

MAREK WITKOWSKI

ZJAWISKO WSPÓLLINIOWOŚCI W STATYSTYCZNYCH BADANIACH PRZYCZYNOWO-SKUTKOWYCH PROCESÓW EKONOMICZNYCH

Dynamiczny rozwój współczesnych metod statystycznych nakłada na badacza obowiązek możliwie precyzyjnego opisu rzeczywistości. Jednym z najbardziej rozległych obszarów zainteresowań tych metod są badania nad identyfikacją prawidłowości w zakresie kształtowania się przyczynowo-skutkowych uwarunkowań zjawisk i procesów społeczno-ekonomicznych.

Mimo jednak stosunkowo szybkiego rozwoju ilościowych metod analizy współzależności bardzo często uzyskiwane w wyniku ich zastosowania rezultaty nie są w pełni zadowalające. Jedną z głównych przyczyn takiego stanu rzeczy stanowi złożony charakter powiązań wzajemnych między badanymi zjawiskami społeczno-ekonomicznymi. Wyrazem tego jest występowanie silnego wewnętrznego skorelowania między nimi, co w teorii statystyki określa się mianem współliniowości. Zjawisko współliniowości występuje zarówno, gdy informacje o zmiennych mają charakter danych przekrojowych, jak i postać szeregów czasowych. W pierwszym przypadku przyczyna współliniowości leży w tendencji do dość proporcjonalnego zmieniania się obserwowanych wielkości ekonomicznych na skutek zmian jednostki obserwacji¹. W drugim natomiast, jak podkreśla Klein, związana jest z tendencją do kształtowania się wielu zjawisk gospodarczych w czasie według takiego samego trendu i wahań okresowych².

Występowanie współliniowości zmiennych stanowi niejednokrotnie poważną przeszkodę w prawidłowym określeniu i efektywnym oszacowa-

¹ M. Gruszczyński, M. Kolupa, E. Leniewska, G. Napiórkowski, *Miary zgodności, metody doboru zmiennych, problemy współliniowości*, Warszawa 1979, s. 187 - 188.

² L. R. Klein, *Wstęp do ekonometrii*, Warszawa 1965, s. 95.

niu wzajemnych zależności, komplikując równocześnie przeprowadzanie analizy numerycznej. Dzieje się tak dlatego, że parametry modelu (współczynniki regresji cząstkowej) odzwierciedlają nie tylko siłę oddziaływania na zmienną objaśnianą tych zmiennych, do których się odnoszą, lecz również pozostałych. Utrudnia to oczywiście proces wydzielania najbardziej istotnych czynników³. Ponadto obniża się wartość poznawcza parametrów przy próbie ich interpretacji ekonomicznej. Po trzecie wreszcie pojawiają się trudności natury numerycznej, związane ze złym uwarunkowaniem macierzy obserwacji na zmiennych objaśniających (tzw. macierzy Gramma). Następuje mianowicie zmniejszenie określoności tej macierzy (tj. zbliżanie się jej wyznacznika do zera), co powoduje w konsekwencji, że błędy średnie szacunku poszczególnych parametrów są bardzo duże, przekraczające nawet, i to znacznie, wartości bezwzględne samych oszacowań. Takie wyniki numeryczne można więc uważać w gruncie rzeczy za bezużyteczne od strony poznawczej. Istnieje bowiem tendencja do nieuzasadnionego wykluczania niektórych istotnych zmiennych z modelu regresji⁴.

Poczynione uwagi i spostrzeżenia wskazują, że współliniowość jest zjawiskiem, którego nie można lekceważyć, jeśli chcemy w sposób możliwie dokładny określić przyczynowo-skutkowe uwarunkowania badanego zjawiska ekonomicznego i czynników je determinujących.

W tym kontekście względnie poprawne rozwiązanie problemu współliniowości wymaga:

- 1) dokonania pomiaru współliniowości,
- 2) określenia zakresu współliniowości na zbiorze zmiennych objaśniających,
- 3) poszukiwania dróg ograniczenia jej ujemnych skutków.

