

Tadeusz Kufel*

E-PROJEKTY W NAUCZANIU METOD IŁOŚCIOWYCH Z WYKORZYSTANIEM OPROGRAMOWANIA GRETL

1. WPROWADZENIE

Coraz częściej w programach studiów pojawiają się jednostki dydaktyczne pod nazwą projekty. Tradycyjne formy nauczania, to jest wykład, ćwiczenia są zastępowane laboratoriami i projektami. Projekty są to samodzielnie wykonywane zadania przez studenta, dla których formą pomocy są dodatkowe materiały dydaktyczne umieszczone w internecie oraz dyżury konsultacyjne pracowników nauczających dany przedmiot. W programie studiów na wielu uczelniach przyjęto zasadę podziału godzin dydaktycznych w przedmiotach ilościowych na trzy części: wykład, laboratorium i projekty. Na przykład, dla 60 godzin ekonometrii podział taki wynosi: 20 godzin – wykład, 20 godzin – laboratorium, 20 godzin – projekty. Taki podział godzin występuje częściej na uczelniach prywatnych i w trybie studiów niestacjonarnych. Zajęcia dydaktyczne określane jako *e-projekty* mają formę kształcenia na odległość i mogą być realizowane w wymiarze 40% godzin dydaktycznych przewidzianych programem studiów¹.

Celem artykułu jest przedstawienie przykładowych zadań realizowanych w ramach projektów z przedmiotów: statystyki, ekonometrii oraz prognozowania i symulacji z wykorzystaniem oprogramowania GRETL². Wiele projektów jest realizowanych z wykorzystaniem komunikacji z wykładowcą (opiekunem) za pomocą poczty elektronicznej. Określenie *e-projekty* wynika z tego, że końcowe

* prof. dr hab., Uniwersytet Mikołaja Kopernika w Toruniu

¹ Rozporządzenie MNiSzW z dnia 25.09.2007 r. dopuszcza wykorzystanie technik kształcenia na odległość do 80% dla jednostek posiadających uprawnienia do nadawania stopnia doktora habilitowanego i do 60% dla jednostek z uprawnieniami do nadawania stopnia doktora, 40% pozostałe jednostki.

² Oprogramowanie GRETL z grupy OPEN SOURCE dostępne jest na stronie internetowej <http://www.kufel.torun.pl> i <http://gretl.sourceforge.net>.

wyniki są przesyłane e-mailem na wskazane konto z nazwą załącznika stanowiącego numer indeksu studenta i do ściśle określonego momentu czasowego.

Forma realizacji zadań z zakresu *e-projekty* jest etapem wstępnym wprowadzenia elementów *e-learningowych* w nauczaniu przedmiotów ilościowych.

2. PROJEKTY ZE STATYSTYKI OPISOWEJ I MATEMATYCZNEJ

Zadania do samodzielnego wykonania obejmujące zagadnienia ze statystyki opisowej i matematycznej mogą być realizowane za pomocą oprogramowania GRETL z wykorzystaniem zainstalowanych baz danych³.

Kilka przykładowych projektów – zadań ze statystyki z wykorzystaniem różnych baz danych zostanie przedstawionych poniżej.

Bazy danych zainstalowane na serwerze *ricardo.ecn.wfu.edu* znajdującym się na Wydziale Ekonomicznym Uniwersytetu Wake Forest (USA) można zainstalować na dysku lokalnym komputera poprzez funkcje **Plik / Pliki baz danych / Instalowanie baz danych...** Bazy te zawierają informacje statystyczne dla Stanów Zjednoczonych, Wielkiej Brytanii, Japonii, Hiszpanii oraz Unii Europejskiej, ale można utworzyć własne bazy danych wykorzystując funkcje zapisu pliku danych jak bazy danych (*.bin). Rysunek 1 zawiera okna programu GRETL dotyczące serwera baz danych, lokalnych baz oraz indeksu zmiennych w otwartej bazie.

Wykorzystując internetowy interfejs komunikacyjny z Bankiem Danych Regionalnych GUS⁴ można zaimportować do Excela, a potem do oprogramowania GRETL szeroki zbiór informacji o jednostkach administracyjnych Polski w podziale na: regiony, województwa, podregiony, powiaty i gminy. BDR zawiera ponad 10 tys. cech statystycznych dla lat 1999 i następnych w układzie danych rocznych, a niewielka ich część jest w układzie danych kwartalnych.

