

Документ подписан простой электронной подписью

Информация о владельце:

ФИО: Макаренко Елена Николаевна

Должность: Ректор

Дата подписания: 10.11.2021 16:54:04

Уникальный программный идентификатор:

c098bc0c1041cb2a4cf926cf171d6715d99a6ae00adc8e27b55cbe1e2dbd7c78

## Приложение 2 к РПД «Эконометрическое моделирование пространственных данных и временных рядов»

### МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Методические указания адресованы аспирантам всех форм обучения.

Учебным планом предусмотрены следующие виды занятий:

- лекции;
- лабораторные занятия.

#### **1. Методические рекомендации по организации работы аспирантов во время проведения лекционных занятий**

В ходе лекционных занятий аспиранту целесообразно вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание изучаемой дисциплины, научные выводы и практические рекомендации, положительный опыт в ораторском искусстве. Рекомендуется задавать преподавателю уточняющие вопросы по теме лекции.

#### **2. Методические рекомендации по подготовке аспирантов к лабораторным занятиям**

Лабораторные занятия способствуют изучению наиболее важных тем учебной дисциплины. углубляются и закрепляются знания аспиранты по рассмотренным на лекциях вопросам, развиваются навыки применения эконометрических методов для решения конкретных задач.

При подготовке к лабораторным занятиям аспирант должен:

- изучить рекомендованную учебную литературу;
- изучить конспекты лекций;
- подготовить ответы на все вопросы по изучаемой теме;
- письменно решить домашние задания, рекомендованные преподавателем при изучении каждой темы.

В процессе подготовки к лабораторным занятиям аспиранты могут воспользоваться консультациями преподавателя.

Вопросы, не рассмотренные на лекциях и лабораторных занятиях, должны быть изучены аспирантами в ходе самостоятельной работы. Цель проведения самостоятельной работы обучающихся – освоение теоретического материала по изучаемой дисциплине, углубление и расширение теоретических знаний; систематизация и закрепление полученных теоретических знаний и практических навыков; формирование умений по поиску и использованию справочной и специальной литературы, а также других источников информации; развитие познавательных способностей и активности обучающихся, творческой инициативы, ответственности и организованности; формирование самостоятельности мышления, способностей к саморазвитию, самообразованию, самосовершенствованию и самореализации; развитие научно-исследовательских навыков; формирование умения применять полученные знания на практике (профессиональной деятельности).

Контроль самостоятельной работы аспирантов над учебной программой курса осуществляется в ходе занятий методом устного опроса или посредством тестирования. В ходе самостоятельной работы каждый аспирант обязан прочитать основную и по возможности дополнительную литературу по изучаемой теме, дополнить конспекты лекций недостающим материалом, выписками из рекомендованных первоис-

точников. Выделить непонятные термины, найти их значение в энциклопедических словарях.

Для подготовки к занятиям, текущему контролю и промежуточной аттестации аспиранты могут воспользоваться электронно-библиотечными системами. Также обучающиеся могут взять на дом необходимую литературу на абонементе университетской библиотеки или воспользоваться читальными залами.

### **3. Методические рекомендации к составлению опорного конспекта**

Можно подготовить опорный конспект в виде развернутого плана ответа на один или несколько вопросов, выносимых на практическое занятие. Составление опорного конспекта представляет собой вид внеаудиторной самостоятельной работы аспиранта по созданию краткой информационной структуры, обобщающей и отражающей суть материала лекции, учебника, других источников информации. Опорный конспект призван выделить главные объекты изучения, дать им краткую характеристику, используя символы, отразить связь с другими элементами. Основная цель опорного конспекта – усилить понимание и облегчить запоминание материала.

#### Основные рекомендации к составлению опорного конспекта:

- объем опорного конспекта - примерно одна страница формата А4 или две страницы формата А5 (формат страницы в тетради), текст должен быть таким, чтобы его можно было воспроизвести за 2-4 минуты;
- структурность – материал желательно располагать малыми логическими блоками, т.е. он должен содержать несколько отдельных пунктов, вместе с тем текст должен быть наглядным и понятным. В его составлении используются различные базовые понятия, термины, знаки (символы) – опорные сигналы. Опорный конспект может быть представлен системой взаимосвязанных геометрических фигур, содержащих блоки концентрированной информации в виде ступенек логической лестницы; рисунка с дополнительными элементами и др.;
- акцентирование – для лучшего запоминания основного смысла, главную идею вопроса целесообразно выделить (шрифтом: другим, жирным, курсивом или использовать разные цветовые варианты оформления текста).

#### Примерный порядок составления опорного конспекта:

- ознакомления с материалом изучаемой темы по тексту лекции, рекомендованной основной и дополнительной литературы;
- выделение главного в изучаемом материале, составление обычных кратких записей;
- подбор к данному тексту опорных сигналов в виде отдельных слов, определенных схем, рисунков, графиков, таблиц;
- продумывание схематического способа представления изученного материала, использование различного шрифта и т.п.;
- составление опорного конспекта.

### **4. Методические рекомендации по решению кейс-задания**

В целях расширения и закрепления полученных знаний при изучении данной дисциплины, аспиранту предлагается ответить на вопросы ситуационных заданий.

Выполнение ситуационного задания преследует главную цель – использовать возможности активного, самостоятельного обучения в сочетании с другими формами учебных занятий и заданий по дисциплине.

Выполнение индивидуального задания в течение семестра позволяет решать следующие задачи обучения:

- глубже изучить отдельные темы учебной дисциплины;
- четко и последовательно выражать свои мысли.

Перед решением кейс-задания должно быть полностью приведено его условие. Само решение следует сопровождать необходимыми расчетами и пояснениями с указанием применяемых формул, анализом и выводами.

В процессе пояснения решения важно учитывать этику, краткость и четкость изложения своей позиции,

профессиональный подход к решению проблем, научный подход, умение формулировать обоснованные выводы.

### 5. Методические рекомендации по выполнению индивидуальных творческих заданий.

Согласно выбранной теме выполняется эконометрическое исследование по следующей схеме.

1. Формулируются конкретные гипотезы, подлежащие теоретическому обоснованию или эмпирической проверке. Объясняется, в чем заключается актуальность данных гипотез с научной точки зрения. При этом необходимо обратить внимание на то, что целью исследования должно быть объяснение экономических явлений. Здесь же следует охарактеризовать контекст исследования с точки зрения проблем экономической или социальной политики, связанных с темой работы. Дать краткий обзор альтернативных точек зрения и/или предложений для решения этих проблем. Объяснить, каким образом результаты расчетов могут быть использованы при оценке существующих предложений, расчете параметров экономической политики, внесению поправок в законодательство и т.п.

Рекомендуется описать и оценить результаты исследований по выбранной теме и известные подходы к ее изучению. Нельзя ни в коем случае ограничиваться только перечнем авторов. В ситуации модификации известной теоретической модели, описывается ее формальная структура. Необходимо обсудить количественные результаты, полученные другими исследователями (по другим регионам, странам и т.п.), объяснить, чем отличается предлагаемый подход, какие результаты рассчитывает получить автор. Обзор литературы тесно связан с постановкой задачи исследования.

2. Подробно описываются источники, структура, методы расчета используемых в анализе переменных. Должны быть указаны достоинства и недостатки используемой в эмпирическом анализе выборочной совокупности. В случае следования (построения) какой-либо теоретической модели необходимо описать математически ее основные параметры, остановиться на связи между моделью и данными, объяснить, в какой мере она отвечает на вопросы, поставленные в первой части работы.

3. Представляется формулировка эконометрической модели, описываются методы расчета зависимых и независимых переменных на основе имеющихся данных и обоснование их включения в модель с точки зрения экономической теории. Проводится анализ описательных статистик. Представляются таблично и графически оформленные согласно правилам результаты эконометрических расчетов, обосновывается на основе соответствующих статистических критериев адекватность построенной модели/моделей. Даются возможные способы интерпретации полученных результатов исследования в свете экономической теории и их практическая значимость. Объясняется в какой мере подтверждаются/опровергаются гипотезы исследования (из первой части). В библиографическом списке указываются основные источники, на которые были сделаны ссылки.

Требования к оформлению задания.

Пояснительная записка оформляется согласно требований ГОСТ 7.32-2001 «Отчет о научно-исследовательской работе. Структура и правила оформления».

### 6. Методические указания по выполнению лабораторных работ

Лабораторная работа 1. Спецификация эконометрических моделей в Eviews.

Традиционной задачей эмпирических исследований в микроэкономике является оценивание кривых Энгеля. Эрнст Энгель установил, что при увеличении дохода семьи доля расходов на питание уменьшается (закон Энгеля). В современных микроэкономических терминах это означает, что эластичность расходов на питание по доходу меньше единицы.<sup>1</sup> При этом имеют в виду, что еда является необходимым товаром, а не предметом роскоши. Зависимость расходов на приобретение некоторого вида товара от доходов называется *кривой Энгеля*. В настоящее время принято, как правило, вместо дохода рассматривать полные расходы.

Файл `expend.wfl` содержит данные, полученные из архива данных журнала *Journal of Applied Econometrics*. Для наших целей мы взяли данные о годовых расходах на питание, отдых и другие товары за период с октября 1986 г. по сентябрь 1987 г. (427 наблюдений). Описание переменных: *f3* – расходы на питание семьи в голландских гульденах (Dfl); *v3* – расходы на отдых, Dfl; *tot3* – полные расходы, Dfl; *prov* – провинция, в которой проживает семья; *reg* – регион проживания; *scl* – социальный класс (1 – нижний класс, ..., 5 – верхний класс); *nahm* – число членов семьи старше 11 лет; *durb* – степень урбанизации (1 – маленькая деревня, ..., 13 – большой город); *nch06* – число детей младше 6 лет в семье; *nch711* – число детей от 7 до 11 лет; *nch1217* – число детей от 12 до 17 лет; *nch18* – число детей старше 18 лет.

Вычислим описательные статистики для всех переменных (в окне группы

**View\Descriptive Stats\Individual Samples**).

Статистика	F3	TOT3	PROV	REG	SCL	NAHM	DURB	NCH06	NCH1217	NCH18	NCH711
Среднее	6890,88	35762,81	7,28	3,37	3,16	2,18	8,26	0,27	0,23	1,12	0,25
Медиана	6838,43	31847,12	7,00	4,00	3,00	2,00	8,00	0,00	0,00	1,00	0,00

<sup>1</sup>Напомним, что эластичность  $y$  по  $x$  - это  $\frac{x}{y} \frac{dy}{dx}$  или  $d[\ln y]/d[\ln x]$ . Она показывает, на сколько процентов изменяется  $y$ , если  $x$  увеличивается на 1%.

Максимум	17270,75	124564,20	13,00	6,00	5,00	5,00	13,00	3,00	2,00	4,00	3,00
Минимум	883,98	8728,86	1,00	1,00	1,00	1,00	3,00	0,00	0,00	0,00	0,00
Стандартное отклонение	2745,74	18053,31	3,02	1,42	1,09	0,88	2,95	0,63	0,54	1,04	0,57
Асимметрия	0,29	1,47	-0,25	-0,03	-0,44	1,01	0,11	2,33	2,29	0,24	2,29
Экссесс	3,23	6,42	2,23	1,67	2,21	3,88	1,70	7,79	7,07	1,96	7,56

Для определения тесноты линейной связи между некоторыми из переменных вычислим корреляционную матрицу (**Open\as Group\View\Covariance Analysis** и опция **Correlation** в появившемся меню) для переменных  $f3$ ,  $tot3$  и  $nahm$ :

Переменные	F3	Tot3	Nahm
F3	1,00	0,50	0,51
Tot3	0,50	1,00	0,33
Nahm	0,51	0,33	1,00

Как и ожидалось, линейная связь имеется между расходами на питание и полными расходами семьи, а также между расходами на питание и числом членов семьи, что выражается в высоких значениях коэффициентов корреляции (0,50 и 0,51 соответственно).

