

AUGUSTYN WOS
Instytut Ekonomiki Rolnej
Warszawa

UWAGI O METODACH BADANIA ELASTYCZNOŚCI PRODUKCJI I PODAŻY ZIEMIOPŁODÓW

Zrozumiałe jest dziś wzrastające stale zainteresowanie problematyką przebiegu reakcji dostosowawczych w sferze produkcji rolnej. Wynika ono nie tyle z chęci poznania przeszłości, ile z dążności do stworzenia takiego systemu oddziaływania na rolnictwo, który byłby możliwie sprawny w przyszłości i którego efekty dałyby się z góry przewidzieć¹. Łatwo zrozumieć, że nie możemy sprawnie przewidywać skutków oddziaływania na produkcję, jeśli nie znamy charakteru reakcji dostosowawczych producentów rolnych. Spotyka się sądy, że wobec występowania dostaw obowiązkowych i całego systemu administracyjnego oddziaływania na rolnictwo, wobec różnorodności celów, jakim polityka w stosunku do rolnictwa była podporządkowana, wobec wybitnej niejednorodności okresu powojennego, itd. nie ma możliwości wykrycia jakichkolwiek prawidłowości i określenia kierunków reakcji producentów na bodźce ekonomiczne. Autor niniejszego artykułu ma w tej sprawie nieco odmienny pogląd. Podstawę do pewnego optymizmu dają wyniki przeprowadzonych badań nad elastycznością produkcji 5 głównych ziemioplodów w Polsce w latach 1952/53—1957/58. Okoliczność ta ośmiela więc do zaprezentowania metod ekonometrycznych, które umożliwiają otrzymanie wiarygodnych — naszym zdaniem — wyników.

Teoria regresji a metoda równań jednoczesnych

Analiza podaży często zawodzi ze względu na różnorodność wpływów czynników ekonomicznych i techniczno-produkcyjnych, które nierzadko wzajemnie krzyżują się i niwelują.

Dotąd rozwinęły się dwa kierunki badania elastyczności produkcji i podaży: jeden — który przyrosty produkcji i podaży usiłuje wyjaśnić dodatkowymi nakładami poszczególnych czynników wytwórczych i drugi — który za punkt wyjścia przyjmuje określone czynniki ekonomiczne, głównie cenowe. Nas interesuje głównie ów drugi punkt widzenia.

W zakresie metodologii badań elastyczności produkcji i podaży panował dotąd niepodzielnie pogląd, że najlepszą i jedynie adekwatną jest tu klasyczna teoria regresji, umożliwiająca rozwiązywanie układu równań metodą najmniejszych kwadratów. Metodę tę z powodzeniem stosowało wielu autorów², doskonaląc ją i przystosowując do badania produkcji i podaży poszczególnych dóbr.

We współczesnej ekonometrii panuje pogląd, że najbardziej poprawne (z metodologicznego punktu widzenia) jest w danym przypadku użycie metody równań jednoczesnych, co oznacza powolne odchodzenia od klasycznej teorii regresji, jako mało

¹ Problem przewidywania rozmiarów produkcji, podaży, cen, etc. nabiera dziś kapitalnego znaczenia. Idzie tu o przewidywanie a nie przepowiadanie. To ostatnie — jak wiadomo — nie ma nic wspólnego z naukowym poznawaniem rzeczywistości.

² Wymienić tu można chociażby nazwiska H. L. Moore'a, L. H. Beana, H. Schultza, J. D. Blacka, M. Ezekiała, E. J. Workinga, F. G. Werrena, F. A. Pearsona, J. M. Caselsa i wielu innych.

precyzyjnej i dającej obciążone estymatory badanej funkcji. Jesteśmy w tej chwili w sytuacji, kiedy jeszcze trudno wydać ostateczny sąd o przewadze którejkolwiek z tych metod. Prace teoretyczne wydają się dawać przewagę metodzie równań jednoczesnych, brak jednakże potwierdzenia tego w praktyce. Użycie jej w konkretnych badaniach empirycznych połączone jest z dużymi nakładami i wymaga skomplikowanej aparatury obliczeniowej. Względy praktyczne przemawiają zatem za metodą pojedynczego równania rozwiązywanego w oparciu o teorię regresji, która jest tańsza, łatwiejsza w użyciu, aczkolwiek mniej precyzyjna. Sądzić wolno, że przyszłość należy do dokładniejszej metody równań jednoczesnych. Jej upowszechnienie zależy dziś głównie od postępu w technicznych środkach prowadzenia badań, jakie postawi się do dyspozycji naukowcom. J. A. Nordin, G. G. Judge i O. Wahby¹, po przeprowadzeniu badań nad popytem na artykuły rolnicze w Stanach Zjednoczonych, przy użyciu obu metod, stwierdzają np., że metoda równań jednoczesnych jest pięć razy droższa od metody najmniejszych kwadratów. Uważa się, że jeśli ilość zmiennych jest większa od 3, to niezbędne jest już użycie maszyn elektronowych. W pewnych przypadkach nie jest wykluczone, że wyniki otrzymane przy użyciu obu omawianych metod będą zasadniczo zbieżne, jak to miało miejsce u K. W. Meinkena².

