

WOJCIECH SROKA
Uniwersytet Rolniczy
Kraków

ZASTOSOWANIE MODELI PRZYCZYNOWO-SKUTKOWYCH DO OCENY CZYNNIKÓW SUKCESU GOSPODARSTW ROLNICZYCH – PORÓWNANIE WYBRANYCH METOD

Zjawisko przyczynowości w naukach ekonomicznych

Jednym z podstawowych celów prowadzonych badań empirycznych jest wyizolowanie zależności przyczynowo-skutkowych pomiędzy zachodzącymi zjawiskami. Większość zjawisk w otaczającym nas świecie nie występuje w odosobnieniu, lecz w różnych związkach. Poznawanie zależności od wieków było i jest jedną z podstawowych metod poznawczych [17]. W naukach ekonomicznych szczególnie istotna wydaje się być problematyka zaszczości oraz zjawisk przyczynowości. Jest ona obecnie przedmiotem wielu dyskusji zarówno na poziomie metafizycznym, jak i modelowym. Problem definicji przyczyny, skutku i związku przyczynowego podejmowany był już przez mędrców starożytnych (Arystoteles, Demokryt z Abdery, Leukippos), jednak do dzisiaj nie ma zgody w tej kwestii. Niektórzy naukowcy, np. Hume, w ogóle wątpią w możliwość ustalenia, co jest przyczyną, a co skutkiem. W ekonomii niejednokrotnie takie rozgraniczenie jest arbitralne. Dotyczy to zarówno zagadnień makro-, jak i mikroekonomicznych [9]. Przykładowo, w badaniach mikroekonomicznych przyjmuje się, że dochód z gospodarstwa rolniczego determinuje jego powierzchnia. Może jednak relacja ta jest odwrotna, tzn. to dochody gospodarstwa decydują o jego powierzchni?

Niemniej jednak, jak píše Osińska [9], wielu ekonomistów w analizach przyczynowych stosuje podejście determinizmu kauzalnego, głoszące, iż każde zjawisko ma przyczynę. Istotne jest tutaj doświadczenie badacza i oparcie na wypracowanych teoriach. Nie należy bowiem zapominać, iż przyczyna wyprzedza skutek i ma większy związek z informacją zawartą w danych niż z tą, która pochodzi z teorii ekonomicznych. Na pierwszy plan wychodzi tutaj operacyjność analiz. Badacza interesuje bowiem predykcja, a znajomość przyczyny ma zwiększać możliwość przewidywania wyniku (skutku). Z uwagi, iż w ekonomii na każde zjawisko (skutek) oddziałuje jednocześnie wiele przyczyn, coraz większego znaczenia w badaniach ekonomicznych nabierają obecnie metody wielowymiarowe

pozwalające na przeprowadzenie różnego rodzaju analiz, w których ocenie przyczynowo-skutkowej poddaje się więcej niż jedną zmienną zależną.

Głównym celem artykułu jest zaprezentowanie wybranych metod wielowymiarowej analizy przyczynowo-skutkowej, a także przedstawienie i ocena przydatności metody równań strukturalnych oraz regresji wielorakiej w analizie czynników sukcesu gospodarstw rolniczych.

Model równań strukturalnych oraz regresja wieloraka jako przykłady modeli przyczynowo-skutkowych

Większość danych statystycznych ma charakter wielowymiarowy. Oznacza to, że zarówno obiekty zbiorowości, jak też badane zjawiska są opisywane za pomocą wielu różnych, zwykle w jakimś stopniu współzależnych cech. Wybór najważniejszych z nich jest niezwykle istotny, ale ich wielość oraz określenie siły wpływu na zmienną objaśnianą są często problemem dla badacza. W badaniach społeczno-ekonomicznych, jak i czasowych, obok prostych metod (w tym analizy regresji, korelacji oraz wariancji), coraz większego znaczenia nabierają metody analizy wielowymiarowej, wśród których należy wyróżnić: analizę głównych składowych, analizę czynnikową, analizę kanoniczną, sieci neuronowe oraz modele strukturalne [2]. W analizach przyczynowo-skutkowych szczególne znaczenie należy przypisać właśnie stosunkowo nowym modelom strukturalnym oraz znajdującej liczne zastosowania regresji wielorakiej.

Analiza regresji jest zapewne najszerzej stosowanym modelem statystycznym, pozwala bowiem na ocenę siły i formy związku między zmiennymi oraz na predykcję jednej zmiennej na podstawie zaobserwowanych wartości skorelowanych z nią innych zmiennych [6]. Regresja wieloraka jest naturalnym rozwinięciem modelu regresji prostej, w której wyróżnia się jedną zmienną zależną y oraz zbiór p zmiennych niezależnych x :

$$Y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_p x_{ip} + \varepsilon_i \quad (1)$$

W równaniu regresji Y_i jest wartością zmiennej wyjaśnianej dla obserwacji i ; x_{ij} ($j=1, 2, \dots, p$) są wartościami p zmiennych wyjaśniających dla obserwacji i ; ε_i jest błędem losowym obserwacji i ; b_0 i b_j są nieznanymi parametrami modelu. Aby jednak mówić o dobrze skonstruowanym modelu regresji wielorakiej, musi zostać spełnionych kilka ważnych warunków. Warunki te stawiane są zmiennym objaśniającym, jak i zmiennej objaśnianej. Najważniejsze warunki poprawnej konstrukcji równania modelu można określić następująco:

- zmienne objaśniające muszą wykazywać związek ze zmienną, którą będą objaśniały;
- zmienne objaśniające powinny cechować się odpowiednim wskaźnikiem własnej zmienności;
- zmienne objaśniające nie mogą być współzależne. Znaczy to tyle, że ich wzajemne współczynniki korelacji muszą wykazywać wartości mniejsze niż korelacji ze zmienną objaśnianą;

- liczba szacowanych parametrów (p) równania nie może przekroczyć liczby obserwacji (n).