Построим линейную регрессионную модель для объяснения логарифма расходов на питание  $lf3$  в зависимости от логарифма полных расходов  $ltot3$ :

$$lf3 = \beta_0 + \beta_1 \cdot ltot3 + u.$$

Результаты оценивания коэффициентов модели методом наименьших квадратов представлены в таблице 3.

Таблица 3

Результаты оценивания регрессионной модели зависимости логарифма расходов на питание от логарифма полных расходов

Dependent Variable: LF3 Method: Least Squares  
Sample: 1 427 Includ observations: 427

	Coefficient	Std. Error	t-Statistic	Prob.
LTOT3	0.549213	0.039579	13.87632	0.0000
C	3.046584	0.410839	7.415526	0.0000
R-squared	0.311799	Mean dependent var		8.741329
Adjusted R-squared	0.310180	S.D. dependent var		0.475891
S.E. of regression	0.395253	Akaike info criterion		0.986094
Sum squared resid	66.39575	Schwarz criterion		1.005095
Log likelihood	-208.5310	Hannan-Quinn criter.		0.993599
F-statistic	192.5523	Durbin-Watson stat		1.988620
Prob(F-statistic)	0.000000			

Оцененная эластичность расходов на питание по доходу (точнее по полным расходам, которые являются прокси-переменной, т.е. переменной заменителем для доходов) равна 0,54. Доверительный 95% интервал для оцененной эластичности имеет вид:

$$\beta = \hat{\beta} \pm t_{\varepsilon/2, n-2} \sqrt{V(\hat{\beta})} = 0,54 \pm 1,960 \cdot 0,04 \text{ или } 0,462 \leq \beta \leq 0,618.$$

Однако в оцененной регрессии не принимается во внимание размер семьи. Это приводит к смещенности оценки, поскольку полные доходы зависят от количества взрослых членов семьи. Поэтому мы, возможно, переоцениваем эластичность по доходу.

Построим регрессию  $lf3$  на переменные  $ltot3$ ,  $nahm$ ,  $nch06$ ,  $nch711$ ,  $scl$ ,  $dur$  и константу. Получим согласно результатам расчетов, представленным в таблице 4, регрессию:

$$lf3 = 4,21 + 0,38 \cdot ltot3 + 0,04 \cdot scl + 0,11 \cdot nch711 + 0,14 \cdot nch06 + 0,18 \cdot nahm - 0,01 \cdot dur.$$

Таблица 4

Результаты оценивания регрессионной модели зависимости логарифма расходов на питание от логарифма полных расходов, социального класса, числа членов семьи младше 6 лет, от 7 до 11 лет, старше 11 лет и степени урбанизации

Dependent Variable: LF3 Method: Least Squares  
Sample: 1 427 Included observations: 427

	Coefficient	Std. Error	t-Statistic	Prob.
LTOT3	0.389599	0.041336	9.425186	0.0000
SCL	0.037727	0.016537	2.281301	0.0230
NCH711	0.115297	0.029163	3.953605	0.0001
NCH06	0.136342	0.026470	5.150799	0.0000
NAHM	0.187000	0.021102	8.861905	0.0000
DURB	-0.012629	0.005763	-2.191474	0.0290
C	4.210594	0.442655	9.512145	0.0000
R-squared	0.497318	Mean dependent var		8.741329
Adjusted R-squared	0.490137	S.D. dependent var		0.475891
S.E. of regression	0.339809	Akaike info criterion		0.695390
Sum squared resid	48.49740	Schwarz criterion		0.761895
Log likelihood	-141.4658	Hannan-Quinn criter.		0.721658
F-statistic	69.25302	Durbin-Watson stat		1.971984
Prob(F-statistic)	0.000000			

Как мы и предполагали, эластичность уменьшилась и стала равна 0,38. Проверим, отличается ли влияние количества детей в возрасте до 6 лет от влияния их количества в возрасте от 7 до 11 лет, т.е. равенство коэффициентов при переменных *nch711* и *nch06*.

Тест Вальда дает результаты, не позволяющие говорить о значимых различиях во влиянии количества детей разных возрастов на величину потребления продуктов.

Wald Test:

Null Hypothesis: C(3)=C(4)

F-statistic	0.271716	Probability	0.602457
Chi-square	0.271716	Probability	0.602183

Построим график зависимости переменной *lf3* от переменной *ltot3* (рис. 6).

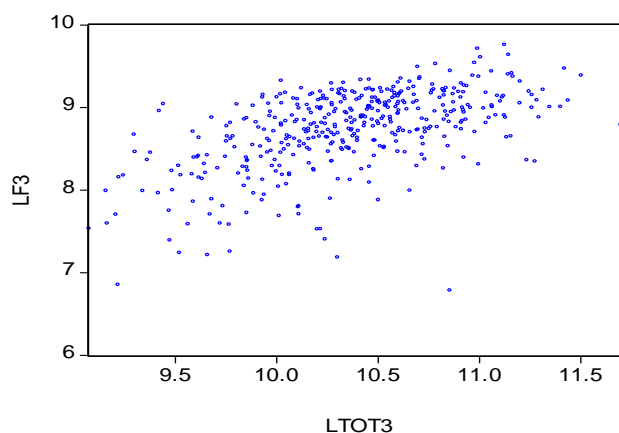


Рис. 6. График зависимости логарифма расходов на питание от логарифма полных расходов

На графике видна параболическая зависимость от переменной *ltot3*, поэтому уточним спецификацию модели, включив в нее квадрат указанной переменной.

Построим регрессию *lf3* на переменные *ltot3*, *ltot3s*, *nahm*, *nch06*, *nch711*, *scl*, *durb* и константу (переменная *ltot3s* получена как  $ltot3s = ltot3 \cdot ltot3$ ). Результаты представлены в таблице 5.

Таблица 5

Результаты оценивания регрессионной модели зависимости логарифма расходов на питание от логарифма полных расходов, его квадрата, социального класса, числа членов семьи младше 6 лет, от 7 до 11 лет, старше 11 лет и степени урбанизации

Dependent Variable: LF3  
Method: Least Squares  
Sample: 1 427  
Included observations: 427

	Coefficient	Std. Error	t-Statistic	Prob.
LTOT3	3.498138	1.066525	3.279938	0.0011
LTOT3S	-0.150076	0.051452	-2.916794	0.0037
SCL	0.031979	0.016510	1.936963	0.0534

NCH711	0.109975	0.028963	3.797111	0.0002
NCH06	0.124841	0.026531	4.705400	0.0000
NAHM	0.183839	0.020944	8.777831	0.0000
DURB	-0.013064	0.005714	-2.286442	0.0227
C	-11.81793	5.512742	-2.143749	0.0326
<hr/>				
R-squared	0.507322	Mean dependent var	8.741329	
Adjusted R-squared	0.499091	S.D. dependent var	0.475891	
S.E. of regression	0.336812	Akaike info criterion	0.679973	
Sum squared resid	47.53227	Schwarz criterion	0.755978	
Log likelihood	-137.1742	Hannan-Quinn criter.	0.709993	
F-statistic	61.63619	Durbin-Watson stat	2.011364	
Prob(F-statistic)	0.000000			

Воспользуемся полученными результатами для оценивания эластичности при различных уровнях дохода *tot3*. Эластичность расходов на питание будет равна:

$$\frac{d\ln f_3}{d\ln tot3} (3,49 \cdot \ln tot3 - 0,15 \cdot \ln^2 tot3) = 3,49 - 2 \cdot 0,15 \cdot \ln tot3.$$

Для среднего дохода (35762 гульдена) она составит  $3,49 - 2 \cdot 0,15 \cdot \ln(35762) = 0,34$ , что меньше единицы, т.е. спрос на продукты питания неэластичен и еда является необходимым товаром.

Построим график зависимости остатков регрессии (переменная *e*) от *tot3* (рис. 7).

График не дает права предполагать независимость ошибок регрессии и регрессора *tot3*, что свидетельствует о гетероскедастичности остатков в построенной модели регрессии.

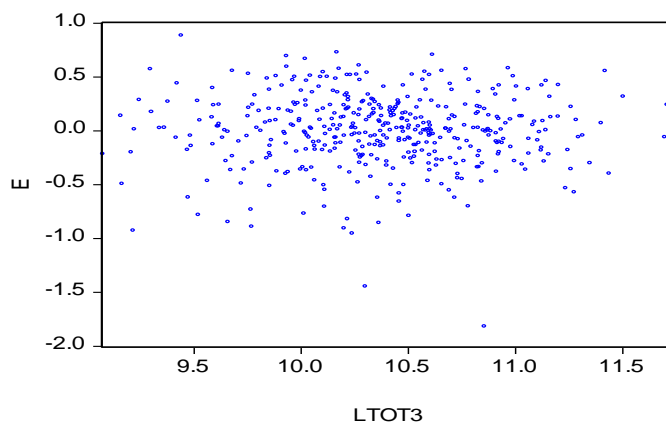


Рис. 7. График остатков регрессионной модели

Этот факт ожидаем, так как ясно, что вариация расходов на питание для семей с различным уровнем дохода будет непостоянна, а именно: для семей с низким уровнем доходов дисперсия расходов на питание больше, а для семей с высоким уровнем доходов дисперсия расходов на питание – меньше.

Тестирование на гомоскедастичность тестом Уайта (в окне уравнения **View\Residual Tests\Heteroskedasticity Tests...White** оставив флажок для опции **Include White cross terms**) показывает (табл. 6), что дисперсия случайной ошибки модели непостоянна и, следовательно, стандартные ошибки построенного уравнения регрессии нуждаются в коррекции.

Таблица 6

Результаты теста Уайта

Heteroskedasticity Test: White

F-statistic	1.632281	Prob. F(33,393)	0.0171
Obs*R-squared	51.47072	Prob. Chi-Square(33)	0.0212
Scaled explained SS	110.9721	Prob. Chi-Square(33)	0.0000

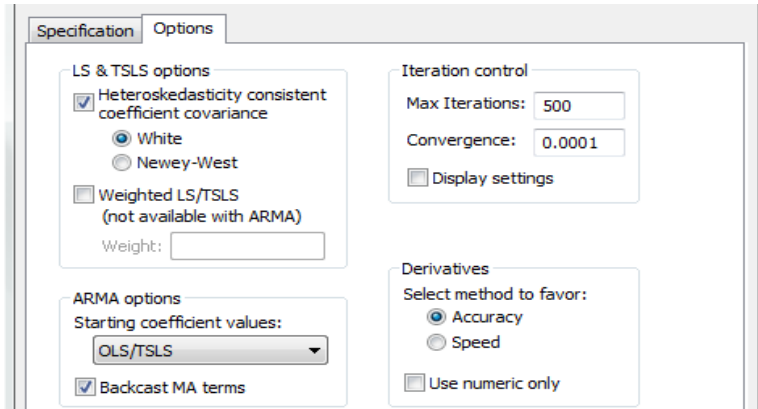
Результаты статистики Дарбина-Уотсона не дают повода усомниться в отсутствии автокорреляции остатков ( $DW = 2,01$ ). Вывод не изменяется и в случае непосредственного тестирования в окне уравнения путем **View\Residual Tests\Serial Correlation LM Test...** и, указав в окне для **Lags to include** количество лагов 1:

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.034320	Prob. F(1,418)	0.8531
-------------	----------	----------------	--------

Obs\*R-squared                      0.035056    Prob. Chi-Square(1)                      0.8515

Для состоятельного оценивания стандартных ошибок коэффициентов (с учетом гетероскедастичности остатков) воспользуемся процедурой Уайта. Нажав кнопку **Estimate** в окне уравнения и выбрав затем вкладку **Options** необходимо отметить **Heteroskedasticity consistent coefficient covariance**:



Результаты представлены в табл. 7.

