dla danych wartości parametrów *a, b* i *c*

Formuła	$c = -5$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,1114	1,1511	1,1126	1,1519
I_F	1,1252	1,1252	1,1265	1,1256
I_F^G	1,0774	1,1323	1,1138	1,0974
μ	0,9635	0,9950	0,9949	0,9638
λ	0,9938	1,0114	0,9937	1,0115
Obciążenie CPI	0,0341	0,0189	–0,0009	0,0547
Formuła	$c = -1$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,1109	1,1510	1,1092	1,1514
I_F	1,1245	1,1252	1,1231	1,1257
I_F^G	1,0787	1,1338	1,1121	1,0994
μ	0,9650	0,9962	0,9963	0,9655
λ	0,9939	1,0115	0,938	1,0114
Obciążenie CPI	0,0323	0,0174	–0,0027	0,0522

Formuła	$c = 0$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,1086	1,1516	1,1112	1,1497
I_F	1,1220	1,1260	1,1256	1,1235
I_F^G	1,0774	1,1349	1,1146	1,0974
μ	0,9661	0,9965	0,9965	0,9654
λ	0,9940	1,0113	0,9937	1,0116
Obciążenie CPI	0,0311	0,0169	–0,0031	0,0524
Formuła	$c = 1$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,1091	1,1516	1,1134	1,1507
I_F	1,1230	1,1258	1,1282	1,1250
I_F^G	1,0779	1,1349	1,1169	1,0988
μ	0,9657	0,966	0,9963	0,9656
λ	0,9939	1,0115	0,9935	1,0115
Obciążenie CPI	0,0314	0,0168	–0,0032	0,0521
Formuła	$c = 2$			
	$a = 0{,}03$ $b = 0{,}3$	$a = 0{,}3$ $b = 0{,}03$	$a = 0{,}03$ $b = 0{,}03$	$a = 0{,}3$ $b = 0{,}3$
I_{La}	1,1089	1,1506	1,1110	1,1470
I_F	1,1225	1,1247	1,1253	1,1212
I_F^G	1,0779	1,1341	1,1148	1,0959
μ	0,9661	0,9970	0,9971	0,9664
λ	0,9940	1,0116	0,9937	1,0116
Obciążenie CPI	0,0310	0,01643	–0,0039	0,0510

Źródło: Obliczenia własne w programie Mathematica 8.0.

Wnioski z badania empirycznego

Na podstawie przeprowadzonego badania empirycznego można stwierdzić, iż wybór formuły indeksu – jako odjemnik we wzorze (10) – do szacowania wielkości obciążenia CPI jest istotny. Nasza analiza wskazała, że zastosowanie indeksu Jevonsa prowadzi do największych wartości obciążenia CPI z tytułu substytucji dóbr, natomiast użycie indeksów superlatywnych (Fishera i Törnqvista) generuje zdecydowanie najmniejsze i podobne wartości obciążenia. Obciążenie rozumiane zgodnie z $BIAS_1$ nie jest stabilne w czasie (por. rys. 1), ale jego wartości są bardzo niskie i nie przekraczają 0,034 punktu procentowego. Jest to wartość mieszcząca się w zakresie obciążenia CPI, jakie uzyskano w innych krajach[13]. Można zatem wnioskować, że relatywnie częsta aktualizacja wag w koszyku dóbr CPI w Polsce skutecznie redukuje efekt substytucji dóbr przy kalkulacji inflacji.