Najprostszą, ale i w pewnym sensie subiektywną metodą wykrywania zjawiska współliniowości jest porównanie wartości współczynników korelacji prostej dla każdej pary zmienionych objaśniających z pewną, z góry ustaloną, dopuszczalną wartością takiego współczynnika. W praktyce zakłada się najczęściej, że zmienne X_i i X_j są współliniowe, jeśli:

$$|r_{ij}| \geq 0,8.$$

Bardziej dokładną miarę efektu współliniowości można uzyskać obliczając wyznacznik macierzy $\mathbf{R}$, której elementami są współczynniki korelacji prostej między zmiennymi objaśniającymi. Wyznacznik ten, jak się okazuje, jest wielkością unormowaną i przyjmuje wartości z prze-

³ G. Kildiszew, A. Frenkel, *Analiza szeregów czasowych i prognozowanie*, Warszawa 1976, s. 99 - 100.

⁴ M. Gruszczyński, M. Kolupa, E. Leniewska, G. Napiórkowski, *Miary zgodności*, s. 115 - 116.

działu $\langle 0, 1 \rangle$, a o silnej współliniowości świadczy bliska zera wartość tego wyznacznika⁵.

Wyznacznik macierzy $\mathbf{R}$ może być również podstawą do testowania hipotezy odnośnie do występowania współliniowości. Wychodzi się wówczas z założenia, że współliniowość to zjawisko odchylania się zmiennych objaśniających od ortogonalności. Wtedy można dokonać weryfikacji hipotezy zerowej H_0 : zmienne objaśniające są niezależne, wobec hipotezy alternatywnej H_1 : zmienne są wzajemnie skorelowane. Sprawdzianem tak postawionej hipotezy zerowej jest statystyka:

$$\lambda^2 = -[n - 1 - \frac{1}{6}(2k + 5)] \cdot \ln |\mathbf{R}|,$$

która przy założeniu prawdziwości H_0 ma w przybliżeniu rozkład o $\frac{1}{2}k(k-1)$ stopniach swobody. Jeśli dla danego poziomu istotności α , $\lambda^2 \geq \lambda_{\alpha}^2$ to H_0 odrzucamy, co oznacza, że w zbiorze zmiennych objaśniających występuje współliniowość (zmienne nie są ortogonalne względem siebie)⁶.

⁵ Ponadto w literaturze podaje się wiele innych podejść do zagadnienia oceny stopnia współliniowości. I tak np. L. R. Klein, *Wstęp*, s. 96, stwierdza, że zjawisko to jest szkodliwe, gdy:

$$|r_{ij}| > R,$$

gdzie: r_{ij} — współczynnik korelacji prostej między zmiennymi objaśniającymi X_i i X_j , R — współczynnik korelacji wielokrotnej między Y a wszystkimi zmiennymi objaśniającymi.

Jako miarę siły współliniowości proponuje się również wyrażenie:

$$\eta = R^2 - \sum_{j=1}^p r_j^2,$$

gdzie: r_j — współczynnik korelacji prostej między zmienną objaśnianą a zmienną objaśniającą X_j .

Por. B. Schips, W. Stier, *Bestimmung der Auswirkung von Multikollinearität zwischen den erklärenden Variablen in Linearen Regressionsmodellen auf Kleinst-Quadrat-Schätzwerte durch Simulation*, Statistische Hefte 1971, nr 2, s. 114-126,

Y. Haitovsky z kolei przyjmuje, że współliniowość jest niebezpieczna jeśli:

$$|R_{ij}| \geq R,$$

gdzie: R_{ij} — współczynnik korelacji częściowej między zmiennymi X_i i X_j .

Par. Y. Haitovsky, *Multicollinearity in Regression Analysis, Comment*, Review of Economics and Statistics 1969, nr 3, s. 486 - 489.