Kluczowym założeniem regresji jest również brak błędu pomiarowego w zmiennych Y i X_j (co zasadniczo różni ją od modeli strukturalnych). Pozostałe założenia odnoszą się do czynnika losowego, tj. braku skorelowania składników losowych, ich rozkładu normalnego, wariancji składników losowych oraz ich wartości oczekiwanej [16]. W analizie regresji parametry modelu b_0 i b_j nie są znane i muszą być oszacowane w próbie danych. Są one tak dobierane, aby suma kwadratów reszt osiągnęła minimum. Ta najbardziej znana i stosowana metoda nosi nazwę metody najmniejszych kwadratów [8].

Bardzo ważne zagadnienie w analizie regresji stanowi ocena rozbieżności między wartościami zmiennej zależnej a wartościami wyliczonymi przez model. Jako miary tej rozbieżności stosowane są: odchylenie standardowe reszt oraz dużo bardziej popularny współczynnik determinacji. Współczynnik determinacji (R^2) jest to liczba z przedziału $\langle 0,1 \rangle$, oznaczająca zależność funkcyjną lub jej brak. Inaczej powiedziawszy, współczynnik determinacji mierzy, jaka część ogólnej zmienności zmiennej zależnej jest wyjaśniana regresją. Zatem im wyższa wartość R^2 , tym lepiej zmienne oddają istotę opisu badanego zjawiska. Niemniej jednak i ta miara nie jest pozbawiona wad, gdyż nie pozwala na porównywanie modeli z różną liczbą zmiennych objaśniających. Okazuje się bowiem, iż wraz z włączaniem do modelu kolejnych zmiennych, jego wartość rośnie [16]. W takiej sytuacji powinno stosować się skorygowany współczynnik determinacji.

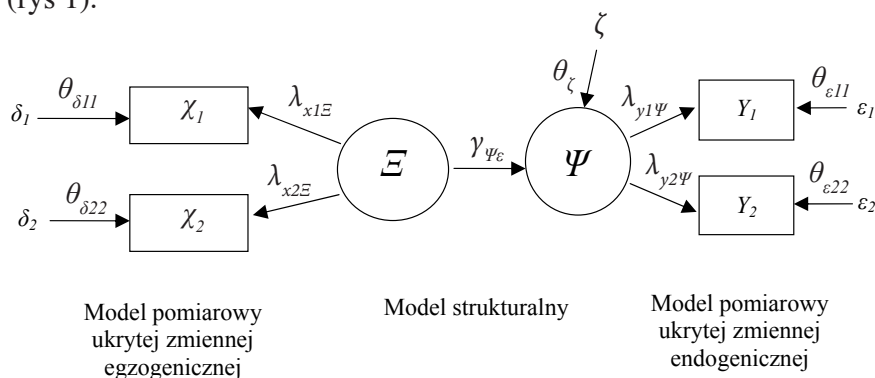
Każdy model należy poddać zarówno weryfikacji merytorycznej, jak i statystycznej. Ta pierwsza odnosi się do zgodności wyliczonych wartości (ich kierunku) z teorią wyjściową, natomiast druga – m.in. do istotności obliczonych współczynników. Pewnym ograniczeniem w analizach wielowymiarowych jest często trudność w interpretacji uzyskanych wyników, będąca konsekwencją stosowania różnych miar. Aby temu zaradzić, wykorzystywane są znormalizowane parametry równania regresji. Oblicza się je odejmując średnią arytmetyczną i dzieląc przez odchylenie standardowe [16].

Złożoność badanej rzeczywistości, w tym trudność przedstawiania pewnych konstruktów teoretycznych (np. sukcesu ekonomicznego) za pomocą jednej zmiennej endogenicznej, skłania do coraz powszechniejszego stosowania w naukach ekonomicznych zaawansowanych metod statystycznych. Do grupy takich technik należy stosunkowo mało rozpowszechniona w Polsce metoda równań strukturalnych [3]. Modele równań strukturalnych są zasadniczo podobne do modeli regresji wielorakiej, są jednak bardziej elastyczne w przypadku opisu interakcji między zmiennymi, zwłaszcza jakościowymi, dla których nie są spełnione klasyczne założenia regresji [12].

Modele te tworzą uogólniony model liniowy i zawierają poza regresją wieloraką także następujące procedury: analizę ścieżek, analizę czynnikową i analizę kowariancji. Model zasadniczo składa się z dwóch części: modelu pomiarowe-

go (pomiaru egzogenicznych oraz endogenicznych zmiennych ukrytych¹) oraz modelu równań strukturalnych. Model regresji stanowi podstawę ogólnego modelu linowego (*general linear model, GLM*) oraz modelu równań strukturalnych (*structural equation model, SEM*).

Fundamentalne narzędzie wykorzystywane podczas modelowania strukturalnego stanowi diagram ścieżkowy, przedstawiający w sposób graficzny związki przyczynowe pomiędzy poszczególnymi zmiennymi. W prezentacji graficznej diagramu zmienne jawne umieszcza się zazwyczaj w ramkach, zmienne ukryte w elipsie/kole, a zmienne reszkowe nie są oznaczane żadną figurą. Zależności między uwzględnionymi w modelu zmiennymi prezentowane są za pomocą strzałek, których grot wskazuje zmienną będącą skutkiem. W przypadku zmiennych reszkowych strzałka wskazuje na zmienną obserwowalną obciążoną błędem (rys 1).



Rys. 1. Diagram przykładowego modelu równań strukturalnych

Źródło: Opracowanie własne na podstawie: [7]

W tabeli 1 przedstawiono symbolikę stosowaną podczas modelowania równań strukturalnych. W modelu pomiarowym ukrytej zmiennej egzogenicznej jawne zmienne egzogeniczne (X_1, X_2), będące obserwowalnymi reprezentantkami ukrytej zmiennej egzogenicznej (Ξ), traktowane są jako zmienne dostarczające o nich informacji. Z tego względu przypisuje się im miano zmiennych wskaźnikowych. Występujące w modelu zmienne reszkowe (ζ_1, ζ_2) dla jawnych zmiennych egzogenicznych (X_1, X_2) reprezentują wpływ popełnionych błędów pomiarowych oraz innych nieuwzględnionych czynników.

