

Modelowanie i szacowanie popytu i substytucji produktów na podstawie mikrodanych

Tomasz Brzęczek*

Streszczenie

Rozważamy szeroki i długi asortyment, typowy dla konglomeratów produkcyjnych, ale szczególnie dla hurtowych oraz detalicznych przedsiębiorstw handlowych. Przedsiębiorstwo stosujące rynkową formułę ceny powinno brać pod uwagę zależności cenowo-popytowe, od których zależą sprzedaż i zysk. Zamiast złożonych rynkowych modeli systemu popytowego czy agregacji wyników z modeli teorii wyboru konsumenta proponujemy i dyskutujemy metodykę szacowania substytucji cenowej towarów bezpośrednio na podstawie grupowania mikrodanych o klientach przedsiębiorstwa. Atutem metodyki jest macierzowy model teoretyczny relacji popytowo-cenowych, w którym ceny są zmiennymi porządkowymi. Model można estymować na podstawie mikrodanych o zakupach klientów dobranych warstwowo-losowo spośród wszystkich klientów przedsiębiorstwa. Do estymacji i weryfikacji modelu zastosowano w artykule metody: wieloczynnikowej analizy wariancji i regresji Poissona. Dane dotyczą ilości sprzedaży produktów klientom wybranego przedsiębiorstwa oraz cen sprzedaży. Wyniki otrzymane obiema metodami są zgodne ze sobą i z teorią mikroekonomii. Weryfikacja założeń metod estymacji wskazała, że nie wszystkie z nich są spełnione w próbie badawczej.

Słowa kluczowe: popyt, substytucja, klient, analiza wariancji, regresja Poissona

Kody JEL: C25, D12

DOI: 10.17451/eko/44/2016/132

* Wydział Inżynierii Zarządzania, Politechnika Poznańska.

1. Wprowadzenie

Cena jest wynikiem gry popytu i podaży tylko na rynku konkurencji doskonałej, gdzie liczba dostawców i nabywców jest duża, a ich wielkości popytu i podaży są zbliżone. Założenia te są spełnione na niektórych rynkach. Natomiast na rynkach produktów przemysłowych zwykle każdy z producentów ma określoną pozycję i segment rynku oraz swoją strategię cenową. Podobnie w zakresie danego asortymentu w handlu hurtowym zwykle panuje oligopol, a w detalicznym – konkurencja monopolistyczna. Liczba linii produktowych i ich zróżnicowanie, czyli szerokość asortymentu, są zwykle określone podobieństwem technologicznym lub użytkowym produktów. Długość, czyli liczba produktów w linii, określona jest zróżnicowaniem gustów nabywców, których segment docelowy określa strategia firmy. W mikroekonomii poziom ceny ma maksymalizować zysk całkowity i wynika z zestawienia funkcji popytu oraz przychodu marginalnego z funkcją kosztów marginalnych. W praktyce zależności opisane przez te funkcje można obserwować jedynie historycznie i to z nieobserwowalnymi błędami losowymi. Dlatego planujący cenę, według marketingu, powinien stosować jedną z trzech formuł: opartą na bieżących lub prognozowanych kosztach jednostkowych i ustalonej marży, opartą na rynku i przewidywalnym poziomie popytu oraz jego elastyczności cenowej albo opartą na benchmarku względem bezpośrednich konkurentów i substytutów. Wszystkie czynniki podanych formuł w sposób dynamiczny podlegają interakcji i są brane pod uwagę w czasie podejmowania decyzji podażyowo-cenowej przez przedsiębiorstwo. Wpierw dostawca przesądza cenę w oparciu o swoje moce produkcyjne, wielkość produkcji i jej koszt przeciętny, a także marżę w relacji do funkcjonalności bezpośrednio konkurujących produktów, czyli na podstawie własnej analizy wartości. Przez bezpośrednio konkurujące produkty rozumiemy różne marki i modele tego samego produktu, czyli tzw. bliskie substytuty. Przez dalsze substytuty rozumiemy różne kategorie produktu o tej samej funkcji. Weryfikacja przez klientów polega na tym, iż decydują oni, czy nabyć produkt od danego dostawcy, znając ceny konkurujących marek i substytutów oraz swoje preferencje wobec nich, a także swój dochód. Proponowana w artykule oryginalna metodyka może dotyczyć zarówno substytucji bliższej, jak i dalszej. Natomiast prezentowane badanie dotyczy dalszej substytucji. W części literaturowej wymieniamy krótko modele i miary substytucji. Na popularności zyskują metody mikroekonometryczne, które umożliwiają symulację popytu indywidualnego klienta, a następnie agregację popytu klientów do wielkości sprzedaży towaru w zależności od jego ceny i ceny towarów substytucyjnych. W artykule stosujemy dwuczynnikową analizę wariancji nabywanych przez klientów ilości, a następnie model licznikowy regresji Poissona liczby sprzedanych produktów. Do merytorycznej weryfikacji szacunków proponujemy teoretyczny model popytu na dwa produkty i ich substytucji w punkcie 4. Jego sformułowanie i estymacja są celem naszych rozważań.

2. Przegląd literatury z zakresu zależności popytowych produktów

W zgodzie z teorią mikroekonomii przyjmijmy definicje zależności popytowej i substytucyjnej. Prawo popytu stwierdza, że nabywana ilość produktu jest malejącą funkcją jego ceny realnej. Definicja oparta jest na pojęciu ceny realnej, a więc implikuje zależność ilości sprzedaży od ceny nominalnej danego produktu (nazwijmy to efektem ceny danego produktu) oraz od cen nominalnych innych produktów. Stymulację popytu przez cenę innego produktu nazywamy substytucją cenową produktów. Wpływ ceny nominalnej danego produktu na wysokość popytu na ten produkt mierzy elastyczność cenowa. Wpływ ceny nominalnej produktu na wysokość popytu na inny produkt mierzy elastyczność mieszana popytu. Jest ona nazywana również elastycznością krzyżową. Wielkość dodatnia wskaźnika oznacza substytucję. Proporcjonalne zmiany cen produktów wywołują natomiast odwrotnie proporcjonalny efekt dochodowy dla popytu. Badanie zależności popytowo-cenowych dotyczy wielu poziomów: rynków szeroko rozumianych dóbr, rynków produktów będących dalszymi substytutami, rynku produktu zróżnicowanego bliskich substytutów (substytucja marek). Tutaj koncentrujemy się na substytucji dalszej towarów w asortymencie przedsiębiorstwa handlowego. Uzasadnijmy to faktem, że przedsiębiorstwo wybrane do badań oferuje produkt zazwyczaj pod jedną marką (obcą lub własną), ale oferuje różne produkty o podobnej funkcji, które są dalszymi substytutami.

W modelowaniu rynkowych zależności popytowo-cenowych stosowano szereg metod ekonometrycznych. Modele jednorównaniowe zwykle uwzględniają wpływ pojedynczego czynnika popytowego, ceny produktu, ceny substytutu albo dochodu, np. model liniowy, wykładniczo-hiperboliczny, Tornquista czy logistyczny (Guzik 2008). Alternatywne podejście to kompletne modele (systemy) zależności popytowych, które są rozwiązaniem zadania maksymalizacji użyteczności konsumenta, czyli stanowią tzw. układ funkcji popytu typu Marshalla (Suchecki 2006, 38–39). Proponowano inne układy popytowe: model rotterdamski, idealnego systemu funkcji popytu (Suchecki 2006). Wydaje się, że najpopularniejszy z nich był liniowy system wydatków (*linear expenditure system*, LES), w którym wydatki na dane dobro są liniową funkcją cen wszystkich rozważanych produktów i wydatków ogólnych (Pollak i Wales 1995, 4). W modelach tych zakłada się określoną postać analityczną funkcji popytu. Na przykład w modelu LES przyjmuje się funkcję Cobba-Douglasa za postać bezpośredniej funkcji użyteczności konsumpcji i z niej wyprowadza liniowe funkcje wydatków. Można w nim wyprowadzić wzory na funkcje elastyczności cenowej, krzyżowej i dochodowej, które zależą od parametrów funkcji użyteczności: minimalnych konsumowanych ilości produktów i udziału produktów w wydatkach na ilości ponad minimalne (Pollak i Wales 1995, 5). Oszacowanie parametrów jest kłopotliwe, gdyż wymaga dokładnych danych o strukturze wydatków i dochodach pojedynczych konsumentów. Zwykle dostępne są dane o wydatkach gospodarstw domowych.