Таблица 7

Результаты регрессионного анализа с использованием процедуры Уайта

Dependent Variable: LF3

Method: Least Squares

Sample: 1 427

Included observations: 427

White Heteroskedasticity-Consistent Standard Errors & Covariance

	Coefficient	Std. Error	t-Statistic	Prob.
LTOT3	3.498138	1.153464	3.032723	0.0026
LTOT3S	-0.150076	0.055281	-2.714797	0.0069
SCL	0.031979	0.016716	1.913049	0.0564
NCH711	0.109975	0.020076	5.477827	0.0000
NCH06	0.124841	0.021937	5.690881	0.0000
NAHM	0.183839	0.023273	7.899304	0.0000
DURB	-0.013064	0.005688	-2.296756	0.0221
C	-11.81793	5.995256	-1.971214	0.0494
R-squared	0.507322	Mean dependent var		8.741329
Adjusted R-squared	0.499091	S.D. dependent var		0.475891
S.E. of regression	0.336812	Akaike info criterion		0.679973
Sum squared resid	47.53227	Schwarz criterion		0.755978
Log likelihood	-137.1742	Hannan-Quinn criter.		0.709993
F-statistic	61.63619	Durbin-Watson stat		2.011364
Prob(F-statistic)	0.000000			

Результаты практически соответствуют предыдущим, однако стандартные ошибки увеличились.

#### Задание для самостоятельной работы.

В файле flat.wfl содержатся данные о стоимости квартир в Москве. Постройте модель стоимости квартиры (или стоимости квадратного метра жилой площади квартиры) в зависимости от имеющихся параметров. Тестируйте гипотезу, что уравнения для цены различны для 1, 2, 3-4 комнатных квартир, т.е. гипотезу, что рынок распадается на рынки однокомнатных, двухкомнатных и трех-четырёхкомнатных квартир.

Лабораторная работа 2. Оценка производственной функции Кобба-Дугласа.

#### Описание файлов данных

В файле TIO2 содержится 16 наблюдений о производстве диоксида титана компанией DuPont Corporation с 1955 по 1970 гг. Переменные:

YEART – год;

PRODT – текущий выпуск;

CUMT – суммарный выпуск в момент  $t-1$ ;

UCOSTT – издержки на единицу продукции, недефлятированные;

DEFL – дефлятор.

В файле POLY содержится 13 наблюдений о производителе полиэтилена с 1960 по 1972 гг.

Переменные:

YEARP – год;

UCOSTP – дефлятированные издержки на единицу продукции;

PRODP – текущий выпуск;

CUMP – суммарный выпуск в момент  $t-1$ .

В файле WEIRD 15 наблюдений над переменными Y, X1 и X2.

В файле NERLOV данные 1955 г. о 145 компаниях по производству электроэнергии.

Переменные:

ORDER – номер наблюдения, упорядочены в порядке возрастания;

COSTS – общие издержки, в млн. долл.;

KWH – выпуск в киловатт-часах, в млрд. кВт-ч.;

PL – уровень зарплаты за час;

PF – стоимость топлива в центах за млн. ед.;

PK – индекс цен рентного дохода от капитала.

В файле UPDATE обновленные данные 1970 г. о 99 компаниях по производству электроэнергии.

Переменные:

OBSNO – номер наблюдения;

COST70 – общие издержки, в млн. долл.;

KWH70 – выпуск в киловатт-часах, в млн. кВт-ч.;

PL70 – цена труда;

PK70 – индекс цен рентного дохода от капитала;

PF70 – индекс цен на топливо;

SL – доля труда в издержках;

SK – доля капитала в издержках;

SF – доля топлива в издержках.

### **Упражнение 1.** Оценка параметров кривой обучения

Выберите один из файлов POLY или TIO2 для работы в этом упражнении. Если вы выбрали TIO2, то вы должны дефлятировать UCOSTT делением на DEFL.

**а)** Создайте переменную LNCP, равную логарифму суммарного выпуска в момент времени  $t-1$ . Также создайте прологарифмированные переменные LNUC и LNY для средних издержек и текущего выпуска соответственно. Постройте график зависимости средних издержек от суммарного выпуска, а также график LNUC от LNCP. Прокомментируйте оба графика и сделайте вывод о виде кривой обучения.

**б)** Используя МНК, оцените параметры кривой обучения в регрессии LNUC на константу и LNCP. Дайте интерпретацию оцененной эластичности кривой обучения. Полученная оценка приемлема? Каков соответствующий наклон кривой обучения? Постройте доверительный интервал для оцененной эластичности кривой обучения, и проверьте гипотезу о равенстве эластичности нулю.

### **Упражнение 2.** Тестирование спецификации кривой обучения

Выберите один из файлов POLY или TIO2 для работы в этом упражнении. Если вы выбрали TIO2, то вы должны дефлятировать UCOSTT делением на DEFL.

**а)** Создайте переменную LNCP, равную логарифму суммарного выпуска в момент времени  $t-1$ . Также создайте прологарифмированные переменные LNUC и LNY для средних издержек и текущего выпуска соответственно. Используя МНК, оцените множественную регрессию LNUC на константу и LNCP и LNY. Используя приемлемый уровень значимости, проверьте нулевую гипотезу о том, что технология производства характеризуется постоянной отдачей от масштаба, т.е. коэффициент при LNY равен нулю. Какова отдача от масштаба для построенной модели и данных?

**б)** Предположим, что вы оцениваете спецификацию кривой обучения как в *упр. 1, б*, с пропущенной переменной LNY. Для оценки величины смещения эластичности кривой обучения постройте указанную регрессию и вспомогательную регрессию LNY на LNCP. Проверьте численно, что аналитическая связь между двумя оценками кривой эластичности существует. Смещение в вашем случае большое или маленькое?

**в)** На основе результатов, полученных в п. *б*, сравните  $R^2$  для уравнения кривой обучения с включенной переменной LNY с коэффициентом детерминации для уравнения без переменной LNY. Почему последний больше? Это всегда так получается? Почему?

**г)** Используя приемлемый уровень значимости, проверьте нулевую гипотезу о том, что оба коэффициента наклона регрессии в п. *а*) равны одновременно нулю (прил. 2). Выполните это тестирование на основе дисперсионного анализа. Сравните результаты тестирования с использованием  $F$ -критерия с результатами отдельных  $t$ -тестов. Результаты взаимно согласованы? Почему?



**Упражнение 3.**  $R^2$  в парной и множественной регрессиях: сюрприз

Для работы в этом упражнении используйте файл данных WEIRD.

- а)** Вычислите парные коэффициенты корреляции для переменных  $Y$ ,  $X_1$  и  $X_2$ . Их значения странные? Почему?
- б)** Прокомментируйте, почему для временных рядов  $R^2$  для множественной регрессии  $Y$  на константу и  $K$  регрессоров, как правило, меньше, чем сумма  $R^2$  для  $K$  парных регрессий  $Y$  на константу и  $X_k$ ,  $k=1, \dots, K$ .
- в)** Теперь постройте МНК парную регрессию  $Y$  на константу и  $X_1$ . Постройте МНК парную регрессию  $Y$  на константу и  $X_2$ . Есть какие-либо неожиданные результаты в  $t$ -статистиках обеих регрессий? Вычислите сумму  $R^2$  для двух регрессий.
- г)** Далее постройте МНК множественную регрессию  $Y$  на константу,  $X_1$  и  $X_2$ . Есть какие-либо неожиданные результаты в  $t$ -статистиках регрессии? Каков  $R^2$  в построенной регрессии? Сравните  $R^2$  с суммой  $R^2$  из п. **в**. Почему результат довольно неожиданный? Как это можно объяснить?

**Упражнение 4.** Повторение классических результатов Nerlove M. в экономии масштаба

Для работы в этом упражнении используйте файл данных NERLOV.

**а)** Сформируйте новые переменные, необходимые для построения регрессии. А именно:  $\text{LNCP3} = \ln(\text{COSTS/PF})$ ,  $\text{LNP13} = \ln(\text{PL/PF})$ ,  $\text{LNP23} = \ln(\text{PK/PF})$  и  $\text{LNKWH} = \ln(\text{KWH})$ . Убедитесь, что наблюдения упорядочены по  $\text{LNKWH}$ .

**б)** Оцените МНК регрессию

$$\ln C^* = \beta_0 + \beta_y \ln y + \beta_1 \ln p_1^* + \beta_2 \ln p_2^*, \quad (2.1)$$

используя переменные, построенные в п. **а**:  $\text{LNCP3} = \ln C^*$ ,  $\text{LNKWH} = \ln y$ ,  $\text{LNP13} = \ln p_1^*$ ,  $\text{LNP23} = \ln p_2^*$ . Nerlove M. получил в 1963 г. следующие результаты для оценок  $\beta_y$ ,  $\beta_1$  и  $\beta_2$  соответственно 0,721, 0,562 и -0,003, со стандартными ошибками 0,175, 0,198 и 0,192 и  $R^2=0,931$ . Ваши расчеты совпадают с результатами Nerlove M.?

**в)** Используя результаты п. **б** и приемлемый уровень значимости, постройте доверительный интервал для  $\beta_y$ . Проверьте гипотезу  $\beta_y=1$ . Отдача от масштаба постоянна? Получите значение отдачи от масштаба на основе  $\beta_y$ . Оцененная отдача возрастает, постоянна или убывает? Экономия от масштаба положительна, равна нулю или отрицательна?

**г)** Спрос на каждый из факторов производства существует, если  $\alpha_i$  положителен,  $i=1,2,3$ . Прокомментируйте оцененное значение  $\alpha_2$  из п. **в**. Оно значимо отличается от нуля? Почему, как вы думаете, Nerlove M. был не удовлетворен оценкой для  $\alpha_2$ ?

**д)** Вычислите и постройте график ошибок уравнения регрессии (2.1). Nerlove M. отмечал, что если график ошибок построить в зависимости от логарифма выпуска, то кривая будет иметь вогнутую форму, ошибки для малых значений выпуска будут положительны, для средних значений – отрицательны, а для больших значений – опять положительны. Подтверждают ли этот факт ваши расчеты? Как такой график зависимости может быть проинтерпретирован? Какова корреляция ошибок с  $\text{LNKWH}$ ? Что это за случай?

**Упражнение 5.** Альтернативные спецификации уравнения для отдачи от масштаба

Для работы в этом упражнении используйте файл данных NERLOV.