Wnioski z badania symulacyjnego

Pierwszym generalnym wnioskiem, jaki można postawić na podstawie badania symulacyjnego, jest to, że poziom całkowitego[14] obciążenia indeksu CPI zdaje się być ujemnie skorelowany z poziomem inflacji (patrz parametry *inf1* oraz *inf2*) szacowanym na zbiorze dóbr D^{st} dostępnych w obu momentach czasowych. Tłumaczyć to można w ten sposób, że imputowane ceny dóbr nowych i znikających z reguły generują inne wskazania cenowych indeksów cząstkowych wyznaczonych dla tych dóbr, niż wynoszą wartości indeksów cząstkowych wyznaczonych dla dóbr pozostałych. W związku z tym np. w sytuacji niewielkiej inflacji na zbiorze D^{st}, ale przy dużych zmianach cen na zbiorach D^{s}_{dis} i D^{t}_{new}, pojawiają się znaczne różnice we wskazaniach nie tylko pomiędzy indeksem Fishera i jego uogólnieniem, ale też pomiędzy indeksem Laspeyresa (wyznaczonym na zbiorze D^{s}) a indeksem I^{G}_{F}. W sytuacji takiej obciążenie w szerokim sensie – rozumiane zgodnie z (19) – będzie odpowiednio duże. Imputacja cen dóbr nowych i znikających ma więc diametralny wpływ na wielkość obciążenia CPI z tytułu ich substytucji.

Drugim generalnym wnioskiem jest fakt, że znak i wielkość korelacji pomiędzy cenami a ilościami dóbr z analizowanego koszyka (parametr c) z reguły niemal nie mają wpływu na wielkość obciążenia CPI (w szerokim sensie). Jak wiadomo [Von der Lippe 2007], gdy ceny i ilości są ujemnie skorelowane, klasyczny indeks cen Laspeyresa I^{st}_{La} wskazuje na większą wartość niż indeks Fishera I_F (liczony na tym samym zbiorze dóbr D^{st}). W powyższej sytuacji obciążenie CPI w wąskim sensie – drugi składnik prawej strony formuły (19) – będzie miało znak dodatni. W sytuacji odwrotnej spodziewamy się ujemnego znaku tego obciążenia. Tymczasem w badaniu symulacyjnym znak całkowitego obciążenia CPI w o wiele mniejszym stopniu, jeśli w ogóle, zależy od parametru c, niż od parametrów a i b, określających ceny imputowanych dóbr nowych i znikających. Wynika to z faktu, że trzeci składnik formuły (19) zawsze prowadził do wartości wielokrotnie większych niż składnik drugi. A zatem finalny znak i wielkość obciążenia CPI w głównej mierze zależy od obciążenia dynamicznego, które z kolei determinują wartości parametrów μ i λ, zależne od cen imputowanych.

---

13 W Wielkiej Brytanii nie stwierdzono tego obciążenia [Marini et al. 2007]), a np. w Kanadzie oszacowano je na poziomie 0,1 punktu procentowego [Crawford 1998].

14 Nie uwzględniono tu składowej obciążenia CPI wynikającej z trafności doboru próby dóbr do koszyka CPI – pierwszy składnik prawej strony formuły (19). W pracy nie analizowano również jakości metod imputacji cen i ich wpływu na ogólne obciążenie szacunków CPI.

Trzecim generalnym wnioskiem jest stwierdzenie, że z reguły (w najbardziej typowej dla zmian cen sytuacji 1) można spodziewać się dodatniej wartości obciążenia CPI w szerokim sensie (patrz tab. 3). Co więcej, nawet jeśli obydwa parametry μ i λ wskazują na wartości mniejsze od jedności (z reguły jednak $\lambda > 1$) to nie jest pewne, że znak obciążenia CPI będzie ujemny. Niewielkie, a czasami nawet ujemne obciążenia CPI pojawiają się w naszym badaniu wtedy, gdy zmiany cen dóbr nowych i znikających są znikome ($a = b = 0{,}03$), przy czym są one mniejsze od przeciętnych zmian cen pozostałych dóbr (patrz tab. 4).

Bibliografia

Abraham K.G. [1995], *Prepared Statement in Cosumer Price Index*, U.S. Government Printing Office, Wasington D.C., United States Senate.

Afriat S.N. [1972], *The Theory of International Comparisons of Real Income and Prices*, w: *International Comparisons of Prices and Outputs*, ed. D.J. Daly, Columbia University Press, New York, s. 13–69.

Balk M. [1995], *Axiomatic Price Index Theory: A Survey*, „International Statistical Review" Vol. 63, s. 69–95.

Białek J. [2012], *Proposition of the general formula for price indices*, „Communications in Statistics: Theory and Methods" Vol. 41, Issue 5, s. 943–952.