Problematiczne jest zdefiniowanie i oszacowanie minimalnych ilości produktów dla konsumentów.

Najlepszym źródłem informacji o zróżnicowaniu preferencji i użyteczności nabywców okazały się mikrodane o ich decyzjach zakupowych. Luce (1959) dość dawno zaproponował model stochastycznej addytywnej użyteczności, który jest podstawą teorii wyboru dyskretnego, czyli wyboru określonej spośród alternatyw w oparciu o jej charakterystyki oraz preferencje i cechy osoby dokonującej wyboru. Parametryzacja modeli okazała się możliwa dzięki metodom mikroekonometrii, z których najczęściej do szacowania popytu indywidualnego stosuje się modele dwu- i wielomianowe. Model wielomianowy ma zmienną zależną określoną na zbiorze alternatyw, które zwykle są wzajemnie wykluczającymi się opcjami wyboru. Tego typu modelem jest warunkowy wielomianowy model logitowy, w którym opisuje się prawdopodobieństwo wyboru każdej z opcji w zależności od parametrów konsumenta i parametrów opcji wyboru (Gruszczyński 2012). Model dobrze opisuje konkurencję marek, czyli substytucję bliską czy doskonałą ze względu na ceny i jakość.

Kolejne prace w tym nurcie stanowią syntezę probabilistycznych modeli wyboru i założenia o deterministycznej funkcji użyteczności. Na przykład Berry, Levinsohn i Pakes (1995) pokazali, że zakładając postać łącznego rozkładu charakterystyk nabywców i atrybutów produktów, jako ich funkcje można szacować elastyczności cenową, krzyżową i dochodową, i to na podstawie danych sprzedażowych zagregowanych dla wszystkich konsumentów. Chib, Seetharaman i Strijnev (2004) omówili sposoby uwzględnienia w modelach wyboru rezygnacji ze wszystkich alternatyw przez konsumenta. Na przykład Mehta (2007) buduje taki hybrydowy model o wysokiej złożoności formalnej. Wykorzystuje systemy popytowe oparte na pośrednich funkcjach użyteczności (*Basic Translog Indirect Utility*, BTL), pochodzące od Pollaka i Walesa (1995), do modelowania zależności popytowo-cenowych między kategoriami produktów. Jednocześnie do wyboru marki produktu stosuje wielomianowy model logitowy. Tak rozbudowany formalnie model wymaga przyjęcia, że znany z góry jest podział na kategorie produktów o popycie współzależnym i niezależnym, a ilości i ceny są sumowane w tej drugiej grupie dla różnych kategorii i marek. To pociąga założenie, że w każdej kategorii produktowej kupowana jest najwyżej jedna marka. Potwierdza się, iż bez założenia postaci funkcji ogólnego wzorca konsumpcji dla wszystkich konsumentów nie mamy legitymowanej teorii ekonomii metody modelowania i estymowania zależności popytowo-cenowych dla produktów. Dlatego wielu badaczy problemu nie zakłada funkcji użyteczności, ale empirycznie bada determinanty popytu indywidualnego. Hopp i Xu (2008) stosują metodykę sieciową analizy czynników popytowych. Często prace te koncentrują się na dynamicznym wpływie promocji cenowej, np. Martinez-Ruiz *et al.* (2008) stosują maszynę wektorów nośnych do oszacowania regresji quasi-parametrycznej dla sprzedaży wybranych kategorii

produktów. Natomiast Bandyopadhyay (2009) stosuje autoregresję wektorową dla danych rynkowych o ilości sprzedaży kategorii produktowych: lodów, posypek i jogurtów.

Probabilistyczne modele wyboru stały się jednak standardem w badaniu wyborów konsumenta i pośrednio można z nich wyprowadzić zależności popytowo-cenowe. Dlatego proponowaną przez nas metodykę porównamy z własnościami tychże modeli:

- 1) Warunkowy wielomianowy model logitowy wymaga dokładnej specyfikacji marek i szczegółowej wiedzy o każdym konsumencie, np. jego dochodzie. Ograniczenie to próbuje się omijać poprzez bardzo mocne założenia. Na przykład uszeregowanie jakościowe produktów jest z góry znane i jednakowe dla wszystkich klientów (ma to sens, gdy dostawca na podstawie swojej wiedzy dzieli marki na wartościowe i premium, albo usługa ma klasę jakości: ekonomiczną i biznesową). Często przyjmuje się teoretyczny typ rozkładu dochodu lub innych charakterystyk klientów. Nasza metodyka zastępuje te założenia próbkowaniem klientów i transakcji z bazy mikro danych przedsiębiorstwa i wymaga mniejszego zakresu informacji o klientach.
- 2) Jedną z nowości proponowanego podejścia jest potraktowanie ceny jako zmiennej porządkowej, co umożliwia próbkowanie i porównania klientów w grupach kupujących produkty po różnych cenach.
- 3) Przejście od modelu wyboru do modelu popytu indywidualnego pewnego okresu wymaga założenia, że konsument wybierze jednokrotnie jeden produkt spośród alternatywnych marek lub klas produktu z oszacowanym w modelu rozkładem prawdopodobieństwa. Dlatego przy dużej liczbie klientów o tych samych cechach wiadomo, że względna częstość wybranych przez nich wariantów produktu jest oszacowana przez rozkład prawdopodobieństwa wyboru tych wariantów w modelu. Model ten nie przystaje do opisu konsumenta dokonującego wielokrotnych zakupów więcej niż jednego produktu i w ilościach różnych od jednostkowych. Nie przystaje, ponieważ wielokrotne wybory tego samego konsumenta musiałyby pozostawać w pewnej zależności od wcześniej wybranych wariantów i ilości. Wynika to z prawa malejącej marginalnej stopy substytucji nakazującej różnicować koszyk konsumpcji w celu maksymalizacji użyteczności. Próbuje się to uwzględnić we wcześniej przytoczonych już badaniach Mehty (2007). My unikamy dodatkowych założeń i rozbudowy formalnej modelu dzięki losowaniu klientów z ich warstw, czy też grup nabywających badany produkt i substytut po różnych kombinacjach ich cen w pewnym okresie. Dzięki temu otrzymujemy losowo wybranych klientów o różnych preferencjach manifestujących się w różnych ilościach nabycia produktu przy danej kombinacji cen produktów. To podejście ma również

ograniczenie, ponieważ wnioskując na podstawie reprezentatywnej próbki klientów, otrzymamy średnie ilości sprzedaży pojedynczemu reprezentantowi danej grupy klientów spośród całej zarejestrowanej i faktycznej ich populacji. Nie można więc opisać wpływu zróżnicowania cech pojedynczego klienta danej grupy na wielkość jego popytu. Dlatego proponowane podejście nie pasuje do opisu dóbr unikatowych lub stosunkowo drogiech dla konsumenta, które kupuje on w ilości jednostkowej na długi okres, czyli co najmniej roczny.

- 4) W modelach wyboru układ cen wariantów produktu wpływa na prawdopodobieństwo ich wyboru łącznie z oceną dodatkowej jakości i wartości, którą droższemu wariantowi przypisuje konsument. Wyższa cena produktu zmniejsza prawdopodobieństwo jego nabycia, ale wyższa jakość tego produktu zwiększa prawdopodobieństwo nabycia przez klientów o wystarczającym dochodzie. Modele wyboru nie ujmują bezpośrednio wpływu ceny na ilość nabywanych sztuk produktu. W naszej metodyce obserwujemy grupy klientów nabywające ten sam produkt po różnym poziomie ceny i w różnych ilościach. Są to różni klienci, ale ich cechy specyficzne powinny się kompensować dzięki losowemu doborowi grup próbki i uśrednieniu wyników. W ten sposób możemy zbadać popyt na produkt i ilości produktu podlegające substytucji na skutek obniżki ceny konkurującego produktu. Badamy wpływ ceny na substytucję ilościową, a nie jakościową, jak w modelu wyboru.

Z porównania naszej metodyki ze standardami literaturowymi wynika, iż jest ona syntezą zasad teorii mikroekonomii i procedury estymacji empirycznej na podstawie mikrodanych. Owe podstawowych zasady mikroekonomii przytaczamy w sekcji 3, a następnie formułujemy w postaci hipotetycznego modelu zależności popytowo-cenowych w sekcji 4.