**а)** Следуя за Nerlove M., разбейте выборку из 145 фирм на 5 подвыборок, в каждой по 29 фирм. Поскольку данные упорядочены по величине выпуска, 29 первых фирм имеют наименьший выпуск и 29 последних фирм имеют наибольший выпуск. Оцените уравнение регрессии (2.1) для каждой из пяти подвыборок. Nerlove M. получил следующие результаты (стандартные ошибки в круглых скобках):

	Параметр				
	$\beta_y$	$\beta_1$	$\beta_2$		$R^2$
Подвыборка 1	0,398 (0,079)	0,641 (0,691)	-0,093 (0,669)		0,512
Подвыборка 2	0,668 (0,116)	0,105 (0,275)	0,364 (0,277)		0,635
Подвыборка 3	0,931 (0,198)	0,408 (0,199)	0,249 (0,189)		0,571
Подвыборка 4	0,915 (0,108)	0,472 (0,174)	0,133 (0,157)		0,871
Подвыборка 5	1,045 (0,065)	0,604 (0,197)	-0,295 (0,175)		0,920

Ваши результаты соответствуют указанным? Как вы прокомментируете различия в результатах?

**б)** На основе оценок  $\beta_y$  из п. **а** вычислите оценки отдачи от масштаба для каждой из пяти подвыборок. Какое образцовое значение отдачи от масштаба соответствует росту выпуска? Как это значение может быть проинтерпретировано? Позволяет ли это говорить об альтернативной спецификации модели?

**в)** Введите новые переменные в уравнение регрессии так чтобы коэффициенты  $\beta_1$  и  $\beta_2$  были одинаковы для всех подвыборок, а константа и  $\beta_y$  оценивались для каждой из подвыборок отдельно. С учетом результатов п. **б** почему такая спецификация приемлема? Оцените такую расширенную модель и сравните пять значений для  $\beta_y$  с полученными Nerlove M.: 0,394 (0,055), 0,651 (0,189), 0,877 (0,376), 0,908 (0,354) и 1,062 (0,169) в круглых скобках – стандартные ошибки. Коэффициенты  $\beta_1$  и  $\beta_2$  по Nerlove M. 0,435 (0,207) и 0,100 (0,196), а  $R^2 = 0,950$ .

**г)** Для каждого из пяти значений  $\beta_y$  из п. **в** вычислите отдачу от масштаба. Что будет происходить с ростом объема выпуска продукции?

д) Сравните результаты в п. в с результатами в п. а. Поскольку результаты п. а составляют частный случай п. в, используя  $F$ -критерий (прил. 2) и приемлемый уровень значимости сформулируйте и проверьте ограничения, налагаемые на уравнение в п. а против альтернативы – регрессии в п. в. Прокомментируйте результат.

е) С учетом факта нелинейности отдачи от масштаба по уровню выпуска, Nerlove M. сформулировал и оценил обобщение модели (2.1), в которую добавил в качестве регрессора переменную  $(\ln y)^2$ . Оцените МНК уравнение регрессии вида:

$$\ln C^* = \beta_0 + \beta_y \ln y + \beta_{yy} (\ln y)^2 + \beta_1 \ln p_1^* + \beta_2 \ln p_2^*.$$

Используя приемлемый уровень значимости, проверьте нулевую гипотезу о том, что отдача от масштаба постоянна, т.е.  $\beta_y=1$ ,  $\beta_{yy}=0$ , против альтернативной гипотезы  $\beta_y \neq 1$ ,  $\beta_{yy} \neq 0$ . Как согласуются результаты  $F$ -критерия (прил. 2) с результатами по  $t$ -критерию для каждого коэффициента в отдельности (прил. 3). Поскольку отдача от масштаба в расширенной модели изменяется в зависимости от уровня выпуска и равна  $r=1/(\beta_y+2\beta_{yy}\ln y)$ , вычислите отдачу от масштаба для медианного уровня LNY в каждой из пяти подвыборок.

**Упражнение 6.** Сравнение отдачи от масштаба по данным 1955 г. и 1970 г.

Для работы в этом упражнении используйте файл данных UPDATE.

а) Используя обновленные для 1970 года данные по 99 фирмам, сформируйте новые переменные, необходимые для построения регрессии (2.1). А именно:  $LNC70 = \ln(\text{COST70/PF70})$ ,  $LN Y70 = \ln(\text{KWH70})$ ,  $LNP170 = \ln(\text{PL70/PF70})$  и  $LNP270 = \ln(\text{PK70/PF70})$ , где  $LNC70$  соответствует  $\ln C^*$ ,  $LN Y70 = \ln y$ ,  $LNP170 = \ln p_1^*$ ,  $LNP270 = \ln p_2^*$ . Вычислите математическое ожидание для KWH70 и сравните его с математическим ожиданием KWH из файла NERLOV для 1955 г. Прокомментируйте результат. Какую отдачу от масштаба в 1970 г. можно ожидать по сравнению с 1955 г.? Почему?

б) Оцените МНК параметры уравнения:

$$LNC70 = \beta_0 + \beta_y LN Y70 + \beta_1 LNP170 + \beta_2 LNP270 + \varepsilon.$$

Постройте доверительный интервал для  $\beta_y$ . Проверьте гипотезу  $\beta_y=1$ . Вычислите отдачу от масштаба. Сравните результат с п. б упр. 4. Каковы различия? Почему?

в) Оцените МНК расширенную регрессию:

$$LNC70 = \beta_0 + \beta_y LN Y70 + \beta_{yy} (LN Y70)^2 + \beta_1 LNP170 + \beta_2 LNP270 + \varepsilon.$$

Отметим, что регрессия п. б – специальный случай расширенной при  $\beta_{yy}=0$ . Используя приемлемый уровень значимости, проверьте нулевую гипотезу о том, что отдача от масштаба не зависит от изменений в выпуске, т.е.  $\beta_{yy}=0$ , против альтернативной гипотезы  $\beta_{yy} \neq 0$ . Используя приемлемый уровень значимости, проверьте нулевую гипотезу о том, что отдача от масштаба постоянна, т.е.  $\beta_y=1$ ,  $\beta_{yy}=0$ , против альтернативной гипотезы  $\beta_y \neq 1$ ,  $\beta_{yy} \neq 0$ . Проинтерпретируйте результаты этих двух различных тестов. Они согласованы между собой?

г) Разбейте выборку из 99 фирм на 5 подвыборок, упорядочив данные по величине выпуска, 20 первых фирм имеют наименьший выпуск и 19 последних фирм имеют наибольший выпуск. Оцените уравнение регрессии п. б для каждой из пяти подвыборок. Вычислите отдачу от масштаба и сравните с результатами Nerlove M. для 1955 г. для каждой из пяти групп: 2,92, 2,24, 1,97, 1,84 и 1,69.

д) Как численно можно проверить утверждение Christensen-Greene о том, что объем производства электричества приближался к точке выпуклости кривой средних издержек? Вы согласны с этим утверждением? Почему?

**Упражнение 7.** Автокорреляция в модели кривой обучения

а) В упр. 1 и 2, используя МНК, оцените две спецификации кривой обучения для одного из файлов данных. На основе статистики Дарбина-Уотсона проверьте гипотезу об автокорреляции первого порядка (прил. 1).

б) Используя процедуру Хилдрета-Лу или Кохрейна-Оркатта, оцените обе модели кривой обучения с учетом наличия автокорреляции первого порядка. Сравните значения эластичностей и их стандартные ошибки с п. а. Что неожиданного вы заметили? Почему?

**Упражнение 8.** Ошибки спецификации модели Nerlove M.

а) Используя данные 1955 г. для 145 компаний из упр. 4 оцените уравнение (2.1) МНК.

б) В (2.1) предполагается, что отдача от масштаба вычисляется для данного уровня выпуска. В файле NERLOV наблюдения упорядочены по возрастанию выпуска. Учитывая этот факт и предполагая, что отдача от масштаба изменяется с изменением выпуска, как должен выглядеть график остатков модели (2.1)? В чем сходство нарушения независимости между ошибками для пространственных данных и автокорреляции первого порядка во временных рядах?

в) Используя результаты п. б и значения статистики Дарбина-Уотсона (прил. 1) проверьте для приемлемого уровня значимости гипотезу о независимости ошибок модели (2.1). Можно ли переоценить модель с помощью процедуры Хилдрета-Лу или Кохрейна-Оркатта или же наличие автокорреляции указывает на фундаментальную проблему в спецификации, которую нельзя решить применением обобщенного метода наименьших квадратов? Обоснуйте свою позицию. Будьте внимательны (вам помогут результаты упр. 5).

**Упражнение 9.** Тестирование равенства коэффициентов для 1955 и 1970 гг.

а) Разбейте выборку из 145 фирм на 5 подвыборок, в каждой по 29 фирм (данные 1955 г. в файле NERLOV). Оцените уравнение регрессии (2.1) для каждой из пяти подвыборок и для всей выборки. Используя тест Чоу, проверьте гипотезу о том, что параметры каждой из пяти подвыборок одинаковы.

б) Разбейте выборку из 99 фирм на 5 подвыборок (данные 1970 г. в файле UPDATE). Оцените уравнение регрессии (2.1) для каждой из пяти подвыборок и для всей выборки. Используя тест Чоу, проверьте гипотезу о том, что параметры каждой из пяти подвыборок одинаковы.

в) Объедините данные для 1955 и 1970 гг. Оцените уравнение регрессии (2.1) для объединенной выборки и проверьте гипотезу о том, что параметры каждой из 10 подвыборок (пять для 1955 г. и пять для 1970 г.) одинаковы. Проверьте гипотезу о том, что параметры для выборки 1970 г. равны параметрам для выборки 1955 г.

г) Оцените регрессию

$$\text{LNC70} = \beta_0 + \beta_y \text{LNY70} + \beta_{yy} (\text{LNY70})^2 + \beta_1 \text{LNP170} + \beta_2 \text{LNP270} + \varepsilon$$

для выборки 1955 г., затем для выборки 1970 г. и, наконец, для объединенной выборки 1955-1970 гг.

Проверьте гипотезу о том, что параметры для выборки 1970 г. равны параметрам для выборки 1955 г.

д) Прокомментируйте результаты тестов в п. а-г. Как они могут помочь в выборе правильной спецификации модели?

### Упражнение 10. Прогнозирование средних издержек

В уравнении  $\ln c_t = \alpha_0 + \alpha_1 \ln n_t$  переобозначим  $\ln c_t = z_t$  и  $\ln n_t = x_t$  и запишем уравнение в виде  $z_t = \alpha_0 + \alpha_1 x_t$ . Если ошибки уравнения независимо распределены по нормальному закону и известно будущее значение  $x_t$ , которое обозначим  $x_\tau$ , то дисперсия прогноза:

$$s_{f,\tau}^2 = V(a_0) + 2x_\tau \text{cov}(a_0, a_1) + x_\tau^2 V(a_1) + s^2 = s^2 \left[ 1 + \frac{1}{T} + \frac{(x_\tau - \bar{x})^2}{\sum_{t=1}^T (x_t - \bar{x})^2} \right],$$

где  $a_0$  и  $a_1$  – оценки МНК для  $\alpha_0$  и  $\alpha_1$ ;  $T$  – количество наблюдений;  $\bar{x}$  – среднее для  $x_t$ ,  $s^2$  – сумма квадратов ошибок регрессии, деленная на  $T-2$ .