Omawiając walory i wady obu metod godzi się podkreślić, że z merytorycznego punktu widzenia jedyną przewagą metody najmniejszych kwadratów stanowi to, że pozwala ona na użycie większej ilości zmiennych objaśniających. Jeśli na zmienną zależną ma wpływ niewielka ilość zmiennych objaśniających, to lepsza jest metoda równań jednoczesnych. Okoliczność ta ma istotne znaczenie przy wyborze metody, której zamierzamy w badaniu użyć.

Względy, o których była mowa wyżej, uniemożliwiają nam jeszcze w tej chwili rozwiniecie badań w oparciu o metodę równań jednoczesnych, niemniej celowe będzie zwrócenie uwagi na główne jej założenia.

W przypadku skorelowania zmiennych objaśniających estymatory równań regresyjnych są obciążone i niezgodne, co zostało dowiedzione przez wielu statystyków-matematyków, między innymi T. Haavelmo, T. C. Koopmansa, W. C. Hooda i innych. W tej sytuacji nie możemy posłużyć się metodą najmniejszych kwadratów i uciec musimy się do metody bardziej skomplikowanej, polegającej na rozwiązywaniu układu równań jednoczesnych, których może być dowolnie wiele w zależności od stopnia złożoności badanych procesów.

Metoda ta została stworzona w roku 1943 przez T. Haavelmo³ i T. C. Koopmansa⁴, oraz szeregu innych statystyków współpracujących z Fundacją Cowlesa⁵, znaną placówką badań ekonometrycznych, związaną swego czasu z Uniwersytetem w Chicago, a następnie przeniesioną do Yale.

Wspomniana grupa uczonych metodę swą przystosowała do badania elastyczności popytu (funkcji popytu) i w pewnym zakresie (zwłaszcza L. Klein) do makroekonomicznych modeli prognoz koniunktury gospodarczej. Niemniej pewne ich spostrzeżenia i dowody matematyczne dają się w zupełności zastosować do konstrukcji funkcji produkcji i funkcji podaży, gdyż ujmują pewne ogólne prawidłowości statystycznego szacowania parametrów funkcji liniowych i parabolicznych.

¹ J. A. Nordin, G. G. Judge i O. Wahby: "Application of Econometric Procedures to the Demands for Agricultural Products", Iowa Agr. Exp. Sta. Res. Bul. 410, lipiec 1954, s. 1029.

² K. W. Meinken: "The Demand and Price Structure for Wheat", USDA, Tech. Bul. 1136, 1955, s. 41.

³ T. Haavelmo: "The Statistical Implications of a System of Simultaneous Equations", *Econometrica*, vol. 11, z. 1/1943, s. 1—12, oraz "The Probability Approach in Econometrics", *Econometrica*, vol. 12, Supplement, 1944.

⁴ T. C. Koopmans: "Statistical Implications of Simultaneous Economic Relations", *Journal American Statistical Association*, 1945 oraz "When Is an Equation System Complete for Statistical Purposes?", Rozdział XVII w zbiorze: "Statistical Inference in Dynamic Models", Cowles Commission Monograph Nr. 10, New York, 1950.

⁵ Zasadnicze wyniki tych prac opublikowane zostały w 2 monografiach wydanych przez Fundację Cowlesa, a mianowicie we wspomnianej wyżej "Statistical Inference in Dynamic Models", oraz wydanej wspólnie przez T. C. Koopmansa i Wm. C. Hooda: "Studies in Econometric Method". Cowles Commission Monograph Nr. 14, New York 1953. Pełny przegląd tych metod daje również L. Klein w swej pracy "A Textbook of Econometrics", Row. Peterson and Co Illinois 1956.

Jest to tym łatwiejsze do wykonania, że zarówno model badania popytu, jak i tym bardziej model prognoz koniunktury zawierają w sobie równanie popytu, jak i równanie produkcji.