3. Teoria zależności popytowo-cenowych a 2-czynnikowa analiza wariacji

Celem rozważań było opracowanie metodyki szacowania popytu klienta i ilościowej substytucji dla porządkowych kategorii ceny konkurujących produktów różnego typu w asortymencie przedsiębiorstwa. Zgodnie z teorią mikroekonomii definiujemy szczegółowo zależność popytowo-cenową badaną w naszej metodyce: podmiot, przedmiot, teoretyczne prawidłowości oraz odpowiednio dobieramy zakres i sposób pozyskiwania oraz przetwarzania danych:

- 1) Przedsiębiorstwo cenodawca ma swoją część rynku konkurencji niedoskonałej, a więc i cenowej. Wielkość rynku przedsiębiorstwa opisujemy pa-

rametrem liczby potencjalnych klientów. W naszym badaniu jest to liczba wszystkich klientów, którzy dokonali od kilku lat jakiegokolwiek zakupu i znajdują się w bazie przedsiębiorstwa. Do badania wybraliśmy próbę warstwowo-losową klientów, którzy w obserwowanym zakresie czasowym nabyli dwa produkty, których zależność popytowo-cenowa jest przedmiotem badania.

- 2) Przedmiotem badania jest oszacowanie relacji popytowej, czyli zależności ilości nabycia i cen nabycia dwóch produktów różnego typu. W dalszym ciągu nazywamy je odpowiednio produktem *A* i produktem *B*. Wybieramy do badania takie dwa produkty, które notują najwyższą sprzedaż i które podejrzewamy o substytucję ze względu na ich definicję oraz konkurencyjne zastosowanie.
- 3) Przyjmujemy, że nie występuje dyskryminacja cenowa klientów. W tym samym czasie za ten sam produkt płacona jest ta sama cena przez wszystkich nabywców. Tym samym typ klienta lub jego budżet zakupowy nie mają wpływu na cenę ani ilość sprzedaży. Dlatego ilość i cena nie pozostają w ujemnym sprzężeniu zwrotnym.
- 4) Przyjmujemy, że przedsiębiorstwo cenodawca kształtuje swoją krótkoterminową politykę cenową w warunkach stałych rynkowych czynników pozacenowych. Konsekwencją tego założenia jest przyjęcie krótkiego zakresu czasowego obserwacji cen i ilości sprzedaży produktów poszczególnym klientom. W badaniu przyjmujemy zakres roczny, w którym transakcje każdego klienta są zagregowane co do ilości, a ceny nabycia są uśrednione dla tych zagregowanych ilości. Chcemy przez to uniknąć problemu wahań typowo sezonowych, które mogłyby kształtować obserwacje ilości i ceny sprzedaży w podokresach roku.

Pokażemy, iż przy odpowiedniej definicji zmiennych i organizacji mikrodatychnych o nabywcach metodą właściwą ze względu na założenia merytoryczne jest 2-czynnikowa analiza wariancji. Model 2-czynnikowy zapisujemy za Staniszem (2007, 313):

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad (1)$$

gdzie: y_{ijk} – zmienna zależna, jest ilością produktu *A* sprzedaną *k*-temu klientowi po *i*-tym poziomie ceny produktu *A*, gdy nabył on również produkt *B*, przypuszczalny substytut, po *j*-tym poziomie jego ceny; μ – średnia ilość produktu *A* nabywana przez klientów dla pełnego zakresu cen; α_i – efekt główny *i*-tego poziomu ceny produktu *A*, czyli odchylenie ilości nabywanej przeciętnie po *i*-tej cenie od średniej ogólnej μ ; β_j – efekt główny *j*-tego poziomu ceny produktu *B*, czyli odchylenie ilości nabywanej przeciętnie od średniej ogólnej μ , gdy cena przypuszczalnego substytutu *B* jest na *j*-tym poziomie; $(\alpha\beta)_{ij}$ – efekt współdziałania (inte-

rakcji) i -tego poziomu ceny A oraz j -tego poziomu ceny B ; ε_{ijk} – losowy błąd doświadczalny o rozkładzie normalnym ze średnią równą 0 i wariancją równą σ^2 .

W podręczniku Stanisza (2007) znajdziemy wszystkie terminy, założenia i implikacje wieloczynnikowej analizy wariancji. Założenia 2-czynnikowej analizy wariancji dotyczą wstępnych własności danych lub przetworzonych danych:

- 1) Zmienna zależna to nabywana przez klienta ilość lub liczba sztuk produktu A . Ma rozkład normalny w każdej grupie klientów, którą wyróżniamy ze względu na kombinacje poziomów ceny nabycia produktu A i ceny nabycia produktu B . Wariancja zmiennej w każdej grupie jest jednakowa. W przeciwnym razie stosujemy odpowiednie przekształcenie zmiennej (Stanisz 2007, 346). Dla rozkładu Poissona wyliczamy pierwiastek zmiennej i poddajemy go analizie wariancji lub, alternatywnie, stosujemy regresję Poissona, jeśli ilość sprzedaży produktu jest zmienną dyskretną o rozkładzie Poissona.
- 2) Zmienne niezależne, nazywane czynnikami, zwykle są zmiennymi nominalnymi lub porządkowymi, które mają przeliczalny zbiór kategorii. W naszym modelu są nimi ceny produktów A i B . Cena jest zmienną dyskretną o dokładności jednego grosza, ale ma relatywny charakter i dlatego przyjmujemy trzy poziomy ceny średniej nabycia produktu A : niski a_1 , średni a_2 i wysoki a_3 , a zakres poziomów jest ustalony tak, aby wyróżnione grupy cenowe cechowała jak najbardziej zbliżona liczba nabywców, a najlepiej równa, jeśli liczba wszystkich nabywców produktu jest podzielna bez reszty przez 3. Podobnie wykonujemy kategoryzację ceny B na poziomy od niskiego b_1 do wysokiego b_3 .
- 3) W analizie wariancji powinno się stosować losowy dobór jednostek doświadczalnych, nazywanych replikacjami badania, którymi w naszym problemie są klienci nabywający obydwie produkty. Trudniej spełnić drugą zasadę randomizacji próby eksperymentalnej, która nakazuje losowo przydzielić jednostki z próby losowej do grup o różnych kombinacjach poziomów zmiennych niezależnych. Żeby tak dobrać próbę nabywców, musielibyśmy jeszcze przed zakupem znać tych klientów, którzy kupią obydwie produkty bez względu na ich ceny, ale w ilości zależnej od tych cen. Następnie spośród nich należałoby wylosować próbę badawczą, a z niej jeszcze raz wylosować klientów, którym przedsiębiorstwo sprzeda produkty po różnych kombinacjach poziomów cen. Taki eksperyment wymagałby dyskryminacji cenowej klientów i utajnienia cen. Sprzedawcy nie postępują tak wobec „stałych” klientów. Dlatego losujemy równą liczbę klientów z każdej grupy kupujących po różnych kombinacjach poziomów cen (to warstwy populacji badanej). Zakładając, że charakterystyka klientów nie wpływa na cenę, nie możemy popełnić błędu selekcji z tytułu przydziału klientów do grup cenowych. Nie wykluczamy okresowych zmian cen dla wszystkich klien-

- tów. Taki dobór warstwowo-losowy zapewnia więc reprezentatywną próbę klientów i ilości sprzedanych im produktów do analizy wariancji. Próba eksperymentalna jest zrównoważona, ponieważ dobieramy losowo równoliczną reprezentację klientów dla każdego układu cen nabycia produktów.
- 4) Sam poziom cen nie jest kontrolowany przez badacza, więc formalnie jest czynnikiem losowym (Stanisz 2007, 258). Jednak w naszej metodyce jest kontrolowany przez przedsiębiorstwo cenodawcę. Kontrola odbywa się w zakresie od ceny niskiej wyznaczonej przez koszty wytworzenia lub koszty niezapadłe, przez cenę średnią do ceny wysokiej, którą wyznaczają rynek i klienci. Stosowany model nazywamy modelem efektów stałych, ponieważ rozważamy pełen zakres cen i nie losujemy ich poziomów, ani zakresów cząstkowych.

