

**Эконометрика, 2018-2019, 3 модуль**  
**Семинары 8-9**  
**4.03.19, 11.03.19**  
**Для Группы Э\_Б2016\_Э\_3**  
**Семинарист О.А.Демидова**

**Модели бинарного выбора**

**Задача 11.1.**

Исследователя интересует зависимость вероятности найти работу от уровня образования индивидуума. Введя в качестве зависимой переменную EMP, равную 1 для работающих и 0 для неработающих и S – количество лет обучения в качестве объясняющей, он оценил логит – модель:

$$P\{EMP_i = 1\} = \frac{1}{1 + \exp\{-Z_i\}}, \quad Z_i = -1.006 + 0.148S$$

(0.242)      (0.018)

Оцените предельный эффект объясняющего фактора для среднего значения переменной  $S = 13.5$ .

**Задача 11.2.**

По наблюдениям для 570 индивидуумов оценена зависимость получения школьником аттестата от обобщенной оценки результатов тестов X. Переменная Y принимает значение 1, если аттестат был получен и 0 в противном случае.

Оцененные модели имеют следующий вид:

Логит:  $P\{Y_i = 1\} = \frac{1}{1 + \exp\{-Z_i\}}, \quad Z_i = -5.004 + 0.1666X$

(0.865)      (0.021)

Пробит:  $P\{Y_i = 1\} = F(Z_i), \quad F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt, \quad Z = -2.7 + 0.53X$

(0.083)      (0.0117)

Дайте экономическую интерпретацию полученным результатам для логит и пробит моделей. Найдите предельный эффект объясняющего фактора в точке  $\bar{X} = 50.15$ .

**Задача 11.3.**

Из 750 обратившихся за ссудой в банк 250 было в ней отказано. Для оценки вероятности получения ссуды были оценены линейная и пробит модели:

$Y = 0.5 + 1.5X$ ,

$P\{Y_i = 1\} = F(Z_i), \quad F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$

$Z = 0.45 + 3X$

где  $Y_i = 1$  для получивших ссуду и 0 иначе, X – доход просителя.

По пробит модели найти предельный эффект дохода в среднем.

**Задача 11.4.**

Для того, чтобы определить, эффективна ли новая методика преподавания микроэкономики, провели следующий эксперимент: протестировали всех студентов по микроэкономике в конце первого и второго семестра. Часть студентов во втором семестре обучали по новой методике, часть по старой. После этого в качестве объясняющей выбрали переменную  $Y$ , равную 1, если результат студента улучшился и 0 в противном случае, а в качестве объясняющих переменных  $X_1$  – результаты теста в первом семестре,  $X_2$  – средний балл по остальным предметам,  $D$  – равную 1, если студент обучался по новой методике и 0, если по старой.

По имеющимся данным оценили пробит- модель:

$$P\{Y_i = 1\} = F(Z_i), \quad F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$$

$$Z = -7.452 + 0.052X_1 + 1.626X_2 + 1.426D.$$

Найдите предельный эффект переменной  $D$  при средних значениях  $\bar{X}_1 = 21.938$ ,  $\bar{X}_2 = 3.117$  (разность вероятностей улучшения результата при  $D = 1$  и  $D = 0$ ).

### Задание 11.3.

Для анализа аудитории, использующей Интернет для учебы по данным для 1314 индивидов были оценены линейная и пробит модели (последняя с предельными эффектами), в которых  $intlear = 1$  при использовании индивидом Интернета для учебы и 0 в противном случае,  $male = 1$  для мужчин и 0 для женщин,  $income$  – заработная плата индивида по основному месту работы,  $age$  – возраст. В скобках указаны соответственно  $t$  или  $z$  - статистики. В чем состоят недостатки линейной модели? Дайте интерпретацию полученным результатам.

