

STUDIA METODOLOGICZNE

Marta WIDŁAK

Metody wyznaczania hedonicznych indeksów cen jako sposób kontroli zmian jakości dóbr¹

Lata 2005—2007 były okresem wyjątkowym dla polskiego rynku mieszkaniowego. Dynamika zmian zachodzących na tym rynku była zjawiskiem nowym dla relatywnie młodego sektora nieruchomości. Konsekwencją tych zmian, a także rosnącej skali zaangażowania systemu bankowego w finansowanie rynku mieszkaniowego jest wzrost zainteresowania i zapotrzebowania na rzetelne analizy. Istotny jest tu zwłaszcza pomiar tendencji cenowych na rynku mieszkań.

W gronie zainteresowanych polskim rynkiem mieszkaniowym toczy się dyskusja dotycząca wyboru metody i wiarygodnego sposobu pomiaru dynamiki cen nieruchomości. Artykuł wpisuje się w ramy tej dyskusji i poszukiwań badawczych. Jest on poświęcony przedstawieniu metod konstrukcji hedonicznych indeksów cen. Metody te wykorzystywane są do pomiaru zjawisk cenowych na rynkach dóbr heterogenicznych. Głównym celem ich stosowania jest uwzględnienie zmian jakości dóbr we wskaźniku dynamiki ich cen (*quality adjustment*). Indeksy hedoniczne najczęściej stosuje się mierząc dynamikę cen nieruchomości, komputerów, sprzętu AGD/RTV, samochodów. Pomimo że metody omówione w tym opracowaniu są od wielu lat stosowane za granicą, w Polsce nie są powszechnie znane i stosowane. Dzieje się tak m.in. ze względu na brak wiedzy dotyczącej tych metod oraz brak odpowiednich źródeł danych.

¹ Za cenne dyskusje i uwagi dziękuję prof. Jackowi Łaszkwowi oraz dr Emilii Tomczyk. Pracę dedykuję Czesławowi Widłakowi.

Motywacją do napisania artykułu była chęć udzielenia odpowiedzi na liczne pytania, kierowane do autorki, dotyczące sposobów konstrukcji hedonicznych indeksów cen mieszkań. Opracowanie to jest kontynuacją prac (Łaszek, Widłak, 2007, 2008) odnoszących się do wiarygodnego pomiaru tendencji cenowych na rynkach mieszkaniowych w Polsce. Jego celem jest przede wszystkim prezentacja i klasyfikacja głównych metod konstrukcji hedonicznego indeksu cen, a tym samym pomiar „czystej” zmiany cen.

CZYM JEST HEDONICZNY INDEKS CEN?

Najbardziej kompleksową odpowiedzią na to pytanie jest definicja podana przez J. Triplett (2006), która brzmi następująco: „**Hedoniczny indeks cenowy** to każdy indeks cenowy, który wykorzystuje funkcję hedoniczną. Funkcja ta, często określana jako model hedoniczny, to relacja pomiędzy cenami różnych rodzajów danego dobra a cechami, które to dobro charakteryzują”.

Przez cechy (charakterystyki/trybuty) dobra należy rozumieć jego elementy składowe opisujące dobro i wpływające na jego cenę. Na przykład w przypadku mieszkania cechami tymi mogą być: liczba pokoi, powierzchnia, lokalizacja, rodzaj kuchni itp. Posługując się różnego rodzaju skalami pomiarowymi cechom tym możemy przyporządkować określone ilości. Określenie „ilość cechy” („ilość charakterystyki/trybutu”) jest tu rozumiane jako konkretny wariant (realizacja) danej cechy dobra. I tak np. ilość cechy *liczba pokoi*, odmierzając na skali ilorazowej, możemy określić jako 1, 2, 3 itd.; ilość cechy *lokalizacja* możemy opisać posługując się skalą porządkową: dobra, przeciętna, zła itd. Wyrażenie „ilości cech” używane w artykule należy zatem rozumieć nie jako liczbę cech (liczbę zmiennych objaśniających) wykorzystanych w modelu hedonicznym, lecz zbiór konkretnych ilości (realizacji/wariantów) wyodrębnionych cech.

Twórcy modeli hedonicznych zakładają, że wartość pewnych dóbr (np. mieszkania) wynika z wartości, jakie mają dla użytkownika poszczególne składowe tych dóbr (np. powierzchnia użytkowa, liczba pokoi, występowanie ogrodu itp.). To założenie pozwala przedstawiać cenę rozpatrywanego dobra jako funkcję jego charakterystyki.

Warto zaznaczyć za J. Triplett, że nie każdy indeks wykorzystujący funkcję regresji jest indeksem hedonicznym (jeśli ta regresja nie jest funkcją hedoniczną), a także nie każdy indeks hedoniczny jest indeksem korzystającym z modelu regresji. Warunkiem koniecznym jest to, aby indeks hedoniczny wykorzystywał funkcję hedoniczną.

Definicja Triplett nie narzuca też sposobu wykorzystania funkcji hedonicznej. Powoduje to, że miano indeksu hedonicznego otrzymuje każda metoda, w której w dowolny sposób użyto funkcji hedonicznej. W praktyce cenowe indeksy hedoniczne są liczone na wiele sposobów. Przedstawienie podstawowych sposobów wyznaczania indeksu cen wykorzystującego model hedoniczny oraz ich klasyfikacja jest właściwym celem tego opracowania. Za podstawowe uzna-

no metody najczęściej opisywane w literaturze przedmiotu (Griliches, 1961; Diewert, 2003, 2006, 2007; Triplett, 2006; Silver, Heravi, 2006, Bover, Izquierdo, 2001 i wielu innych) oraz uznane i przedstawiane przez międzynarodową grupę statystyków pracującą nad określeniem międzynarodowych standardów metodologicznych w zakresie wyznaczania indeksów cen (ILO, 2004).

Początki powstawania hedonicznych modeli cen datowane są na lata 20. i 30. XX w., kiedy to związane były głównie z pracami ekonomistów rolnictwa. Pierwsze próby wyrażenia ceny dóbr jako funkcji jego atrybutów były podejmowane przez: Waughę (1928) i dotyczyły cen warzyw, Vaila (1932) — cen nawozów sztucznych, Haasa (1922) — cen ziemi, Taylora (1916) — cen bawełny oraz Courta (1938), przy czym Court był pierwszym, który opublikował swoje prace (Griliches, 1988). Rozwój formalnych metod analizy hedonicznej wiąże się jednak z okresem późniejszym. Są to lata 60. i 70. XX w., które przyniosły prace Rosena i Grilichesa.

Klasycznym przykładem pierwszych zastosowań modeli hedonicznych jest praca Grilichesa opublikowana w 1961 r. i dotycząca rynku samochodów, a także modele cen komputerów opracowane przez Chowa w 1967 r. Na rynku mieszkaniowym pierwszy hedoniczny indeks cen został opublikowany przez the U.S. Census Bureau w roku 1974. Lata 90. XX w. oraz obecna dekada przyniosły kolejne prace i coraz częstsze zastosowania modeli hedonicznych cen. Między innymi znacznie rozwinął się nurt badań zajmujący się wykorzystaniem modeli hedonicznych w konstruowaniu dopasowanych jakościowo indeksów cen na rynkach dóbr wysoce heterogenicznych. Literatura tematu jest bardzo bogata, a wśród badaczy zajmujących się obecnie tą dziedziną statystyki wyróżnić należy m.in. wkład Diewerta, Silvera, Heraviego i Quigleya.

Kwestii dostosowania indeksów cenowych do zmian jakości badanych dóbr (*quality adjustment*) zaczęto poświęcać szczególną uwagę mniej więcej od połowy lat 90. XX w. Dostosowanie indeksów cenowych do zmian jakości dóbr (kontrola zmian jakości lub korekta uwzględniająca zmianę jakości) jest szeroko podkreślane zarówno w kontekście wskaźników CPI, jak i wskaźników sektorowych, służących pomiarom zmian cen dóbr heterogenicznych. Pomiar ceny dobra w dwóch okresach — bieżącym i bazowym — często przeprowadzany jest na odmiennych jakościowo reprezentantach badanego dobra. Sytuacje te są charakterystyczne dla rynków dóbr heterogenicznych (np. mieszkania) lub szybko zmieniających się (np. komputery). Dostosowanie wskaźników cen do zmian jakości badanych dóbr ma umożliwić pomiar „czystej” zmiany cen, wynikającej z innych przyczyn niż różnice jakościowe.

W przypadku koszykowych wskaźników cen (jak np. CPI) problem zmiany jakości pojawia się przede wszystkim w sytuacji zamiany modeli reprezentantów tworzących koszyk. W przypadku, gdy stary model dobra został wycofany z obrotu i trzeba zastąpić go nowym, z reguły lepszym jakościowo, modelem reprezentanta. W takiej sytuacji zmiana ceny rozpatrywanego reprezentanta może wynikać zarówno z różnic jakościowych starego i nowego modelu dobra, jak i z

innych przyczyn ekonomicznych. Indeks cenowy dostosowany do zmian jakości powinien odzwierciedlać tylko tę część zmiany cen, która nie wynika z różnic jakościowych modeli reprezentantów.