4. Teoretyczny model popytu i substytucji cenowej dwóch produktów

Weryfikację merytoryczną modelu analizy wariancji sprzedaży przeprowadzamy w oparciu o teorię ekonomii. W tym celu, bazując na prawidłowościach ekonomii, formułujemy teoretyczny model wyników, w którym zakłada się występowanie zależności popytowej między produktami A i B , wraz z efektem substytucji i dochodowym. Wyprowadzony model wynikowy przedstawiony jest w Tabeli 1. Dzięki opracowanej metodyce spodziewamy się oszacować zależność średniej ilości popytu pojedynczego klienta od kombinacji cen. Oczekiwane na podstawie teorii kierunki wzrostu popytu klienta na produkt A przedstawia Tabela 1. Pionowe, czarne strzałki ilustrują prawo popytu dla produktu A , a dokładnie wzrost popytu w górę kolumny wraz z obniżką ceny nominalnej produktu A (jednocześnie obniża się cena realna, ponieważ w danej kolumnie cena produktu B się nie zmienia). Oszacowaniem tego efektu jest ciąg parametrów α_i , które powinny spełniać nierówność $\alpha_1 > \alpha_2 > \alpha_3$. Efekt występuje tylko w pewnym zakresie cen, gdy dwa pierwsze albo dwa ostatnie parametry w nierówności są równe ($\alpha_1 = \alpha_2 > \alpha_3$ albo $\alpha_1 > \alpha_2 = \alpha_3$). Efekt zupełnie nie występuje, gdy nie można odrzucić hipotezy o równości wszystkich trzech parametrów i właśnie taką hipotezę zerową testuje wieloczynnikowa analiza wariancji (Stanisz 2007, 312). Im większe różnice parametrów, tym elastyczniejszy cenowo jest popyt na produkt A . Zauważmy też, że parametry α_i oznaczają odchylenia średnich brzegowych od średniej ogólnej i stąd sumują się do 0. Dlatego co najmniej jeden i co najwyżej dwa parametry w naszym zadaniu są ujemne, co nie przeczy podanej wyżej hipotetycznej nierówności. Dodając średnią ogólną μ do parametrów α_i otrzymamy średnie brzegowe sprzedaży względem poziomu ceny produktu A i warunek prawa popytu zapiszemy wtedy analogicznie:

$$\mu + \alpha_1 > \mu + \alpha_2 > \mu + \alpha_3 \quad (2)$$

Tabela 1. Oczekiwane kierunki wzrostu popytu i ilości sprzedaży produktu A klientowi nabywającemu substytuty cenowe A i B po różnych kombinacjach poziomów cen

Cena produktu A	Cena produktu B		
	Niska b_1	Średnia b_2	Wysoka b_3
Niska a_1	↑	←	→
Średnia a_2		↑	→
Wysoka a_3			→

Legenda:

- ↑ prawo popytu na produkt A,
- ↔ substytucja cenowa, zastępowanie produktu B produktem A,
- ← efekt dochodowy spadku cen obydwu produktów,
- skumulowany efekt dochodowy i substytucji.

Źródło: opracowanie własne.

W Tabeli 1 poziome strzałki o czarnym konturze ilustrują wzrost wielkości popytu i ilości sprzedaży produktu A w prawą stronę wiersza, czyli wraz ze wzrostem ceny produktu B. Jest to efekt substytucji droższego produktu B. Oszacowaniem tego efektu jest ciąg parametrów β_j , które powinny spełniać nierówność podwójną $\beta_1 < \beta_2 < \beta_3$, albo chociaż nierówność pojedynczą dla dwóch z trzech parametrów. W analizie wariancji testujemy $H_0: \beta_1 = \beta_2 = \beta_3$, czyli brak efektu substytucji. Jeśli odrzucimy tę hipotezę, to im większe różnice parametrów, tym silniejszy jest efekt substytucji. Dodając średnią ogólną μ do parametrów β_j otrzymamy średnie brzegowe sprzedaży względem poziomu ceny produktu B i warunek efektu substytucji zapiszemy wtedy analogicznie:

$$\mu + \beta_1 < \mu + \beta_2 < \mu + \beta_3 \quad (3)$$

Jak wiemy z mikroekonomii, jednoczesne zmiany cen produktów w tym samym kierunku kompensują efekt substytucji i pozostaje czysty efekt dochodowy, czyli zmiany siły nabywczej dochodu. Zilustrowany jest on grubą szarą strzałką wzdłuż głównej przekątnej w kierunku lewego górnego rogu Tabeli 1. Przejście od cen wysokich a_3b_3 do cen niskich a_1b_1 zwiększy sprzedaż produktu A. Na drugiej przekątnej wzrostowi jednej ceny odpowiada spadek drugiej i odwrotnie, stąd kumulują się efekty substytucji i dochodowy w kierunku cen a_1b_3 . Dla dóbr normalnych **najniższa** cena produktu A i **najwyższa** produktu B wywołają dodatni efekt substytucji i ujemny efekt dochodowy dla wielkości popytu na produkt A, ale sumaryczny efekt powinien być dodatni. Szacunkiem efektów ukośnych w Tabeli 1 jest ciąg iloczynów $(\alpha\beta)_{ij}$, które nazywa się w analizie wariancji efektem interakcji czynników, czyli w naszym przypadku cen produktów. W analizie

wariancji rozstrzyga się jedynie, czy wszystkie iloczyny są równe (H_0), czy też co najmniej jeden z iloczynów jest różny od pozostałych (H_1). Odrzucenie H_0 oznacza wystąpienie efektów ukośnych, czego również się spodziewamy po badanych zależnościach cenowo-popytowych.

5. Organizacja danych i dobór próby badawczej

W wybranym przedsiębiorstwie handlu hurtowego zgromadzono mikrodane o sprzedaży produktów w grupie artykułów uniwersalnych dla rolnictwa w 2011 roku w następującym zakresie:

- 1) Dane przekrojowe dla wszystkich 1950 zarejestrowanych przez przedsiębiorstwo klientów, z których 181 kupiło w badanym roku oba wybrane produkty.
- 2) Zmienną objaśnianą jest liczba sprzedanych danemu klientowi szpul folii do sianokiszonki o różnym kolorze oraz o zbliżonych wymiarach i powierzchni całkowitej (produkt A). Jednostką pomiarową jest liczba sprzedanych szpul, gdyż klienci nabywają całe szpule. Był to produkt o najwyższej sprzedaży w wybranej grupie artykułów. Zmienna jest dalej oznaczona zapisem FOLIA_IL.
- 3) Pierwsza zmienna objaśniająca to poziom ceny produktu A , któremu przypisaliśmy $i = 1, 2, 3$ zakresy, a mianowicie: NISKI a_1 (średnia cena nabycia w przedziale od 200,14–218,29 zł), ŚREDNI a_2 (od 218,3–233,83) i WYSOKI a_3 (od 233,84–288,62). Długości przedziałów są różne, ale dzięki temu liczba klientów w przedziałach różni się o co najwyżej jednego i jest to odpowiednio: 60, 60 i 61. Zmienną oznaczyliśmy przez FOLIA_CE.
- 4) Drugą zmienną objaśniającą jest poziom ceny produktu B stanowiącego prawdopodobny substytut. Spośród 10 produktów o najwyższej sprzedaży w 2011 roku w wybranej grupie artykułów definicyjnie substytutem folii do kisenia siana jest siatka do owijania i przechowywania suchego siana. Zmienną oznaczyliśmy przez SIATKA_CE. Jej również przypisaliśmy trzy poziomy $j = 1, 2, 3$, mianowicie: NISKI b_1 (średnia cena nabycia w przedziale od 217,81–237,31 zł), ŚREDNI b_2 (237,32–261,66) i WYSOKI b_3 (261,67–360). Długości przedziałów cenowych znowu dobrano tak, aby zapewnić równomierny rozkład liczby klientów między przedziały: 60, 60 i 61.
- 5) Otrzymaliśmy dziewięć kombinacji poziomów obu cen $a_i b_j$: od $a_1 b_1$ do $a_3 b_3$. Mogliśmy badać ilość sprzedaży wszystkich 181 klientów, ale wtedy liczba klientów w poszczególnych grupach cenowych wahała się od 8 w grupie $a_1 b_3$ do 34 w grupie $a_1 b_1$. Własności metody analizy wariancji są statystycznie lepsze, gdy licznosc grup jest równa i dlatego do każdej grupy cenowej wylosowaliśmy po ośmiu klientów z danej warstwy cenowej. Daje nam tą łączną próbę warstwowo-losową o liczbie 72 klientów.