Interesującą miarę natężenia współliniowości zaproponował H. Theil, *Zasady ekonometrii*, Warszawa 1979, s. 183. Jest nią wyrażenie:

$$\mu = R^2 - \sum_{j=1}^p (R^2 - \tilde{R}_j^2),$$

gdzie $\tilde{R}_j^2$ — współczynnik determinacji między zmienną Y a zmiennymi objaśniającymi bez zmiennej X_j .

⁹ D. E. Farrar, R. R. Glauber, *Multicollinearity in Regression Analysis, The Problem Revisited*, Review of Economics and Statistics 1967, nr 49, s. 92 - 107.

Stwierdzenie faktu występowania współliniowości jest dopiero pierwszym krokiem zmierzającym do rozwiązania problemu. W dalszym ciągu konieczne wydaje się wyodrębnienie tych zmiennych, które są współliniowe. Do tego celu można użyć zaproponowanego przez Farrara i Glaubera statystycznego testu istotności. Służy on do weryfikacji hipotezy $H_0: R_i^2 = 0$ wobec hipotezy alternatywnej $H_1: R_i^2 > 0$, a jako kryterium oceny zakresu współliniowości czynników-argumentów wprowadza się statystykę⁷:

$$F_i = (r_{ii} - 1) \frac{n - k}{k - 1},$$

gdzie:

R_i — jest współczynnikiem korelacji wielokrotnej między zmienną objaśniającą i a pozostałymi zmiennymi objaśniającymi,

r_{ii} — jest i -tym elementem diagonalnym macierzy $\mathbf{R}^{-1}$.

Statystyka ta, przy założeniu prawdziwości H_0 , ma rozkład F -Fischera o $n - k$ i $k - 1$ stopniach swobody. Zmienna X_i podlega wpływowi współliniowości, jeśli przy danym poziomie istotności α , $F_i \geq F_\alpha$. Postępowanie takie prowadzi w konsekwencji do wyodrębnienia takiego zespołu zmiennych, w stosunku do którego można twierdzić, że istotnie charakteryzuje się współliniowością.

W stosunku do tak wyróżnionych zmiennych można przeprowadzić z kolei badanie wzajemnych zależności poprzez zastosowanie testu t -Studenta, określonego poniższą funkcją testową⁸:

$$t_{ij} = \frac{R_{ij} \sqrt{n - k}}{\sqrt{1 - R_{ij}^2}},$$

gdzie: R_{ij} — jest współczynnikiem korelacji cząstkowej między zmiennymi X_i i X_j .

Określona w ten sposób funkcja testowa służy do sprawdzania hipotezy zerowej $H_0: R_{ij} = 0$ wobec hipotezy alternatywnej $H_1: R_{ij} \neq 0$ i przy założeniu prawdziwości H_0 ma ona rozkład t -Studenta o $n - k$ stopniach swobody. Jeśli dla danego α , $t_{ij} \geq t_\alpha$, to hipotezę o niezależności zmiennych X_i i X_j należy odrzucić. Powyższy test istotności, jak łatwo zauważyć, stanowi kryterium pomijania niektórych zmiennych, dla których stopień wewnętrznej korelacji jest wysoki. Należy go jednak stosować z pewną dozą ostrożności, odrzucenie bowiem pewnej liczby zmiennych zabija prowadzoną analizę pod względem merytorycznym. W praktyce nader często zdarza się, że nieistotność wyeliminowanych czynników jest rezultatem wysokiego skorelowania ich z innymi. Stąd można pominąć pewne zmienne, o dużym, choć pośrednim wpływie na zmienną objaśnianą. Postępowanie takie okazuje się jednak nieskuteczne, gdy wszy-

⁷ D. E. Farrar, R. R. Glauber, *Multicollinearity*, s. 92 - 107.

⁸ A. Zeliaś, *Badanie współliniowości w modelach ekonometrycznych*, Przegląd Statystyczny 1977, nr 2, s. 217 - 218.

stkie czynniki-argumenty objęte są współliniowością (a o tym informuje test F).