Na rynkach dóbr wysoce heterogenicznych i szybko zmieniających się, takich jak mieszkania czy komputery, trudno mówić o reprezentantach i budować koszyk na nich oparty. Różnice jakości są immanentną cechą pomiaru cen na tego typu rynkach. Na przykład na rynku komputerów rozwój technologii powoduje, że zmiany jakości są częste i następują szybko. Wówczas zmiany modeli reprezentantów są praktycznie ciągłe i korekta jakości powinna być dokonywana stale. W przypadku rynku mieszkań sporadycznie dokonywane są transakcje tymi samymi obiektami z okresu na okres, a tylko takie transakcje nie podlegają problemowi różnic jakości. Na rynku praktycznie nie występują dwa identyczne mieszkania² i w dwóch różnych okresach jednostki mieszkaniowe objęte obrotem (a więc takie, dla których można ustalić cenę rynkową) zawsze będą różniły się od siebie jakościowo.

Reasumując, w przypadku indeksów cenowych dóbr, które nie ulegają szybkim zmianom i są w miarę jednorodne, sporadycznie zachodzi potrzeba dokonania korekty zmian jakości. Inaczej jest w przypadku dóbr wysoce zróżnicowanych i podlegających ciągłym ulepszeniom, dla których korekta ta powinna być stałym elementem wyznaczania indeksu cen. Indeks taki powinien z definicji uwzględniać różnice pomiędzy każdym dobrem z okresu t i $t+1$. Jedną z metod korekty wskaźników cen o zmiany jakości badanych dóbr wykorzystuje informacje płynące z funkcji hedonicznej. Skoro funkcja ta określa relację pomiędzy ceną dobra a poszczególnymi jego cechami, to, teoretycznie, znając tę funkcję, cechy dobra i jego cenę obserwowaną na rynku można wyznaczyć „czystą” zmianę ceny. W praktyce problem jest bardziej złożony i stąd różne sposoby wykorzystania funkcji hedonicznej oraz informacji o cechach i cenach do wyznaczania „czystej” dynamiki cen, nazywane metodami wyznaczania hedonicznych indeksów cen.

KLASYFIKACJA METOD KONSTRUKCJI HEDONICZNYCH INDEKSÓW CEN

W myśl przyjętej definicji indeks hedoniczny cen to indeks wykorzystujący funkcję hedoniczną. Z reguły funkcją tą jest model regresji hedonicznej.

W praktyce funkcję hedoniczną wykorzystuje się do konstrukcji indeksów cenowych na kilka sposobów. Ze względu na stopień wykorzystania informacji

² Heterogeniczność mieszkań wynika z różnic pomiędzy wieloma atrybutami, takimi jak: powierzchnia użytkowa, wiek, typ, standard, lokalizacja i cechy tej lokalizacji (np. skład społeczny danego osiedla, dostęp do terenów zielonych, odległość od szkół, centrów komunikacji i handlu). Na rynku mieszkaniowym nie istnieją dwa identyczne mieszkania. Jeśli nawet ich cechy fizyczne są takie same, to mieszkania zawsze różnią się lokalizacją (ogólną lub szczegółową) i jej cechami.

pochodzącej z funkcji hedonicznej cen, metody konstrukcji indeksu możemy podzielić na dwie grupy — **metody pośrednie** oraz **bezpośrednie**.

W **pierwszej grupie** znajdują się metody, które w niewielkim zakresie korzystają z zależności określonej funkcją hedoniczną. W metodach tych badana jest zmiana cen ustalonego koszyka reprezentantów, tzw. metoda reprezentantów (*matched models*). Modelu hedonicznego używa się wyłącznie w celu dopasowania jakościowego reprezentantów, którzy zmieniają się w ustalonym koszyku dóbr, np. gdy dany reprezentant produktu został wycofany ze sprzedaży lub gdy pojawił się nowy reprezentant, który nie był uwzględniany wcześniej, lub gdy starego reprezentanta zastępujemy nowym, nieco odmiennym jakościowo typem dobra³. W tych metodach model regresji służy do dopasowania jakości reprezentanta nowego do starego, tak by nadal porównywano dwa podobne obiekty. Technicznie dopasowania jakości (*quality adjustment*) dokonuje się wykorzystując jedną z dwóch metod:

- 1) imputacji (*imputation method*),
- 2) hedonicznego dopasowania jakości (*hedonic quality adjustment method*).

Grupę metod pośrednich w literaturze określa się na wiele sposobów. W literaturze anglojęzycznej najczęściej nazywa się je metodami hedonicznego dopasowania jakościowego (*hedonic quality adjustment*). Ze względów językowych użycie tej nazwy nie jest wskazane w opracowaniach polskich⁴. Dlatego w podanej klasyfikacji przyjąłem określenie podane przez CPI Manual — **metody pośrednie** (*indirect*). Nazwa ta wiąże się z faktem, że metody z tej grupy wykorzystują informację z modeli hedonicznych w ograniczonym zakresie, tylko w przypadku wybranych reprezentantów, a sam indeks nie pochodzi bezpośrednio z modelu regresji.

Jeszcze inne nazwy spotykane w literaturze to: metody złożone, hybrydowe lub czyste (*composite, hybrid, pure*). Są dwa uzasadnienia dla tych nazw. Po pierwsze metody pośrednie łączą ceny teoretyczne, wyznaczone za pomocą modeli hedonicznych, z cenami rzeczywistymi gromadzonymi przez urzędy statystyczne (reprezentanci podlegający zmianie i reprezentanci, którzy jej nie podlegają). Po drugie metody te łączą tradycyjne schematy konstrukcji agregatowych indeksów cen z modelem hedonicznym, za pomocą którego kontroluje się zmiany jakościowe w koszyku reprezentantów. Czasem metody z tej grupy określane są całościowo jako metody imputacji, gdyż wykorzystują funkcje hedoniczne do przypisania cen odbitych (*shadow prices*) określonym produktom.

W metodach **drugiej grupy** indeks wyliczany jest bezpośrednio w modelu regresji hedonicznej (jako wartość współczynnika regresji stojącego przy jednej ze zmiennych objaśniających) lub informacje z tego modelu służą do wyznacza-

³ W literaturze anglojęzycznej zabieg zastępowania reprezentantów z jednoczesnym dopasowaniem jakościowym potocznie określa się jako „łatanie” (*patching*).

⁴ Czytelnik literatury anglojęzycznej spotka się z określeniem *hedonic quality adjustment* co najmniej w dwóch kontekstach — jako nazwa jednej grupy metod hedonicznych lub jako określenie zabiegu wyznaczenia „czystej” zmiany cen, dokonywanego także przez zastosowanie metod z drugiej grupy.

nia cen odbitych wszystkich (a nie tylko wybranych) reprezentantów dóbr, aby następnie posłużyć do międzyokresowych porównań poziomów cen. Można zatem powiedzieć, że stopień wykorzystania informacji pochodzącej z funkcji hedonicznej jest w metodach bezpośrednich zdecydowanie większy niż w metodach pośrednich. Metody z tej grupy określane są jako **metody bezpośrednie** (*direct*), gdyż oszacowanie funkcji hedonicznej warunkuje wyznaczenie dynamiki cen dóbr (dynamika ta nie mogłaby być wyznaczona bez znajomości funkcji hedonicznej). Metody bezpośrednie można podzielić na metody:

- 1) ze zmiennymi zero-jedynkowymi czasu — dalej zwana w skrócie ZZC (*time dummy variable method*),
- 2) cen charakterystyk (*characteristic price method*),
- 3) imputacji (*imputation method*).

Zastosowanie tych metod nie wymaga tworzenia koszyka reprezentantów, choć go nie wyklucza. Główną zaletą metod bezpośrednich jest fakt, że w łatwy sposób pozwalają wyznaczyć indeks cen, wykorzystując wszystkie dostępne dla danego okresu dane. Wykorzystanie wszystkich danych zalecane jest w przypadku rynków bardzo zmiennych i zróżnicowanych produktów. Dla tego typu rynków najczęściej nie można ustalić stałej próby reprezentantów. Wystarczy pomyśleć o rynkach komputerów, sprzętu RTV, AGD czy rynku mieszkań. Ustalenie stałego koszyka reprezentantów dla tych dóbr wydaje się być praktycznie niemożliwe, a każdy taki koszyk obciążony jest istotnym błędem niereprezentatywności. Z tego powodu metod pośrednich nie wykorzystuje się do wyznaczania indeksu cen na rynku mieszkaniowym. Podział metod konstrukcji hedonicznych indeksów cenowych przedstawia wykres.

Definiując sposób wykorzystania modeli hedonicznych w konstrukcji indeksów cen, Triplett (2003) zwraca uwagę, że metody pośrednie wykorzystują funkcję hedoniczną szacowaną na danych z innego okresu niż okres, którego dotyczy indeks, a metody bezpośrednie przeciwnie — wykorzystują funkcję regresji estymowaną odrębnie dla każdego okresu, dla którego chcemy wyznaczyć indeks cen. Nie jest to jednak wyznacznik klasyfikacji zaproponowanej w artykule. Warunku tego nie spełnia bezpośrednia metoda imputacji, której Triplett nie uwzględnia w swojej pracy, a która jest prezentowana w CPI Manual. Zgodnie z definicją tam podaną metoda ta w jednym z wariantów wymaga jednorazowego oszacowania modelu regresji wykorzystywanego następnie do porównań międzyokresowych.