Tabela 2. Statystyki każdej grupy klientów wyróżnionej ze względu na kombinacje poziomów cen nabycia: średnia liczba szpul folii nabywana przez klienta danej grupy; wariancja liczby szpul nabywanych przez klienta danej grupy; liczba klientów wylosowanych do grupy

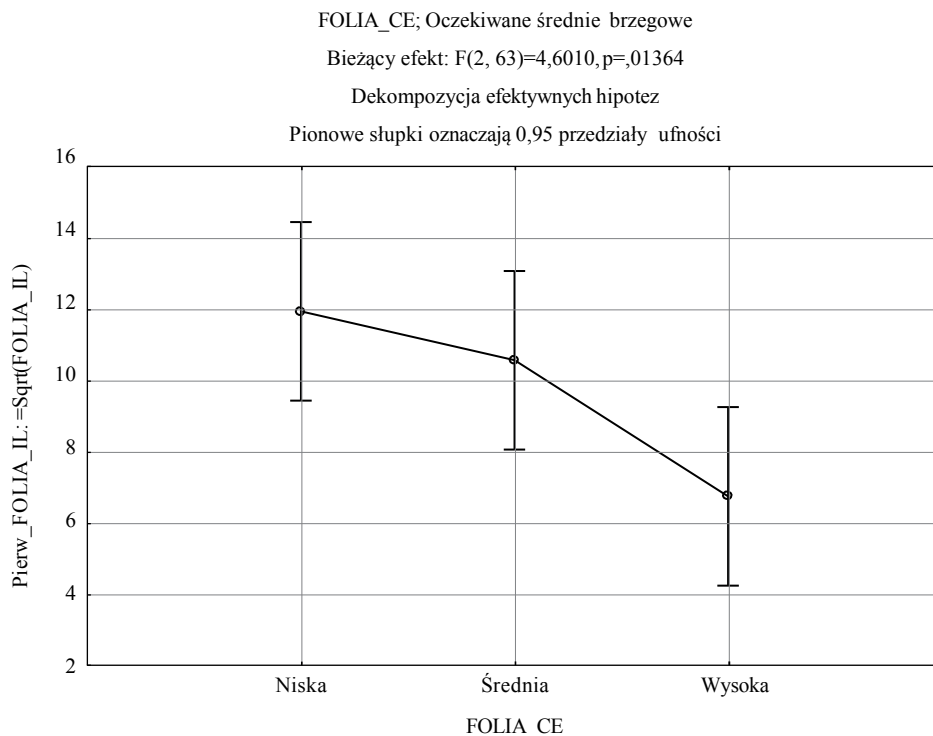
Cena folii	Cena siatki								
	Niska b_1			Średnia b_2			Wysoka b_3		
Niska a_1	123,75	4276	8	329,38	297599	8	181,75	26951	8
Średnia a_2	108,75	6683	8	87,63	1951	8	194	29175	8
Wysoka a_3	18,75	239	8	128,25	8136	8	54,88	7823	8

Źródło: obliczenia własne.

W Tabeli 2 przedstawiamy statystyki opisowe dla ilości sprzedaży folii poszczególnym klientom w grupach kupujących po różnych kombinacjach cen. Jest to wstępna weryfikacja modelu hipotetycznego na podstawie średnich ilości sprzedaży pojedynczemu klientowi z różnych grup. W kolumnie ceny niskiej b_1 średnia sprzedaż rośnie w górę kolumny, a dla kolumn b_2 i b_3 dwa na trzy porównania ilości potwierdzają wzrost ilości ku górze kolumny, np. w kolumnie b_2 w górnej komórce jest najwyższa ilość sprzedaży 329,38; a więc: $128,25 < 329,38$ i $87,63 < 329,38$. Potwierdza się prawo popytu. Podobnie prowadzimy porównanie średnich ilości sprzedaży w kolejnych komórkach danego wiersza. W każdym wierszu dwa na trzy porównania potwierdzają istnienie efektu substytucji, czyli wzrost średniej sprzedaży w prawą stronę wiersza np. dla wiersza ceny niskiej a_1 zachodzi relacja: $123,75 < 329,38$ i $123,75 < 181,75$. W pełni potwierdzony jest efekt dochodowy, gdyż średnia ilość popytu indywidualnego rośnie na głównej przekątnej w kierunku lewego górnego rogu. Potwierdzony jest także skumulowany efekt substytucji i dochodowy na drugiej przekątnej.

6. Wyniki analizy wariancji i ocena istotności statystycznej

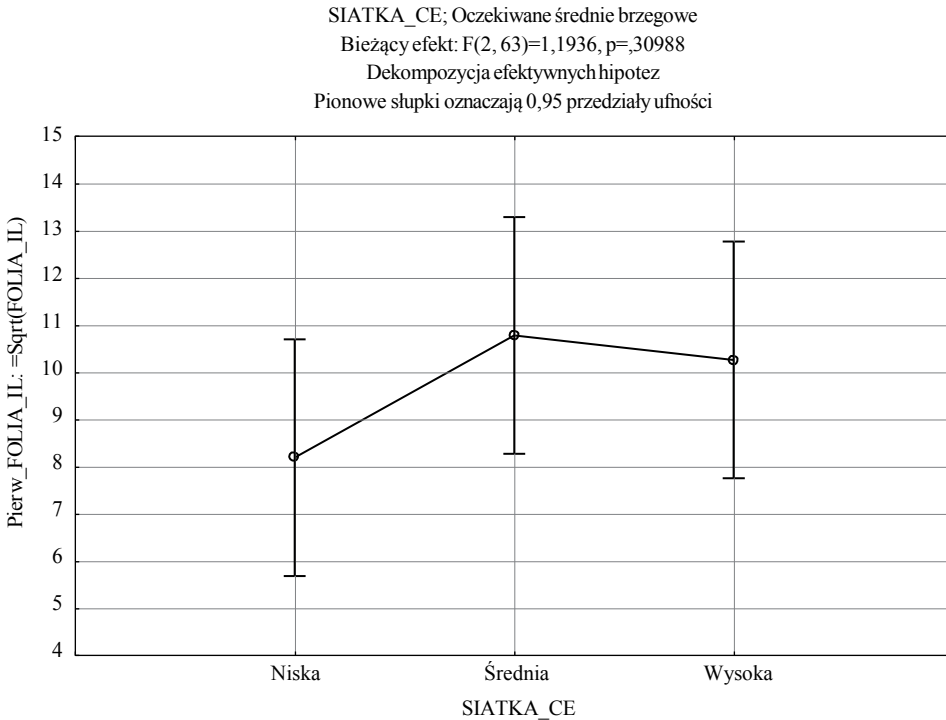
Na podstawie Tabeli 2 weryfikujemy również założenie analizy wariancji o równości wariancji grupowych sprzedaży. Obserwujemy wyraźny wzrost wariancji grupowych wraz ze wzrostem średnich grupowych. Ilość sprzedaży ma więc rozkład Poissona. Dla lepszego dopasowania danych o takim rozkładzie do analizy wariancji pierwiastkujemy zmienną zależną (Stanisz 2007, 346). Dlatego dalszej analizie wariancji poddajemy pierwiastek ilości sprzedaży folii, który oznaczono „Pierw_FOLIA_IL” w programie STATISTICA, który posłużył do wyznaczenia wyników.



Rycina 1. Weryfikacja prawa popytu

Źródło: opracowanie własne.

Na Rycinie 1 widzimy szacunki punktowe (kółka) i przedziałowe (słupki) średniej brzegowej pierwiastka ilości sprzedaży dla trzech poziomów ceny folii. Ilustracją jest malejąca krzywa łamana. Spełniony jest warunek prawa popytu (2). Im mniej zachodzą na siebie pionowe słupki przedziałów ufności, tym istotniejszy statystycznie jest nasz wniosek. Do oceny poziomu istotności efektu używa się testu F , który jest ilorazem wariancji średnich brzegowych (ma ona 2 stopnie swobody, ponieważ są 3 średnie brzegowe i wyznaczona dla nich średnia ogólna) i wariancji błędu wewnątrzgrupowego (ma ona 63 stopnie swobody, ponieważ wyliczono ją dla 72 klientów, dla których wyznaczono też 9 średnich grupowych). Statystyka F wynosi 4,6 i jest istotna z prawdopodobieństwem błędu około 0,01, co potwierdza efekt popytowy.