W celu ograniczenia ujemnych skutków współliniowości podejmowano próby wykorzystania w badaniach empirycznych szeregu metod np. standaryzacji zmiennych, selekcji zmieninych, metody głównych składowych, metody reszt czy też odchyleń od trendu⁹. Nie zawsze jednak przynosiły one spodziewane rezultaty. Metody te bowiem (oprócz metody głównych składowych) dawały jedynie możliwość pewnego ograniczenia współliniowości, co przy znacznym jej natężeniu i to w całym zespole zmiennych objaśniających z reguły nie było wystarczające i prowadziło do występowania w dalszym ciągu dużego skorelowania tych zmiennych między sobą. W metodzie głównych składowych natomiast zasadnicze trudności tkwią w sensownej interpretacji ekonomicznej użytych komponentów, gdyż są one zmiennymi, wprowadzicie ortogonalnymi względem siebie, ale również zmiennymi ukrytymi, bezpośrednio nieobserwowalnymi, którym nie można z reguły nadać żadnego miana.

Z punktu widzenia procesu identyfikacji znaczne korzyści może oddać, jak się okazało, tzw. metoda regresji grzbietowej (ridge regression)¹⁰. Z uwagi na to, że nie jest to metoda zbyt często stosowana w praktyce, konieczne wydaje się przedstawienie ramowego jej zarysu. Generalnie rzecz biorąc metoda regresji grzbietowej należy do metod regresji obciążonej, cechą których jest poszukiwanie takich systemów estymacji, które pozwalają otrzymywać bardziej efektywne estymatory kosztem ich obciążenia.

Metoda regresji grzbietowej (krawędziowej) opiera się na (przyjęciu zasady, że w przypadku współliniowości lepiej jest posługiwać się estymatorami obciążonymi, ale za to o wariancji mniejszej od wariancji odpowiednich estymatorów, danych klasyczną metodą najmniejszych kwadratów. Realizacja tej zasady polega na wprowadzeniu dodatkowej macierzy W , określonej jako:

$$W_{k,k} = \begin{bmatrix} w & 0 & 0 & \dots & 0 \\ 0 & w & 0 & \dots & 0 \\ 0 & 0 & w & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & w \end{bmatrix},$$

przy czym $w \in (0,1)$.

⁹ M. Gruszczyński, M. Kolupa, E. Leniewska, G. Napiórkowski, *Miary zgodności*, s. 204-205; A. Zeliaś, *Badanie współliniowości*, s. 217-225. Bardzo interesującą propozycję pozbywania się współliniowości zaproponował również Z. Hellwlg, *Zastosowanie przekształcenia ortogonalnego do wyznaczania dopuszczalnych wartości zmiennych objaśniających w modelach ekonometrycznych*, Przegląd Statystyczny 1974, nr 3, s. 307 - 315.

¹⁰ A. E. Hoerl, R. W. Kennard, *Biased Estimation for Nonorthogonal Problems*, Technometrics 1970, nr 12, s. 55-82; D. W. Marquardt, R. D. Snee, *Ridge Regression in practice*, American Statistician 1975, vol. 28/1, s. 3 - 20.

Przy tym założeniu estymator grzbietowy dany jest relacją:

$$\alpha_G = (X_*' X_* + W)^{-1} X_*' y_*,$$

gdzie:

X_* — macierz $(n \cdot k)$ obserwacji na zmiennych objaśniających w postaci znormalizowanej,

y_* — wektor $(n \cdot 1)$ obserwacji na zmiennej objaśnianej w postaci znormalizowanej,