$$INTLEAR = -0.78 - .013 AGE - 4.53 \cdot 10^{-10} INCOME - 0.073 MALE$$

(18.48)    (-11.45)                      (-0.96)                      (-3.05)

$$P\{INTLEAR_i = 1\} = F(Z_i), \quad F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-t^2/2} dt$$

$$Z = 1.01 - 0.044 AGE - 1.51 \cdot 10^{-9} INCOME - 0.23 MALE$$

(7.25)    (-10.82)                      (-0.98)                      (-3.08)

Marginal effects after probit

variable	dy/dx	Std. Err.	z	P>z	[ 95% C.I. ]	
age	-.0147444	.00133	-11.08	0.000	-.017353	-.012136
income	-5.07e-10	.00000	-0.98	0.328	-1.5e-09	5.1e-10
male*	-.0783057	.02536	-3.09	0.002	-.128015	-.028597

(\*) dy/dx is for discrete change of dummy variable from 0 to 1

4. По 1000 наблюдений Винни-Пух оценил логистическую модель  $\mathbb{P}(Y_i = 1) = F(\beta_0 + \beta_1 X_i)$ , где  $X_i$  — количество времени в часах, проведённое в гостях, а  $Y_i$  — факт застревания при выходе.

Оценки параметров равны  $\hat{\beta}_0 = 1$ ,  $\hat{\beta}_1 = 2$ , с оценкой ковариационной матрицы

$$\begin{pmatrix} 0.25 & 0.1 \\ 0.1 & 0.36 \end{pmatrix}.$$

- Проверьте значимость отдельных коэффициентов при уровне значимости 5%;
- Найдите предельный эффект времени, проведённого в гостях, на вероятность застрять при выходе для полчасового визита;
- Найдите максимально возможный предельный эффект.

### Эндогенность. Инструментальные переменные

Критика М.Фридменом стандартной функции потребления, раздел 8.5.

1) (Доугерти, 8.7) В некоторой экономике дисперсия переменного дохода составляет 0.5 от дисперсии постоянного дохода, склонность к потреблению товаров кратковременного пользования за счет постоянного дохода составляет 0.6, а расходы на товары длительного пользования отсутствуют. Каким будет значение мультипликатора, полученного на основе построения «наивной» регрессионной зависимости потребления от дохода, и каково его истинное значение?

2) (Доугерти, раздел 8)