Boskin M.J. et al. [1996], *Toward a More Accurate Measure of the Cost of Living*, Final Raport to the Senate Finance Committee from the Advisory Commision to Study the Consumer Price Index.

Crawford A. [1998], *Measurement Biases in the Canadian CPI: An Update*, „Bank of Canada Review" Spring, s. 38–56.

Cunningham A.W. [1996], *Measurement biases in price indexes: an application to the UK's RPI*, Bank of England, Working Paper Series 47.

Diewert W.E. [1974], *Applications of Duality Theory*, w: *Frontiers of Quantitative Economics*, Vol. 2, eds. M.D. Intriligator, D.A. Kendrick, North-Holland, Amsterdam, s. 106–171.

Diewert W.E. [1976], *Exact and Superlative Index Numbers*, „Journal of Econometrics" Vol. 4, s. 115–145.

Diewert W.E. [1992], *Fisher Ideal Output, Input and Productivity Indexes Revisited*, „Journal of Productivity Analysis" Vol. 3, s. 211–248.

Diewert W.E. [1993], *The economic theory of index numbers: a survey*, w: *Essays in index number theory*, Vol. 1, eds. W.E. Diewert, A.O. Nakamura, Amsterdam, s. 177–221.

Diewert W.E. [1996], *Comment on CPI biases*, „Business Economics" Vol. 31, s. 30–35.

De Haan [2002], *Generalised Fisher Price Indexes and the Use of Scanner Data in the Consumer Price Index (CPI)*, „Journal of Official Statistics" Vol. 18, No. 1, s. 61–85.

Ducharme L.M. [2000], *The Canadian CPI and the Bias Issue: Present and Future outlooks*, „Estadistica Espanola" Vol. 42, No. 145, s. 25–41.

Hałka A., Leszczyńska A. [2011], *Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski*, „Gospodarka Narodowa" nr 9, s. 51–75.

Marini G., Piergallini A., Scaramozzino P. [2007], *Inflation bias after the Euro: evidence from the UK and Italy*, „Applied Economics" Vol. 39, s. 461–470.

Narodowy Bank Polski [2003], *Strategia polityki pieniężnej po 2003 roku*.

Nordhaus W.D [1998], *Quality Change in Price Indexes*, „Journal of Economic Perspectives" Vol. 12, s. 59–68.

Samuelson P.A., Swamy S. [1974], *Invariant economic index numbers and canonical duality: Survey and synthesis*, „American Economic Review" Vol. 64, s. 566–593.

Silver M., Heravi S. [2007], *Why elementary price index number formulas differ: Evidence on price dispersion*, „Journal of Econometrics" No. 140, s. 874–883.

White A.G. [1999], *Measurement Biases in Consumer Price Indexes*, „International Statistical Review" Vol. 67, No. 3, s. 301–325.

Von der Lippe P. [2007], *Index Theory and Price Statistics*, Peter Lang, Frankfurt, Germany.

CONSUMER PRICE INDEX MEASUREMENT BIAS

Summary

The article takes an in-depth look at the Consumer Price Index (CPI), which is widely used as a basic measure of inflation.

In practice, the author says, when measuring the CPI economists usually use the so-called Laspeyres price index, which does not take into account changes in the structure of consumption resulting from price changes in a given time interval. The problem is that the Laspeyres index can be biased due to commodity substitution, the author says.

The article discusses potential sources of the CPI bias and the application of two approaches for calculating the index. The first approach is connected with superlative indices, while the second uses the so-called generalized Fisher price index.

The article presents the results of empirical and simulation studies conducted by the author. An empirical analysis for the 2010–2013 period points to the existence of a small and unstable CPI substitution bias, according to the author. The simulation study, in turn, makes it possible to conclude that imputations of prices of new and disappearing goods are crucial for CPI bias calculations, Białek says. Moreover, the sign and value of the correlation between prices and quantities do not generally influence the CPI substitution bias in the broad sense, the author notes.

Keywords: Consumer Price Index (CPI), superlative index, Laspeyres index, Fisher index

JEL classification codes: E17, E21, E30