α_G — wektor $(k \cdot 1)$ estymatorów parametrów regresji grzbietowej. Natomiast macierz wariancji-kowariancji tych estymatorów ma postać:

$$\text{var}(\alpha_G) = \sigma_u^2 (X_*' X_* + W)^{-1} X_*' X_* (X_*' X_* + W)^{-1},$$

gdzie σ_u^2 jest wariancją składnika losowego, której estymatorem jest wyrażenie:

$$s^2 = \frac{1}{n-k} (y_*' y_* - y_*' X_* \alpha_G).$$

Podstawową kwestią w metodzie regresji grzbietowej jest dobór określonych wartości w . Z reguły postępowanie w tym zakresie ma charakter iteracyjny i polega na kolejnym ustalaniu w począwszy np. od 0,05, a następnie badaniu, przy której kolejnej iteracji oceny parametrów się stabilizują. Jako kryterium tej stabilizacji można założyć, by różnice między nimi nie były większe od 0,01, to jest by

$$|a_{i,l+1} - a_{i,l}| \leq 0,01,$$

gdzie l jest numerem kolejnej iteracji¹¹.

Wykorzystanie metody regresji grzbietowej gwarantuje prawie całkowitą eliminację współliniowości, co z punktu widzenia wymogów procesu identyfikacji ma właśnie podstawowe znaczenie. W prezentowanym ujęciu metoda ta nie uwzględnia natomiast zróżnicowania wpływu współliniowości na wariancje estymatorów, gdyż do elementów głównej przekątnej macierzy $(X_*' X_*)$ dodaje się zawsze stałą wartość w . Pozbawia się w ten sposób pewnej interesującej informacji o działaniu zmiennych.

Warto również zauważyć, że przy zastosowaniu omawianej metody nie można przeprowadzić formalnego testu istotności zmiennych, ponieważ nie jest znany potrzebny rozkład zmiennej losowej (wprowadzanie macierzy W daje bowiem rozkład niecentralny, którego kształtu nie sposób oszacować). Z uwagi jednak na fakt, że wszystkie zmienne występują w postaci znormalizowanej¹², ich znaczenie w opisie badanego

¹¹ A. E. Hoerl, R. W. Kennard, *Biased*, s. 71 - 32.

¹² Normalizacja polega, jak wiadomo, na wystandaryzowaniu zmiennych za pomocą średniej arytmetycznej i odchylenia standardowego, wg wzoru:

$$x_{*i} = \frac{x_i - \bar{x}_i}{\sigma_{x_i}}.$$

zjawiska można określić bezpośrednio na podstawie bezwzględnych wartości stojących przy tych zmiennych objaśniających ocen parametrów regresji.

Zasygnalizowany problem współliniowości szczególnie ostro wystąpił w procesie dokonywania identyfikacji czynników określających poziom kosztów handlowych. Badania te prowadzone przez Zakład Statystyki i Demografii AE w Poznaniu objęły między innymi swym zakresem przedsiębiorstwa handlu detalicznego artykułami spożywczymi, działające w środowisku miejskim. Jako potencjalne czynniki kosztotwórcze wybrano zmienne reprezentujące różne aspekty działalności ekonomicznej przedsiębiorstw handlowych, a mianowicie:

- 1) rozmiary świadczonych usług, których symptomem jest wartość sprzedaży detalicznej, X_1 ,
- 2) czynnik pracy w postaci:
 - liczby zatrudnionych ogółem, X_4 ,
 - liczby pracowników operatywnych, X_5 ,
- 3) czynnik bazy materialno-technicznej, tj.
 - powierzchnię lokali sklepowych, X_6 ,
 - powierzchnię sal sprzedażowych, X_7 ,
 - liczbę sklepów detalicznych, X_2 ,
 - liczbę punktów sprzedaży, X_3 ,
- 4) czynnik kapitałowy, za którego symptom przyjęto wartość zapasu końcowego, X_8 ,

Przy wyborze takich a nie innych determinant kierowano się z jednej strony ogólnymi przesłankami natury ekonomicznej, z drugiej natomiast dostępnością odpowiednio wiarygodnego materiału empirycznego¹³. Otrzymano w ten sposób osiem cech wyjściowych, dla których (macierz korelacji prostej przedstawia tabela 1.

Tabela 1

Macierz korelacji między zmiennymi objaśniającymi poziom kosztów przedsiębiorstw handlu detalicznego artykułami spożywczymi w 1972 r.