Zostało dowiedzione¹, że asymptotycznie nieobciążone oceny parametrów funkcji liniowej można otrzymać stosując model równań jednoczesnych (simultane równań), który umożliwi rozwiązywanie układu równań strukturalnych względem odpowiednich współczynników. Dochodzenie to otrzymało nazwę metody **maksymalnego prawdopodobieństwa** (*Maximum Likelihood Approach*). Szacunki maksymalnego prawdopodobieństwa uważane są za statystycznie zgodne i poprawne, wobec czego używa się ich często w zastosowaniach statystycznych. Stosując tylko pewne operacje matematyczne, zmierzające do określenia maksimum tzw. funkcji prawdopodobieństwa, można stosunkowo łatwo wyprowadzić szukane równanie, o ile tylko badana rzeczywistość nie jest zbyt skomplikowana. Metody zapewniające maksymalnie prawdopodobne szacunki oparte na jednoczesnym rozwiązywaniu wszystkich równań strukturalnych określane są jako szacunki maksymalnego prawdopodobieństwa przy pełnej informacji, albo krótko — **metody pełnej informacji**. Stosuje się je jednakże bardzo rzadko, gdyż ilość prac obliczeniowych jest tu przerażająco wielka i w żadnej mierze nie można obejść się bez skomplikowanych maszyn elektronicznych.

Z tych bardzo praktycznych względów częściej stosuje się metodę maksymalnego prawdopodobieństwa przy ograniczonej informacji zwaną krótko **metodą ograniczonej informacji**². W tym przypadku ustala się równania po kolei, nie uwzględniając informacji odnoszących się do zmiennych wszystkich pozostałych równań, a wybrane grupy równań traktuje się jako jedno równanie. Uważa się, że metoda ta zapewnia statystycznie zgodne estymatory współczynników regresji, które są bardziej wiarygodne niż wyniki uzyskane przy pomocy innych metod wykorzystujących te same informacje.

Granice i możliwości stosowania metody najmniejszych kwadratów

Nie trzeba chyba wyjaśniać powodów, dla których nie możemy jeszcze dziś rozwinąć badań nad elastycznością popytu i podaży w oparciu o metodę ograniczonej, a tym bardziej pełnej informacji. Jedyne co pozostaje — to posłużenie się tradycyjnym modelem pojedynczego równania, rozwiązywanego metodą najmniejszych kwadratów. Używając tej metody, winniśmy jednakże zdawać sobie w pełni sprawę z jej wad i braków, a nade wszystko winniśmy jasno określić, w jakich okolicznościach może ona być z powodzeniem zastosowana.

Metodę najmniejszych kwadratów stosować można bez zastrzeżeń jedynie wówczas, kiedy zmienne objaśniające nie są ze sobą skorelowane. W każdym innym przypadku otrzymamy współczynniki obciążone pewnym błędem systematycznym. Metoda najmniejszych kwadratów nie daje wówczas możliwości oszacowania wpływu netto badanej zmiennej objaśniającej na zmienną zależną. Otrzymany parametr równania regresji ujmuje w pewnym sensie najczęściej **wpływ brutto**, to znaczy wpływ badanej zmiennej objaśniającej i tego wszystkiego, co jest z nią związane. W rzeczywistości gospodarczej bardzo rzadko zdarza się by A nie było skorelowane z innymi zmiennymi posiadającymi wpływ na X . Dlatego też winniśmy dążyć do ustalenia tzw. wpływów głównych, pamiętając jednocześnie, że żaden z nich nie jest wolny od wpływu czynników dodatkowych. Gdybyśmy zgodzili się co do tego, że z punktu widzenia potrzeb i celu przeprowadzanego badania wystarczy nam określenie wpływu brutto, to metodę najmniejszych kwadratów uznać można by za adekwatną i wystarczającą.

G. S. Shepherd pisze np.: „Jest właściwie rzeczą niemożliwą określić wpływ netto A na X . Mówiąc o wpływie netto A na X , po wyłączeniu wpływu netto B , mamy w gruncie rzeczy na myśli wpływ netto A i wszystkiego co jest skorelowane z A przez przypadek lub w inny sposób, **poza** B . Mówiąc o wpływie netto A na X po wyłączeniu B i C , mamy w gruncie rzeczy na myśli wpływ netto A oraz wszyst-

¹ Por. wyżej cytowane prace Koopmansa i Hooda.

² Założeniom tej metody z punktu widzenia potrzeb badania elastyczności popytu wiele uwagi poświęcił ostatnio Z. Pawłowski w szeregu kolejnych artykułów opublikowanych w „Przeglądzie Statystycznym”.

kiego, co skorelowane z A przez przypadek lub w inny sposób, poza B i C , i tak dalej przy dalszych zmiennych” — i nieco dalej w konkluzji: „...współzależności rozumiane dosłownie netto — są nieosiągalne, gdyż określić czysty wpływ netto możemy tylko wówczas, kiedy uwzględnimy wszystkie pozostałe wpływy — a tych są dosłownie setki. Jest to praktycznie niemożliwe. Niemniej z praktycznego punktu widzenia można osiągnąć dość dokładne przybliżenie do czystego wpływu netto. Rozważania ekonomiczne i opublikowane badania z dziedziny ekonomiki rolnictwa wskazują, że w większości wypadków można uzyskać dostatecznie dokładne wyniki posługując się stosunkowo ograniczoną ilością zmiennych. Prawie wszystkie te badania stosują tylko dwie lub trzy zmienne niezależne. Bo choć współzależność jest powszechna, to jednak jej ilościowa waga maleje szybko po uwzględnieniu ważniejszych zmiennych, a tych właśnie jest zazwyczaj tylko kilka”¹.