Literatura traktująca o indeksach hedonicznych jest bardzo bogata, a jej lektura uwidacznia brak standardów w kwestii nazewnictwa i klasyfikacji stosowanych metod. Na przykładzie rynku mieszkaniowego, w zależności od przyjętego kryterium, można podać kilka klasyfikacji hedonicznych indeksów cen. Za kryteria te można przyjąć:

- sposób konstrukcji koszyka (np.: *repeat sales*, *mix-adjustment*, średnia ważona, reprezentatywna nieruchomości, wagi w indeksie cen charakterystyki);
- sposób wykorzystania funkcji hedonicznej w formule indeksu — łącznie jako uzupełnienie metody reprezentantów (metody pośrednie) oraz traktując mo-

del hedoniczny jako samodzielny sposób na wyznaczenie indeksu (metody bezpośrednie);

- segmentację rynku mieszkaniowego (indeksy nowych mieszkań, indeksy cen mieszkań z rynku wtórnego, indeksy domów itp.).

Opisany w artykule podział metod konstrukcji hedonicznych indeksów cenowych odnosi się do kryterium drugiego. Okazuje się, że nawet w wąskim kręgu badaczy najbardziej zasłużonych dla rozwoju omawianej dziedziny pojawiają się znaczne rozbieżności w nazewnictwie i klasyfikacji indeksów hedonicznych. Na

przykład określenie „metoda imputacji” (*imputation method*) jest używane przez Triplett i ILO dla dwóch różnych rodzajów metod, z kolei Silver i Heravi używają dokładnie tego samego określenia dla metody określonej tu jako „metoda cen charakterystyk”. Co więcej z rozbieżnościami tymi często trudno się spierać. Przecież w podanym przykładzie wszystkie trzy metody za pomocą modelu regresji imputują ceny dóbr. Rozbieżności te oraz brak wyraźnych standardów i jasnej klasyfikacji metod stanowią bardzo dużą przeszkodę dla badaczy zapoznających się z problematyką hedonicznych indeksów cen. Jednakże międzynarodowe forum zajmujące się opracowaniem standardów statystyki gospodarczej rozpoczyna obecnie prace nad stworzeniem podręcznika dotyczącego kompilacji indeksów cenowych na rynku mieszkań.

METODY WYZNACZANIA HEDONICZNYCH INDEKSÓW CEN

Najczęściej stosowana postać funkcji hedonicznej to model regresji, zatem dla celów dalszych wyjaśnień będę posługiwała się zamiennie określeniem funkcja i regresja hedoniczna. Ogólnie model regresji hedonicznej można zapisać w następujący sposób:

$$p_{i,t} = a_0 + \sum_j a_j z_{i,j,t} + \varepsilon_{i,t} \quad (1)$$

gdzie:

- p — cena dobra,
- a — współczynnik regresji,
- z — wektor cech opisujących dobro,
- i — i -te dobro,
- j — j -ta cecha dobra,
- t — indeks czasu (stąd: $z_{i,j,t}$ oznacza ilość j -tej cechy dla i -tej jednostki dobra pochodzącego z okresu t , z_i — oznaczenie to wykorzystywane w dalszej części opracowania oznacza wektor cech charakteryzujących i -te dobro),
- ε — składnik losowy.

W zdecydowanej większości badań modele hedoniczne są modelami semilogarytmicznymi⁵, dlatego w dalszej części artykułu wzory będą podawane dla tej postaci modelu:

$$\ln p_{i,t} = a_0 + \sum_j a_j z_{i,j,t} + \varepsilon_{i,t} \quad (2)$$

⁵ Uzasadnienie i przykłady dotyczące rynku mieszkaniowego można znaleźć praktycznie w każdej pracy empirycznej. Dla polskiego rynku mieszkaniowego uzasadnienie semilogarytmicznej (wykładniczej) postaci funkcyjnej można znaleźć w już przywołanej pracy Łaszek i Wiślak.

Współczynniki modelu regresji a_j interpretuje się jako przeciętne ceny poszczególnych cech badanego dobra, zwane dalej cenami charakterystyki (cech/ atrybutów). Przyjmuje się, że ceny te odzwierciedlają nieobserwowalne, rynkowe wyceny wartości każdej charakterystyki (cech/ atrybutów) dobra, a więc są wynikiową działania sił popytu i podaży. Nieobserwowane ceny charakterystyki dobra często określane są jako ceny implikowane lub domniemane (*implicit prices*). Założmy, że postać funkcji hedonicznej jest prawidłowa, tzn. poprawna jest specyfikacja modelu, zależność funkcyjna między zmienną objaśnianą i zmiennymi objaśniającymi oraz nie występują inne problemy ekonometryczne związane najczęściej ze współliniowością zmiennych czy heteroskedastycznością składnika losowego. Przy takich założeniach składnik losowy ε_t z równania (1) obrazuje niedoszacowanie lub przeszacowanie rynkowej ceny dobra względem ceny średniej wynikającej z cech tego dobra (jego jakości mierzonej wartością poszczególnych cech charakterystyki, takich jak: wielkość mieszkania, liczba łazienek, nasłonecznienie, typ kuchni itp.) oraz tego, jak średnio cechy te wyceniane są przez rynek.

Metody pośrednie

Pośrednie metody szacowania hedonicznych indeksów cenowych mają zastosowanie dla indeksów wyznaczanych dla koszyka reprezentantów (metoda reprezentantów). Gdy badamy zmiany wartości ustalonego koszyka reprezentantów, modele regresji hedonicznej są wykorzystywane wyłącznie do korekty zmian jakości koszyka, która to zmiana zachodzi pod wpływem zmiany poszczególnych reprezentantów. Istnieje kilka wariantów zastosowania regresji hedonicznej w trakcie korekty zmiany cen uwzględniającej zmianę jakości dóbr.

Metoda imputacji (pośrednia)

Triplett wyjaśnił, że celem tej metody jest imputacja (oszacowanie, przypisanie) cen reprezentantów, gdy:

- znajdują się w koszyku, lecz nie można ustalić notowań ich cen (dobra „stare”, wycofywane z obrotu rynkowego),
- są dobrami nowymi, dopiero co pojawiającymi się w koszyku (takie, dla których wcześniej nie można było ustalić notowań cen, gdyż nie występowały na rynku),
- następuje zamiana reprezentanta „starego” na „nowy”.

Inaczej mówiąc, w metodzie tej za pomocą modelu regresji hedonicznej imputowane są ceny wybranych reprezentantów, dla których nie można ustalić ceny rzeczywistej w jednym z rozpatrywanych okresów, a które składają się na koszyk reprezentantów. Od imputacji czy inaczej przypisania cen wzięła się nazwa metody. Niektórzy badacze nazywają tę metodę metodą złożoną lub hybrydową, dla pokreślenia, że funkcja hedoniczna jest tu wykorzystywana tylko do wyznaczenia szacunkowych cen niektórych reprezentantów, a dynamika po-

zostałych jest wyznaczana konwencjonalnie — przez porównanie ich cen rzeczywistych.

Założmy, że w okresie $t+1$ pojawia się nowy reprezentant dobra, oznaczmy go indeksem n . Dobra tego nie było w koszyku w poprzednim okresie t , zatem, aby móc wyznaczyć dynamikę jego ceny, należy cenę rzeczywistą z okresu $t+1$ porównać z szacunkową ceną tego dobra z okresu t . Hedoniczny model cen (2), oparty na danych z okresu t , posłuży do imputacji brakującej ceny. Cenę imputowaną dobra n zapiszemy zatem następującym wzorem⁶:

$$\hat{p}_{n,t} = \exp\left(a_{0,t} + \sum_j a_{j,t} z_{n,j,t+1}\right) \quad (3)$$

Wówczas wskaźnik zmiany cen dobra n w okresie $t+1$ w porównaniu do okresu t wynosi:

$$Index_n^{(1)} = \frac{p_{n,t+1}}{\hat{p}_{n,t}} \quad (4)$$

Dla przypadku, gdy w okresie $t+1$ dla innego dobra m , będącego w koszyku reprezentantów w poprzednim okresie t , nie możemy ustalić ceny rzeczywistej (np. gdy zostało ono wycofane z obrotu), metodę imputacji stosujemy analogicznie do pierwszego przykładu. Funkcję regresji hedonicznej szacujemy na podstawie danych pochodzących z okresu $t+1$ i wyznaczamy wartość tej funkcji dla ilości cech reprezentanta m . Szacowanie regresji na danych pochodzących z okresu bieżącego niesie problem praktyczny, a mianowicie późniejszą publikację indeksu, ze względu na czas poświęcony przygotowaniu nowego modelu. Dla zachowania symetrii informacji indeks agregatowy powinien zawierać szacowane zmiany cen reprezentantów pojawiających się i znikających z rynku.

Algorytm ten stosujemy analogicznie w przypadku, gdy nowe dobro n ma zastąpić w koszyku reprezentantów stare dobro m , o nieco odmiennej jakości.

Agregatowy wskaźnik cen tworzymy wykorzystując indeksy indywidualne, przy czym dla niektórych reprezentantów są to indeksy opisane wzorem (4), a dla pozostałych — wskaźniki cen rzeczywistych. Należy oczywiście pamiętać, że jeśli dla symetryczności indeksu włączane są zarówno ceny modeli n , jak i m , to wówczas należy zwrócić uwagę na odpowiednie rozdysponowanie wag między nimi.