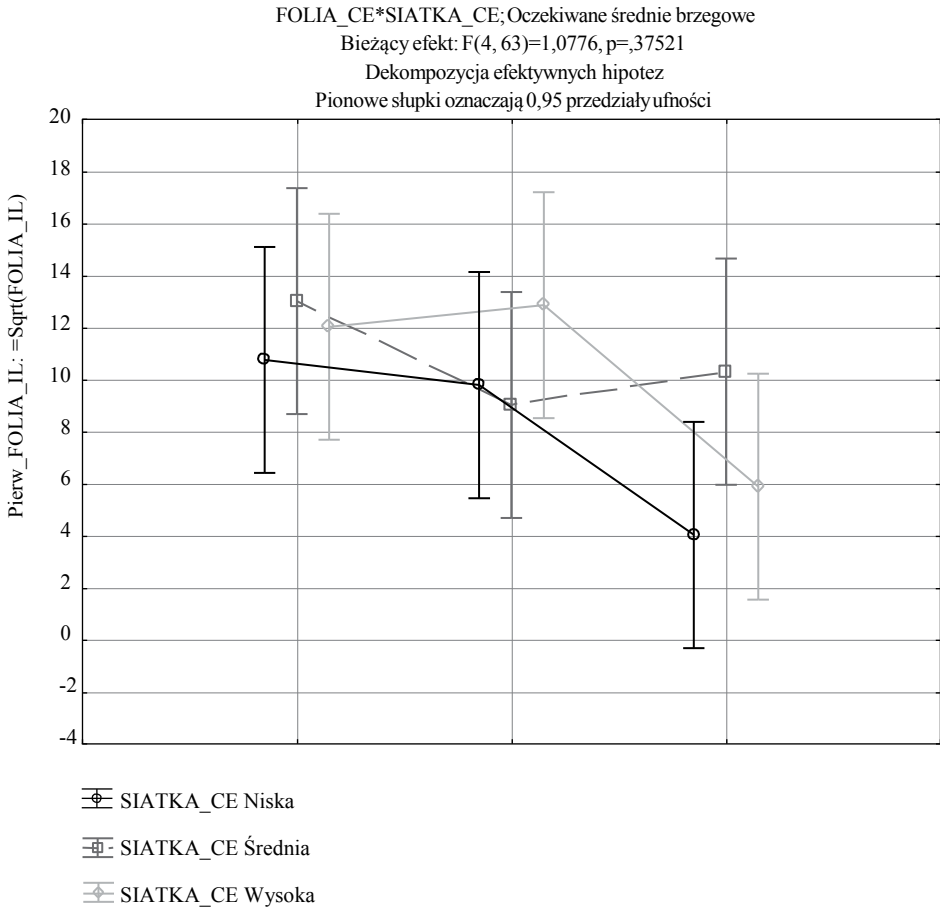


Rycina 2. Weryfikacja substytucji

Źródło: opracowanie własne.

Na Rycinie 2 widzimy oczekiwane średnie brzegowe pierwiastka ilości sprzedaży (kółeczka) dla trzech poziomów ceny hipotetycznie substytucyjnej siatki. Dwie pierwsze średnie oraz pierwsza i trzecia spełniają warunek substytucji (3), co przejawia się rosnącym odcinkiem krzywej łamanej na wykresie dla wzrostu ceny z niskiej do średniej. Efekt w całym zakresie cenowym jest istotny z prawdopodobieństwem błędu około 0,31; a więc nie mamy podstaw do odrzucenia hipotezy o braku substytucji.

Wyniki dla interakcji przedstawiono na Rycinie 3. Łamane pokazują prawo popytu na folię przy różnej cenie siatki. Przebieg łamanej to nic innego jak porównanie średnich grupowych w danej kolumnie Tabeli 2, co omówiliśmy na koniec sekcji 5. Interakcje średnich wykazują oczekiwany porządek, ale są one istotne na poziomie błędu aż około 0,38. Wysoki błąd może być spowodowany faktem, że pierwiastkując dane o ilości sprzedaży, zmniejszyliśmy ich zmienność, ale również różnice średnich grupowych w próbie.



Rycina 3. Weryfikacja interakcji cen

Źródło: opracowanie własne.

Co do założeń analizy wariancji, to sprawdzimy jeszcze, czy rozkład zmiennej zależnej w grupach jest normalny i czy jednorodna jest wariancja. Wiadomo jednak, że test F używany w ANOV-ie jest odporny na odchylenia danych od tych założeń, jeśli próby są równoliczne, jak w naszym badaniu (Stanisz 2007, 342). Odnosnie do normalności rozkładów grupowych, to sprawdziliśmy wykresy normalności i zauważalne odchylenie wykazuje tylko rozkład w dwóch grupach: dla niskiej ceny folii i średniej ceny siatki i odpowiednio dla kombinacji cen wysokiej i średniej. Natomiast test Levene'a potwierdził, iż pomimo spierwiastkowania ilości sprzedaży folii wariancje sprzedaży w grupach cenowych są istotnie różne. Współczynnik determinacji R^2 w modelu wyniósł 20% i taką część całkowitej zmienności ilości sprzedaży wyjaśnia poziom cen produktów. Podsumowujemy w Tabeli 3 istotne różnice średnich (są one podane na końcu strzałki różnicą średnich dla grupy na końcu i grupy na początku strzałki). Dzięki analizie wariancji

widzimy, jaki jest stopień istotności oszacowanych efektów. W sekcji 7 zweryfikujemy ponownie efekty popytu i substytucji ilości sprzedaży folii, szacując na wyjściowych danych regresję Poissona, a więc przy założeniu różnej wariancji ilości sprzedaży klientom w grupach nabywających po różnych cenach.

Tabela 3. Istotnie statystycznie (poziom podany w nawiasie) różnice średniego pierwiastka liczby szpul folii nabywanej przez klientów

Cena folii	Cena siatki		
	Niska b_1	Średnia b_2	Wysoka b_3
Niska a_1	4,87 (0,12) 5,19 (0,01)	2,59 (0,15)	8 (0,01)
Średnia a_2	3,82(0,04)		
Wysoka a_3			

Źródło: obliczenia własne.

7. Regresja Poissona

Ze względu na heteroskedastyczność wariancji grupowych w próbce ilości sprzedaży klientom uzupełnieniem metodycznym zaprezentowanej analizy zależności cenowo-popytowych jest model regresji Poissona. W modelu tym wariancja ilości produktu kupowanej przez klientów jest równa średniej ilości produktu nabywanej przez klienta z danej grupy. Te parametry popytu klienta są równe dla danej grupy klientów, ale ich poziom może być inny w poszczególnych grupach. To inaczej niż w analizie wariancji. Kupowana przez klienta liczba sztuk produktu A po i -tej kategorii ceny produktu A ($i = 1, 2, 3$) oraz j -tej kategorii ceny produktu B ($j = 1, 2, 3$) to zmienna losowa Y_{ij} o rozkładzie Poissona i następującej funkcji prawdopodobieństwa nabycia liczby sztuk y_{ij} :

$$P(Y_{ij} = y_{ij}) = \frac{\lambda_{ij}^{y_{ij}}}{y_{ij}!} e^{-\lambda_{ij}} \quad (4)$$

gdzie wartość oczekiwana i wariancja rozkładu popytu klienta wynoszą λ_{ij} dla kombinacji ceny i -tej oraz j -tej. Przyjmujemy też za Gruszczynskim (2012) standardową, wykładniczą zależność wartości oczekiwanej liczby sztuk od zmiennych objaśniających, czyli cen substytucyjnych produktów A i B . Uwzględniając interakcje cen, formułujemy następujący model regresji Poissona o podstawie danej liczbą e :

$$\lambda_{ij} = e^{\mu + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 x_3 + \beta_2 x_4 + \gamma_1 x_1 x_3 + \gamma_2 x_1 x_4 + \gamma_3 x_2 x_3 + \gamma_4 x_2 x_4} \quad (5)$$

gdzie:

μ – wyraz wolny wykładnika, potęga e^μ podaje wartość oczekiwaną liczby sztuk produktu A nabywanej przez klienta w pełnym zakresie cen; α_1, α_2 – parametry wpływu odpowiednio ceny niskiej i średniej produktu A na oczekiwaną liczbę sztuk produktu A nabywaną przez klienta; iloczyny $100\alpha_1$ i $100\alpha_2$ podają odpowiednio dla ceny niskiej i średniej stopę przyrostu oczekiwanej liczby sztuk względem średniej ogólnej e^μ ; x_1, x_2 – zmienne kodujące odpowiednio kategorie ceny niskiej ($i = 1$) i średniej ($i = 2$) produktu A , oraz pośrednio kategorię niejawną ceny wysokiej ($i = 3$). Kodowanie jest następujące: $x_1 = 1$ i $x_2 = 0$, jeśli $i = 1$; albo $x_1 = 0$ i $x_2 = 1$, jeśli $i = 2$; albo $x_1 = -1$ i $x_2 = -1$, jeśli $i = 3$. Jest to tak zwane kodowanie z sigma ograniczeniami (Stanisz 2007, 677); β_1, β_2 – parametry wpływu odpowiednio ceny niskiej i średniej produktu B na oczekiwaną liczbę sztuk produktu A nabywaną przez klienta; iloczyny $100\beta_1$ i $100\beta_2$ podają odpowiednio dla ceny niskiej i średniej stopę przyrostu oczekiwanej liczby sztuk względem średniej ogólnej e^μ ; x_3, x_4 – zmienne kodujące odpowiednio kategorie ceny niskiej ($j = 1$) i średniej ($j = 2$) produktu B , oraz pośrednio kategorię niejawną ceny wysokiej ($j = 3$). Kodowanie jest następujące: $x_3 = 1$ i $x_4 = 0$, jeśli $j = 1$; albo $x_3 = 0$ i $x_4 = 1$, jeśli $j = 2$; albo $x_3 = -1$ i $x_4 = -1$, jeśli $j = 3$; $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ – parametry wpływu interakcji, czyli kombinacji niereferencyjnych poziomów cen produktu A i produktu B , odpowiednio niskiej i niskiej, niskiej i średniej, średniej i niskiej oraz średniej i średniej.