а) Используя данные из *упр. 1*, оцените МНК  $\alpha_0$  и  $\alpha_1$ , а также их дисперсии и ковариацию и  $s^2$ .

б) Вычислите на основе результатов п. а прогнозируемые средние издержки и дисперсию ошибки прогноза с учетом приведенной выше формулы для уровня накопленной суммы выпуска в 2 раза большего, чем его значение для последнего  $T$ -го наблюдения (сначала вычислите выпуск, а затем его прологарифмируйте).

в) Постройте доверительный интервал для прогнозного значения из п. б. Прокомментируйте результат.

г) Как влияет увеличение объема выборки на величину дисперсии ошибки прогноза? Как влияет разность между  $x_\tau$  и  $\bar{x}$ ?

Лабораторная работа 3. Оценка предложения труда и дискриминации по заработной плате.

### Описание файлов данных

В файле CPS78 содержится 550 наблюдений о случайно выбранных работниках в USA (май 1978 г.). Переменные:

ED – количество лет образования;

SOUTH – фиктивная переменная, равна 1, если работник проживает на юге, иначе 0;

NONWH – фиктивная переменная, равна 1, если работник не белый, иначе 0;

HISP – фиктивная переменная, равна 1, если работник латиноамериканец, иначе 0;

FE – фиктивная переменная, равна 1 для женщин, и 0 для мужчин;

MARR – фиктивная переменная, равна 1 для семейных, иначе 0;

MARRFE – фиктивная переменная, равна 1 для замужних женщин, иначе 0;

EX – число лет стажа работы (= AGE-ED-6);

EXSQ – квадрат числа лет стажа работы;

UNION – фиктивная переменная, равна 1, если имеется профсоюз на работе, иначе 0;

LNWAGE – логарифм средней часовой заработной платы;

AGE – возраст в годах;

NDEP – число детей до 18 лет в семье;

MANUF – фиктивная переменная, равна 1, если работает в обрабатывающей промышленности, иначе 0;

CONSTR – фиктивная переменная, равна 1, если работает в строительстве, иначе 0;

MANAG – фиктивная переменная, равна 1, если работа управленческая или административная, иначе 0;

SALES – фиктивная переменная, равна 1, если работает в торговле, иначе 0;

CLER – фиктивная переменная, равна 1, если работает чиновником, иначе 0;

SERV – фиктивная переменная, равна 1, если работает в сфере услуг, иначе 0;

PROF – фиктивная переменная, равна 1, если профессионально-технический работник, иначе 0.

В файле CPS85 содержится 534 наблюдений о случайно выбранных работниках в USA (май 1985 г.). Описание переменных соответствует приведенному выше.

### Упражнение 1. Обзор фактов

*а)* Используя данные из CPS78, вычислите математическое ожидание и стандартное отклонение для переменной LNWAGE. Вычислите геометрическое среднее LNWAGE путем потенцирования среднего. Вычислите математическое ожидание и стандартное отклонение для переменных ED и EX.

*б)* Каково выборочное среднее переменной NONWH? Из 550 наблюдений в файле сколько индивидов небелые и не латиноамериканцы? Каково выборочное среднее переменной HISP? Из 550 наблюдений в файле сколько латиноамериканцев? Какова доля женщин в выборке? Сколько в выборке женщин?

*в)* Отсортируйте данные по полу. Каково среднее и стандартное отклонение количества лет образования для женщин и мужчин? Среднее по зарплате? Отсортируйте данные по расе и сравните для каждой группы среднее и стандартное отклонение количества лет образования, среднее по зарплате.

*г)* Повторите пп. *а–в* для файла данных CPS85. Какие изменения произошли в пропорциях по полу и расе в 1985 году по сравнению с 1978 годом? Используя индекс потребительских цен, равный 1 для 1978 г. и 1,649 для 1985 года, вычислите среднюю реальную зарплату для всей выборки, по полу, по расе. Какая группа индивидов имеет более высокую зарплату. Какие изменения произошли в человеческом капитале в 1985 г. по сравнению с 1978 г.?

*д)* Проверьте гипотезу о нормальном распределении переменной LNWAGE. Проверьте гипотезу о нормальном распределении переменной WAGE. Прокомментируйте результаты.

**Упражнение 2.** Заработная плата и образование: альтернативные спецификации с фиктивными переменными  
Выберите один из файлов данных CPS78 или CPS85.

*а)* Оцените МНК простую модель регрессии LNWAGE на константу и ED. Каково значение коэффициента наклона? Как можно интерпретировать этот коэффициент? Постройте и прокомментируйте 95%-й доверительный интервал для этого коэффициента (прил. 3). Каково значение  $R^2$ ? Прокомментируйте.

*б)* Спецификация модели в п. *а* предполагает, что смещение и угол наклона прямой одинаковы для всех индивидов. Предположим, что эффект образования отличается для состоящих в профсоюзе и не состоящих в профсоюзе индивидов, т.е. для первых  $WAGE = \alpha_N e^{\beta ED} e^\varepsilon$  и для вторых соответственно  $WAGE = \alpha_U e^{\beta ED} e^\varepsilon$ , где  $\alpha_N$  и  $\alpha_U$  – коэффициенты пропорциональности. После логарифмирования с учетом фиктивной переменной UNION для членства в профсоюзе получим регрессию:

$$LNWAGE = \alpha + \alpha' UNION + \beta ED + \varepsilon.$$

Оцените последнюю регрессию МНК. Проинтерпретируйте оценки для  $\alpha$  и  $\alpha'$ , и их связь с  $\alpha_N$  и  $\alpha_U$ . Проверьте гипотезу о том, что доходы членов и не членов профсоюза одинаковы на 5%-ом уровне значимости.

*в)* Альтернативная спецификация модели п. *а*. Создадим новую переменную  $NONU = 1 - UNION$ . Оцените МНК регрессию:

$$LNWAGE = \alpha_0 NONU + \alpha_1 UNION + \beta ED + \varepsilon.$$

Проинтерпретируйте оценки для  $\alpha_0$  и  $\alpha_1$ , и их связь с  $\alpha_N$  и  $\alpha_U$  из п. *б*. Какова связь между  $\alpha_0$  и  $\alpha_1$  и  $\alpha$  и  $\alpha'$ ? Почему? Ваши численные оценки подтверждают эту связь? Проверьте гипотезу о том, что доходы членов и не членов профсоюза одинаковы на 5%-ом уровне значимости. Сравните результаты тестирования с п. *б*.

Сравните отдачу от образования и  $R^2$  с полученными в п. *б*. Почему так получилось?

*г)* Возможный вариант спецификации уравнения регрессии

$$LNWAGE = \alpha^* + \alpha_0^* NONU + \alpha_1^* UNION + \beta ED + \varepsilon.$$

Что ошибочно в такой спецификации? Почему оценки модели нельзя найти МНК?

*д)* Еще одна возможность оценить МНК регрессию вида

$$LNWAGE = \alpha^0 + \alpha^1 NONU + \beta ED + \varepsilon.$$

Проинтерпретируйте оценки  $\alpha^0$ ,  $\alpha^1$  и  $\beta$ . Какова связь между оценками  $\alpha^0$ ,  $\alpha^1$  и  $\beta$  и оценками  $\alpha$ ,  $\alpha'$  и  $\beta$  в п. *б*. Почему?

**Упражнение 3.** Фиктивные переменные взаимодействия: доходы одиноких и женатых мужчин и замужних женщин

Выберите один из файлов данных CPS78 или CPS85.

*а)* Используя МНК, оцените простую модель человеческого капитала без учета пола и супружеского статуса:

$$LNWAGE = \alpha + \beta_1 ED + \beta_2 EX + \beta_3 EXSQ + \varepsilon.$$

Прокомментируйте полученные оценки коэффициентов.

*б)* С учетом пола оцените расширенную регрессию вида

$$LNWAGE = \alpha + \alpha_F FE + \beta_1 ED + \beta_2 EX + \beta_3 EXSQ + \varepsilon.$$

Прокомментируйте полученные оценки коэффициентов. Проверьте гипотезу о том, что доходы мужчин и женщин одинаковы на 5%-ом уровне значимости.

*в)* С учетом супружеского статуса оцените расширенную регрессию вида

$$LNWAGE = \alpha + \alpha_F FE + \alpha_M MARR + \beta_1 ED + \beta_2 EX + \beta_3 EXSQ + \varepsilon.$$

Прокомментируйте полученные оценки коэффициентов  $\alpha$ ,  $\alpha_F$  и  $\alpha_M$ . Проверьте гипотезу о том, что супружеский статус не влияет на доходы. Проверьте гипотезу о том, что влияние супружеского статуса на доходы одинаково для мужчин и для женщин.

*г)* В теории человеческого капитала предполагается, что влияние супружеского статуса на доходы зависит от пола, в частности, женатые женщины получают доходы меньше, чем женатые мужчины. Спецификация регрессии п. *в* не позволяет протестировать этот факт. Почему? Создайте фиктивную переменную взаимодействия  $INFMAR = FE \cdot MARR$ . Эта переменная по значениям совпадает с переменной  $MARRFE$ , имеющейся в выборке. Проверьте это, создав вспомогательную переменную  $CHEK = INFMAR - MARRFE$ , и убедившись, что все ее значения равны нулю.

*д)* Оцените МНК регрессию:

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_M \text{MARR} + \alpha_{FM} \text{INFMAR} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Дайте интерпретацию оценкам коэффициентов  $\alpha$ ,  $\alpha_F$ ,  $\alpha_M$  и  $\alpha_{FM}$ . Подтверждают ли оценки теорию? Почему? Сформулируйте и проверьте гипотезу, о том, что эффект супружеского статуса не зависит от пола. Как вы можете прокомментировать полученный результат.

**Упражнение 4.** Исследование профиля "опыт-зарплата"

Выберите один из файлов данных CPS78 или CPS85.

*а)* Оцените основную модель вида

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Какие знаки должны иметь коэффициенты  $\beta_2$  и  $\beta_3$  согласно теории. Ваши оценки согласуются с теорией? Какое значение EX максимизирует LNWAGE? Вычислите число лет опыта работы и предполагаемый возраст, в котором LNWAGE максимально для индивидов с 8, 12 и 16 годами образования.

*б)* В координатах  $\exp(\text{LNWAGE})$  и AGE, используя оценки уравнения из п. *а*, постройте график профиля возраст-зарплата для средних по выборке значений FE, UNION, NONWH, HISP и ED. Прокомментируйте график. В чем некорректность вашей интерпретации?

*в)* Создайте переменную  $\text{EDEX} = \text{ED} \cdot \text{EX}$ . Оцените регрессию:

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \beta_4 \text{EDEX} + \varepsilon.$$

Какой знак ожидается у коэффициента  $\beta_4$ ? Почему? Найдите значения  $\text{EX}^*$ , максимизирующие доходы индивидов с 8, 12 и 16 годами образования. Вычислите соответствующий возраст таких индивидов. Сравните результаты с полученными в п. *а*.

*г)* В координатах  $\exp(\text{LNWAGE})$  и AGE, используя оценки уравнения из п. *в*, постройте графики профилей возраст-зарплата для индивидов с 8, 12 и 16 годами образования для средних по выборке значений FE, UNION, NONWH и HISP. Прокомментируйте графики.

*д)* Оставьте в выборке только женщин. Повторите пп. *в* и *г* для одиноких и замужних женщин отдельно (удалите переменную FE из регрессии). Какие различия в профилях для замужних и одиноких женщин вы получили? Почему? Ваши ожидания совпали с вашими эмпирическими оценками?