Ciekawe rozwiązanie, tzw. indeks z podwójną imputacją (*double imputation* lub *complete hybrid*) proponują Silver i Heravi oraz Pakes. W indeksie tym dla dobra n lub m ceny z obydwu okresów t i $t+1$ (okresu występowania danego

⁶ Przy czym Triplett (2006) proponuje dodawać do szacunkowej ceny mieszkania wartość połowy wariancji resztowej, w celach korekty obciążenia związanego z logarytmiczną postacią funkcyjną modelu. To samo rozwiązanie proponuje ILO (2004, s. 126).

modelu i okresu, w którym go brakuje) zastępuje się cenami imputowanymi. Indeks ten można zatem zapisać wzorem:

$$Index_n^{(2)} = \frac{\hat{p}_{n,t+1}}{\hat{p}_{n,t}} \quad (5)$$

Uzasadnieniem dla takiej formuły jest interpretacja składnika losowego funkcji hedonicznej. Jedni interpretują składnik losowy jako miarę występującego na rynku przewartościowania lub niedowartościowania ceny dobra względem jego wartości dla konsumenta. Różnice ceny i wartości wynikają z różnic w preferencjach nabywców. Przy takiej interpretacji stosowanie metody z podwójną imputacją nie jest poprawne. Inni, jak Silver, Heravi czy Pakes twierdzą, że składnik losowy odzwierciedla głównie błąd specyfikacji funkcji hedonicznej, a więc zawiera w sobie informację o, pominiętych w równaniu regresji, istotnych zmiennych cenotwórczych. Tym uzasadniają słuszność stosowania metody podwójnej imputacji cen. Zachęcają do korzystania z formuły podwójnej imputacji tam, gdzie specyfikacja funkcji regresji jest niewystarczająca. Triplett ma przeciwstawne stanowisko. Uważa, że błąd specyfikacji modelu powinien być eliminowany poprawną specyfikacją.

Mechanizm dostosowania wskaźników zmian cen do zmian jakościowych dóbr wykorzystywany w metodzie imputacji (lub podwójnej imputacji) najprościej wyjaśnić na przykładzie sytuacji, gdy w okresie $t+1$ dobro n zastępuje w koszyku dobro m . Wówczas, dzięki zastosowaniu metody imputacji, wskaźnik zmiany ceny dobra n opiera się na tym samym zestawie cech z_n . Gdybyśmy nie stosowali metody imputacji, wówczas cena dobra n zostałaby porównana z ceną dobra m . Dobro n może być substytutem dobra m , ale wektory z_n i z_m będą zawsze różne.

Metoda hedonicznego dopasowania jakości

Metoda hedonicznego dopasowania jakości jest podobna do metody imputacji. Model hedoniczny służy do wyznaczenia składnika korygującego zmiany jakościowe dobra. Następnie, przy pomocy tego wskaźnika koryguje się wskaźnik zmiany cen.

Oznaczmy wskaźnik korekty jakościowej przez $A(h)$. Załóżmy, że dobro n zastępuje w okresie $t+1$ dobro m , znajdujące się w koszyku w okresie t . Dobro nowe n różni się od dobra starego m i te różnice możemy przedstawić właśnie za pomocą składnika korygującego $A(h)$. Wykorzystując oznaczenia użyte w równaniu (2), Triplett proponuje następującą wartość składnika korygującego jakość dobra n :

$$A(h) = \exp \left(\sum_j a_{j,t} \frac{z_{n,j,t+1}}{z_{m,j,t}} \right) \quad (6)$$

ILO podaje nieznacznie różniący się składnik korygujący:

$$A(h) = \exp\left(\sum_j a_{j,t}(z_{n,j,t+1} - z_{m,j,t})\right) \quad (7)$$

Skorygowaną o różnice jakościowe cenę dobra n wyznaczamy następnie ze wzoru:

$$\hat{p}_{n,t} = p_{m,t} A(h) \quad (8)$$

Korzystając z tego równania obliczymy hedoniczny wskaźnik zmiany ceny dobra n w okresie $t+1$:

$$Index_n^{(3)} = \frac{p_{n,t+1}}{\hat{p}_{n,t}} \quad (9)$$

Analogicznie do metody imputacji, wyliczoną za pomocą równania (9) dynamikę ceny zmieniającego się dobra podstawiamy następnie do indeksu agregatowego.

Addytywną korektę jakości proponują Baaschera i LaCroixa (za: Triplett). Wówczas indeks wyznaczamy ze wzoru (9), a cenę imputowaną dobra n zapisujemy następująco:

$$\hat{p}_{n,t} = p_{m,t} + \left[\exp\left(\sum_j a_{j,t} z_{n,j,t+1}\right) - \exp\left(\sum_j a_{j,t} z_{m,j,t}\right) \right] \quad (10)$$

Przedstawiony opis metod dotyczy przypadku, gdy nowe dobro n sprawdzamy do porównywalności z dobrem starym m . Oczywiście algorytm ten stosuje się także w drugą stronę, czyli gdy pozostawiamy w koszyku stare dobro m , lecz jego cena jest ceną dobra n . Jednak ze względów praktycznych stosowane jest głównie podejście pierwsze. Wystarcza bowiem oszacować równanie regresji dla okresu bazowego t , a następnie stosować je dla wielu okresów przyszłych. Wskazane jest przy tym sprawdzanie zmian, jakie mogą zajść w modelu regresji poprzez szacowanie modeli hedonicznych dla nowych okresów i porównywanie wyników z modelu referencyjnego, a w razie znacznych różnic — zmiana tego modelu.

Triplett przeprowadza dowód formalny na to, że metoda hedonicznego dopasowania jakości nie różni się niczym od pośredniej metody imputacji wówczas, gdy regresje hedoniczne w obydwu przypadkach szacowane są na tym samym zbiorze danych. Z reguły jednak, w odróżnieniu od pozostałych metod, w metodzie dopasowania jakościowego model regresji jest szacowany na danych pochodzących z innych okresów lub na danych o innej długości okresu, a nawet na

danych pochodzących z innych badań. Tylko w tym podejściu możliwe jest oszacowanie funkcji hedonicznej i wyznaczenie indeksu uwzględniającego dwa różne zbiory danych. Jest to źródłem praktycznych korzyści i dlatego też metoda ta jest najczęściej wykorzystywana przez urzędy statystyczne.

Metody bezpośrednie

Metody bezpośrednie mogą być stosowane dla indeksów opartych na stałym koszyku reprezentantów lub mogą być stosowane dla wszystkich dostępnych danych z danego okresu. W drugim przypadku stają się jednocześnie sposobem wyznaczania indeksu agregatowego. W przypadku rynków produktów niejednorodnych (dla których ustalenie stałego, tradycyjnego koszyka reprezentantów jest praktycznie niemożliwe) często wykorzystuje się metody bezpośrednie do wyznaczenia indeksu opartego na specyficznym rodzaju koszyka reprezentantów.

Metoda ze zmiennymi zero-jedynkowymi czasu (ZZC)

Ogólnie postać modelu ze zmiennymi zero-jedynkowymi czasu (ZZC) można zapisać następującym równaniem:

$$\ln p_{i,t} = a_0 + \sum_{j=1}^J a_j z_{i,j,t} + \sum_{\tau=2}^T b_{\tau} D_{\tau} + \varepsilon_{i,t} \quad (11)$$

gdzie:

D — zmienna zero-jedynkowa (0-1) czasu; $D_{\tau} = 1$, gdy obserwacja pochodzi z okresu $\tau = t$ oraz $D_{\tau} = 0$, gdy $\tau \neq t$ (np. $D_2 = 1$ dla obserwacji z okresu $t = 2$ i $D_2 = 0$ dla obserwacji z innych okresów),

J — liczba cech dobra uwzględniona w równaniu,

T — liczba okresów, z których pochodzą obserwacje uwzględnione w równaniu.

Metodę tę stosuje się w dwóch różnych wariantach, jako:

- sąsiadujących okresów (*adjacent period approach*),
- wszystkich okresów (*pooled lub multi-pooled regression approach*).

W pierwszym wariantcie model wyznaczany jest dla danych z dwóch sąsiadujących ze sobą okresów i zawiera tylko jedną zmienną 0-1 czasu. W wariantcie drugim model regresji szacowany jest dla więcej niż dwóch następujących po sobie okresów i zawiera o jeden mniej niż liczba okresów zmiennych 0-1 czasu. Jak wynika z własności modelu semilogarytmicznego (np. Gruszczyński, Podgórska, 1996), wartość pierwszej pochodnej logarytmu zmiennej objaśnianej po zmiennej objaśniającej czasu reprezentuje średnią stopę wzrostu zmiennej objaśnianej. Dlatego dla niewielkiej wartości dynamiki w czasie można założyć, że wartość b_{τ} jest średnią stopą wzrostu ceny mieszkania. Przy większej dynamice

należy korzystać z dokładnego wzoru uzyskanego przez policzenie wartości funkcji eksponentialnej po obydwu stronach równania. Wówczas w wariancie modelu dla dwóch sąsiadujących ze sobą okresów, $\tau = 1$ i $\tau = 2$, gdzie $\tau = 1$ jest okresem bazowym (zgodnie ze wzorem (11) w modelu pominięta jest zmienna D_1), dokładna wartość wskaźnika dynamiki cen (indeks hedoniczny) określona jest wzorem:

$$Index = \exp(b_2) \quad (12)$$

Indeks ten jest skorygowanym jakościowo indeksem hedonicznym cen dobra. Wartość indeksu pokazuje, jak zmienił się ogólny poziom cen po uwzględnieniu innych zmiennych cenotwórczych dobra — jego charakterystyk z. Dekompozycja ceny na zmienne czasowe i inne cenotwórcze zmienne, jakimi są cechy dobra, a także interpretacja współczynników regresji *ceteris paribus* sprawiają, że wyrażenie z równania (12) interpretujemy jako wskaźnik „czystej” zmiany cen, a więc niezależnej od efektów różnic jakościowych badanych dóbr.