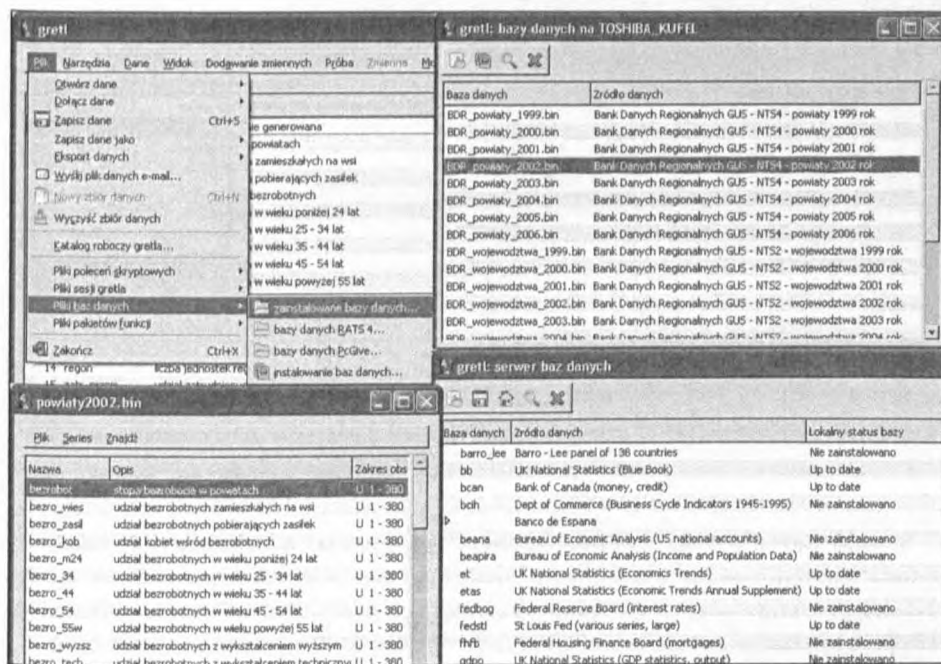
³ Procedury tworzenia baz danych w oprogramowaniu GRETL przedstawiono w referacie: Tadeusz Kufel, *Budowa internetowych baz danych i udostępnianie przykładów z podręczników ekonometrii w oprogramowaniu GRETL*, XVI Konferencja Dydaktyczna pt. *Unifikacja programów nauczania metod ilościowych*, 11–12 czerwca 2007 w Łodzi, http://www.kufel.torun.pl/kufel_2007_slajd_lodz.pdf.

⁴ Bank Danych Regionalnych GUS – BDR (<http://www.stat.gov.pl>) jest największym w Polsce uporządkowanym zbiorem informacji o sytuacji społeczno-gospodarczej, demograficznej, społecznej oraz stanie środowiska, opisującym województwa, powiaty oraz gminy jako podmioty systemu organizacji społecznej i administracyjnej państwa, a także regiony i podregiony stanowiące elementy nomenklatury jednostek terytorialnych do celów statystycznych. Bank Danych Regionalnych gromadzi, systematycznie uzupełnia i aktualizuje informacje statystyczne o poszczególnych jednostkach podziału terytorialnego.

Bank Danych Regionalnych GUS przeniesiony do baz GRETLA zawiera prawie tysiąc cech z zakresu:

- stanu ludności i ruchu naturalnego,
- rynku pracy, wynagrodzeń i świadczeń społecznych,
- rolnictwa i leśnictwa, transportu i łączności,
- stanu i ochrony środowiska, gospodarki mieszkaniowej i komunalnej,
- przemysłu, budownictwa, handlu i turystyki,
- szkolnictwa podstawowego, ponadpodstawowego i wyższego,
- ochrony zdrowia, kultury i sztuki,
- dochodów i wydatków budżetów jednostek administracyjnych,
- inwestycji i środków trwałych,