Analogiczną konstrukcję posiada model pomiarowy ukrytej zmiennej endogenicznej, gdzie zmienna (Ψ) przedstawiana jest za pomocą jawnych zmiennych endogenicznych (Y_1, Y_2), będących jej reprezentantami. Podobnie jak w poprzednim modelu pomiarowym, zmienne reszkowe ($\varepsilon_1, \varepsilon_2$) informują o błędach pomiaru. Odzworowanie zależności między tak przedstawioną egzogeniczną (Ξ) a endogeniczną (Ψ) zmienną ukrytą, będące zasadniczym ele-

¹ Zmienna ukryta jest wielkością, której nie można bezpośrednio mierzyć. Zazwyczaj są to zmienne opisujące wszelakie konstrukty teoretyczne, tj. konstruowane przez badacza desygnaty pojęć teoretycznych.

mentem modelowania strukturalnego, dokonywane jest przy użyciu modelu strukturalnego [7]. Z punktu widzenia celu modelowania, to właśnie model strukturalny dostarcza najważniejszych informacji.

Tabela 1

Zmienne i parametry modelu równań strukturalnych	
Symbol	Znaczenie
Zmienne modelu	
Ψ (Psi)	Ukryta zmienna endogeniczna
ζ (Zeta)	Zmienne resztkowe (błędu) dla ukrytych zmiennych endogenicznych
Y (Ypsilon)	Jawne zmienne endogeniczne, zmienne pomiarowe dla ukrytych zmiennych endogenicznych
ε (Epsilon)	Zmienne resztkowe (błędu) dla jawnych zmiennych endogenicznych
Ξ (Ksi)	Ukryte zmienne egzogeniczne
χ (Chi)	Jawne zmienne egzogeniczne, zmienne pomiarowe dla ukrytych zmiennych egzogenicznych
Parametry modelu	
γ (Gama)	Współczynniki reprezentujące wpływ ukrytych zmiennych egzogenicznych na ukryte zmienne endogeniczne
λ (Lambda)	Ładunki czynnikowe określające siłę związku między zmiennymi ukrytymi i jawnymi
θ (Theta)	Korelacje/kowariancje między zmiennymi błędu a zmiennymi jawnymi lub ukrytymi

Źródło: Opracowanie własne na podstawie: [12].

W procesie analizy jawne zmienne egzogeniczne oraz endogeniczne często poddawane są procedurze standaryzacji. Dla standaryzowanych wartości zmiennych wskaźnikowych układ równań modelu strukturalnego zaprezentowanego na rysunku 1 przybiera następującą postać:

$$\Psi_k = \gamma_{\Psi\Xi} \zeta_k + \theta_{\zeta} \zeta_k \quad \text{„model strukturalny”}$$

$$\left. \begin{aligned} u_{k\chi 1} &= \lambda_{\chi 1\Xi} \zeta_k + \theta_{\delta 11} \delta_{k1} \\ u_{k\chi 2} &= \lambda_{\chi 2\Xi} \zeta_k + \theta_{\delta 22} \delta_{k2} \end{aligned} \right\} \text{„model pomiarowy ukrytej zmiennej egzogenicznej”}$$

$$\left. \begin{aligned} u_{kY 1} &= \lambda_{Y 1\Psi} \Psi + \theta_{\varepsilon 11} \varepsilon_{k1} \\ u_{kY 2} &= \lambda_{Y 2\Psi} \Psi + \theta_{\varepsilon 22} \varepsilon_{k2} \end{aligned} \right\} \text{„model pomiarowy ukrytej zmiennej endogenicznej”}$$

Parametr k określa numer obserwacji, a zmienne ukryte traktowane są jako zmienne posiadające „ukrytą” wartość cechy. Estymacja parametrów określających siły i kierunki poszczególnych związków dokonywana jest w modelu równań strukturalnych na podstawie informacji empirycznych. Najczęściej informa-

cje te zapisywane są w postaci macierzy korelacji. Istota estymacji parametrów polega na znalezieniu takich wartości parametrów, dla których wyznaczone korelacje teoretyczne przyjmują możliwie najbliższe wartości korelacji empirycznych. W praktyce estymacja odbywa się w drodze rozwiązania układu równań jedną z metod iteracyjnych. W opracowaniu korzystano z uogólnionej metody najmniejszych kwadratów, a następnie metody największej wiarygodności [15].

Następnym etapem budowy modelu jest jego weryfikacja, podczas której dokonuje się analizy sensowności oszacowanych parametrów, a następnie weryfikacji dobroci modelu. W tym miejscu należy wspomnieć o roli stopni swobody w identyfikacji modeli strukturalnych. Stopnie swobody są definiowane jako liczba unikalnych korelacji lub kowariancji w macierzy wyjściowej danych, pomniejszonych o liczbę szacowanych współczynników [1]. Badacz najczęściej dąży do maksymalizacji stopni swobody, przy jednoczesnym dobrym dopasowaniu modelu do danych. Każdy kolejny oszacowany parametr „zużywa” stopnie swobody, a ich liczba nie może być niższa niż zero [11]. Najczęściej spotykanym wskaźnikiem dopasowania modelu do danych rzeczywistych jest wartość statystyki χ^2 . Test ten dotyczy hipotezy zerowej, głoszącej, że reszty standaryzowane macierzy empirycznej i teoretycznej wynoszą 0, tzn. że ograniczenia nałożone przez badacza wynikające z przyjętego modelu są trafne [12]. Oceniając dobroć modelu, można posłużyć się również indeksem pierwiastka średniokwadratowego błędu aproksymacji RMSEA (wartości mniejsze niż 0,05 informują o dobrym dopasowaniu). Oprócz tych wskaźników na uwagę zasługują indeks GFI Joerskoga, indeks dobroci dopasowania AGFI Joerskoga oraz unormowany indeks Bentlera-Bonetta, określające stopień dopasowania empirycznej i teoretycznej macierzy korelacji. Przyjmuje się, iż dany model jest dobrze dopasowany, gdy wskazane indeksy przyjmują wartość powyżej 0,9. W literaturze można znaleźć również inne wskaźniki pozwalające ocenić dobroć dopasowania modelu, jednak ich zastosowanie jest dużo węższe [7].