Gdy stanąć zatem na stanowisku absolutnej precyzji, to żaden wpływ netto przy użyciu metody najmniejszych kwadratów nie da się określić. W tym celu musimy użyć bardziej skomplikowanej, omawianej poprzednio, metody równań jednoczesnych. Jeśli jednak z różnych powodów pozostać musimy przy metodzie najmniejszych kwadratów, pamiętać winniśmy, że uzyskane parametry mogą być obciążone błędem systematycznym.

Tak czy inaczej wielu autorów uważa, że w dziedzinie badań ekonomiczno-rolnych metoda najmniejszych kwadratów znajduje stosunkowo najszersze zastosowanie. Pogląd taki reprezentuje K. A. Fox², G. S. Shepherd³, R. Foote⁴, a także L. R. Klein⁵.

Chcąc użyć w naszych badaniach tradycyjnej metody regresji, musimy wobec tego udowodnić, że:

1. Użyte przez nas **zmienne są egzogeniczne** (to jest takie, które wprawdzie wpływają na zmienną zależną, jednak odwrotnie, oddziaływanie zmiennej zależnej na zmienne objaśniające bądź w ogóle nie występuje, bądź też jest tak słabe, że praktycznie można je pominąć), lub

2. że zmienna zależna jest funkcją wartości zmiennych objaśniających z poprzednich okresów, natomiast nie jest funkcją wartości zmiennych endogenicznych, łącznie współzależnych ze zmienną zależną. Mówimy o nich jako o **opóźnionych wartościach zmiennych endogenicznych**.

Chcemy w szczególności zwrócić uwagę na drugi przypadek, gdyż — naszym zdaniem — stwarza on niesłychanie rozległe perspektywy dla zastosowań metody najmniejszych kwadratów. Jak się okazuje, z dużą łatwością można skonstruować taki model badania elastyczności produkcji, podaży czy popytu, w którym wystąpią opóźnione zmienne endogeniczne, wobec czego układ będzie rozwiązalny przy użyciu metody najmniejszych kwadratów, a otrzymane parametry będą dawały przybliżony obraz rzeczywistości⁶. Przykład może wyjaśnić lepiej cały ten problem. Weźmy pod uwagę produkcję i podaż ziemniaków. Wiadomo powszechnie, że zasiewy i ceny ziemniaków są wzajemnie ze sobą skorelowane w tym sensie, że zasiewy są funkcją ceny, a cena jest funkcją zasiewów. Zmienne użyte w tym sensie są endogeniczne i równanie nie może być rozwiązane metodą najmniejszych kwadratów (otrzymane parametry nie będą statystycznie zgodne). Gdy jednak tylko dowiedzimy, że zasiewy są funkcją ceny z poprzedniego okresu i same z kolei determinują cenę w okresie przyszłym, to użyte tu zmienne są opóźnionymi (odroczone) zmiennymi endogenicznymi i w tym sensie nie są ze sobą wzajemnie

¹ G. S. Shepherd: *Agricultural Prices Analysis*, Ames, Iowa 1957, The Iowa State College Press. wyd. IV.

² K. A. Fox: *The Analysis of Demand for Farm Products*, USDA. Tech. Bull. 1081. wrzesień 1953.

³ G. S. Shepherd, op. cit.

⁴ J. Friedman i R. Foot: *Computational Methods for Handling Systems of Simultaneous Equations Applications to Agriculture*, w „*Agriculture Handbook*” nr 94, USDA. listopad 1955.

⁵ K. R. Klein: *Single Equation Vs. Equation System Methods of Estimation in Econometrics*, *Econometrica*, vol. 28, z. 4/1960, s. 871.

⁶ „Wiemy obecnie, że ocena nasza będzie jedynie wówczas statystycznie zgodna i poprawna, kiedy ilości skonsumowane oraz czynniki kształtujące popyt możemy uznać za zmienną egzogeniczną, lub opóźnioną zmienną endogeniczną” — pisze w cytowanej tu pracy G. S. Shepherd.

skorelowane. Przypadek ten jest bardzo prosty i ustalenie charakteru poszczególnych zmiennych nie napotyka na żadne trudności. W praktyce badawczej bardzo często mają miejsca przypadki znacznie bardziej złożone i wówczas odpowiedź bynajmniej nie jest prosta i jednoznaczna.