Zmienne objaśniające	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	1,000	0,888	0,898	0,975	0,971	0,946	0,956	0,862
X_2	0,888	1,000	0,971	0,842	0,834	0,879	0,893	0,685
X_3	0,898	0,971	1,000	0,865	0,858	0,914	0,927	0,696
X_4	0,975	0,842	0,865	1,000	0,998	0,955	0,964	0,894
X_5	0,971	0,834	0,858	0,998	1,000	0,954	0,963	0,896
X_6	0,946	0,879	0,914	0,955	0,954	1,000	0,990	0,838
X_7	0,956	0,893	0,927	0,964	0,963	0,990	1,000	0,835
X_8	0,862	0,685	0,696	0,894	0,896	0,838	0,835	1,000

Źródło: Obliczenia własne wykonane w UOPI Akademii Ekonomicznej w Poznaniu.

¹³ Z. Zakrzewski, *Ekonomika handlu wewnętrznego*, Warszawa 1972, s. 372-395; E. Bogacka-Kisiel, *Czynniki kształtujące koszty przedsiębiorstw handlowych*, Zeszyty Naukowe WSE Wrocław, 1973, nr 67, s. 37.

Analizując tę tablicę widać wyraźnie, że stopień skorelowania przyjętych zmiennych jest bardzo wysoki, co sugeruje występowanie współliniowości zmiennych. Fakt ten został potwierdzony przez obliczony wyznacznik macierzy $\mathbf{R}$, który wyniósł

$$\det \mathbf{R} = 0,128876 \cdot 10^{-6},$$

a więc praktycznie nieistotnie różnił się od zera. Zastosowanie testu Farrara-Glaubera pozwoliło na zobiektywizowanie poczynionego sprostowania, gdyż okazało się, że

$$\lambda^2 = 859,425 > \lambda_{0,05}^{28} = 41,337,$$

co oznacza odrzucenie H_0 , a więc przyjęcie, że w zbiorze zmiennych objaśniających występuje zjawisko współliniowości. Okazało się również, że jej zakresem objęte były wszystkie przyjęte w badaniu zmienne objaśniające. Obiektywnej miary w tym zakresie dostarczył test F -Fischera (patrz tabela 2). Wobec tego zastosowano metodę regresji grzbietowej.