Regresja wieloraka oraz metoda równań strukturalnych w ocenie czynników sukcesu ekonomicznego gospodarstw – praktyczne wykorzystanie

Podstawą analizy były oryginalne dane źródłowe zebrane za pomocą kwestionariusza wywiadu standaryzowanego. Badaniem objęto 179 gospodarstw rodzinnych prowadzących produkcję w 6 mezoregionach geograficzno-gospodarczych Karpat Polskich, tj. Bieszczadach, Beskidzie Niskim, Beskidzie Sądeckim, na Podhalu, Beskidzie Makowskim oraz Beskidzie Żywieckim. Gospodarstwa zostały dobrane na podstawie wcześniej sformułowanych kryteriów, które miały wyselekcjonować podmioty przodujące, tj. najlepsze w danym mezoregionie, osiągające dochody pozwalające na godziwe opłacenie pracy rolnika oraz jego rodziny, a także dodatnią akumulację kapitału².

² Gospodarstwa przodujące dobrano metodą ekspercką. Na podstawie obiektywnie sformułowanych kryteriów pracownicy ośrodków doradztwa rolniczego wytypowali podmioty, które można uznać za przodujące.

Do wyznaczenia listy zmiennych objaśniających sukces ekonomiczny gospodarstw oraz określenia ich wzajemnych zależności wykorzystano zarówno rozważania teoretyczne, jak również badania empiryczne, przeprowadzone przez innych autorów [15, 10, 13]. Do wstępnej analizy zakwalifikowano możliwie szeroki wachlarz zmiennych, przy czym założono, że każdy ze zidentyfikowanych czynników sukcesu (zasoby ziemi, zasoby pracy, intensywność organizacji, powiązanie z rynkiem, transfery budżetowe, akumulacja, kapitał ludzki, chłonność na postęp³) będzie opisywany za pomocą co najmniej 2 zmiennych jawnych. Dobór zmiennych, szczególnie w przypadku zastosowanego podejścia confirmacyjnego, jest bardzo ważnym zagadnieniem, ponieważ wyodrębnione zmienne pierwotne decydują o jakości analiz. Jak podaje Zakrzewska [19], *dobre zmienne to zmienne proste i łatwo interpretowalne*. Zmienne mogą przyjmować charakter miar względnych i bezwzględnych, przy czym często podkreśla się wyższość tych pierwszych, tj. wskaźników struktury i natężenia. W przypadku analizy gospodarstw rolniczych, stosowanie miar względnych umożliwia analizę podmiotów o różnej wielkości bez preferowania tych największych, stąd w analizie wykorzystano głównie różnego rodzaju wskaźniki natężenia i struktury. Pierwotna lista zmiennych była znacznie szersza, lecz część zmiennych nie spełniała zakładanych warunków, bądź niedostatecznie dobrze wyjaśniała badane zjawisko. W tabeli 2 przedstawiono więc wyłącznie zmienne wykorzystane w modelach.

Chcąc dokonać kwantyfikacji zmiennej „sukces ekonomiczny”, sięgnięto do literatury, według której termin ten (z łac. *successus*) oznacza pomyślny wynik jakiegoś przedsięwzięcia – triumf. Przez sukces można również rozumieć osiągnięcie takiego poziomu celów, by były one oceniane pozytywnie przez opiniodawców (kierowników gospodarstw). W przypadku rolnictwa większość naukowców jest zgodna, iż głównym celem gospodarstw rolniczych staje się maksymalizowanie różnych kategorii nadwyżki ekonomicznej, w tym przede wszystkim dochodu rolniczego [4, 18, 20].

W skali mikroekonomicznej dochody stanowią cele gospodarowania, gdyż wpływają na poziom życia rodzin chłopskich, a także określają rozmiary popytu na dobra i usługi pozarolnicze [18]. Kategoria ta może jednak wykazywać duże wahania w czasie i nie przedstawia w sposób pełny stopnia realizacji celów gospodarowania (np. podnoszenia statusu ekonomicznego). Woś [18] twierdzi, iż w skali mikroekonomicznej o poziomie rozwoju gospodarstwa, a więc także o pewnym statusie ekonomicznym rodziny rolniczej, informuje również wartość majątku. Obecnie przyjmuje się, że głównym celem działania przedsiębiorstwa jest maksymalizacja jego wartości; w przypadku gospodarstwa rolniczego będzie się to przejawiać m.in. wzrostem wartości majątku ogółem (nie tylko produkcyjnego). Można zatem założyć, iż zdematerializowanym wymiarem sukcesu ekonomicznego będzie wartość majątku gospodarstwa. W niniejszej analizie do estymacji poszczególnych parametrów oraz oceny stopnia dopasowania poszczególnych modeli zastosowano pakiet

³ Szerzej w publikacji, która ukazała się na łamach *Zagadnień Ekonomiki Rolnej* w 2010 roku [14].

Modelowanie równań strukturalnych (SEPATH) oraz regresji wielorakiej, zawarty w programie Statistica 10.0.

Do badania siły i kierunku wpływu poszczególnych zmiennych na sukces ekonomiczny gospodarstw zastosowano podejście zasobowe. Według tej koncepcji, o sukcesie będzie decydować potencjał ekonomiczny gospodarstw w postaci odpowiednio dobranych zasobów, ich konfiguracji, a także zdolności do ich wykorzystania. Przeprowadzona analiza dotyczy zarówno zasobów materialnych, tj. powierzchni gospodarstwa, zasobów pracy itp., jak też niematerialnych, obrazujących m.in. powiązanie gospodarstwa z rynkiem, intensywność produkcji itp.