Ekonometryczny model badania funkcji produkcji i podaży ziemiopłodów

Biorąc za punkt wyjścia powyższe uwagi, możemy podjąć próbę skonstruowania pewnego układu równań objaśniających w zakresie zachowania się produkcji i podaży ziemiopłodów pod wpływem niektórych czynników ekonomicznych. Równania te stanowią statystyczno-matematyczny zapis prostych zależności ekonomicznych dotyczących funkcji produkcji i podaży poszczególnych ziemiopłodów w warunkach drobnotowarowego rolnictwa.

Założeniem tego modelu jest, że badanie odbywa się w skali masowej, a nie pojedynczego gospodarstwa rolnego. Podaż pojedynczego producenta bowiem ma tak minimalny wpływ na rynek, że w żadnej mierze nie jest w stanie oddziaływać na kształtowanie się cen.

Proponowany model uwzględnia wpływ tylko najważniejszych czynników, wprowadzając do każdego równania składnik resztowy (u_t).

W badanej funkcji produkcji głównych ziemiopłodów uwzględnić trzeba — naszym zdaniem — następujące zmienne:

X_t = zasiewy danego ziemiopłodu,

P_t = cena bieżąca uzyskiwana za dany produkt przez rolników, rozumiana jako średnia ważona cena skupu nadobowiązkowego i wolnego rynku (uwzględniająca ceny dostaw obowiązkowych, kontraktacji i wolnego rynku)¹,

P_t^1 = średnia ważona cena skupu nadobowiązkowego i wolnego rynku (jak wyżej) przewidywana na okres następny, lub $P_t^1 + \theta$ — na dowolną ilość okresów naprzód,

R_t = relacja ceny danego produktu do indeksu cen rolnych, wyrażająca jego siłę konkurencyjną w stosunku do ogółu ziemiopłodów².

Q_t = produkcja danego ziemiopłodu (ilość),

M_t = przewidywany plon badanego ziemiopłodu.

D_t = popyt na dany produkt na rynku lokalnym (nie obejmuje to popytu zaspokajanego centralnie przez aparat handlu uspołecznionego),

C_t = jednostkowe koszty produkcji badanego ziemiopłodu. Zmienna ta (lub ściślej relacja kosztów wytworzenia do ceny zbytu: $P_t - C_t$) ma znaczenie przy badaniu elastyczności produkcji w dłuższych okresach czasu; w okresach krótkich zjawisko to odgrywa zwykle niewielką rolę,

¹ Cena ta może mieć zastosowanie wyłącznie przy badaniu procesów dostosowawczych w krótkich okresach czasu (z roku na rok). Badając elastyczność produkcji w dłuższych okresach należy użyć — naszym zdaniem — średniej ważonej ceny skupu (uwzględniającej obok dostaw nadobowiązkowych, kontraktacji i wolnego rynku także dostawy obowiązkowe). Cena ta bowiem pozwala uwzględnić efekt dochodowy, tzn. fakt, że oddziaływanie na produkcję w dłuższych okresach odbywa się wyłącznie poprzez regulowanie dochodów.

² Zmienna ta może wystąpić również w postaci indeksu cen rolnych. W razie potrzeby można jej uniknąć, deflując P_t , lub P_t^1 przez odpowiedni indeks cen. Wówczas szeregi cen wystąpią jako „ceny realne”, aczkolwiek to pojęcie ma tu znaczenie wyłącznie umowne. Nie idzie tu o realną siłę nabywczą złotych, ale — jak określono wyżej — o siłę konkurencyjną względem cen innych ziemiopłodów. W naszych badaniach zamierzamy posłużyć się oszacowanym wcześniej relatywnym indeksem cen skupu nadobowiązkowego i kontraktacji obliczonym przy użyciu cen porównywalnych. W przypadku średniej ważonej ceny skupu, użyć należy relatywnego indeksu cen skupu ogółem, obliczonego przy użyciu cen porównywalnych. Por. A. Woś: W sprawie indeksu cen rolnych, Zagadnienia Ekonomiki Rolnej nr 6/1960, s. 75 (vide: indeks IV i VI).

S_t = podaż danego produktu, rozumiana jako suma sprzedaży poza dostawami obowiązkowymi, traktowana z punktu widzenia badanego rejonu, a więc rynku lokalnego,

u_t = składnik resztowy ujmujący również wpływ czynników przyrodniczo-klimatycznych.