**Упражнение 5.** Дискриминация по заработной плате рабочих, не являющихся членами профсоюза

*а)* Для файла данных CPS78 вычислите и сохраните выборочные средние для LNWAGE, FE, UNION, NONWH, HISP, ED, EX и EXSQ. Оцените параметры уравнения:

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Коэффициент  $\alpha_U$  статистически значим? Каково его потенцированное значение и как оно интерпретируется?

*б)* Отсортируйте файл данных по переменной UNION. Вычислите средние для LNWAGE, FE, UNION, NONWH, HISP, ED, EX и EXSQ отдельно для состоящих в профсоюзе и не состоящих в профсоюзе. Прокомментируйте различия для переменных LNWAGE, FE, NONWH, HISP, ED, EX.

*в)* Для отсортированных данных п. *б* оцените отдельно для состоящих в профсоюзе и не состоящих в профсоюзе регрессию вида:

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Сравните оценки коэффициентов для двух регрессий. В частности, для состоящих в профсоюзе и не состоящих в профсоюзе различия в доходах больше для женщин, чем для мужчин при фиксированных значениях других переменных? Как меняются различия в доходах для состоящих в профсоюзе и не состоящих в профсоюзе в зависимости от расы?

Подтверждается ли гипотеза о том, что влияние уровня образования на логарифм доходов меньше для членов профсоюза?

Сравните профили возраст-зарплата для состоящих в профсоюзе и не состоящих в профсоюзе при средних по выборке значениях FE, NONWH, HISP, ED из п. *а*. Прокомментируйте результаты.

*г)* Используя тест Чоу проверьте гипотезу о совпадении коэффициентов регрессий для состоящих в профсоюзе и не состоящих в профсоюзе. Проинтерпретируйте результаты.

*д)* Оценим различия в зарплате для состоящих в профсоюзе и не состоящих в профсоюзе. Вычтем среднее по выборке LNWAGE для не состоящих в профсоюзе из среднего по выборке LNWAGE для состоящих в профсоюзе. Для каждой из независимых переменных из п. *в* вычтем средние по подвыборкам для состоящих в профсоюзе и не состоящих в профсоюзе и умножим на веса – оценки коэффициентов регрессии п. *в* для состоящих в профсоюзе и сложим. Полученную сумму вычтем из разности в средних значениях LNWAGE. Потенцируем полученное значение разности. Как может быть проинтерпретировано это значение? Насколько оно велико?

*е)* Повторите п. *д* для уравнения  $\ln W^u - \ln W^n = \hat{\beta}^n (\bar{z}^u - \bar{z}^n)$ , где индексы соответствуют  $u$  – член профсоюза,  $n$  – не член профсоюза,  $\ln W$  – логарифм заработной платы,  $z$  – объясняющие переменные. Сравните и проинтерпретируйте три альтернативные оценки различий в оплате труда работников, состоящих и не состоящих в профсоюзе в пп. *а*, *д* и *е*. Какая из оценок предпочтительнее и почему?

*ж)* Прodelайте пп. *а–е* для данных файла CPS85. Премия за членство в профсоюзе выросла или уменьшилась? Прокомментируйте результаты оценок.

**Упражнение 6.** Дискриминация в зарплате по полу: измерение и интерпретация

а) Для файла данных CPS78 оцените уравнение

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Проинтерпретируйте значение коэффициента  $\alpha_F$  и его потенцированное значение. Проверьте гипотезу о том, что пол не влияет на уровень зарплаты.

б) Отсортируйте файл данных CPS78 на две подвыборки для женщин и мужчин. Сколько наблюдений в каждой из подвыборок? Вычислите средние для переменных LNWAGE, NONWH, HISP, UNION, ED, EX, EXSQ отдельно для мужчин и для женщин. Вычислите разности в средних. В среднем мужчины более или менее образованы, чем женщины? Какова разница в опыте работы? Для кого переменная EX является более обоснованной мерой обучения в процессе работы? Почему?

в) Используя отсортированные данные из п. б, оцените отдельно для мужчин и женщин регрессию:

$$\text{LNWAGE} = \alpha + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Сравните результаты. Эффект образования на зарплату отличается в зависимости от пола? Эффект расы? Проинтерпретируйте результаты.

г) Используя приемлемый уровень значимости, проверьте тестом Чоу совпадение регрессий в п. в. Прокомментируйте результат.

д) Примените процедуру Blinder-Оахаса для оценки дискриминации по полу в зарплате по  $\bar{\ln} y^* - \bar{\ln} y_* = \hat{\beta}^* (\bar{X}^* - \bar{X}_*) + \bar{X}_* \cdot \Delta \hat{\beta}$ .

Проверьте правильность разложения, вычислив разницу в средних для LNWAGE для мужчин и для женщин. Какая доля разницы в средней зарплате обусловлена гендерными различиями? Сравните с результатами в пп. в, а.

е) Примените процедуру Blinder-Оахаса для оценки дискриминации по полу в зарплате по формуле  $\bar{\ln} y^* - \bar{\ln} y_* = \hat{\beta}_* (\bar{X}^* - \bar{X}_*) + \bar{X}^* \cdot \Delta \hat{\beta}$ . Какая доля разницы в средней зарплате обусловлена гендерными различиями? Сравните с результатами в п. д.

ж) Повторите пп. а–е для данных из файла CPS85. Прокомментируйте гендерные различия в доходах в 1985 г. по сравнению с 1978 г. Изменился ли уровень дискриминации? Почему?

з) (факультативно). Для замужних и одиноких женщин уравнения зарплаты различны, так как замужние женщины на время декретного отпуска выбывают из состава рабочей силы. Используя процедуру Blinder-Оахаса, оцените дискриминацию по зарплате для замужних и одиноких женщин (воспользуйтесь файлами CPS78 и CPS85). Прокомментируйте результаты.

**Упражнение 7.** Гетероскедастичность в уравнении регрессии для зарплаты

Выберите один из файлов данных CPS78 или CPS85.

а) Оцените уравнение

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Вычислите обычные и с поправкой на гетероскедастичность стандартные ошибки. Последние больше первых? Что это означает? Почему?

б) Используйте процедуру обобщенного МНК (ОМНК) для оценки коэффициентов модели п. а. Для этого получите вектор квадратов остатков модели, и постройте регрессию квадратов остатков на переменные FE, UNION, NONWH, HISP, ED, EX и EXSQ. Выберите значимые переменные и получите веса ОМНК. Оцените модель п. а ОМНК. Прокомментируйте результат.

в) Воспользуйтесь одной из стандартных процедур пакета для коррекции стандартных ошибок на гетероскедастичность. Прокомментируйте результаты. Сравните их с п. б, сделайте выводы.

**Упражнение 8.** Эндогенность и оценки двухшагового метода наименьших квадратов

Выберите один из файлов данных CPS78 или CPS85.

а) Оцените уравнение регрессии обычным МНК:

$$\text{LNWAGE} = \alpha + \alpha_F \text{FE} + \alpha_U \text{UNION} + \alpha_N \text{NONWH} + \alpha_H \text{HISP} + \beta_1 \text{ED} + \beta_2 \text{EX} + \beta_3 \text{EXSQ} + \varepsilon.$$

Создайте переменную квадрат возраста AGESQ и используйте переменные AGE и AGESQ в качестве инструментов для EX и EXSQ. Оцените регрессию с помощью инструментальных переменных, воспользовавшись командой `ivreg LNWAGE NONWH HISP UNION ED (EX EXSQ= AGE AGESQ)`.

Сравните результаты двух регрессий. Сделайте выводы.

Познакомьтесь с опциями команды `ivreg`.

б) Воспользуйтесь тестом Хаусмана для выбора лучшей регрессии из п. а. Прокомментируйте результаты теста. Сделайте выводы.

Лабораторная работа 4. Модель Солоу.

Описание файлов данных

1) Данные содержатся в файле Growth. Используются данные из архива журнала *Journal of Applied Econometrics*. Единицей наблюдения является страна, даны результаты наблюдений для 121 страны. Используются следующие переменные:

NUM – номер страны в базе данных;

NOIL (фиктивная) – 1 для страны, не добывающей нефть, 0 – для добывающей;

INTER (фиктивная) – 1 для страны с хорошим качеством данных, 0 – в противном случае;

OECD (фиктивная) – 1 для страны OECD, 0 – в противном случае;

GDP60 – ВВП на душу населения в 1960 г.;

GDP85 – ВВП на душу населения в 1985 г.;

GDPGRO – средний рост ВВП на душу населения с 1960 г. по 1985 г. (в %);

POPGRO – средний рост работоспособного населения с 1960 г. по 1985 г. (в %);

IONY – средняя доля инвестиций (включая государственные) в общем объеме ВВП с 1960 г. по 1985 г. (в %);

SCH – средняя доля работоспособного населения, имеющая полное среднее образование с 1960 г. по 1985 г. (в %);

LIT – доля людей среди населения старше 15 лет, умеющих читать и писать в 1960 г.

2) Данные содержатся в файле Klein и содержат 22 наблюдения с 1920 по 1941 гг. по США. Используются следующие переменные:

year – год;

C – потребление;

P – прибыль акционерных компаний;

Wp – доходы населения;

I – инвестиции;

K1 – акционерный капитал предыдущего года;

X – ВВП;

Wg – государственные доходы;

G – государственные расходы;

T – косвенные налоги и чистый экспорт.

Используя файл данных Growth выполните упражнения 1-3.

**Упражнение 1.** Обзор фактов и описательные статистики переменных

*a)* Вычислите суммарные статистики всех переменных. Проверьте, имеют ли смысл ваши результаты.

*б)* Вычислите корреляционную матрицу всех переменных. Дайте интерпретацию наиболее важных результатов. Соответствуют ли они тому, что вы ожидали?

**Упражнение 2.** Анализ стационарного состояния

Пусть  $g = 2\%$  и  $\delta = 3\%$ .

*a)* Оцените уравнение

$$\ln(\text{GDP85}) = \pi_0 + \pi_1 \ln s_k + \pi_2 \ln s_h + \pi_3 \ln(n + g + \delta), \quad (5.1)$$

используя данные по всем странам, за исключением тех, для которых пропущены наблюдения какой-либо независимой или зависимой переменной.

*б)* Исходная модель Солоу не включает человеческий капитал. Оцените уравнение (5.1), удалив переменную  $\ln \text{SCH}$ . Сравните с результатом, полученном в п. *a*. В чем состоит основное различие? Объясните это различие, используя теорию.

*в)* Структурная форма

$$\ln y(t) = \frac{1}{1 - \alpha - \beta} [\alpha \ln s_k + \beta \ln s_h - (\alpha + \beta) \ln(n + g + \delta)]$$

накладывает некоторое линейное ограничение на параметры  $\pi_1, \pi_2, \pi_3$  приведенной формы. Что это за ограничение? Протестируйте (на 5%-ом уровне значимости) выполнимость этого ограничения. Оцените вновь уравнение (5.1), используя это ограничение. Сравните ваш результат с результатом, полученным в п. *a*. Воспользуйтесь командами `constraint define 1 ...` и `cnsmreg ...`, `c(1)` пакета Stata.

*г)* Выразите структурные параметры  $\alpha$  и  $\beta$  через  $\pi_1, \pi_2$  и постройте их оценки.