Analogicznie, jeśli model (11) szacujemy na podstawie danych z kilku (dwóch lub więcej) okresów, poszczególne wskaźniki dynamiki cen są wartościami funkcji eksponentialnej dla oszacowanych wartości współczynników b_τ .

W metodzie ZZC wszystkich okresów współczynniki regresji a_j odzwierciedlają przeciętne wartości wycen rynkowych poszczególnych cech dobra pochodzące ze wszystkich rozpatrywanych okresów. Sprowadza się to do przyjęcia założenia, że w szacowanych okresach wyceny rynkowe poszczególnych cech dobra są stałe. Założenie to jest podstawowym źródłem krytyki metody ZZC (Conniffe, Duffy, 1999 i in.).

Metoda cen charakterystyk

Odmienne do metody ZZC, w tej metodzie przyjmowane jest założenie, że ceny implikowane charakterystyk dobra są zmienne z okresu na okres. Zwolennicy tego podejścia twierdzą, że zamiana ceny całkowitej dobra jest wywołana zmianą cen implikowanych (wyceny charakterystyk dobra, a_j). Na przykład pod wpływem mody nabywcy mieszkań zaczynają przedkładać aneksy kuchenne nad kuchnie zamknięte. Wówczas cena implikowana zmiennej w wariancie „kuchnia zamknięta” będzie maleć, co może doprowadzić do spadku ceny całkowitej mieszkania (*ceteris paribus*). Inny przykład dotyczy komputerów — malejące ceny pamięci komputerowej powodują obniżkę cen samych komputerów. Według tego założenia w tej metodzie wykorzystuje się ceny implikowane w ważonej, konwencjonalnej formule indeksu cenowego.

Dla porządku rozwinę kwestię nazewnictwa tej metody. I tak Griliches, który jako pierwszy podał ideę tworzenia indeksów hedonicznych, nazywa tę metodę indeksem cen charakterystyki (*price-of-characteristics index*, przez innych auto-

rów określaną też jako *characteristic price index*, *price index for characteristics*, *hedonic index of characteristics prices*). Inne spotykane w literaturze anglojęzycznej nazwy tej metody to: bezpośrednia metoda charakterystyk (*direct characteristics method*), alternatywna bezpośrednia metoda zmiany cen (*alternative direct measure of the price change*) czy metoda pojedynczego okresu (*single period method*). Ze względu na teoretyczne i statystyczne własności towarzyszące tej metodzie w literaturze indeks wyliczony tą metodą określa się jako superlatywny i idealny indeks hedoniczny (*superlative and exact hedonic index*). Silver i in. (2003, 2006) metodę tę określili bardziej ogólnie jako metodę imputacji (*hedonic imputation*) i rzeczywiście może ona być uznana za szczególny przypadek metody imputacji.

Uporządkowanie nazw jest ważną częścią procesu klasyfikowania metod hedonicznych, pozwala uniknąć rozbieżności i niezrozumienia wśród badaczy, a wreszcie ułatwia dalsze studia tematu. Do określenia tej metody będę stosowała nazwę zaproponowaną wcześniej — metoda cen charakterystyk. Metoda ta wymaga, aby dla każdego z badanych okresów szacowany był odrębny model regresji hedonicznej zapisany wzorem (2). Interpretacja współczynników regresji stojących przy poszczególnych zmiennych objaśniających równania (2) pozwala na zbudowanie indeksu cen wykorzystującego tradycyjne formuły Laspeyresa i Paasche’ego (w konsekwencji także superlatywne indeksy Fishera, Tornqvista, Walsha). Wówczas jest to agregatowy indeks cech badanego dobra, a jego wagami są reprezentujące rynek ilości tych cech. Korzystając z modelu zapisanego równaniem (2) wskaźnik zmiany cen między okresem t i $t+1$ zapiszemy jako:

index cen charakterystyk typu Laspeyresa:

$$Index_L = \frac{\exp \sum_j a_{j,t+1} q_{j,t}}{\exp \sum_j a_{j,t} q_{j,t}} \quad (13)$$

index cen charakterystyk typu Paasche’ego:

$$Index_P = \frac{\exp \sum_j a_{j,t+1} q_{j,t+1}}{\exp \sum_j a_{j,t} q_{j,t+1}} \quad (14)$$

gdzie q_j — waga ilościowa j -tej cechy; q_j może być sumą ilości cech dla poszczególnych i -tych jednostek produktu lub może wyrażać średnią ilość danej cechy w badanej populacji. Na przykład, jeśli badamy indeks zmiany cen dwóch mieszkań o powierzchni 30 m² i 40 m², to wówczas q_1 może wynosić 70 m² lub 35 m². Dla zmiennych typu jakościowego (np. rodzaj kuchni) wagi odzwierciedlają udział danej cechy w badanej próbie.

Stosując tę formułę przyjmujemy, że nabywcy kupując np. mieszkanie płacą tak naprawdę za pewien zbiór właściwości tego mieszkania, a nie za samo mieszkanie. Kupują oni pewien określony zbiór ilości (wariantów) cech mieszkania: 60 m² powierzchni użytkowej, jasną kuchnię, balkon, centralną lokalizację budynku, dobrą komunikację z centrum, bliskość szkół itp.

Metoda imputacji (bezpośrednia)

Metoda ta prezentowana jest przez ILO. Korzystając z równania (2) wyznacza się model regresji hedonicznej na danych z okresu $t+1$ (lub t w zależności od tego, czy wyznaczamy indeks o stałym czy bieżącym okresie referencyjnym), a następnie poprzez podstawienie do oszacowanego równania ilości charakterystyk dóbr z okresu t ($t+1$) imputuje się tym dobrom ceny, które korespondują z okresem $t+1$ (t). Oszacowane w ten sposób ceny dóbr określa się jako ceny odbite. Nazwę nasuwa interpretacja tak wyznaczonych cen — są one prawdopodobnymi (przypisanymi) cenami z okresu $t+1$ (t) dóbr sprzedanych w okresie t ($t+1$). Indeks wyznaczony za pomocą tej metody zapisujemy poniższymi wzorami:

$$Index_{LA} = \frac{\sum_{i=1}^N \frac{1}{N} \hat{p}_{i,t+1}(z_{i,t})}{\sum_{i=1}^N \frac{1}{N} p_{i,t}(z_{i,t})} \quad (15)$$

lub

$$Index_{PA} = \frac{\sum_{i=1}^N \frac{1}{N} p_{i,t+1}(z_{i,t+1})}{\sum_{i=1}^N \frac{1}{N} \hat{p}_{i,t}(z_{i,t+1})} \quad (16)$$

gdzie:

$\hat{p}(z_i)$ — cena odbita dobra i , będąca funkcją wektora jego charakterystyk z_i ,

$p(z_i)$ — rzeczywista cena obserwowana dobra i o wektorze charakterystyk z_i .

Na przykład $\hat{p}_{i,t+1}(z_{i,t})$ oznacza cenę odbitą wyliczoną przez podstawienie do równania regresji hedonicznej, oszacowanej na próbie danych z okresu $t+1$, charakterystyk z_i dóbr sprzedanych w okresie t . Korzystając z przekształconego równania (2) cenę odbitą dobra i w okresie $t+1$ możemy wyznaczyć ze wzoru:

$$\hat{p}_{i,t+1} = \exp \left(a_{0,t+1} + \sum_j a_{j,t+1} z_{i,j,t} \right) \quad (17)$$

Jak podaje ILO, indeks może być oparty na średnich cenach rzeczywistych i odbitych (wzory (15) i (16)) lub może być średnią z indywidualnych indeksów cen badanych dóbr, co zapisujemy jako:

$$Index_{LB} = \frac{1}{N} \sum_{i=1}^N \frac{\hat{p}_{i,t+1}(z_{i,t})}{p_{i,t}(z_{i,t})} \quad (18)$$

lub

$$Index_{PB} = \frac{1}{N} \sum_{i=1}^N \frac{p_{i,t+1}(z_{i,t+1})}{\hat{p}_{i,t}(z_{i,t+1})} \quad (19)$$

Metoda ta, podobnie jak dwie pierwsze z grupy metod bezpośrednich, stosowana jest zarówno dla wszystkich obserwacji z danego okresu (całej badanej populacji), jak i dla ustalonego koszyka reprezentantów. Bezpośrednia metoda imputacji różnić się będzie wówczas od metody pośredniej tym, że jest wyznaczana dla wszystkich reprezentantów, a nie tylko dla tych dóbr, które wypadły lub/i pojawiły się w koszyku. W bezpośredniej metodzie imputacji kontrola zmian jakości dóbr w porównywanych okresach odbywa się przez porównanie ceny rzeczywistej i odbitej wyznaczonej dla tego samego wektora charakterystyk z_i . Indeksy przedstawione wzorami (15)–(19) obrazują zatem „czystą” (dla tych samych wektorów z_i , a więc bez różnic jakościowych) zmianę ceny między okresami t i $t+1$. Podstawowa różnica między zastosowaniem równań (15), (16) i (18), (19) polega na tym, że w pierwszym podejściu model regresji hedonicznej musi być wyznaczany na danych z okresu bieżącego, a w drugim — na danych z okresu bazowego. Jeśli przyjmiemy jeden okres bazowy, wówczas stosuje się wzory (16) i (19), wyznaczając jeden model regresji i wykorzystując go dla bieżących okresów.