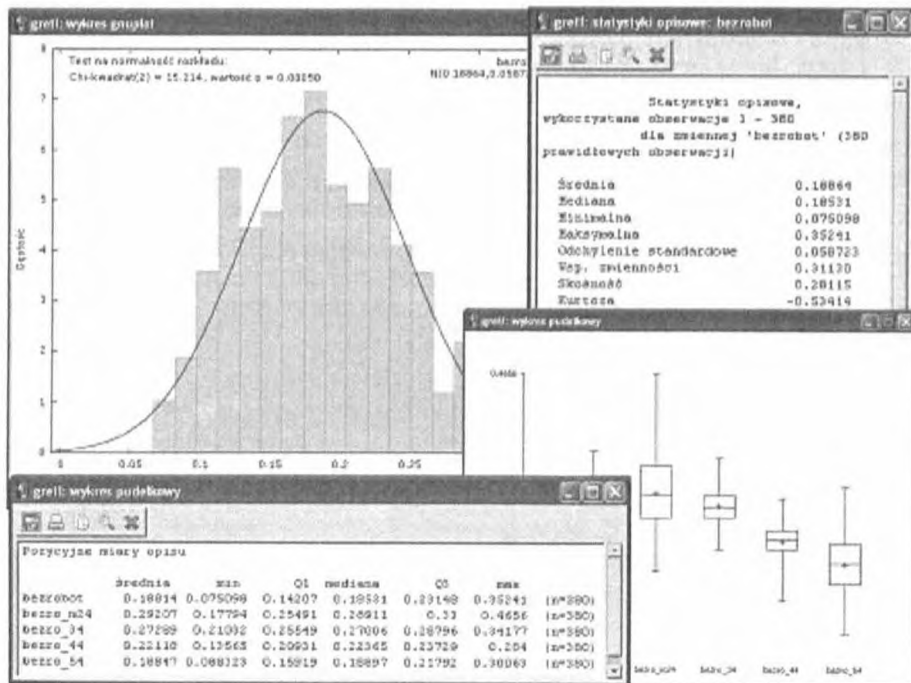
w przekrojach jednostek powiatowych (NTS4) i wojewódzkich (NTS2).



Rys. 1. Okna programu GRETL: serwer baz danych, lokalne bazy danych, indeks zmiennych otwartej przykładowej bazy

Zadanie do samodzielnego wykonania z zakresu statystyki opisowej i matematycznej może dotyczyć danych przekrojowych dla wybranego jednego roku w układzie jednostek administracyjnych, np. powiatów czy województw. Zrealizować można zagadnienia dotyczące wyznaczania miar opisowych i pozycyjnych, co przedstawia rysunek 2. Wyniki zaprezentowane w oknach rysunku 2 przedstawiają klasyczne i pozycyjne miary opisowe. Duża ilość cech w przygo-

towanych bazach (*.bin) z informacjami statystycznymi dotyczącymi powiatów (ponad 1000 cech dla 7 lat) pozwala przydzielić niepowtarzalne zadania ze statystyki opisowej dla bardzo dużych grup studenckich.



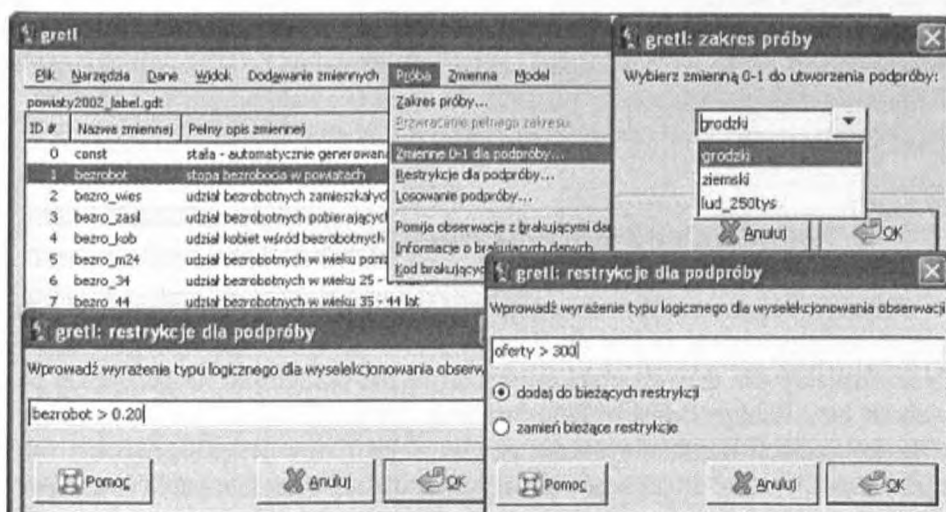
Rys. 2. Okna programu GRETL: klasyczne miary opisowe, pozycyjne miary opisowe, rozkłady częstości, wykresy pudełkowe

Każde przydzielone zadanie zawiera informacje: o zbiorowości, badanym okresie, analizowanych cechach i uzupełnione jest wykazem miar, które należy oszacować i zinterpretować, wraz ze wskazaną formą prezentacji wyników: tabel i typów wykresów.