*д)* Добавьте в регрессионное уравнение п. *a* фиктивные переменные NOIL и OECD и проверьте их значимость. Проверьте, является ли линейная спецификация (5.1) разумной, добавляя квадраты независимых переменных и перекрестные члены.

*е)* Согласно "золотому правилу накопления капитала", доли инвестиций  $s_k, s_h$  должны выбираться таким образом, чтобы в стационарном состоянии величина  $c = (1 - s_k - s_h)y$  была максимальной. Теоретически оптимальное значение величины  $s_k$  определяется как  $s_k^* = \alpha$ . Используя оценки, полученные в п. *в*, проверьте, удовлетворяют ли в среднем инвестиции в физический капитал золотому правилу.

**Упражнение 3.** Рост ВВП

*a)* Оцените уравнение

$$\ln \frac{\text{GDP85}}{\text{GDP60}} = \pi_0 + \pi_1 \ln s_k + \pi_2 \ln s_h + \pi_3 \ln(n + g + \delta) + \pi_4 \ln \text{GDP60}, \quad (5.2)$$

и интерпретируйте результат.

Исходная модель Солоу не включает человеческий капитал. Оцените уравнение (5.2), удалив переменную  $\ln \text{SCH}$ . Сравните результаты. Объясните различия.

б) Структурная форма

$$\ln \frac{y(t)}{y(0)} = (1 - e^{-\lambda t}) \left\{ \frac{1}{1 - \alpha - \beta} [\alpha \ln s_k + \beta \ln s_h - (\alpha + \beta) \ln(n + g + \delta)] - \ln y(0) \right\}$$

ет некоторое линейное ограничение на параметры  $\pi_1, \pi_2, \pi_3$  приведенной формы. Что это за ограничение? Протестируйте (на 5%-ом уровне значимости) выполнимость этого ограничения. Оцените вновь уравнение (5.2), используя это ограничение. Сравните ваш результат с результатом, полученным в п. а.

в) Используя результат п. б, постройте оценки структурных параметров  $\lambda, \alpha, \beta$ . Проинтерпретируйте результаты. Сравните ваши оценки параметров  $\alpha$  и  $\beta$  с оценками, полученными в предыдущих упражнениях.

з) Добавьте фиктивные переменные NOIL и OECD в уравнение п. а и проверьте их значимость. Проверьте, является ли линейная спецификация (5.2) разумной, добавляя квадраты независимых переменных и перекрестные члены.

д) Оцените уравнение (5.2) отдельно для стран-членов OECD и для стран не членов OECD и проинтерпретируйте результаты. Проверьте, совпадают ли коэффициенты уравнения (5.2) (за исключением свободного члена) для стран-членов OECD и для стран не членов OECD.

е) Выберите наилучшее с вашей точки зрения уравнение и постройте 95%-й доверительный интервал для скорости сходимости  $\lambda$ . Проинтерпретируйте результат.

#### Упражнение 4. Модель Клейна

Используйте данные из файла Klein.

а) Выполните идентификацию параметров приведенной формы модели обычным МНК и двухшаговым МНК. Сравните результаты. Прокомментируйте различия. Объясните причину различий в результатах расчетов двумя методами.

б) В чем отличия двухшагового МНК от трехшагового МНК? Выполните идентификацию параметров приведенной формы модели трехшаговым МНК. Сравните результаты расчетов с оценками двухшагового МНК. В чем отличия, почему?

Лабораторная работа 5. Модель Бокса-Дженкинса и модели волатильности.

1. Исходные данные. Индекс объема выпуска промышленной продукции в РФ с 1991 г. по 1995 г. после выделения неслучайной составляющей имеет вид:

Год	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
1991	-12,1	-9,0	-0,8	-5,4	-0,9	-2,6	13,7	20,7	16,9	3,9	-3,7	14,9
1992	-9,2	-8,2	-4,6	-3,7	-5,6	-5,7	0,9	-11,0	-3,5	3,3	8,2	18,1
1993	-14,9	-2,1	5,6	6,3	1,9	7,0	3,1	6,6	11,4	11,7	13,6	11,6
1994	-23,9	-17,3	-17,1	-20,3	-16,6	-10,9	-10,4	-6,9	-6,1	-0,9	16,4	3,3
1995	4,4	-17,4	-18,1	-8,4	8,7	3,0	2,8	3,9	13,9	17,9	20,9	10,6

а) Найдите АКФ и ЧАКФ ряда, его первых и вторых разностей. б) какие модели для указанного ряда вы можете предложить. Обоснуйте свое мнение. в) идентифицируйте несколько подходящих моделей ряда (не более трех). Выберите наилучшую из моделей. Обоснуйте свой выбор.

2. Исходные данные. В файле sp500 ([www.econ.kuleuven.ac.be/gme/](http://www.econ.kuleuven.ac.be/gme/)) даны ежедневные значения индекса Standard&Poors 500 начиная с января 1981 г. по апрель 1991 г., вычисленные как изменения прологарифмированного индекса. Всего 2783 наблюдения. Основываясь на этих данных, постройте модель прогнозирования изменения логарифма индекса S&P500. Оцените качество построенной модели.

3. В файле output ([www.stern.nyu.edu/~wgreene/Text/econometric-analysis.htm](http://www.stern.nyu.edu/~wgreene/Text/econometric-analysis.htm)) имеются данные о денежном агрегате M1, ВВП и дефляторе цен в США поквартально с 1950 г. по 1983 г. Подберите ARIMA модель и модель регрессии. Изучите вопрос о включении в модель фактора времени.

4. На сайте [www.rts.ru](http://www.rts.ru) имеется информация об итогах торгов акциями ОАО РАО ЕЭС России (код EESR). Для данных о стоимости акций за последние шесть месяцев получите доходности указанного актива, постройте графики ряда доходностей, АКФ, ЧАКФ, а также графики квадратов доходностей и АКФ и ЧАКФ для квадратов доходностей. Постройте адекватную модель для доходностей актива.

5. Сгенерируйте ряд длиной 500 наблюдений в соответствии с моделью ARCH(3):  $\sigma_t^2 = 1 + 0,3u_{t-1}^2 + 0,25u_{t-2}^2 + 0,15u_{t-3}^2$ . По сгенерированному ряду оцените модель ARCH(3). Сравните оценки с истинными параметрами модели. Сравните динамику оценки условной дисперсии и ее истинных значений. Постройте модель GARCH(1, 1). Сделайте выводы об ее адекватности исходным данным.

Лабораторная работа 6. Модель CAPM.



### Описание файлов данных

Следующие файлы данных содержат 120 помесечных наблюдений с января 1978 по декабрь 1987 года по следующим 17 компаниям:

Нефтедобывающие: MOBIL и TEXACO.  
Производство компьютеров: IBM, DEC, DATGEN.  
Электроэнергетика: CONED, PSNH.  
Лесная промышленность: WEYER, BOISE.  
Электронная промышленность: MOTOR, TANDY.  
Авиаперевозки: PANAM, DELTA.  
Банки: CONTIL, CITCRP.  
Продукты питания: GERBER, GENMIL.

В файле MARKET содержатся данные о доходности рыночного портфеля, в файле RKFREE данные о доходности безрискового актива – 30 дневных US облигаций.

В некоторых файлах некоторые ряды данных имеют другие, отличные от указанного периоды наблюдений:  
CONOCO, временной ряд в файле EVENTS, с января 1976 по сентябрь 1981 (только 69 наблюдений);  
APM, с декабря 1977 по декабрь 1987 (121 наблюдение);  
EVENTS, с января 1976 по декабрь 1985;  
GOLD, с января 1976 по декабрь 1985.

Каждый файл данных содержит месячные доходности финансовых активов.

Файл APM содержит дополнительные данные для построения регрессии по Ross' Arbitrage Pricing Model. Переменные с декабря 1977 по декабрь 1987:

POIL - (\$) номинальная цена за баррель нефти,  
FRBIND – индекс промышленного производства, 1972=100, ряд скорректирован на сезонность;  
CPI – индекс потребительских цен, (1967=100).

Файл EVENTS содержит данные с января 1976 по декабрь 1985 гг. Переменные:

DATE – дата, первые две цифры – год, следующие две – месяц;  
GPU - коммунальное хозяйство;  
DOW - химическая компания Dow;  
DUPONT - компания DuPont;  
MARK76 - рыночная доходность;  
RKFR76 – безрисковая доходность;  
DATE1 – формат даты, как у переменной DATE, но только 69 наблюдений с 1/76 по 9/81.

Файл GOLD содержит ряды данных с января 1976 по декабрь 1985. Переменные:

GOLD - % изменения цен на золото в течение месяца;  
MARK76 - рыночная доходность;  
RKFR76 - безрисковая доходность.

### Упражнение 1. Начало работы с данными, вычисление описательных статистик

а) Используя данные из файла MARKET, постройте график доходностей для последних 36 месяцев.

б) Далее, используя данные из market и RKFREE, получите переменные премии за риск ( $r_p - r_f$ ) для любой компании по вашему выбору и рыночной премии ( $r_m - r_f$ ). Вычислите средние значения  $r_p$ ,  $r_m$ ,  $r_f$ ,  $r_p - r_f$ ,  $r_m - r_f$ . Отметим, что доходности являются месячными. Для перевода средних месячных доходностей в годовые можно воспользоваться формулой  $r_{\text{год}} = (1 + \bar{r})^{12} - 1$ , где  $\bar{r}$  – средняя месячная доходность. Вычислите годовые доходности для  $r_p$ ,  $r_m$ ,  $r_f$ ,  $r_p - r_f$ ,  $r_m - r_f$ . Полученные значения правдоподобны?

в) Постройте графики премии за риск для выбранной компании и для рынка за 36 месяцев с января 1985 года по декабрь 1987 года. Какие выводы вы можете сделать? Как вы думаете,  $\beta$ -показатель компании больше или меньше единицы? Почему?

г) Вычислите дисперсию и стандартное отклонение для доходности компании и рынка для периода в 36 месяцев. Вычислите коэффициент парной корреляции  $\rho_{jm}$ . Вычислите  $\beta$  для компании с учетом формулы

$$\rho_{jm} = \hat{\beta}_j \frac{\hat{\sigma}_m}{\hat{\sigma}_j}. \text{ Совпадают ли значение } \beta \text{ с ожидаемым в п. в?}$$

### Упражнение 2. Оценка $\beta$ методом наименьших квадратов (МНК)

Из списка отраслей выберите компанию с относительно высоким уровнем риска и компанию с относительно маленьким уровнем риска. Компании должны принадлежать разным отраслям. Разделите вашу выборку на первую половину (январь 1978 – декабрь 1982) и вторую половину (январь 1983 – декабрь 1987) и выберите одну из них, с которой будете работать.