PORÓWNANIE METOD — WYBRANE ASPEKTY EMPIRYCZNE I TEORETYCZNE

Hedoniczne indeksy cen są z powodzeniem stosowane na świecie począwszy od lat 60. XX w. Metody pośrednie wykorzystywane do korekty jakości zmian koszyka stosuje się w wielu przypadkach w statystyce publicznej zarówno w indeksach CPI, jak i PPI. Metody bezpośrednie dedykowane są konstrukcji indeksów sektorowych, głównie dla rynków dóbr wysoce heterogenicznych i szybko zmieniających się, np. rynek mieszkań oraz rynek produktów IT. Mimo popularyzacji metod hedonicznych nadal istnieje wiele nierozwiązanych i spornych kwestii natury teoretycznej i empirycznej związanych z ich stosowaniem. Powstają prace badawcze, a zakres problemów, których dotyczą jest bardzo szeroki. Na podstawie doświadczeń z badań własnych oraz przeglądu literatury

jedynie zasygnalizowano, z jakiego typu problemami przyjdzie zmierzyć się badaczom podejmującym pracę nad hedonicznymi indeksami cen. Wiele z rozważanych problemów nie zostało do końca rozwiązanych i stanowi otwarte pole badawcze.

DOSTĘPNOŚĆ DANYCH

Podstawową trudnością w konstrukcji indeksów hedonicznych jest dostępność danych. Oszacowanie dobrej jakości modelu regresji wymaga bowiem dużych zbiorów obserwacji. Zbiory te powinny zawierać dokładną informację o cechach badanych dóbr, a dokładniej, o wszystkich cechach, które determinują cenę dobra. Na przykład modele cen mieszkań, poza fizycznymi charakterystykami samego mieszkania (pow. użytkowa, rodzaj kuchni, dodatkowe pomieszczenia itp.), powinny zawierać tzw. zmienne lokalizacyjne (np.: odległość od ośrodków edukacyjno-kulturowych, terenów zielonych, centrów handlowych, centrów dużych miast, wskaźnik struktury społecznej, infrastruktura drogowa i wiele innych). Dopiero kompletna informacja o cechach fizycznych nieruchomości i cechach jej lokalizacji może odzwierciedlić rzeczywistość kształtowania się cen i umożliwić pełną specyfikację modelu⁷. Oczywiście podstawą szacowania i weryfikacji modeli hedonicznych jest wiedza ekonomiczna. Konwencjonalna metoda badania dynamiki cen ustalonych reprezentantów, w przeciwieństwie do metod wykorzystujących modele hedoniczne, nie stawia tak wysokich wymagań co do zasobów danych i wiedzy. W praktyce dostęp do odpowiednich zbiorów danych decyduje o wyborze metody. Jak podaje Triplett, spośród metod pośrednich urzędy statystyczne najczęściej wybierają metodę hedonicznego dopasowania jakości, gdyż w przeciwieństwie do pozostałych umożliwia ona oszacowanie modelu na innym zbiorze danych niż zbiór uwzględniony w konstrukcji indeksu. Dostępność wiarygodnych danych jest jedną z podstawowych barier w stosowaniu metod hedonicznych i z pewnością jeszcze długo nią pozostanie.

Metoda reprezentantów a metody hedoniczne

Bezpośrednie metody wyznaczania indeksów hedonicznych powstały jako alternatywa dla konwencjonalnej metody reprezentantów i dlatego literatura skupiająca się na analizie różnic między tymi metodami jest dość obszerna. Podnoszone kwestie dotyczą teorii, zastosowań i analizy wyników. Jednocześnie porównania przeprowadzane są zarówno pomiędzy metodami koszykowymi z konwencjonalnymi metodami kontroli zmian jakościowych⁸, jak i korektami hedonicznymi (metody pośrednie) oraz z metodami bezpośrednimi, głównie metodą ZZC lub

⁷ Więcej na temat dostępności i rodzaju danych wykorzystywanych w hedonicznym modelowaniu cen mieszkań można znaleźć w pracach Łaszek, Widłak (2007, 2008) oraz Tomczyk, Widłak (2010).

⁸ Na przykład porównanie bezpośrednie, metody typu overlap, oceny ekspertów i inne.

metodą cen charakterystyk. Analizowane są także różnice zastosowań modeli bezpośrednich dla reprezentantów i dla danych niedopasowanych.

Najczęściej przedmiotem porównań są różnice formalne, różny sposób ważenia danych i dopasowania jakościowego oraz aspekty praktyczne (takie jak dostępność danych) i wreszcie różnice wyników (gdy indeksy oparte są na tych samych próbach danych, np. Aizcorbe, Pho (2005)).

W kwestii formalnej częstym pytaniem jest, czy hedoniczny indeks cen w formule metody ZZC jest zgodny z teorią indeksów statystycznych⁹? Teoria ta podaje, że statystyczny indeks cenowy powinien być wyznaczany na podstawie jednej z określonych formuł: Laspeyresa, Paasche'ego, Fishera lub ich dalszych modyfikacji.

W pewnych warunkach odpowiedź na to pytanie jest pozytywna i zależy przede wszystkim od postaci funkcyjnej regresji hedonicznej. Triplett przeprowadza formalny dowód. Zakładając, że funkcja hedoniczna dana równaniem (11) zawiera tylko jedną zmienną zero-jedynkową czasu, przekształca on estymator metody najmniejszych kwadratów (MNK) dla współczynnika b_2 stojącego przy tej zmiennej. Przekształcenie to pokazuje, że indeks ten jest geometryczną średnią cen dobra z okresu t i $t+1$ podzieloną przez wyrażenie dotyczące dopasowania jakościowego. Jeśli dodatkowo estymator regresji będzie wyznaczony WMNK¹⁰, to wówczas indeks ten będzie odpowiednio ważoną średnią geometryczną cen z okresu t i $t+1$ podzieloną przez ważne wyrażenie dopasowania jakościowego. Dla liniowej funkcji hedonicznej formuła indeksu po wyprowadzeniu okazuje się arytmetyczną średnią cen z okresu badanego i od której odejmowane jest liniowe wyrażenie *quality-adjustment*, będące liniowym indeksem ilości wyznaczonym dla charakterystyk dobra. Podobnie dla innych postaci funkcji regresji hedonicznej uzyskujemy inne postacie konwencjonalnych formuł indeksów cenowych.

Wykorzystując to wyprowadzenie Triplett pokazuje dalej, że indeks z metody ZZC (postać log-liniowa), szacowanej dla ustalonego koszyka reprezentantów, jest równoważny co do wartości indeksowi cen ze zwykłej metody reprezentantów w formule nieważonego indeksu geometrycznego. To samo można pokazać dla innej niż geometryczna formuły indeksu i innej niż liniowa zależności funkcyjnej ujętej w modelu regresji hedonicznej. Zatem, gdy bierzemy pod uwagę dokładnie tych samych reprezentantów (koszyk pozostaje stały w okresie bazowym i badanym), wyniki indeksu hedonicznego i metod konwencjonalnych są takie same. Stąd wniosek, że metody hedoniczne będą znajdowały zastosowanie przede wszystkim w przypadku tych produktów, dla których trudno jest wyznaczyć stały koszyk reprezentantów.

⁹ Index Number Theory — teoria indeksów statystycznych, określająca statystyczne i matematyczne własności wskaźników. Podsumowania tych własności dokonał E. Diewert. Zaproponował on także dziewięć testów sprawdzających poszczególne własności. Dobry indeks statystyczny powinien spełniać te warunki.

¹⁰ Ważona MNK; wagami są liczba lub wartość transakcji (*sales-weighted regression*, Triplett, 2006, s. 56).

Ciekawe wyniki dają badania szacujące wpływ wyboru metody kontroli jakości na wartość indeksu. Triplett i McDonald (Triplett, 2006) przeprowadzili badania indeksów cen lodówek. Badania te polegały na dekompozycji różnicy wartości pomiędzy indeksem wyznaczonym konwencjonalnie jako średnia arytmetyczna cen oraz hedonicznym indeksem cen w formule ZZC. Autorzy podzielili tę różnicę na część wynikającą z różnicy zastosowanej formuły (arytmetyczna a geometryczna) oraz część wynikającą z odmiennego podejścia do kontroli zmian jakościowych (metoda reprezentantów a model hedoniczny). Wyniki analizy pokazały, że tylko 10% całkowitej różnicy między wskaźnikami było spowodowane odmienną formułą indeksu, a pozostałe 90% wynikało z odmiennego sposobu dopasowania jakości.

Inną ważną różnicą pomiędzy metodą reprezentantów i hedonicznym indeksem cen w formule ZZC jest zakres wykorzystania dostępnych danych i informacji w nich zawartych.