W celu rozszerzenia analizowanych zagadnień można wprowadzić dodatkowo zagadnienia badań dla podprób zbiorowości, np. dla powiatów grodzkich czy ziemskich, dla powiatów o wysokiej czy niskiej stopie bezrobocia. Funkcje wyboru podprób przedstawia rysunek 3, a poniżej przedstawiono przykłady tych poleceń w języku skryptowym.

```
smpl grodzki --dummy
smpl bezrobocie > 0.20
smpl oferty > 300 --restrict
```

Oprogramowanie GRETL pozwala zrealizować podstawowe testy parametryczne: dla jednej średniej, wariancji i wskaźnika struktury, dla dwóch średnich, wariancji i wskaźników struktury. Wprowadzanie informacji do okien dialogowych wymaga tylko wskazanie zmiennych i sformułowanie hipotezy zerowej. Podobnie można wykonać testy nieparametryczne dla: testów znaków, testu sumy rang Wilcoxona, testu rangowanych znaków Wilcoxona. Te wszystkie testy realizują funkcje *Narzędzia/Testy parametryczne* oraz *Narzędzia/Testy nieparametryczne*. W takiej ułatwionej procedurze obliczeniowej ważnym elementem jest sformułowanie hipotez statystycznych oraz podjęcie właściwej decyzji weryfikacyjnej.



Rysunek 3. Okna programu GRETL: wyboru podpróby za pomocą zmiennych 0-1 i wyrażeń typu logicznego dla wybranych zmiennych.

Zagadnienia analiz współzależności można wykonać z wykorzystaniem oprogramowania GRETL. Należą do nich: współczynnik korelacji liniowej Pearsona, współczynnik korelacji kolejnościowej Spearmana, tau-Kendalla, budowa tablic korelacyjnych i testu niezależności chi-kwadrat.

W celu tworzenia tablic korelacyjnych należy w oprogramowaniu GRETL utworzyć zmienne dyskretne za pomocą formuł logicznych. Przykładowa formuła kodowania przedziałów dla dochodów budżetów powiatowych na 1 mieszkańca w przedziałach po 1000 zł z wartościami od 1 do 4 ma postać (polecenie *genr* w konsoli lub w oknie dialogowym *definiowanie nowej zmiennej*)

```
genr d_new = (do-  
chod>=0)+(dochod>=1000)+(dochod>=2000)+(dochod>=3000)  
discrete d_new .
```

Tylko zmienne dyskretne pozwalają utworzyć tablicę krzyżową (korelacyjną). Definiowanie zmiennej dyskretnej wykonuje się w oknie *Edycji atrybutów* (F2). Realizacja funkcji *Widok/Tabela krzyżowa* wyznacza liczebności tabeli korelacyjnej i szacuje statystykę testu niezależności *chi-kwadrat*.