Przypomnijmy, że badanym produktem A jest folia do kisenia siana. Jednostką sprzedaży jest jedna szpula, za którą podana jest cena. Podobne są jednostki pomiarowe dla substytucyjnego produktu B , którym jest siatka do przechowywania suchego siana. Dla opisanych w Tabeli 2 danych oszacowaliśmy model (5) metodą największej wiarygodności. Wpierw weryfikujemy istotność statystyczną szacunku modelu (5) za pomocą testu ilorazu wiarygodności (Gruszczynski 2012, 83). Iloraz wiarygodności okazał się istotnie wyższy dla szacunku modelu (5) niż dla zagnieżdżonych w nim modeli (na poziomie 0,01), w tym dla szacunku modelu z wyrazem wolnym. Współczynnik dopasowania R^2 McFaddena wynosi 0,29 dla otrzymanego szacunku modelu (5). Występuje nadmierny rozrzut (zob. Stanisz 2007, 793), ponieważ istotnie statystycznie jest odchylenie obserwacji od wielkości oszacowanych w modelu (5). Odchylenie, nazywane dewiancją, ma rozkład chi-kwadrat o liczbie stopni swobody danej różnicą liczby obserwacji w próbie i liczby parametrów modelu (5). Badana próba ma 63 stopnie swobody i taka jest wartość oczekiwana dewiancji. Lewostronny obszar krytyczny dla testu istotności dewiancji (poziom istotności 0,05) rozciąga się od $[0; 43,5]$ a prawostronny od $[87; \infty]$. Tymczasem empiryczna wielkość dewiancji wynosi w próbie 9534. Model jest dość słabo dopasowany, a lepsze dopasowanie może dać model re-

gresji ujemnej dwumianowej. Jednak istotność statystyczną estymatorów można ocenić w warunkach nadmiernego rozrzutu. Oceniamy istotność szacunków w warunkach nadmiernego rozrzutu, co wymaga zwiększenia średnich błędów ocen parametrów mnożnikiem 12,3; czyli pierwiastkiem z ilorazu odchylenia i stopni swobody. Błędy szacunków podstawiamy do testów istotności Walda. Szacunki parametrów i poziom ich istotności podane zostały w Tabeli 4. Efekty nieistotne traktujemy łącznie z tymi istotnymi jako całość modelu, ponieważ jak już wspomniano, szacunek modelu (5) ma istotnie wyższy iloraz wiarygodności w porównaniu ze wszystkimi zagnieżdżonymi w nim modelami.

Tabela 4. Szacunek MNW modelu (5) regresji Poissona liczby szpul folii sprzedanych klientowi

Parametr	Szacunek	Parametr	Szacunek
μ	**4,67	γ_1	0,07
α_1	**0,60	γ_2	0,15
α_2	0,14	γ_3	0,40
β_1	** -0,52	γ_4	** -0,71
β_2	*0,37		

* szacunek istotny na poziomie błędu 0,1; ** szacunek istotny na poziomie błędu 0,05

Źródło: obliczenia własne

Na podstawie szacunków parametrów w Tabeli 4 szacujemy wartość oczekiwaną liczby szpul folii nabywanych przez klienta po dowolnej kombinacji cen. Na przykład dla kombinacji cen wysokich obu produktów a_3b_3 obliczymy z modelu (5):

$$\hat{\lambda}_{33} = e^{(-1)+0,15 \times (-1) \times (-1) + 0,40 \times (-1) \times (-1) - 0,71 \times (-1) \times (-1)} \approx 54,88 \quad (6)$$

Wyliczenie (6) pokazuje znaną własność regresji Poissona, mianowicie szacunki warunkowych wartości oczekiwanych są równe średnim grupowym w próbie obserwacji (zob. Tabela 2), jeśli model ma co najmniej tyle parametrów, ile jest w próbie grup.

Interpretacja parametrów podana jest pod równaniem (5) i została oparta na własności stałej stopy przyrostu funkcji wykładniczej. Posłuży ona do oceny zgodności z hipotetycznym modelem zależności popytowo-cenowych z Tabeli 1. Parametry α potwierdzają dodatni efekt popytowy obniżania ceny szpuli folii w całym zakresie cenowym. Efekt ten jest wysoce istotny dla ceny niskiej. Parametr β_1 jest wysoce istotny i ujemny, co potwierdza substytucję folii siatką, gdy cena tej ostatniej jest niska. Gdy cena siatki jest średnia, istotny na poziomie 0,1

parametr β_2 jest dodatni. Jest to zgodnie z pierwszą nierównością w podwójnej nierówności parametrów β , którą sformułowaliśmy jako warunek występowania substytucji w całym zakresie cenowym w sekcji 4. Relacja wartości bezwzględnej szacunków tych parametrów pokazuje, że blisko cały efekt substytucji zachodzi w zakresie cen niskich i średnich siatki. W interpretacji efektów pomaga wyrażenie szacunku warunkowej wartości oczekiwanej liczby szpul folii nabywanych przez klienta za pomocą procentowej różnicy od szacunku średniej ogólnej, który w próbie wynosi 136,4. Wyniki przedstawione są w Tabeli 5 i potwierdzają hipotetyczny model z tabeli 1 z wyjątkiem porównania wielkości sprzedaży dla cen a_2b_2 i a_1b_2 oraz porównania dla cen a_2b_3 i a_1b_3 . Na przykład dla kombinacji cen niska–niska wyliczenie odchylenia względnego wygląda następująco:

$$(e^{4,67 + 0,60*1 - 0,52*1 + 0,07*1*1} - 136,4)/136,4 \approx -9\% \tag{7}$$

Tabela 5. Odchylenie szacunków brzegowych względem szacunku ogólnej wartości oczekiwanej nabywanej przez klienta liczby szpul folii (%)

Cena folii	Cena siatki		
	Niska b_1	Średnia b_2	Wysoka b_3
Niska a_1	-9	141	33
Średnia a_2	-20	-36	42
Wysoka a_3	-86	-6	-60

Źródło: obliczenia własne

W Tabeli 5 zauważamy znacząco podwyższony popyt na folię dla kombinacji cen powyżej głównej przekątnej tabeli. Aż o 141% średniej wyższy jest on dla niskiej ceny folii i średniej ceny siatki. Te wyniki potwierdzają silną substytucję siatki folią, gdy cena folii jest o jeden poziom niższa, czyli np. jest niska, gdy cena siatki jest średnia albo jest średnia, gdy cena siatki jest wysoka. Widać też, że dla wysokiej ceny siatki obniżanie ceny folii do niskiej nie zwiększa popytu na folię, ale możliwe, że wystąpi wtedy silny efekt dochodowy w postaci wzrostu popytu na siatkę. Dlatego szukając ceny maksymalizującej sprzedaż klientowi, powinno się brać pod uwagę wpływ kombinacji cen na wyniki ilości i kwoty sprzedaży obydwu produktów. Wymagałoby to powtórzenia analizy dla ilości sprzedaży siatki w roli zmiennej zależnej. Na tym etapie możemy natomiast stwierdzić, która kombinacja cen przyniesie najwyższą sprzedaż samej folii. Dlatego dla każdego poziomu ceny folii wybieramy taką wysokość ceny siatki, dla której popyt na folię jest najwyższy. W Tabeli 5 kursywą zaznaczono odchylenia procentowe popytu dla wybranych kombinacji cenowych. Dla średniej ceny folii najwyższy popyt jest mniejszy o 99% średniej ogólnej od najwyższej wielkości popytu dla

ceny niskiej ($42 - 141 = -99$). Dla ceny wysokiej najwyższy popyt jest mniejszy o 149% średniej ogólnej niż dla ceny niskiej. Takie są relatywne różnice średnich liczb szpul: 194 i 128,25 od najwyższej średniej dla ceny niskiej, czyli od 329,28. Po wyrażeniu różnic w procencie średniej ogólnej, która wynosi 136,4. Dla każdej z wybranych kombinacji cenowych szacunek średniej liczby szpul nabywanych przez klienta mnożymy przez wielkość środkową ceny w odpowiednim dla danej wielkości popytu zakresie cenowym. W ten sposób wyliczamy sprzedaż folii odpowiednio dla jej ceny na poziomie niskim, średnim i wysokim:

$$329,28 \times 209,15 = 68\,868,91;$$

$$194,00 \times 226,06 = 43\,855,64;$$

$$128,25 \times 261,23 = 33\,502,75.$$

Stąd wiemy, że najwyższa średnia sprzedaż folii jednemu klientowi to 68 868,91; gdy cena folii jest niska, a cena siatki – średnia.