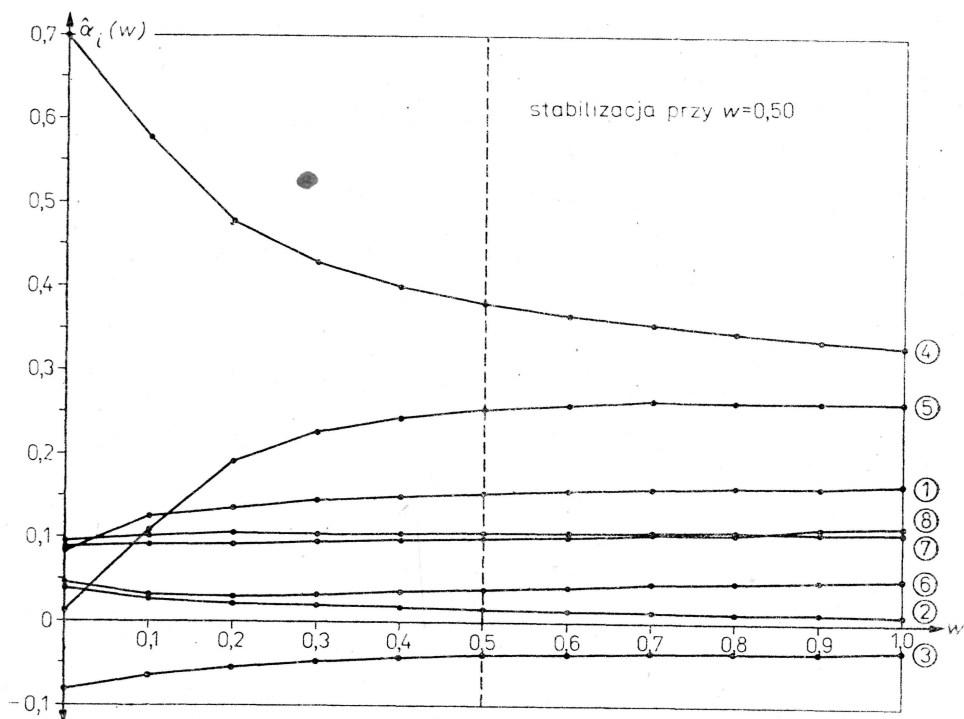
Tabela 2

Określenie zakresu współliniowości na zbiorze zmiennych objaśniających (test F)

Zmienne objaśniające	Wartość sprawdzianu hipotezy – F_t
Wartość sprzedaży detalicznej – X_1	211,4
Liczba punktów sprzedaży – X_2	135,8
Liczba sklepów detalicznych – X_3	200,2
Liczba zatrudnionych ogółem – X_4	3507,0
Liczba pracowników operatywnych – X_5	3080,0
Powierzchnia lokali sklepowych – X_6	369,6
Powierzchnia sal sprzedażowych – X_7	567,0
Wartość zapasu końcowego – X_8	34,3
Wartość krytyczna sprawdzianu hipotezy	

Źródło: Obliczenia własne na podstawie macierzy odwrotnej $\mathbf{R}^{-1}$

Jak się okazało, oceny parametrów ustabilizowały się dla wielkości $w=0,50$ (patrz wykres). Ich wartości przedstawia tabela 3. Jak wynika z badań najbardziej istotnymi zmiennymi kosztotwórczymi okazały się zmienne reprezentujące czynnik pracy (tj. liczba zatrudnionych ogółem i liczba pracowników operatywnych), na drugim miejscu występuje wartość sprzedaży detalicznej, a na trzecim powierzchnia sal sprzedażowych. Do grupy zmiennych o nieznacznym wpływie na koszty należą



Zmiany parametrów regresji grzbietowej przy różnych wartościach w dla badanych przedsiębiorstw handlowych

Tabela 3

Parametry opisowe modelu kosztów przedsiębiorstw handlu detalicznego artykułami spożywczymi oszacowane metodą regresji grzbietowej

Parametr α_i	Ocena parametru α_i	Błąd średni szacunku parametru $S(\alpha_i)$	$\frac{\alpha_i}{S(\alpha_i)}$
α_1	0,1525	0,0037	41,2162
α_2	0,0162	0,0025	6,4800
α_3	0,0426	0,0034	12,5294
α_4	0,3906	0,0027	144,6667
α_5	0,2483	0,0028	88,6785
α_6	0,0368	0,0039	9,4359
α_7	0,1060	0,0049	21,6327
α_8	0,0999	0,0011	90,8182
$n=57$		$w=0,50$	$S_u=0,1471$

Źródło: Obliczenia własne wykonane w UOPI Akademii Ekonomicznej w Poznaniu.

zmienne reprezentujące stan bazy materialno-technicznej (tj. stan sieci handlowej i powierzchnia lokali sklepowych ogółem).

Oceniając otrzymane wyniki z czysto merytorycznego punktu widzenia wydaje się, że mieszczą się one w ramach ogólnych rozważań teo-