- а) Используя Stata, выбранные 60 наблюдений, и уравнение  $r_j - r_f = \alpha_j + \beta_j (r_m - r_f) + \varepsilon_j$ , оцените параметры уравнения МНК для каждой из фирм двух отраслей экономики. Оценки  $\beta$  соответствуют вашим ожиданиям и интуиции? Поясните.
- б) Для одной из компаний постройте графики фактической премии за риск и полученной по модели, а также график ошибок модели. Какие промежутки времени или даты соответствуют большим значениям ошибки?
- в) Для каждой из компаний проверьте гипотезу  $\alpha=0$  против альтернативы  $\alpha \neq 0$  на 95%-ом уровне значимости (прил. 3). Отклонение этой гипотезы говорит о не соответствии CAPM (Capital Asset Pricing Model) реальности? Почему?
- г) Для каждой из компаний постройте 95%-й доверительный интервал для  $\beta$  (прил. 3). Проверьте гипотезу о том, что риск компании совпадает со средним уровнем риска на рынке, т.е.  $\beta=1$ , против альтернативы  $\beta \neq 1$ . Вы нашли что-то необычное?
- д) Для каждой из компаний вычислите долю рыночного, недиверсифицируемого риска. У.Шарп считает типичным значение этой доли в 30%. Ваши вычисления соответствуют утверждению У.Шарпа?
- е) В вашей выборке высокие значения оценки  $\hat{\beta}$  соответствуют высоким значениям  $R^2$ ? Прокомментируйте ваш случай.

### Упражнение 3. Специальный случай с золотом

- а) Используя данные из файла GOLD с января 1976 по декабрь 1979 и CAPM, сгенерируйте переменные премий за риск для золота и рынка и оцените значение  $\beta$  для золота. Вычислите 95%-й доверительный интервал для  $\beta$  (прил. 3). Ваши оценки соответствуют здравому смыслу? Почему возможности такого актива могут быть привлекательны для инвестора, желающего снизить риск путем диверсификации? Что можно сказать о предполагаемой доходности такого актива?
- б) Оцените  $\beta$  для GOLD, используя данные с января 1980 по декабрь 1985 гг. Вычислите 95% доверительный интервал для  $\beta$  (прил. 3). Каковы изменения? Прокомментируйте возможные факторы спроса и предложения, приводящие к изменению значения  $\beta$ .

### Упражнение 4. Обратная регрессия

- а) Используя данные с января 1983 по декабрь 1987 из файла DELTA, а также данные для  $r_m$  из файла MARKET и для  $r_f$  из файла RKFREE, постройте переменные премии за риск для Delta Airlines  $Y_t$  и для рынка в целом  $X_t$ . Предположим, что вместо «корректной» CAPM  $Y_t = \alpha + \beta X_t + \varepsilon_t$  специфицирована «некорректная» обратная регрессия  $X_t = \delta + \gamma Y_t + v_t$ . Покажите, что  $\delta = -\alpha/\beta$ ,  $\gamma = 1/\beta$  и  $v_t = (-1/\beta)\varepsilon_t$ .
- б) Оцените параметры в уравнении  $X_t = \delta + \gamma Y_t + v_t$ . Обозначим найденные оценки для  $\delta$  и  $\gamma$  как  $d$  и  $g$ . Каково значение  $R^2$ ? Для приемлемого уровня значимости проверьте гипотезу о равенстве нулю  $\gamma$ . Найдите оценки  $\beta$  и  $\alpha$  как  $b_x = 1/g$  и  $a_x = -d/g$ .
- в) Оцените параметры корректного уравнения  $Y_t = \alpha + \beta X_t + \varepsilon_t$ . Обозначим найденные оценки  $b_y$  и  $a_y$  для  $\beta$  и  $\alpha$  соответственно. Каково значение  $R^2$ ? Для приемлемого уровня значимости проверьте гипотезу о равенстве нулю  $\beta$ .
- г) Какую из оценок следует предпочесть  $b_x$  или  $b_y$ ?

### Упражнение 5. Использование CAPM для создания портфелей

Из списка выберите две компании с относительно высоким уровнем риска и две компании с относительно маленьким уровнем риска. Компании должны принадлежать разным отраслям.

- а) Вычислите стандартное отклонение и среднее для доходностей каждой из выбранных компаний с января 1983 по декабрь 1987 гг. Результаты совпадают с вашими ожиданиями? Почему? В какую из компаний вы советуете инвестировать?
- б) Сконструируйте три альтернативных портфеля (1 миллион долларов всего). Портфель 1: 50% в компании с низким уровнем риска и 50% в компании с высоким уровнем риска. Портфель 2: 50% в каждой из компаний с низким уровнем риска. Портфель 3: 50% в каждой из компаний с высоким уровнем риска. Вычислите коэффициент корреляции между доходностью компаний в каждом из трех портфелей. Прокомментируйте значения и дайте интерпретацию этих корреляций. Для каждого из трех портфелей вычислите среднее и стандартное отклонение доходностей для рассматриваемого периода времени. Есть какие-нибудь сюрпризы?
- в) Какой из трех портфелей более оправдан с точки зрения снижения несистематического риска инвестиций? Почему?
- г) Оцените CAPM для каждого из трех портфелей. Проверьте гипотезу  $\beta=1$ . Какой из портфелей имеет меньшую долю несистематического риска?
- д) Для портфеля 1 сравните  $R^2$  из регрессии в п. г с  $R^2$  для отдельных регрессий для этих двух компаний. Вы ожидаете, что  $R^2$  из портфельного уравнения будет больше, чем из отдельных уравнений? Почему? Проинтерпретируйте результаты.

### Упражнение 6. Оценка стабильности $\beta$

Выберите две отрасли.

- а) Разделите выборку на две части: первая с января 1978 по декабрь 1982, вторая с января 1983 по декабрь 1987. Используя CAPM и приемлемый уровень значимости, протестируйте гипотезу, что для каждой компании параметры  $\alpha$  и  $\beta$  одинаковы для двух частей выборки, т.е. параметры постоянны для каждой их компаний по времени.

(Воспользуйтесь тестом Чоу, для этого используйте регрессию по всей выборке с фиктивной переменной для второй части выборки и переменными взаимодействия с фиктивной переменной, а также командой `test` пакета Stata).

б) Для каждой отрасли проверьте гипотезу о том, что параметры  $\alpha$  и  $\beta$  одинаковы для всех фирм в отрасли с января 1978 по декабрь 1987 гг.

в) Для каждой отрасли проверьте гипотезу о том, что параметры  $\alpha$  и  $\beta$  одинаковы для всех фирм в отрасли и совпадают со значениями в обеих временных интервалах с января 1978 по декабрь 1982 и с января 1983 по декабрь 1987 гг. Что вы можете сказать об устойчивости параметров между компаниями и в течение времени?

#### **Упражнение 7.** Нарушение предположений МНК

а) Для любой из фирм постройте уравнение CAPM. Проверьте гипотезу о гомоскедастичности дисперсии остатков регрессии. Прокомментируйте результаты.

б) Для любой из фирм постройте уравнение CAPM. Проверьте гипотезу об автокоррелированности остатков регрессии с помощью теста Дарбина-Уотсона (прил. 1). Переоцените модель, воспользовавшись процедурой Хилдрета-Лу или Кохрейна-Оркатта. Прокомментируйте полученные результаты.

в) Для любой из фирм постройте уравнение CAPM. Получите остатки регрессии. Проверьте гипотезу о нормальном распределении этих остатков, воспользовавшись критерием Колмогорова-Смирнова. Сделайте выводы.

#### **Упражнение 8.** Некоторые случаи при оценке CAPM

Два случая: 1) эффект аварии на атомной электростанции, произошедшей в 28 марта 1979 г., на доходность активов General Public Utilities, которой принадлежал завод. 2) эффект войн покупателей DuPont и Dow Chemical в июне-августе 1981 г. в попытке завладеть компанией Conoco. Данные находятся в файле EVENTS.

а) Постройте переменные для премии за риск для компаний GPU, DUPONT и DOW и рынка в целом с января 1976 по декабрь 1985 гг. и для CONOCO с января 1976 по сентябрь 1981 гг. Вычислите математические ожидания доходностей.

б) Оцените на основе CAPM  $\alpha$  и  $\beta$  для General Public Utilities для периода времени с января 1976 по декабрь 1985 г., но исключив из выборки наблюдение для апреля 1979 г. На основе полученной регрессии вычислите прогнозируемое для апреля 1979 года значение премии за риск и ошибку прогноза. Как вы проинтерпретируете полученный результат?

в) Постройте фиктивную переменную с именем TMIDUM, которая равна 1 только для апреля 1979 г. и нулю для всех остальных месяцев. Включите эту переменную в модель CAPM в качестве объясняющей и получите оценки МНК. Сравните значение оценки коэффициента при переменной TMIDUM со значением ошибки из п. б) для апреля 1979 г. Сравните оценки коэффициентов  $\beta$  для этих двух регрессий. Почему они равны? Протестируйте гипотезу о равенстве нулю коэффициента при переменной TMIDUM. Катастрофа оказалась значимым событием?

г) Используя CAPM и как можно более длинный период времени, оцените  $\alpha$  и  $\beta$  для трех компаний DuPont (покупатель), Dow Chemical (неудачник), Conoco (цель покупки) с учетом того, что для Conoco данные оканчиваются в сентябре 1981 г. Для каждой компании вычислите среднее значение ошибок регрессии для месяцев июня, июля, августа и сентября 1981 г.

д) Постройте фиктивные переменные для DuPont, Dow Chemical и Conoco, которые равны 1 для июня, июля, августа и сентября 1981 г. и равны 0 для остальных месяцев. Используя CAPM, оцените расширенную регрессию для января 1976-декабря 1985 гг. для DuPont и Dow Chemical и январь 1976 – сентябрь 1981 гг. для Conoco, включив фиктивные переменные. Сравните оценки коэффициентов при этих переменных с ошибками, полученными для каждого из месяцев в п. г. Сравните значения  $\beta$ . Прокомментируйте результаты. Как вы можете проинтерпретировать знаки оцененных коэффициентов, в частности, кто из держателей акций выиграл, а кто проиграл? Проверьте для всех компаний гипотезу о равенстве нулю коэффициентов при фиктивных переменных.

**Упражнение 9.** Выполните самостоятельно на основе модели CAPM оценку стоимости активов на примере предприятия реального сектора экономики в соответствии с направленностью программы.

Лабораторная работа 7. Векторные авторегрессии.

1. Сгенерируйте в EViews временной ряд, подчиняющийся авторегрессии первого порядка с коэффициентом 0,99. Проверьте полученный ряд на стационарность с помощью различных тестов.

2. В файле lab7 имеются два временных ряда денежного агрегата M2 и инфляции. Протестируйте ряды на причинность по Грейнджеру, на коинтеграцию по Йохансену. Сделайте выводы. Постройте адекватную модель векторной авторегрессии.

#### **7. Методические рекомендации по подготовке к сдаче зачета**

Зачет осуществляется в рамках завершения изучения дисциплины (модуля) и позволяет определить качество усвоения изученного материала, а также степень сформированности компетенций.

На зачет выносятся материал в объеме, предусмотренном рабочей программой учебной дисциплины.

Экзаменатору предоставляется право задавать обучающимся вопросы в рамках пройденного материала, а также, помимо теоретических вопросов, предлагать задачи практико-ориентированной направленности по программе данного курса.

Зачет, как правило, проводится в устной форме.

На подготовку к ответу отводится 40 минут.

Результат зачета выражается словами «зачтено», «не зачтено».

За нарушение дисциплины и порядка обучающиеся могут быть удалены с зачета.

Подготовка обучающегося к зачету включает в себя три этапа:

- самостоятельная работа в течение семестра;
- непосредственная подготовка в дни, предшествующие зачету по темам курса;
- подготовка к ответу на вопросы, указанные в рабочей программе дисциплины.

При подготовке к зачету обучающимся целесообразно использовать материалы лекций, основную и дополнительную литературу.