Analiza przyrostów procentowych popytu i ceny oraz elastyczności popytu jest możliwa dla każdej pary kombinacji cenowych. Przeanalizujemy przejście od kombinacji a_2b_1 do kombinacji a_1b_2 . Elastyczność cenową wyliczymy, analizując wpływ obniżenia ceny z poziomu a_2 do a_1 . Cenę folii obniżono z 226 na 214 (są to wielkości środkowe przedziałów cenowych), czyli o -5,5% względem ceny średniej dla tych zakresów. Odpowiednio przyrost popytu obliczamy tylko dla obniżki ceny folii, czyli między kombinacjami a_2b_1 i a_1b_1 . Ilość wzrosła o $-9 - (-20) = 11$ procent średniej ogólnej. Stąd elastyczność cenowa w tym zakresie popytu wynosi $11/(-5,5) = -2$. Natomiast elastyczność krzyżowa wynika z porównania względnego przyrostu popytu między kombinacjami a_1b_1 i a_1b_2 w stosunku do względnego przyrostu ceny z poziomu b_1 do b_2 . Ilość sprzedaży folii zwiększa się o $141 - (-9) = 151$ procent średniej ogólnej, a cena siatki wzrasta o $249,49 - 227,31 = 22,18$, czyli o 9,3%. Dlatego elastyczność krzyżowa dla tego zakresu ceny B wynosi $151/9,3 = 16,2$; czyli wskazuje na silną substytucję w tym zakresie cenowym.

8. Podsumowanie

W artykule przedstawiono oryginalną metodykę modelowania zależności popytowo-substytucyjnej dwóch produktów. Skonstruowano w Tabeli 1 macierzowy model hipotetyczny, który zdefiniowano szczegółowo według prawidłowości mikroekonomii i dzięki temu jego wyniki mają jednoznaczną interpretację ekonomiczną: efektów popytu, substytucji i dochodowego. Analityczną postać modelu dopasowano do założeń metodycznych w 2-czynnikowej analizie wariancji oraz regresji Poissona. Odpowiednio zorganizowano również dane. Model analizy czynnikowo-cenowej ilości sprzedaży ma prostszą formę matematyczną i jest łatwiejszy w interpretacji, ale wymaga rozkładu normalnego zmiennej zależnej i jej równej wariancji w każdej grupie obserwacji wyróżnionej ze względu na poziomy

czynników. W celu spełnienia tych założeń stosuje się przekształcenia zmiennej zależnej. Tak postąpiono również w tym artykule w badaniach wysokości sprzedaży i cen dwóch produktów różnego rodzaju, wybranych z asortymentu przedsiębiorstwa handlowego. Zastosowanie przekształcenia powoduje lepsze dopasowanie danych do założeń, ale w badanej próbie nie jest to pełne dopasowanie. W razie niejednorodności wariancji ilości sprzedaży w grupach klientów przydatną metodą estymacji jest model regresji Poissona. Takie własności okazały się mieć dane, które posłużyły do weryfikacji przydatności prezentowanej metodyki. Otrzyma-ny szacunek regresji Poissona ma wyższy współczynnik determinacji niż model ANOVA i ma szacunki parametrów istotne nawet w warunkach nadmiernego rozrzutu ilości sprzedaży. W dalszych pracach należy rozważyć zastosowanie innych metod estymacji przedstawionego hipotetycznego modelu, które mogą się okazać lepiej dopasowane do badanej tu próbki obserwacji. Jednak otrzymane oboma metodami oszacowania modelu hipotetycznego są istotne i spójne dla przebadanej warstwowo-losowej próby klientów wybranego przedsiębiorstwa. W szczególności istotne i merytorycznie spójne z teorią ekonomii i modelem hipotetycznym są szacunki efektów: prawa popytu i efektu dochodowego. Natomiast istotny szacunek popytowych efektów substytucji otrzymano w metodzie Poissona. Jak pokazano w ostatniej sekcji, szacunki można zastosować do wyliczenia elastyczności cenowej i krzyżowej popytu lub do wyboru kombinacji cen, dla której sprzedaż produktu jest najwyższa.

Bibliografia

- Bandyopadhyay, Subir. 2009. „A dynamic model of cross-category competition: Theory, tests and applications”. *Journal of Retailing* 85 (4): 468–479.
- Berry, Steven, James Levinsohn i Ariel Pakes. 1995. „Automobile prices in market equilibrium”. *Econometrica* 63 (4): 841–890.
- Chib, Siddhartha, P.B. Seetharaman i Andrei Strijnev. 2004. „Model of brand choice with a no purchase option calibrated to scanner panel data”. *Journal of Marketing Research* 41 (2): 184–196.
- Gruszczyński, Marek. red., Monika Bazyl, Monika Książek, Marcin Owarczuk, Adam Szulc, Arkadiusz Wiśniowski i Bartosz Witkowski. 2012. *Mikroekonometria. Modele i metody analizy danych indywidualnych*. Warszawa: Wolters Kluwer.
- Guzik, Bogusław. 2008. *Podstawy ekonometrii*. Poznań: Wydawnictwo Akademii Ekonomicznej w Poznaniu.
- Hopp, Wallace J. i Xiaowei Xu. 2008. „A static approximation for dynamic demand substitution with applications in a competitive market”. *Operations Research* 56 (30): 630–645.
- Luce, R. Duncan. 1959. *Individual Choice Behaviour: A Theoretical Analysis*. New York: Wiley.
- Martínez-Ruiz, María Pilar, Alejandro Mollá-Descals, Miguel Ángel Gómez-Borja, José Luis Rojo-Álvarez. 2008. „Using Support Vector Semiparametric Regression to estimate the effects of pricing on brand substitution”. *International Journal of Market Research* 50 (4): 533–557.
- Mehta, Nitin. 2007. „Investigating consumers’ purchase incidence and brand choice decisions across multiple product categories: A theoretical and empirical analysis”. *Marketing Science* 26 (2): 196–217.
- Pollak Robert A. i Terrence J. Wales. 1995. *Demand System Specification and Estimation*. Oxford: Oxford University Press.
- Stanisz, Andrzej. 2007. *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny*. Tom 2. Kraków: StatSoft.
- Suchecky, Bogdan. 2006. *Kompletne modele popytu*. Warszawa: PWE.

Products demand and substitution modelling and estimation for microdata

Abstract

Let us consider wide and deep assortment, which is typical for a multiproduct producer or for a wholesaler. Commodity price affects its other commodities quantities sold to the customer. We mean price-demand relations, that should be taken into account by the oligopolist that prices his commodities using market formula. Therefore we propose our theoretical model of substitution that holds own price effect, price substitution effect and income effect. We define price as discrete variable with low, middle and high range of values and model theoretically demand for two substitutes concerning their prices relation. The model is appropriate for count sales data of individual customers. We use ANOVA method for individual demand analysis in groups of customers. Customers are selected by random stratified sampling. Our model is verified with 2-factor ANOVA of main and price interaction effects. Feasible method of our model verification is also Poisson regression, which we estimate in the paper. Results of both estimation methods are coherent and approve our theoretical model of demand relations between substitutes.

Keywords: demand, substitution, customer, ANOVA, Poisson regression.

JEL Codes: C25, D12

DOI: 10.17451/eko/44/2016/132