Hedoniczny indeks cen w formule ZZC, w przeciwieństwie do metody reprezentantów z reguły, opiera się na wszystkich danych dostępnych dla danego okresu. Dlatego uważa się, że dokładniej informuje on o rzeczywistych procesach cenowych. Można powiedzieć, że z biegiem czasu szereg wartości indeksu wyznaczonego z zero-jedynkowej zmiennej czasu nie dezaktualizuje się, gdyż model uwzględnia zmienność struktury produktu na rynku i w tym sensie jest bardziej aktualny od indeksów opartych na stałych koszykach reprezentantów.

Trzeba pamiętać, że wszystkie zalety indeksu hedonicznego wyznaczanego za pomocą którejkolwiek z metod bezpośrednich zachodzą, gdy poprawnie wyspecyfikowany został model hedoniczny. Funkcja hedoniczna przez błędy specyfikacji lub brak odpowiednich danych może pomijać ważne czynniki wpływające na cenę, a niebędące cechami samego produktu (np.: w przypadku komputerów może to być serwis sklepu, w przypadku mieszkań — ocena firmy deweloperskiej lub pośrednika, jeśli mowa o mieszkaniu z rynku wtórnego). Ponadto, w modelach regresji występuje także błąd szacunku współczynników regresji. Potencjalne błędy specyfikacji modelu oraz błędy szacunku cen implikowanych mogą przemawiać na niekorzyść indeksów hedonicznych wobec metod konwencjonalnych. Kwestia ta, łącznie z trudnością uzyskania danych, sprawia, że bardzo często instytuty statystyczne decydują się pozostać przy metodach koszykowych, a funkcje hedoniczne wykorzystują do rozwiązywania zastępowalności reprezentantów wypadających z koszyka (metody pośrednie). Niestety, w przypadku niektórych rynków zastosowanie metod koszykowych jest praktycznie niemożliwe.

Metoda ZZC a metoda cen charakterystyk

Spośród metod bezpośrednich najczęściej stosuje się metodę ZZC oraz metodę cen charakterystyk. Większość kluczowych zagadnień dotyczących metod bezpośrednich odnosi się do porównania tych dwóch metod.

Wielu teoretyków argumentuje, że indeks cen charakterystyk z kilku powodów jest bardziej właściwy od hedonicznego indeksu cen z metody ZZC. Po pierwsze metoda cen charakterystyk ma zdecydowaną przewagę w łatwości nadania interpretacji ekonomicznej. Indeks wyznaczony tą metodą pokazuje czystą zmianę cen dobra zobrazowaną zmianami cen implikowanych w danym okresie do cen implikowanych z okresu referencyjnego. Po drugie, w myśl teorii indeksów statystycznych, indeks cen charakterystyki jest bliski wskaźnikowi idealnemu¹¹ (*exact, ideal*). Indeksy cenowe typu Laspeyresa, najczęściej używane w indeksach CPI, są uznawane za dobre przybliżenie i górną granicę idealnych wskaźników cen (za dolną granicę idealnych wskaźników cen uznaje się indeksy typu Paasche'ego). Dodatkowo, używając metody cen charakterystyk można skonstruować indeksy superlatywne, czyli takie, które nie wyróżniają żadnego z dwóch okresów — bazowego i bieżącego. Zgodność z teorią statystycznych indeksów cen oraz możliwość nadania interpretacji ekonomicznej hedonicznego indeksu cen charakterystyk są wymieniane jako jego główne zalety w porównaniu do metody ZZC.

Kolejną zaletą metody cen charakterystyk jest możliwość zastosowania jej także do konstrukcji wskaźników ilości (Józwiak, Podgórski, 2000) w odniesieniu do charakterystyk badanych dóbr. Griliches (1961) wykorzystał stworzone w ten sposób wskaźniki zmian ilości charakterystyk do korekty zmian jakościowych koszyka dóbr. Podzielił on indeks cen ofertowych samochodów (składnik CPI) przez wskaźnik zmian ilości wyznaczony dla charakterystyk opisujących samochody, uzyskując w ten sposób skorygowany o zmianę jakości indeks cenowy.

Innym praktycznym argumentem przemawiającym za stosowaniem metody cen charakterystyk jest fakt, że w metodzie ZZC ostateczna formuła indeksu (arytmetyczna czy geometryczna) determinowana jest przez zależność funkcyjną zmiennych objaśnianej i objaśniających w modelu hedonicznym. Niektórzy statystycy chcieliby natomiast z góry narzucić formułę indeksu. Inaczej jest w przypadku indeksu cen charakterystyk, gdzie kształt formuły indeksu nie zależy od postaci zależności funkcyjnej ceny i atrybutów dobra.

Jednak najpoważniejszym zarzutem wobec metody ZZC pozostaje kwestia założenia dotyczącego stałości wycen atrybutów produktu we wszystkich uwzględnionych w modelu okresach. W rzeczywistości, zwłaszcza w przypadku niektórych produktów (np. komputerów), ceny poszczególnych składowych dobra zmieniają się dość dynamicznie z okresu na okres. Indeksy cen charakterystyk nie są przedmiotem tej krytyki, gdyż przyjmują odwrotne założenie.

¹¹ W myśl teorii ekonomii idealny indeks cenowy to taki, który odzwierciedla minimalną zmianę kosztów, jakie musi ponieść konsument, aby pomimo zmiany cen dóbr nie zmienił się jego poziom użyteczności czerpanej z konsumpcji. Ta teoretyczna konstrukcja, znana jako indeks kosztów utrzymania (ang. *Cost-of-Living Index*), została zaproponowana na początku XX w. przez rosyjskiego ekonomistę A. A. Konusa. Przyjmuje się, że obserwowalne indeksy cenowe typu Laspeyresa i Paasche'ego są odpowiednio górną i dolną granicą wartości teoretycznego, idealnego indeksu kosztów utrzymania.

Conniffe i Duffy opisują doświadczenia przy tworzeniu indeksu cen mieszkań (*ESRI Index*). Zwracają oni uwagę, że wybór pomiędzy metodami ZYC i metodą cen charakterystyk jest przede wszystkim związany z problemem stałości parametrów. Wybierając indeks cen charakterystyk przyjmujemy, że źródłem zmienności cen dobra między okresem t i $t+1$ jest nieobserwowalna bezpośrednio na rynku zmiana wycen jego parametrów. Stosując indeks z metody ZYC zakładamy, że wyceny cech dobra nie zmieniają się we wszystkich okresach wziętych do modelu. Jeśli zatem wyceny te w rzeczywistości zmieniły się (np. wskutek zmiennych gustów lub kosztów produkcji), to założenie o stałości wycen w modelu ZYC powoduje dodatkowe obciążenie składnika losowego. Z tych powodów Conniffe i Duffy nazywają metodę ZYC wszystkich okresów „metodą z ograniczeniami” (*constrained hedonic index*). Pewnym rozwiązaniem tego dylematu może być stosowanie metody ZYC sąsiadujących okresów w miejsce metody wszystkich okresów. Obszerne badania tego zagadnienia przeprowadzili także Silver i in. (2006) oraz Diewert i in. (2007).

Choć metodę ZYC krytykuje się za założenie stałości cen implikowanych, to, jak zauważa Triplett, krytycy nie sformalizowali dowodu będącego odpowiedzią na pytanie, czy założenie stałości współczynników regresji wpływa na wartość współczynnika stojącego przy zmiennej 0-1 czasu, a więc docelową wartość indeksu cenowego? Natomiast porównanie rezultatów badań empirycznych daje rozbieżne wnioski. Dla przykładu Triplett zestawiając wartości indeksu mierzono obydwiema metodami pokazuje, że powyższa krytyka ma uzasadnienie. Z pracy Conniffe i Duffy wynika, że liczenie indeksu metodą ZYC oraz metodą cen charakterystyk daje praktycznie identyczne rezultaty. W zastosowaniu empirycznym dla polskiego rynku mieszkaniowego Tomczyk i Wiślak (2010) dochodzą do podobnych wniosków, jak Conniffe i Duffy.

Bezpośrednia metoda imputacji a metoda cen charakterystyk

Blizsze spojrzenie na metodę cen charakterystyk ujawnia, że jest ona szczególnym przypadkiem metody imputacji. Przy odpowiednich przekształceniach, a zatem wyznaczając indeks o stałej podstawie, np. według wzoru (15) — za wektor charakterystyk przyjmując średnie ilości cech z ustalonego okresu bazowego oraz podstawiając zarówno do mianownika, jak i licznika ceny imputowane — otrzymamy indeks podany wzorem (13).

Podstawową różnicą pomiędzy metodą cen charakterystyk i metodą bezpośredniej imputacji jest sposób ważenia obserwacji czy inaczej sposób agregacji danych. W bezpośredniej metodzie imputacji, podobnie jak w metodzie ZYC, wagi dla każdej obserwacji są takie same. W metodzie cen charakterystyk przyjmuje się określony system wag i najczęściej jest nim średnia lub suma cech dla poszczególnych jednostek dobra. Z tej różnicy wypływa wniosek, że o ile metoda cen charakterystyk nie musi odnosić się do całej populacji (regresje mogą być szacowane na próbach, a wagi są ustalane jednorazowo i stałe), o tyle

dwie pozostałe metody bezpośrednie powinny być szacowane na całej populacji (wszystkie obserwacje) z danego okresu lub reprezentatywnej próbie obserwacji.