retycznych z zakresu ekonomiki handlu. Czynnikiem pracy bowiem w działalności handlowej, a szczególnie w grupie branż spożywczych, jest decydujący dla realizacji sprzedaży detalicznej, a co za tym idzie, koszty osobowe stanowią zasadniczą część kosztów globalnych. W badanych przedsiębiorstwach ich udział stanowi od 35 do 40% kosztów globalnych. Duże znaczenie wartości sprzedaży detalicznej jest także zrozumiałe, gdyż zwiększenie jej rozmiarów pociąga za sobą wzrost ogólnej sumy kosztów, przyczynia się natomiast zwykle do spadku kosztu jednostkowego. Wiąże się to w głównej mierze z silnym oddziaływaniem sprzedaży na tzw. koszty zmienne. Znaczna rola powierzchni sal sprzedażowych związana jest z kolei ze specyfiką (branży. Przedmiotem obrotów w branżach spożywczych są artykuły o specyficznych właściwościach fizykochemicznych, szybko psujące się, często o krótkim okresie przydatności do spożycia. Składowanie tych towarów na zapleczu nie jest więc racjonalne, istotne jest natomiast właściwe ich eksponowanie w odpowiednio wyposażonych w środki techniczne salach sprzedażowych. W takim ujęciu powierzchnia sal sprzedażowych stosunkowo wiernie charakteryzuje daną jednostkę handlową, stając się synonimem majątku trwałego zaangażowanego w danej jednostce handlowej¹⁴. Specyfika branży stanowi również przyczynę niezbyt istotnego oddziaływania wartości zapasu końcowego na koszty badanych przedsiębiorstw handlowych. Przedmiotem obrotów w handlu artykułami spożywczymi są w znacznej mierze artykuły o dużej częstotliwości zakupu, zagwarantowanie więc ciągłości sprzedaży wymaga częstego uzupełniania zapasów, ale równocześnie ograniczenia do minimum czasu przechowywania większości towarów z uwagi na ich podatność na szybkie psucie się. Nie bez wpływu na nieistotny charakter wartości zapasu końcowego jako czynnika kosztotwórczego pozostał również fakt, że w badanym okresie system finansowy umożliwiał dość wysokie finansowanie zapasów środkami własnymi, bez konieczności korzystania z kredytów bankowych.

Podsumowując, uzyskane w toku przeprowadzonej analizy wyniki wydają się wskazywać na dużą użyteczność metody regresji grzbietowej do identyfikacji czynników kosztotwórczych. Nie należy jednak przeceniać zbytnio tego spostrzeżenia, z uwagi na stosunkowo wąski zakres jej praktycznego, jak dotychczas, wykorzystania. Dla potwierdzenia szerszych możliwości jej stosowalności są prowadzone przez Zakład Statystyki i Demografii AE w Poznaniu dalsze studia weryfikacyjne tej metody na autentycznym materiale empirycznym.

¹⁴ Należy podkreślić, że znaczenie powierzchni sal sprzedażowych zwiększa się w miarę rozpowszechniania nowych form sprzedaży detalicznej. Staje się ona wówczas dominującym czynnikiem, którego zadaniem jest możliwie najpełniejsza charakterystyka sieci detalicznej, W. Kozłowska, M. Kłoc, *Sposoby oceny wydajności sklepów*, Biblioteka IHW, Warszawa 1971, s. 7 - 10.

COLLINEARITY PHENOMENON IN STATISTICAL CAUSE-EFFECT
STUDIES OF ECONOMIC PROCESSES

Summary

The elaboration is devoted to the problems of variables collinearity and its repercussions for a study process of cause-effect conditions of socio-economic phenomena. The study is carried out with the use of statistical methods of interdependence analysis. It appears that collinearity is a crucial obstacle for correct estimation of mutual relations between economic variables.

The solution of that problem needs several measures. Their synthesis has been accomplished in the paper. There have been discussed the following problems: 1) measurement of collinearity (Glauber-Farrar test), 2) range of collinearity on a set of explanatory variables (F-test, t-test), 3) search for ways of reduction adverse effects of collinearity.

In the second part of the paper has been presented one of the methods of reducing collinearity i.e. the ridge regression method. That method is generally founded on such basis that it is better to use biased estimators but with variance smaller than variance of those estimators obtained with the help of the CLSM. The method has been implemented for an analysis of a real statistical date. The figures regarded costs and their determinants in state-owned foodstuffs retail trading enterprises.

The study confirmed large utility of the ridge regression method in a process of identification of cost-creating factors, but considering narrow scope of its hitherto applications further verification studies are indispensable.

Thumaczył Tadeusz Kowalski