Tabela 2

Specyfikacja zmiennych użytych w analizie czynników sukcesu gospodarstw

Egzogeniczne zmienne jawne (obserwowalne)	Endogeniczne zmienne jawne (obserwowalne)
Zasoby pracy własnej (FWU/gosp.)	Wartość dochodu z gosp. rolniczego (zł/gosp.)
Zasoby pracy własnej (FWU/100 ha UR)	Wartość dochodu z gosp. rolniczego na osobę pełnozatrudnioną (zł/1 FWU)
Powierzchnia UR	Wartość środków trwałych (zł/gosp.)
Powierzchnia gruntów leśnych	
Wartość produkcji towarowej (zł/1 ha UR)	
Wartość produkcji kontraktowanej (zł/ ha UR)	
Inwestycje netto (zł/ 1000 zł wartości środków trwałych)	
Stopa inwestycji (relacja wartości inwestycji brutto do dochodu z gospodarstwa rolniczego powiększonego o wartość amortyzacji) (%)	

Źródło: Opracowanie własne.

Przeprowadzone analizy miały na celu przede wszystkim wskazanie ograniczeń i słabości stosowania regresji wielorakiej w porównaniu do metody równań strukturalnych, stąd niektóre aspekty (np. dotyczące analizy reszt, efektu katalizy, koincydencji itp.) nie zostały omówione. Niemniej jednak, przy budowie modeli zostały one uwzględnione, podobnie jak wszystkie z warunków stawianych zmiennym objaśniającym i objaśnianym. Z badanej próby wyłączono również obserwacje odstające oraz wykazujące braki danych.

Wyboru postaci analitycznej modeli i zmiennych objaśniających dokonano na podstawie statystycznej analizy związków między nimi. Dobór zmiennych przeprowadzono za pomocą metody regresji krokowej „w przód”. Zakłada ona kolejne (krokowe) dołączanie do listy tych zmiennych objaśniających uwzględnionych w modelu, które mają najistotniejszy wpływ na zmienną zależną.

W tabeli 3 przedstawiono wyniki modelowania dla trzech zmiennych objaśnianych, opisujących sukces ekonomiczny gospodarstw. W każdym przypadku wybrano równanie, które charakteryzuje się najlepszym dopasowaniem oraz istotnością wszystkich parametrów (zadowalający poziom p). Uzyskane modele poddano weryfikacji merytorycznej i statystycznej. Wyniki badań wskazują, iż modele cechuje istotność statystyczna, o czym świadczą zadowalające wartości statystyki F . Ponadto, wartości współczynników determinacji wskazują na dość dobre dopasowanie przedstawionych modeli.

Tabela 3

Wyniki oszacowania modeli wpływu poszczególnych zmiennych na kategorie obrazujące sukces ekonomiczny gospodarstw

Zmienne objaśniające	Zmienna objaśniana		
	Y_1 (dochód z gosp. rolniczego)	Y_2 (dochód z gosp. rolniczego/FWU)	Y_3 (wartość majątku gospodarstwa)
	Wartość współczynnika β		
X_1 (pow. UR)	0,723 ($p=0,0002$)	0,450 ($p=0,0001$)	0,281 ($p=0,0241$)
X_2 (FWU/100 ha)	- 0,149 ($p=0,0001$)	- 0,423 ($p=0,0001$)	- 0,441 ($p=0,0003$)
X_3 (wartość dopłat ogółem/ha UR)	0,260 ($p=0,0001$)	0,263 ($p=0,0002$)	0,251 ($p=0,0126$)
X_4 (wartość prod. towarowej/ha UR)	0,273 ($p=0,003$)	0,304 ($p=0,0001$)	-
X_5 stopa inwestycji	-	-	0,131 ($p=0,0002$)
X_6 DJP / 100 ha UR	-	0,142 ($p=0,0484$)	0,364 ($p=0,003$)
R	0,835	0,784	0,751
R^2	0,797	0,715	0,524
Skorygowany R^2	0,789	0,704	0,504
Statystyka F	(4,163) 75,336	(5,163) 52,125	(5,163) 33,991

Źródło: Opracowanie własne.

W pierwszym modelu, opisującym sukces za pomocą zmiennej „dochód z gospodarstwa rolniczego”, zmienną w największy sposób nań oddziałującą okazała się powierzchnia użytków rolnych ($\beta=0,723$). Poza tym istotne okazały się również: wartość produkcji towarowej ($\beta=0,273$), wartość dopłat ogółem/ha UR ($\beta=0,260$) oraz zasoby pracy ($\beta=-0,149$). Pierwsze trzy zmienne wykazywały wartość dodatnią, co oznacza, iż wzrost którejkolwiek z nich o jedno odchylenie standardowe prowadził do wzrostu dochodu z gospodarstwa rolniczego o β jej odchylenia standardowego. Odwrotną sytuację zauważono w przypadku zmiennej zasoby pracy (FWU/100 ha UR) – ich zmniejszenie powinno przyczynić się do zwiększenia dochodu z gospodarstwa rolniczego. Ograniczając się wyłącznie do interpretacji kluczowej zmiennej, tj. powierzchni UR, należy

wskazać, iż jej wzrost o jedno odchylenie standardowe będzie prowadzić do wzrostu dochodu z gospodarstwa rolniczego o 0,723 odchylenia standardowego. Podobnie interpretowane są pozostałe zmienne.

Oceniając drugi z modeli, tj. obrazujący sukces gospodarstwa za pomocą zmiennej „dochód z gospodarstwa rolniczego w przeliczeniu na jedną osobę rodzinnej siły roboczej”, należy podkreślić dobre dopasowanie modelu oraz zadowalającą wartość statystyki F . Również w przypadku tego równania zmienną mającą największy wpływ na zmienną objaśnianą była powierzchnia użytków rolnych ($\beta=0,450$), jednak w stosunku do poprzedniego modelu wyraźnie straciła na znaczeniu. Istotnymi okazały się również zmienne X_2 , X_3 , X_4 oraz X_6 . Na uwagę zasługuje wysoka ujemna wartość zmiennej X_2 – FWU/100 ha UR, która informuje, iż zmniejszenie zasobów pracy o jedno odchylenie standardowe będzie prowadzić do wzrostu dochodu przeliczanego na 1 osobę zatrudnioną o 0,423 odchylenia standardowego.