Podsumowanie

Przedstawione metody hedoniczne są podstawowymi sposobami korekty indeksów cenowych przy zmianach jakości badanych dóbr. Metody te powinny być stosowane szczególnie do badania dynamiki cen dóbr wysoce heterogenicznych. Dla tego typu dóbr trudno badać dynamikę cen wykorzystując tradycyjną metodę reprezentantów, gdyż w każdym z badanych okresów potencjalni reprezentanci znacznie różnią się od siebie pod względem jakości. Na rynkach tego typu produktów do konstrukcji indeksów cenowych wykorzystuje się przede wszystkim bezpośrednio metody hedoniczne. Metody pośrednie stosowane są jako dopełnienie metody koszykowej w określonych sytuacjach, w których zachodzi konieczność imputacji brakujących cen lub korekty zmian jakości wybranych reprezentantów.

Kryterium zaproponowanego w artykule podziału metod na pośrednie i bezpośrednie jest sposób i zakres wykorzystania informacji z modelu hedonicznego. W metodach bezpośrednich indeks nie mógłby zostać wyznaczony bez oszacowania modelu regresji, natomiast w metodach pośrednich informacja z modelu nie jest konieczna do wyznaczenia indeksu, pełni ona rolę dodatkową jako narzędzie dostosowań indeksu do zmian jakości lub uzupełnienia brakujących obserwacji. W metodach bezpośrednich indeks jest wyznaczany bezpośrednio w równaniu regresji hedonicznej lub bezpośrednio z tego równania brane są ceny implikowane charakterystyk podstawiane następnie do formuły wskaźnika cen (Laspeyresa bądź Paasche'ego). Metody bezpośrednie mogą, ale nie muszą, być stosowane dla indeksów wykorzystujących koszyk reprezentantów; metody pośrednie są wyłącznie dopełnieniem metody koszykowej.

Indeks hedoniczny z metody ZZC powstał jako alternatywa dla indeksu opartego na metodzie koszykowej. Jak pisze Triplett, brak jest formalnych różnic pomiędzy tymi indeksami wówczas, gdy badany jest ustalony zestaw reprezentantów. Stąd wniosek, że metoda ZZC będzie stosowana przede wszystkim na tych rynkach, na których ustalenie stałego koszyka reprezentantów wydaje się być prawie niemożliwe. Największą przeszkodą w stosowaniu metod pośrednich jest jednak dostępność danych, obciążenie wyników indeksu związane z błędami specyfikacji modeli hedonicznych, a w praktyce niechęć do stosowania bardziej wymagających niż konwencjonalne metod konstrukcji indeksów statystycznych.

Wskaźnikowi cen charakterystyk, w przeciwieństwie do wskaźnika z metody ZZC, łatwo jest nadać interpretację ekonomiczną. Ponadto, wykorzystując metodę cen charakterystyk można utworzyć indeksy superlatywne. W metodzie cen charakterystyk, przeciwnie niż w często stosowanej metodzie ZZC wszystkich

okresów, przyjmuje się założenie o zmienności wyceny cech badanych dóbr, co jest zgodne z teoretycznym modelem wyceny hedonicznej. Dobrym rozwiązaniem problemów związanych z założeniem stałości parametrów w metodzie ZZC jest stosowanie wariantu sąsiadujących okresów. Na szczególną uwagę zasługuje fakt, że w niektórych badaniach empirycznych wyniki indeksów z metody ZZC wszystkich i sąsiadujących okresów są zbieżne.

Dla rynku mieszkaniowego najczęściej stosowane są indeksy wyznaczone z metody ZZC oraz indeks cen charakterystyk. Metoda ta jest specyficzną formą metody koszykowej, śledzi ona zmianę cen koszyka cech dobra. Ceny reprezentantów (cech badanego dobra) nie są bezpośrednio obserwowane na rynku, lecz pochodzą z modelu hedonicznego (ceny implikowane). Bardzo ciekawym indeksem cen mieszkań, wykorzystującym metodę cen charakterystyk i dokładnie reprezentującym rynek poprzez zastosowanie jego stratyfikacji, jest indeks mieszanego dopasowania jakości. O indeksie tego typu można powiedzieć, że odzwierciedla on zmianę cen stałego koszyka reprezentującego rynek (koszyk nie ulega zmianom jakościowym — jest stały), a model hedoniczny służy do imputacji cen reprezentantów. W podejściu tym ważne jest prawidłowe ustalenie sztucznej reprezentacji rynku badanego dobra i podobnie jak w przypadku koszyków tradycyjnych reprezentacja ta powinna być rewidowana. Metoda mieszanego dopasowania umożliwia wyznaczenie dynamiki nawet wtedy, gdy w badanym okresie brak jest obserwacji reprezentujących dany segment rynku. Sytuacje te są charakterystyczne dla pierwotnego rynku nieruchomości.

mgr Marta Widlak — NBP, Instytut Ekonomiczny

LITERATURA

- Aizcorbe A., Pho Y. (2005), *Differences in Hedonic and Matched-Model Price Indexes: Do the weights matter?*, Department of Commerce, Bureau of Economic Analysis, Working Paper Series, WP2005-06, Washington
- Bover O., Izquierdo M. (2001), *Ajustados de calidad en los precios: metodos hedonicos y consecuencias para la contabilidad nacional*, „Economic Studies”, No. 70, Banco de Espana, Madryt
- Conniffe D., Duffy D. (1999), *Irish house price indices — methodological issues*, The Economic and Social Review 30
- Diewert E. (2003), *Hedonic Regressions. A Consumer Theory Approach*, [w:] C. R. Feenstra, M. D. Shapiro (red.), *Scanner Data and Price Indexes*, „NBER Book Series in Income and Wealth”, University of Chicago Press
- Diewert E. (2006), *Adjacent period dummy variable hedonic regressions and Bilateral Index Number Theory*, „Annales d'economie et de statistique”, nr 79/80, Paryż
- Diewert E. (2007), *Hedonic imputation versus time dummy hedonic indexes*, „Discussion Paper” No. 007-07, Department of Economics, University of British Columbia, Vancouver
- Diewert E., Heravi S., Silver M. (2007), *Hedonic Imputation versus Time Dummy Hedonic Indexes*, „IMF Working Paper”, No. 07/234

- Griliches (1961), *Hedonic Price Indexes for Automobiles: An Econometric of Quality Change*, [w:] *The Price Statistics of the Federal Government*, NBER, Chicago
- Griliches (1988), *Hedonic price indexes and the measurement of capital and productivity: some historical reflections*, „NBER Working Papers Series”, No. 2634, Cambridge
- Gruszczyński M., Podgórska M. (red.) (1996), *Ekonometria*, SGH, Warszawa
- ILO (2004), *Consumer Price Index Manual: Theory and Practice*, International Labour Organisation, Geneva
- Józwiak J., Podgórski J. (2000), *Statystyka od podstaw*, PWE, Warszawa
- Łaszek J. A., Widłak M. (2007), *Indeksy rynkowych cen mieszkań jako narzędzie badania rynku*, „Finansowanie Nieruchomości”, nr 3(12), Warszawa
- Łaszek J. A., Widłak M. (2008), *Badanie cen na rynku mieszkań prywatnych zamieszkałych przez właściciela z perspektywy banku centralnego*, „Bank i Kredyt”, nr 8, Warszawa
- Silver M., Heravi S. (2003), *The measurement of quality-adjusted price changes*, [w:] C. R. Feenstra, M. D. Shapiro (red.), *Scanner Data and Price Indexes*, „NBER Book Series in Income and Wealth”, University of Chicago Press
- Silver M., Heravi S. (2006), *The Difference Between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes*, „IMF Working Paper”, nr 6/181, Paryż
- Tomczyk E., Widłak M. (2010), *Konstrukcja i własności hedonicznego indeksu cen mieszkań dla Warszawy*, „Bank i Kredyt”, nr 1 (2010), Warszawa
- Triplett J. E. (2006), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, OECD, Paryż

SUMMARY

The aim of the article is to present and classify construction methods of hedonic price indexes being a way to control quality good changes in price indexes. Main differences between methods are pointed as well as conclusions drawn from their comparison. Theoretical literature and international empirical surveys are base of the study. The cited examples concern mainly hedonic price indexes of dwellings. Construction methods of hedonic price indexes may be classified as direct and indirect (by the classify criterion of using hedonic function). Direct methods are used for high heterogenic goods, where it is very difficult to fix basket of representatives. Indirect methods are used as additional to basket approach in the price dynamics survey.

РЕЗЮМЕ

Целью статьи является презентация и классификация методов конструкции гедонических показателей цен, которые считаются способом контроля изменений качества товаров в показателях цен. В статье представлены основные расхождения между методами и сформулированы выводы после их сопоставления. Основой для статьи была литература и международные эмпирические исследования. Названные примеры касаются главным образом гедонических показателей цен квартир.

Учитывая в качестве критерия разделения способ использования гедонической функции, методы конструкции гедонических показателей цен можно разделить на косвенные и прямые. На рынке товаров высоко гетерогенных, для которых трудно определить целевой набор (корзину) представителей, используются прямые методы. Косвенные методы используются в качестве дополнения корзинного подхода в обследовании динамики цен.