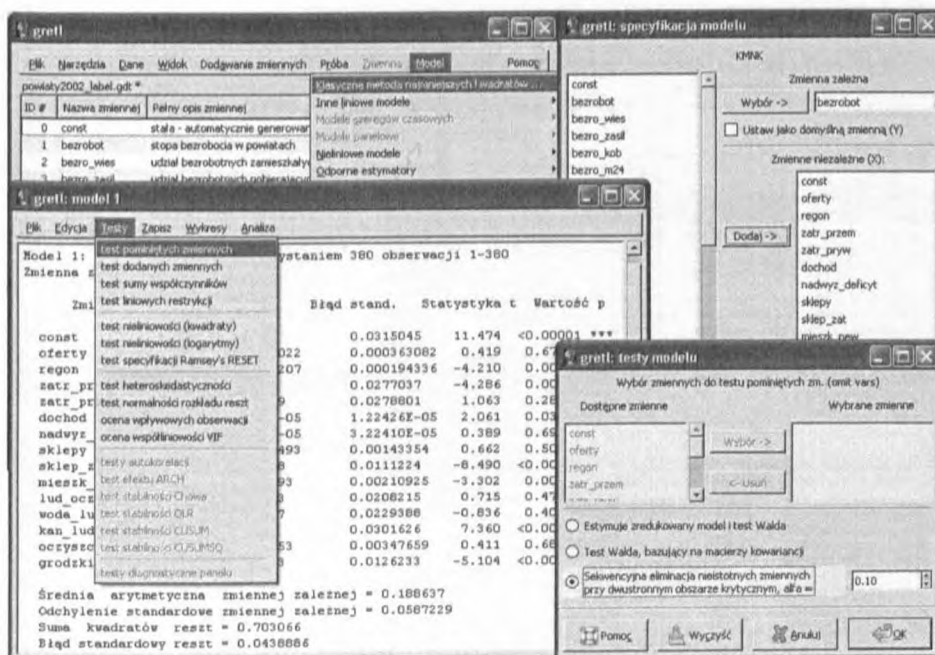
W ramach przedmiotu statystyki można realizować zadania związane z prostymi równaniami regresji. Budowę równania regresji można wykonać w oparciu o wykresy punktowe XY, a pełne wyniki estymacji pozwalają dokonać całościowej oceny równania.

Zbiór informacji statystycznych dostępny w zbudowanych bazach GRETLa prezentujących informacje z Banku Danych Regionalnych GUS w przekrojach powiatów i województw dla lat od 1999 do 2006 pozwala przydzielić zadania – projekty o podobnej skali trudności dla dużej grupy studenckiej.

3. PROJEKTY Z EKONOMETRII I PROGNOZOWANIA

Projekty z ekonometrii można podzielić na dwie grupy w zależności od typu danych statystycznych, to jest dla danych przekrojowych i dla szeregów czasowych. Projekty dla danych przekrojowych można realizować na podstawie danych z baz Banku Danych Regionalnych GUS. Kategorii endogenicznych w Banku Danych Regionalnych dla przekroju powiatów jest kilkanaście i zaliczyć do nich można: stopa bezrobocia, wskaźnik zatrudnienia, przeciętne wynagrodzenie, produkcja rolna, budownictwo (inwestycje oddane do użytku – np. mieszkania), oczyszczalnie ścieków, produkcja przemysłowa, turystyka, itp. Duża ilość kategorii ekonomicznych pozwala zaplanować projekty z zakresu budowy jednorównaniowych modeli ekonometrycznych, a także projekty z zakresu budowy wielorównaniowych współzależnych modeli ekonometrycznych.

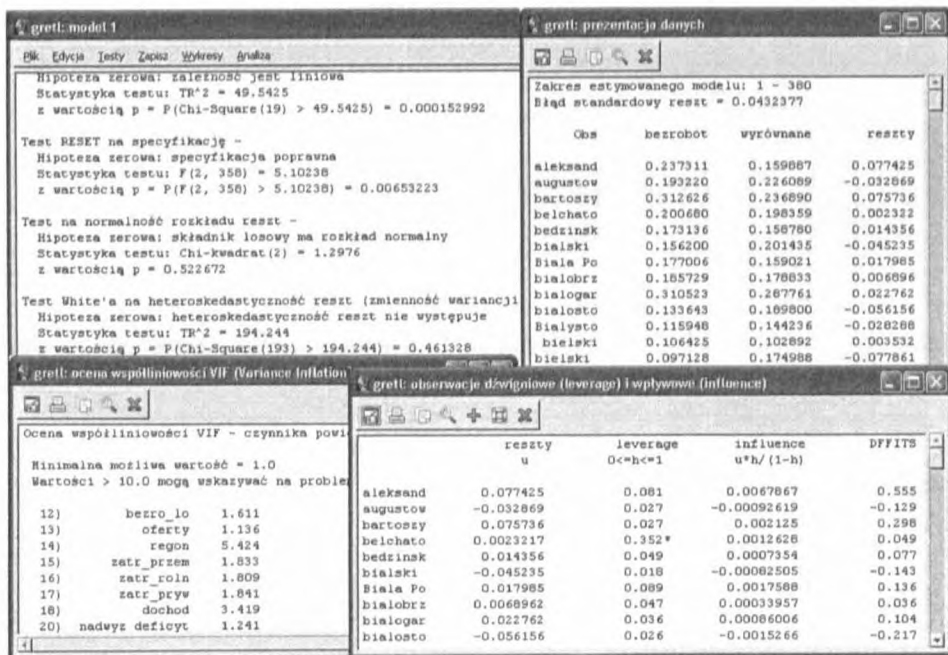
W oprogramowanie GRETLa w sposób intuicyjny można przeprowadzić całą procedurę budowy modelu ekonometrycznego, poprzez etapy: specyfikacji, estymacji, weryfikacji i wnioskowania, a dla szeregów czasowych także predykcji. Rysunek 4 przedstawia okna dialogowe i okna wyników z kolejnymi etapami budowy modelu ekonometrycznego szacowanego KMNK.