Biorąc pod uwagę wyliczone wyżej zmienne proponowany układ równań regresji mógłby otrzymać następujący zapis:

$$X_t = b_{11}P_{t-\theta} + b_{12}X_{t-\theta} + b_{13}P_t^1 + b_{14}R_{t-\theta} + b_{15} + u_t \quad \dots \quad (1)$$

$$P_t = b_{21}(S_t - D_t) + b_{22}P_{t-\theta} + b_{23}R_t + b_{24} + u_t \quad \dots \quad (2)$$

$$P_t^1 = b_{31}P_{t-\theta} + b_{32}(\Delta P_{t-\theta}) + b_{33} + u_t^2 \quad \dots \quad (3)$$

$$Q_t = b_{41}X_t + b_{42}M_t + b_{43} + u_t \quad \dots \quad (4)$$

$$S_t = b_{51}Q_t + b_{52}P_t + b_{53}S_{t-\theta} + b_{54}R_t + b_{55} + u_t \quad \dots \quad (5)$$

Przedstawione wyżej równania wystarczają — naszym zdaniem — do określenia elastyczności produkcji i podaży produktów w danym okresie czasu (t). Zakładamy przy tym, że współzależności między danymi zmiennymi są liniowe. Zwrócić trzeba uwagę, że wprowadzenie współczynników b_{12} , b_{22} i b_{53} oznacza, iż zakładamy pewną inercję zasiewów, ceny i podaży, utrzymując, że jest ona nie tylko funkcją zmieniających się czynników ekonomicznych, ale także absolutnego poziomu tychże z okresów poprzednich.

Wszystkie współczynniki b są wielkościami stałymi dla danego równania, a ich suma = 1 ($b_1 + b_2 + \dots + b_n + u_t = 1$).

Wszystkie użyte w niniejszym układzie pięciu równań zmienne są zmiennymi **egzogenicznymi**, lub **opóźnionymi wartościami zmiennych endogenicznych**. Stwierdzenie tego faktu jest dla nas bardzo ważne, gdyż upoważnia do użycia prostej, a przy tym **praktycznie możliwej do zastosowania** metody najmniejszych kwadratów.

Zdawać trzeba sobie jednak sprawę, że nie wszystkie z określonych wyżej zmiennych będą mogły być w praktycznym badaniu uwzględnione; niektóre musimy pominąć ze względu na brak zasługujących na zaufanie danych statystycznych.

Szacowanie parametrów charakteryzujących wpływ badanych zmiennych objaśniających na zmienną zależną nie przedstawia merytorycznie trudności, dopóki zależności mają charakter liniowy.

Regresja krzywolinijna i metoda graficznej aproksymacji krzywej produkcji i podaży

Najprostszą w użyciu i dającą dobre — według naszego rozeznania — wyniki, w przypadku regresji liniowej, jest metoda korelacji wielorakiej zaprezentowana przez M. Ezekieła³.

Ponieważ metoda ta jest powszechnie znana nie będziemy jej tu bliżej omawiać. Problem staje się nieporównanie bardziej skomplikowany, gdy regresja nie jest liniowa. W ekonometrii istnieje szereg poglądów na możliwość i sposób szacunku parametrów funkcji krzywoliniowych. Jednym z pierwszych, który zajął się tym problemem był właśnie M. Ezekiel. Prace swe rozpoczął on mniej więcej 40 lat

¹ Równanie to — jak łatwo zauważyć — uwzględnia wpływ czynników rynkowych, przyjmując, że zarówno spożycie wewnętrzne danego produktu jak i spisanie jest constans i zawiera się w zmiennej $b_{12}X_{t-\theta}$. Gdyby założenie to okazało się nie możliwe do przyjęcia, należałoby dodatkowo uwzględnić takie zmienne jak ludność rolnicza i stan inwentarza żywego.

² Równanie to możemy zapisać także jako: $P_t^1 = P_{t-1}^1 + \beta (P_{t-1} - P_{t-1}^1)$, gdzie: $0 < \beta < 1$. Tej postaci formuły na cenę przewidywaną używa M. Nerlove w swej pracy: „The Dynamic of Supply Estimation of Farmers Response to Price”, Baltimore 1958, The Johns Hopkins Press, s. 52—53.

³ M. Ezekiel: Methods of Correlation Analysis, New York 1953, John Wiley and Sons, Inc.

temu niejako na marginesie studiów nad wieloraką korelacją liniową¹. Zasluga Ezekieła było przełamanie impasu, jaki istniał w teorii korelacji wielorakiej i regresji. Zaproponował on wówczas **metodę kolejnych przybliżeń**, którą sam autor charakteryzuje następująco: „...ustala się cząstkową regresję liniową, a następnie zmienną zależną koryguje się względem odchylen od średnich wszystkich zmiennych niezależnych z wyjątkiem jednej, natomiast wykres korelacji lub układ odpowiednich punktów wykonuje się na podstawie owych wartości poprawionych i zmiennej niezależnej. Zmienną zależną z kolei koryguje się względem wszystkich pozostałych zmiennych niezależnych z wyjątkiem jednej. Skorygowane zmienne stanowią z kolei wartości zmiennych, oraz pierwsze przybliżenie krzywej określonej względem tej właśnie zmiennej. Krzywych tych używa się jako podstawy dla skorygowania czynnika zależnego o wpływ krzywoliniowej aproksymacji wszystkich zmiennych niezależnych z wyjątkiem jednej, eliminując po kolei każdą z nich. Krzywe aproksymacji drugiego stopnia ustalane są przy pomocy w ten sposób skorygowanych wartości, między wartościami wszystkich po kolei branych zmiennych niezależnych. Operację tę przeprowadza się dla każdej zmiennej niezależnej po kolei, uzyskując pełny zestaw pierwszych przybliżeń do krzywych regresji netto. Na krzywych tych dokonuje się dalszych poprawek aż do momentu dopóki poszczególne krzywe regresji przestaną wykazywać jakiegokolwiek odchylenia².