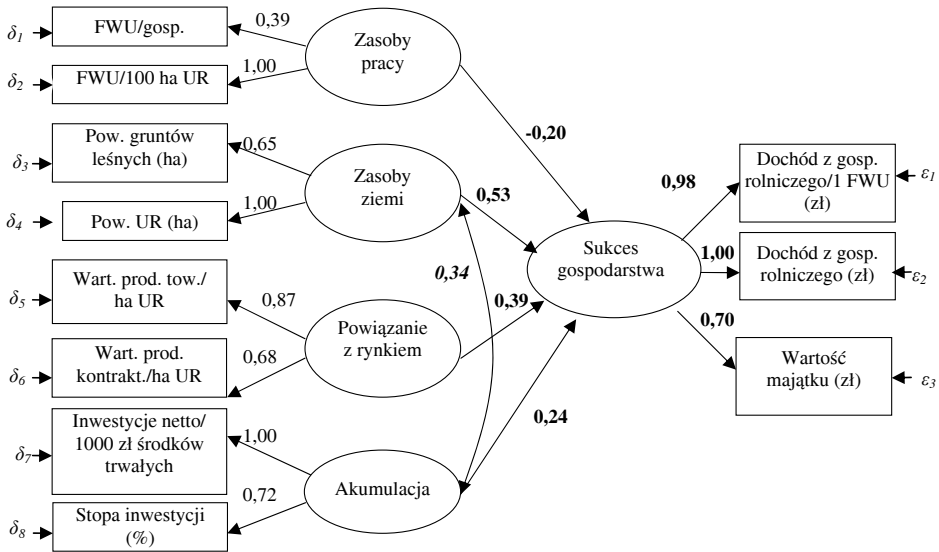
Trzeci model, w którym sukces ekonomiczny przedstawiony jest z wykorzystaniem zmiennej „majątek gospodarstwa”, wykazuje najslabsze dopasowanie, jednak poszczególne mierniki dobroci dopasowania modelu do danych rzeczywistych są wystarczająco wysokie. Najwyższą wartość wśród zmiennych objaśniających miała zmienna zasoby pracy (FWU/100 ha UR), nieco mniejszą obsada zwierząt (DJP/100 ha UR), a najmniejszą stopa inwestycji.

Podsumowując wyniki poszczególnych modeli, należy podkreślić, iż zmienne objaśniające dobrze wyjaśniały zmienność zmiennej zależnej, a wskazane zależności były poprawne merytorycznie. Z przeprowadzonych analiz wynika, że na sukces ekonomiczny wpływało wiele z zaproponowanych zmiennych, jednak ich zestaw i siła oddziaływania były zróżnicowane w zależności od przyjętej zmiennej objaśnianej. Powstaje więc pytanie, która zmienna najlepiej obrazuje sukces ekonomiczny? W literaturze przedmiotu można znaleźć różne odpowiedzi, przy czym autorzy podkreślają subiektywność przy ich wyborze. Dobrym rozwiązaniem wydaje się więc ujęcie wszystkich zmiennych łącznie, co możliwe jest przy wykorzystaniu modelu równań strukturalnych.

Zasadniczą i chyba najtrudniejszą częścią modelowania strukturalnego jest budowa modelu teoretycznego. Jest to szczególnie ważne, gdyż przyjęcie nieprawdziwych lub nieadekwatnych założeń spowoduje, iż badacz wyciągnie pochopne lub błędne wnioski. Stąd w pierwszej fazie badań na podbudowie teoretycznej testowano szereg układów wyodrębnionych zmiennych wskaźnikowych dla różnej liczby zmiennych ukrytych. W konsekwencji wyspecyfikowano model, który w najlepszy możliwy sposób odzwierciedla badaną rzeczywistość. Przedstawiony model bazuje na czterech ukrytych zmiennych egzogenicznych, opisujących czynniki sukcesu, oraz jednej ukrytej zmiennej endogenicznej, tj. sukcesie ekonomicznym (rys. 2). Istota metody SEM polega na odzwierciedleniu wzajemnych zależności pomiędzy bezpośrednio nieobserwowalnymi zmiennymi ukrytymi, poprzez zastosowanie do tego celu ich obserwowalnych reprezentantek, tj. odpowiednich wskaźników pomiarowych. O badanych czynnikach sukcesu gospodarstw należy mówić jako o zmiennych bezpośrednio

niemierzalnych, a za ich wskaźniki pomiarowe należy przyjąć przedstawione wcześniej zmienne pierwotne (wskaźnikowe).

Cechą charakterystyczną, a zarazem dużą zaletą technik modelowania zaliczanych do SEM jest to, że uwzględniają występowanie pojęć bezpośrednio niemierzalnych, nazwanych także ukrytymi. Ułatwia to znacząco kwantyfikowanie poszczególnych czynników sukcesu, jak też kategorii „sukces gospodarstwa”. Dzięki temu możliwe jest pełniejsze wnioskowanie na temat wpływu np. czynnika ziemi na sukces ekonomiczny. Wydaje się bowiem, że praktykowane często w literaturze sprowadzanie czynnika ziemi do jednej zmiennej – powierzchni UR – jest dużym uproszczeniem rzeczywistości. Dotyczy to także kategorii „sukces gospodarstwa”. Wyniki analizy czynników sukcesu gospodarstw za pomocą metody regresji jasno ukazują, iż w zależności od przyjętej zmiennej zależnej wyniki są różne.



Rys. 2. Oszacowany model czynników sukcesu gospodarstw

Źródło: Opracowanie własne.