Rysunek 4. Okna programu GRETL: specyfikacji modelu, wyniki oszacowanego modelu i testu pominiętych zmiennych.

Na rysunku 4 okno o nazwie *gretl: testy modelu* zawiera automatyczną iteracyjną procedurę sekwencyjnej eliminacji nieistotnych zmiennych, tzw. procedurę eliminacji *a posteriori*, dla wskazanego poziomu istotności. Procedura automatycznej eliminacji jest sprawdzana dodatkowo testem *F* na istotność całego zbioru wyeliminowanych zmiennych i wskazuje ona na pełną poprawność wyeliminowanych nieistotnych zmiennych.

Oprogramowanie GRETL zawiera wiele testów weryfikujących przydatność modelu ekonometrycznego do praktycznego wykorzystania. Wyniki testów są prezentowane w osobnych oknach z wszystkimi pomocniczymi modelami, ale GRETL posiada oprócz właściwego okna wyników, krótkie podsumowania, które prezentują hipotezę zerową, obliczoną statystykę testu oraz najważniejszą informację z punktu widzenia wnioskowania, empiryczny poziom istotności – wartość *p*. Wskazanie wartości *p* umożliwia szybkie podjęcie decyzji weryfikacyjne. Należy dodać, że program GRETL zawiera także wmontowane tablice statystyczne z wartościami krytycznymi rozkładu: normalnego, *t*-Studenta, *chi-kwadrat*, *F* oraz Durbina-Watsona. Ponadto możliwa jest ocena współliniowości zmiennych objaśniających za pomocą wskaźnika *VIF* i ocena wpływowości i dźwigniowych obserwacji, a także identyfikacja nietypowych obserwacji. Rysunek 5 przedstawia przykładowe okna wyników testów i elementów diagnostyki modelu.



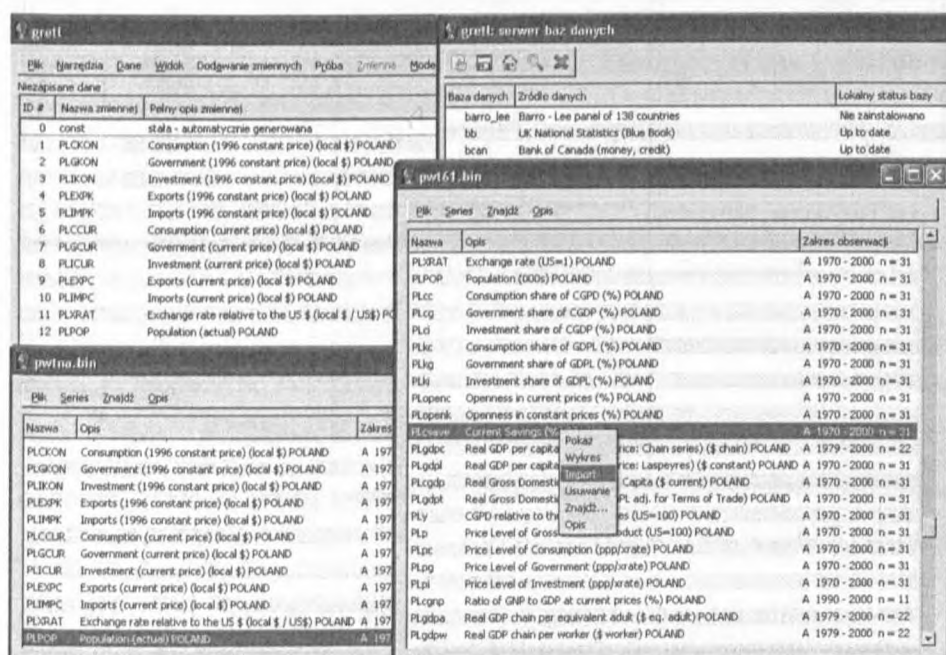
Rysunek 5. Okna programu GRETL: wyniki testów diagnostycznych modelu ekonometrycznego, oceny współliniowości i obserwacji dźwigniowych.