W metodzie Ezekieła kryje się pewna, istotna naszym zdaniem, niekonsekwencja. Polega ona na tym, że zaprzęga się całą skomplikowaną aparaturę matematyczną po to, aby później otrzymane linie dopasowywać z konieczności odręcznie do wszystkich krzywni w analizie regresyjnej. Nawet przy użyciu najbardziej precyzyjnej metody matematycznej nie otrzymamy takiej linii krzywej, która prawidłowo wyjaśniałaby wszystkie zmiany w badanych zmiennych. Louis Bean zajmując się tym problemem w latach dwudziestych doszedł do przekonania, że skomplikowana procedura matematyczna, z małym uszczerbkiem dla precyzji badań, może być zastąpiona metodą graficzną. Bean stwierdził, że zarówno linie proste, jak i krzywe w analizie regresji należy ustalić odręcznie. Metoda Beana okazuje się prosta i przejrzysta. Jest ona wskazana — naszym zdaniem — wszędzie tam, gdzie same zmienne nie są absolutnie wiarygodne i w związku z tym nie zasługują na zbyt drobiazgowo opracowanie. Sytuacja taka ma w dużym stopniu miejsce w naszych warunkach, co skłaniałoby raczej do użycia w badaniach graficznej metody Beana³. Metodę swą sprawdził Bean również w roku 1929 na materiale rzeczywistym, badając elastyczność uprawy ziemniaków w Stanach Zjednoczonych i otrzymał zasługujące na zaufanie wyniki⁴.

Dla wyjaśnienia założeń metody graficznej Beana weźmy pod uwagę najprostszy przypadek korelacji wielorakiej, gdzie zmienna zależna X_1 determinowana jest przez zmienne objaśniające: X_2 i X_3 . W tym przypadku musimy się posłużyć dwoma układami współrzędnych, z których jeden przedstawia regresję X_1 względem X_2 niezależnie od wpływu X_3 , a drugi regresję X_1 względem X_3 niezależnie od wpływu X_2 . Zwracamy uwagę na określenie „niezależnie od...”, gdyż jest to główne założenie całej metody graficznej.

W badaniu zjawisk ekonomicznych najczęściej zdarza się, że punkty na obu wykresach nie wykazują wyraźnej prawidłowości, co jest zrozumiałe, gdyż mamy tu do czynienia z zakłócającym wpływem innych zmiennych. Nam zależy na określeniu wpływu netto każdej ze zmiennych. Najprostszy sposób ustalenia wpływu netto polega na wybraniu dwu takich lat, w których wartości zmiennej objaśniającej, np. X_2 są takie same. Wówczas różnica między nimi może być w przybliżeniu uznana za wynik wpływu drugiej zmiennej, powiedzmy X_3 . Linia łącząca te dwa lata stanowiłaby przybliżony szacunek wpływu X_2 na X_1 niezależnie od wpływu X_3 . Należy dążyć do tego, aby w analogiczny sposób połączyć liniami pomocniczymi

¹ M. Ezekiel: A Method of Handling Curvilinear Correlation for Any Number of Variables, Journal of the American Statistical Association, vol. 19, 1924.

² M. Ezekiel: Methods of Correlation Analysis, New York 1941, John Wiley and Sons, Inc., s. 223.

³ L. H. Bean: Applications of a Simplified Method of Graphic Curvilinear Correlation, Bureau of Agricult. Econ., USDA, kwiecień 1929 oraz

A. Simplified Method of Graphic Curvilinear Correlation, Journal of the American Statistical Association, vol. 24, z. 4/1929, s. 386—397.

⁴ L. H. Bean: The Farmers Response to Price, Journal of Farm Economics, z. 3/1929, s. 368—385.

wszystkie punkty figurujące na wykresie. Otrzymamy w ten sposób szereg linii przebiegających zwykle w jednym kierunku i wykazujących podobne nachylenie. Między tymi liniami możemy przeciągnąć teraz pewną linię główną, która uwzględniłaby kierunek, nachylenie, oraz długość poszczególnych linii pomocniczych. Będzie to pierwsze przybliżenie szukanej przez nas linii regresji netto.