W przedstawionym modelu występowały relatywnie silne oraz statystycznie istotne związki między ukrytymi zmiennymi egzogenicznymi a ich zmiennymi wskaźnikowymi. W przypadku zmiennej „zasoby pracy”, największe znaczenie dla jej opisu miały zasoby pracy wyrażone w FWU/100 ha UR (1,00), mniejsze natomiast – zasoby przeliczone na gospodarstwo (FWU/gosp.) (0,39). Wartość współczynnika ścieżkowego równa 1,0 jest stałą przyjętą przez model i oznacza, że dana zmienna była w 100% wyjaśniana przez zmienną „zasoby pracy”. W konsekwencji, wartość ładunku czynnikowego drugiej ze zmiennych była równa współczynnikowi korelacji z tą pierwszą. Przyjęcie wartości stałej dla którejkolwiek zmiennej nie obniża jednak wartości modelu. Analizując zależności między zmienną „zasoby ziemi” a jej zmiennymi wskaźnikowymi, zauwa-

żono, iż największe znaczenie należało przypisać powierzchni użytków rolnych (1,00). Nieco mniejsze znaczenie w badanych gospodarstwach miała natomiast powierzchnia gruntów leśnych⁴ (0,65). W przypadku pozostałych zmiennych ukrytych („powiązanie z rynkiem” oraz „akumulacja”) należy stwierdzić, iż były one mocno skorelowane z ich wskaźnikami pomiarowymi. Interpretacja tych zależności jest identyczna jak w przypadku omówionych już zasobów pracy i ziemi, zrezygnowano zatem ze szczegółowego ich opisu.

Analiza wykazała, iż model pomiarowy zmiennej endogenicznej, tj. kategorii „sukces gospodarstw”, był również dość dobrze definiowany przez zmienne wskaźnikowe, co świadczy o dobrym stopniu wyjaśniania ich zmienności. Największe znaczenie dla jej opisu miał dochód z gospodarstwa rolniczego (1,00), nieco mniejsze dochód z gospodarstwa rolniczego w przeliczeniu na 1 osobę pełnozatrudnioną (0,94), najmniejsze zaś – wartość majątku (0,65).

Analizując główny problem modelowania, tj. wpływ ukrytych zmiennych egzogenicznych (czynników sukcesu) na zmienną endogeniczną („sukces gospodarstw”), należy stwierdzić, iż w sposób istotny na sukces gospodarstw oddziaływały „zasoby pracy”, „zasoby ziemi”, „powiązanie z rynkiem” oraz „akumulacja”. Wyniki modelowania wskazują, że na sukces gospodarstw zmienne te mają zróżnicowany wpływ. Determinantą sukcesu, która miała największe znaczenie, okazały się „zasoby ziemi” (wzrost tej zmiennej o jedno odchylenie standardowe spowoduje wzrost zmiennej „sukces gospodarstw” o 0,53 jej odchylenia standardowego), nieco mniejsze znaczenie dla rozwoju miały „powiązanie z rynkiem” (0,39), „akumulacja” (0,24) oraz „zasoby pracy” (-0,20). Ujemna wartość zmiennej „zasoby pracy” informuje, że zmniejszanie zasobów pracy będzie przyczyniać się do sukcesu gospodarstw.

Wobec merytorycznie zasadnych i istotnych wyników modelowania postanowiono przejść do następnego etapu, tj. weryfikacji oszacowanego modelu. W świetle wartości Chi-kwadrat i poziomu prawdopodobieństwa stwierdzić można było niedostateczne dopasowanie macierzy teoretycznej do wyników empirycznych. Wskaźnik Chi-kwadrat uplasował się na poziomie 88,13, przy prawdopodobieństwie $p=0,0079$. Nie należy jednak poprzestać na analizie dopasowania modelu na podstawie jedynie wartości Chi-kwadrat [12]. Analiza wartości absolutnych wskaźników dopasowania, tj. GFI (0,94) oraz AGFI Joreskoga (0,89), wykazała akceptowalny poziom dopasowania modelu. Podobnie wartość wskaźnika RMSEA Steigera-Linda (0,09) oraz unormowanego indeksu Bentlera-Bonetta (0,84) pozwalała na uznanie modelu za w miarę dobrze dopasowany.

⁴ Większość z badanych gospodarstw uzyskiwała również przychody ze sprzedaży drewna, stąd do analiz włączono także powierzchnię gruntów leśnych.

Podsumowanie i wnioski

Modele równań strukturalnych są zasadniczo podobne do modeli regresji wielorakiej, jednak w przeciwieństwie do nich mogą one być stosowane w opisie zależności nieliniowych, skorelowanych zmiennych niezależnych, skorelowanych składników losowych oraz zmiennych ukrytych, będących wynikiem oddziaływania wielu zmiennych wskaźnikowych. Ułatwia to znacząco kwantyfikowanie kategorii bezpośrednio niemierzalnych oraz złożonych, takich jak „sukces gospodarstw”. Wadą metody równań strukturalnych jest mocno rozbudowany aparat obliczeniowy, czego konsekwencją jest brak możliwości „objęcia” przez badacza wszystkich operacji obliczeniowych. W porównaniu do metody regresji, w SEM mierniki dopasowania modelu nie dają również jednoznacznej oceny jakości modelu, a także nie można ich porównywać z innymi miarami, np. R^2 . W efekcie, modele strukturalne w dużej mierze oparte są na założeniach poprawności modelu, a standardowa teoria statystyczna nie pozwala na test tych założeń. Należy zwrócić także uwagę na dużą czasochłonność prowadzonych analiz oraz trudności z raportowaniem i interpretacją wyników. Mimo tych zastrzeżeń, obecnie uznaje się, że SEM jest jedną z najbardziej obiecujących technik statystycznych, umożliwiającą pełniejszą analizę badanych zjawisk i relacji między zmiennymi.