Formułując polecenia do projektu można wskazać na konieczność przeprowadzenia pełnej analizy wyboru modelu ekonometrycznego z spośród kilku oszacowanych modeli za pomocą kryteriów informacyjnych AIC, BIC i HQC.

Kolejną grupę projektów może być związana z danymi w formie szeregów czasowych, które pozwolą oprócz zbudowania modelu ekonometrycznego oszacowanie prognozy i wyznaczenie błędów prognoz.

Bardzo przydatnym zbiorem danych z szeregami czasowymi może być baza danych makroekonomicznych dla wielu krajów, która tworzona jest na Uniwersytecie w Pensylwanii (USA). Baza o nazwie Penn World Table⁵ została przeniesiona do GRETLA i jest dostępna na stronie serwera ricardo.ecn.wfu.edu. Baza ta, obejmując informacje dla około 180 państw, dostępna jest w dwóch wersjach: *pwt61.bin* (24 cechy) z dużą ilością deflatorów cen, a baza *pwtna.bin* – narodowe rachunki posiada 12 cech wyrażonych w cenach bieżących i stałych dla każdego państwa, co przedstawia rysunek 6.

⁵ Pełna informacja o bazie dostępna jest na stronie: <http://pwt.econ.upenn.edu>. Aktualnie dostępna jest baza w wersji 6.2 dla 188 państw dla lat 1950-2004 i co najmniej 25 makroekonomicznych danych o częstotliwości rocznej dla każdego państwa. Do bazy GRETLA przeniesiono wersję 6.1 dla lat 1950-2000.

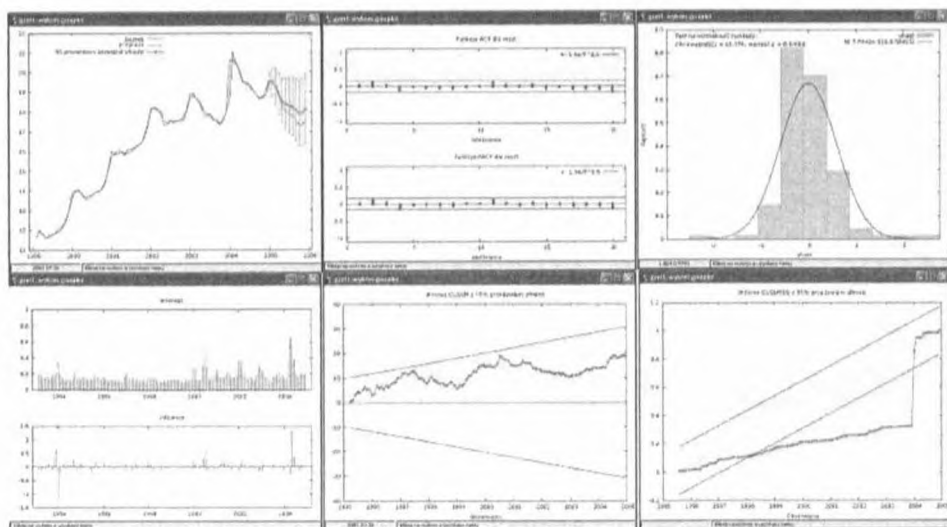


Rysunek 6. Okna programu GRETL: serwer baz danych *ricardo.ecn.wfu.edu* wraz z otwartymi bazami *pwt61.bin* i *pwtna.bin*.

Zaletą pracy z bazami typu *.bin jest to, że do programu GRETL przenosi się tylko potrzebne informacje z danymi (funkcja *Import* – patrz rys. 6). Wczytanie całej bazy, która zawiera kilka tysięcy szeregów czasowych, bardzo utrudniałoby znalezienie i wybór właściwych kategorii ekonomicznych.