W tej sytuacji ustalenie wpływu netto X_2 na X_1 staje się już proste, bo jeśli przeprowadzona linia główna przedstawia wpływ X_2 na X_1 , to pionowe odległości poszczególnych punktów od tej linii w górę i w dół powinny odzwierciedlać wpływ netto X_2 na X_1 . Następny etap polega więc na zmierzeniu owych różnic, czyli reszt, które następnie odkłada się na drugim wykresie względem drugiej zmiennej objaśniającej. Na wykresie tym ustala się linię zerową (na osi rzędnych), a następnie odkłada pionowe odległości poszczególnych punktów otrzymanych we wstępnej aproksymacji. Wówczas linia przeprowadzona przez punkty otrzymane na drugim wykresie odzwierciedla wpływ netto X_2 na X_1 . Jeśli pierwsza linia aproksymacji została ustalona prawidłowo i jeśli zmienna zależna X_1 jest całkowicie determinowana przez zespół dwu uwzględnionych zmiennych objaśniających — to linia otrzymana na drugim wykresie powinna przechodzić dokładnie przez wszystkie punkty. Jeśli powstanie jednak pewien rozrzut, to oznacza to, że albo źle wyznaczono pierwszą linię główną, albo, że konieczne jest uwzględnienie dalszych zmiennych objaśniających.

Jedną z wielkich zalet graficznej metody analizy regresyjnej — poza prostotą — jest to, że pozwala ona badaczowi śledzić to wszystko, co działo się w okresie objętym analizą. Metoda matematyczna nie daje takich możliwości. Tu liczby wrzuca się niejako do kotła i otrzymuje się gotowy wynik, nie śledząc tego wszystkiego, co miało miejsce wewnątrz procesu obliczeniowego. Nawet skromna praktyka badawcza pozwala stwierdzić, jak wielkie to ma znaczenie, zwłaszcza wówczas, kiedy występują czynniki zakłócające.

Z drugiej jednak strony merytoryczna ocena wyników otrzymanych przy pomocy metody graficznej jest trudniejsza, gdyż wymaga równorzędnego uwzględnienia co najmniej trzech kryteriów: 1) stopnia skupienia punktów wokół linii regresji netto, 2) stopnia zgodności wyników z ogólną teorią ekonomiczną, 3) zgodności wyników z tym, co już wiadomo na temat danego zjawiska. Żadne z tych kryteriów samo przez się nie wystarcza.

Drugą zaletą metody graficznej jest to, że daje ona dużą oszczędność pracy i środków, zwłaszcza w przypadku regresji krzywoliniowej. Jak obliczono — jest ona dwu- a nawet trzykrotnie tańsza niż zwykła metoda matematyczna. Przy metodzie matematycznej badacz musi najpierw obliczyć prostoliniową korelację wieloraką, a następnie każdą regresję musi zbadać matematycznie pod kątem jej krzywoliniowości. Ustalenie krzywej maksymalnie odpowiadającej punktom wymaga nieraz wielu skomplikowanych zabiegów. Pozytywny efekt jest ten, że otrzymuje się współczynnik w wyrażeniu liczbowym, co przy metodzie graficznej jest trudne do osiągnięcia.

Ponadto, jak stwierdzono, krzywe ustalone metodą graficzną na ogół lepiej pasują do danych empirycznych, aniżeli krzywe ustalone matematycznie. Metoda matematyczna nie jest też całkiem wolna od subiektywizmu. Tu badacz również powodując się własnym sądem — wybiera najlepszy kształt linii krzywej, co nie może nie mieć wpływu na wyniki. Uzyskany rezultat w dużym stopniu zależy od tego, jak dobrano kształt krzywej. Metoda graficzna jest w wyższym jeszcze stopniu subiektywna. Jej przewaga w tym względzie polega jednak na tym, że wszelkie niedokładności linii wykreślonych odręcznie jawnie występują na wykresach, podczas gdy niedokładności metody matematycznej są bardzo trudne do rozszyfrowania.

Metoda graficzna przydatna jest najbardziej wszędzie tam, gdzie: 1) ilość obserwacji jest niewielka, 2) gdzie ilość niezbędnych do uwzględnienia zmiennych nie jest zbyt duża (najwyżej 3—4), 3) gdzie korelacja jest istotna, co znacznie zmniejsza ryzyko subiektywnej oceny przy wykreślaniu krzywych aproksymacji. Te okoliczności występują bardzo często w badaniach ekonomicznych. Prowadzone dotychczas w Polsce prace nad elastycznością produkcji podają się wskazywać na dużą przydatność metody graficznej wszędzie tam, gdzie regresja jest jawnie krzywoliniowa. W przypadku regresji liniowej pierwszeństwo należy dać metodzie matematycznej, która w tym przypadku jest w miarę obiektywna, pamiętając jednakże, że daje wyniki obciążone częściowo błędem systematycznym, o czym mówiliśmy wyżej.