Wyniki przeprowadzonych badań potwierdzają dużą przydatność zastosowanego aparatu badawczego, przy czym należy zwrócić uwagę, iż w zależności od sposobu pomiaru zmiennej endogenicznej „sukces gospodarstw”, wyniki modelowania były zróżnicowane, tj. poszczególne zmienne objaśniające w różnym stopniu wpływały na zmienną objaśnianą. W przypadku modelu regresji wykazano, iż w sytuacji, gdy „sukces gospodarstw” będzie przedstawiany za pomocą zmiennej „dochód z gospodarstwa rolniczego” lub „dochód z gospodarstwa rolniczego w przeliczeniu na osobę pełnozatrudnioną”, za najważniejszy czynnik sukcesu gospodarstwa należało uznać powierzchnię użytków rolnych, natomiast jeśli „sukces” będzie mierzony poziomem majątku gospodarstwa, to najważniejszym czynnikiem rozwoju okazują się zasoby pracy. Metoda ta nie pozwala więc na jednoznaczne wskazanie głównej determinanty sukcesu gospodarstw.

Łączne ujęcie wymienionych powyżej zmiennych objaśnianych oraz zastosowanie ich jako wskaźników pomiarowych ukrytej zmiennej „sukces gospodarstw” znacznie zobiektywizowały prowadzone rozważania. Wyniki modelu równań strukturalnych wskazują, iż głównym czynnikiem sukcesu gospodarstw była ich powierzchnia, a następnie kolejno powiązanie z rynkiem, akumulacja oraz zasoby pracy. Ujęcie większej liczby zmiennych oraz wystarczająca jakość modelu wskazują na wyższość modeli równań strukturalnych (nad metodą regresji) w określaniu zależności przyczynowo-skutkowych.

Literatura:

1. Aaker D., Bagozzi R.: Unobservable variables in Structural Equation Modeling with application in industrial selling. *Journal of Marketing Research*, vol. XVI, 1979.
2. Balicki A.: Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne. Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk 2009.
3. Cwalina W.: Zastosowanie modelowania równań strukturalnych w naukach społecznych [w:] *Czytelnia Statsoft*. Wydawnictwo Statsoft Polska, Kraków 2000.
4. Czyżewski A., Wiśniewska J.: Metody pomiaru dochodów rolniczych w krajach UE. *Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu*, nr 1, Poznań 1996.
5. Dautzenberg K.: Erfolgsfaktoren von landwirtschaftlichen Unternehmen mit Marktfruchtanbau in Sachsen-Anhalt. Institut für Agrarentwicklung in Mittel und Osteuropa (IAMO), Halle 2005.
6. Konarski R.: Modele równań strukturalnych. PWN, Warszawa 2009.
7. Korol J.: Modele równań strukturalnych i sieci neuronowe w modelowaniu rozwoju zrównoważonego. *Akademickie Wydawnictwo Ekonomiczne*, Gorzów Wielkopolski 2005.
8. Kukuła K.: Elementy statystyki w zadaniach. PWN, Warszawa 2010.
9. Osińska M.: Ekonometryczna analiza zależności przyczynowych. *Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika*, Toruń 2008.
10. Petersen V.: Unternehmensführung und Unternehmenserfolg. *Agrarwirtschaft*, nr 2, 2005.
11. Sagan A.: Wybrane problemy identyfikacji i pomiaru struktur ukrytych. *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie*, nr 543, Kraków 2000.
12. Sagan A.: Zmienne jakościowe w modelowaniu strukturalnym [w:] *Analiza i prognozowanie zjawisk rynkowych o charakterze niemetrycznym* (red. E. Gantar). *Wydawnictwo Akademii Ekonomicznej w Katowicach*, Katowice 2003.
13. Schultze M.: Erfolgsfaktoren landwirtschaftlicher Unternehmen mit Marktfruchtanbau. Eine empirische Analyse mit dem Partial Least Squares – Verfahren. *Dissertation zur Erlangung des akademischen Grades doctor agriculturarum*. Martin-Luther-Universität Halle-Wittenberg, 2008.
14. Sroka W., Dacko M.: Ocena czynników rozwoju przodujących gospodarstw rolniczych z wykorzystaniem drzew regresyjnych typu C&RT. *Zagadnienia Ekonomiki Rolnej*, nr 2, 2010.
15. Sroka W. : Wykorzystanie modelu równań strukturalnych do oceny czynników rozwoju przodujących gospodarstw karpaccich. *Zeszyty Naukowe SERiA*. Warszawa 2010.
16. Stanisław A.: Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 3. *Analizy wielowymiarowe*. Wydawnictwo Statsoft Polska, Kraków 2007.
17. Szaleniec M.: Sieci neuronowe i regresja wieloraka – czyli jak określić złożoność w badaniach naukowych? [w:] *Zastosowania statystyki i data mining w badaniach naukowych* (red. J. Wątroba). *Wydawnictwo Statsoft Polska*, Kraków 2008.
18. Woś A.: Rolnictwo i sektor żywnościowy w 2001 roku. *IERiGŻ*, Warszawa 2002.
19. Zakrzewska M.: Analiza czynnikowa w budowaniu i sprawdzaniu modeli psychologicznych. *Wydawnictwo Naukowe UAM*, Poznań 1994.
20. Zegar J.: Strategia polskiego rolnictwa po akcesji do Unii Europejskiej. *Zagadnienia Ekonomiki Rolnej*, nr 2, 2003.

WOJCIECH SROKA
University of Agriculture
Kraków

APPLICATION OF CAUSE AND EFFECT MODELS
TO EVALUATE FARMS' SUCCESS FACTORS
– THE COMPARISON OF SELECTED METHODS

Summary

The paper presents the application of multi-dimensional cause and effect models. The author focused on the presentation of methodical assumptions and the practical application of the regression analysis along with the method of structural equations as part of the assessment of factors behind the success of farms. A database was analyzed which was prepared based on a questionnaire/interview conducted in 163 leading farms in the region of the Carpathian Mountains. The paper also presents a set of variables explaining and representing factors that hypothetically have the greatest impact on the success of farms: labor, land, production technology, market relations, budget transfers, human capital, progress absorption and accumulation possibilities of farms.