Powyższe bazy mogą być podstawą do budowy prognoz z różnych typów modeli trendu, np. wielomiany zmiennej czasowej, modele ADL, ARMA, ARIMA, a przede wszystkim do budowy modeli przyczynowo-skutkowych z wykorzystaniem do predykcji.

Budowa dynamicznych modeli ekonometrycznych pozwala w pełni wykorzystać wszystkie elementy diagnostyczne modelu oprogramowania GRETL. Oprogramowanie GRETL, oprócz szerokich diagnostycznych elementów, zawiera wiele metod estymacji z zakresu uogólnionych metod najmniejszych kwadratów, ale do najważniejszych zalet należy zaliczyć graficzną prezentację wyników. Przykłady możliwych wykresów w etapie diagnostycznym modelu przedstawia rysunek 7.



Rysunek 7. Okna programu GRETL: przykłady wykresów związanych z diagnostyką modelu

Opracowanie założeń projektu z ekonometrii dla studentów kierunku, np. zarządzanie, ekonomia może dotyczyć tylko budowy jednorównaniowych modeli ekonometrycznych z pełną diagnostyką modeli, predykcją i z jej błędami. Dla kierunku informatyka i ekonometria można wprowadzić dodatkowe tematy związane z testami stopnia integracji procesów (test ADF, ADF-GLS, KPSS), modelami VAR, VECM, a także modelami ARCH i GARCH.

4. PODSUMOWANIE

Podsumowując przydatność oprogramowania GRETL do nauczania ekonometrii, dla użytkownika, który pierwszy raz rozpoczyna budować modele ekonometryczne, należy w pierwszej kolejności wskazać na jego bezpłatną dostępność jako programu z grupy *open source*, a także na interfejs graficzny użytkownika w języku polskim⁶.

Dodatkowym atutem jest także bezpośredni dostęp z poziomu oprogramowania GRETL do internetowych baz danych statystycznych, w których informacje ekonomiczne dotyczą: Stanów Zjednoczonych, Kanady, Japonii i krajów Europy zachodniej, a także możliwość zbudowania własnych baz danych.

Coraz częściej wykładowcy przedmiotów: statystyka, ekonometria, prognozowanie i symulacji zaczynają wspomagać nauczanie tych przedmiotów narzędziami

⁶ Tłumaczenie wykonywane jest przez Tadeusza Kufela i Pawła Kufela (UMK Toruń).

dziami informatycznymi, a oprogramowanie GRETŁ okazało się bardzo dobrym pomocnikiem w nauczaniu tych przedmiotów. Ponadto, bardzo przydatnym narzędziem okazało się także w realizacji zadań w formie projektów – prac do samodzielnego wykonania przez studenta.

LITERATURA

- Adkins L. (2007), *Using GRETŁ for Principles of Econometrics*, 3rd edition, <http://www.learneconometrics.com/gretl.html>.
- Baiocchi G., Distaso W. (2003), GRETŁ: Econometric Software for the GNU Generation, *Journal of Applied Econometrics*, vol. 18, s. 105-110.
- Cottrell A., Lucchetti R. 'Jack' (2008), *Gretl User's Guide, GNU Regression, Econometrics and Time Series*, <http://gretl.sourceforge.net>.
- Kufel, T. (2007), *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETŁ*, WN PWN, Warszawa.
- Mixon J.W., Smith R.J. (2006), Teaching Undergraduate Econometrics with GRETŁ, *Journal of Applied Econometrics*, vol. 21, s. 1103-1107.
- Mixon J.W., Smith R.J. (2006), *A Visual Guide to Using GRETŁ*, <http://csob.berry.edu/faculty/economics/gretlguide/GretlGuide.html>.
- Osińska M. (2006), *Ekonometria współczesna*, TNOiK, Toruń.
- Piłatowska M. (2006), *Repetitorium ze statystyki*, Wydawnictwo Naukowe PWN, Warszawa.

Tadeusz Kufel

E-PROJECTS IN TEACHING QUANTITATIVE METHODS WITH USE OF GRETŁ PACKAGE

The purpose of the paper is to present possibilities of assignments e-project in statistics, econometrics and forecasting courses. In GRETŁ there is a big database of statistical information for United States, Great Britain, Japan and European countries. There is also another database for Polish districts in 1999–2006, which contains over 2000 variables.

This huge number of data ensure that student's assignments for statistics and econometrics courses can be used for large number of students at the same level of difficulty.