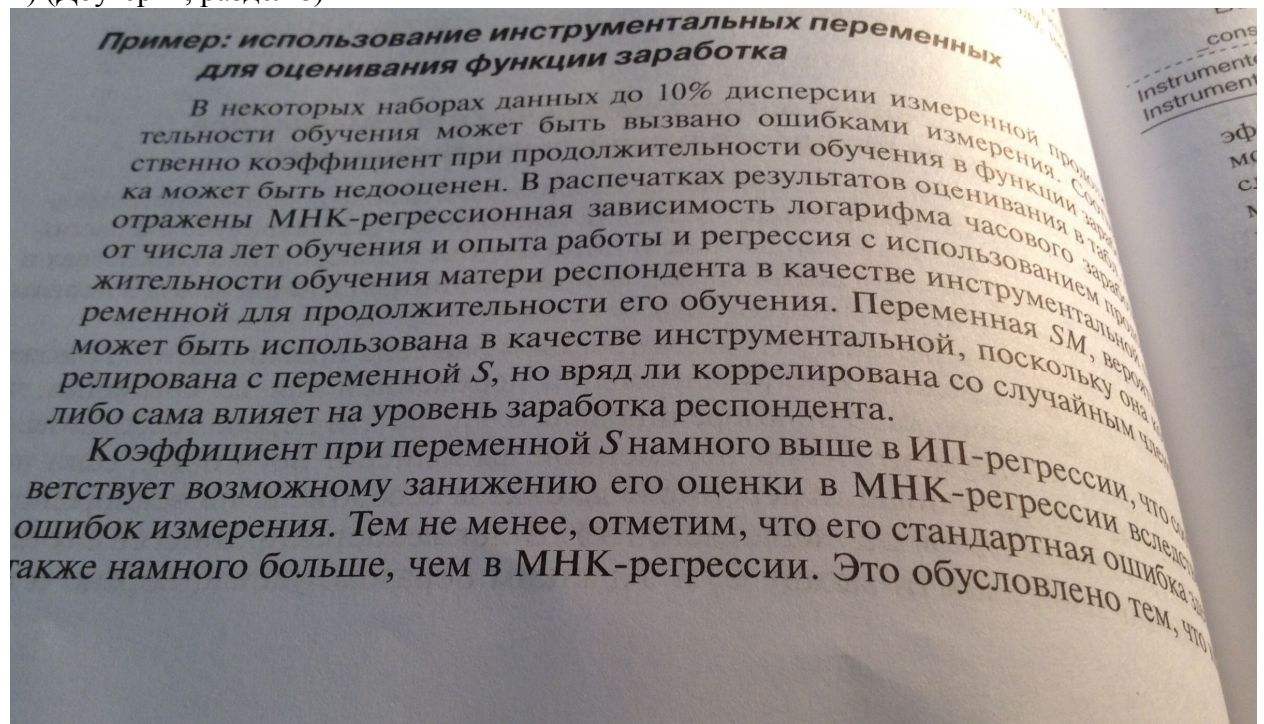


Таблица 8.2

reg LGearn S EXP

Source	SS	df	MS				
Model	50.9842581	2	25.492129				
Residual	135.723385	537	.252743734				
Total	186.707643	539	.34639637				

LGearn	Coef.	Std. Err.	t	P >  t	[95% Conf. Interval]
S	.1235911	.0090989	13.58	0.000	.1057173 .141465
EXP	.0350826	.0050046	7.01	0.000	.0252515 .0449137
_cons	.5093196	.1663823	3.06	0.002	.1824796 .8361596

Number of obs = 540  
F(2,537) = 100.86  
Prob > F = 0.0000  
R-squared = 0.2731  
Adj R-squared = 0.2704  
Root MSE = .50274

reg LGearn EXP (S=SM)  
instrumental variables (2SLS) regression

Source	SS	df	MS				
Model	46.9446075	2	23.4723038				
Residual	139.763036	537	.260266361				
Total	186.707643	539	.34639637				

LGearn	Coef.	Std. Err.	t	P >  t	[95% Conf. Interval]
S	.1599676	.0252801	6.33	0.000	.1103076 .2096277
EXP	.0394422	.0058092	6.79	0.000	.0280306 .0508537
_cons	-.0617062	.4061769	-0.15	0.879	-.8595966 .7361841

Number of obs = 540  
F(2,537) = 28.38  
Prob > F = 0.0000  
R-squared = 0.2514  
Adj R-squared = 0.2486  
Root MSE = .51016

Instrumented: S  
Instruments: EXP SM



Таблица 8.4

. ivreg LGEARN EXP ASVABC MALE ETHBLACK ETHHISP  
(S=SM SF SIBLINGS LIBRARY)  
Instrumental variables (2SLS) regression

Source	SS	df	MS				
Model	64.4915831	6	10.7485972				
Residual	122.21606	533	.229298424				
Total	186.707643	539	.34639637				
LGEARN	Coef.	Std. Err.	t	P >  t	[95% Conf.		
S	.111379	.0476886	2.34	0.020	.0176984	Inter	
EXP	.0258798	.0081187	3.19	0.002	.0099313	Inter	
ASVABC	.0092263	.007991	1.15	0.249	-.0064714	Inter	
MALE	.2619787	.0429283	6.10	0.000	.1776492	Inter	
ETHBLACK	-.0121846	.0822942	-0.15	0.882	-.1738454	Inter	
ETHHISP	.0457639	.0955115	0.48	0.632	-.1418612	Inter	
_cons	.2258512	.3887468	0.58	0.562	-.5378125	Inter	

Number of obs = 540  
F(6,533) = 37.7  
Prob > F = 0.000  
R-squared = 0.346  
Adj R-squared = 0.338  
Root MSE = .478

Instrumented: S  
Instruments: EXP ASVABC MALE ETHBLACK ETHHISP SM SF SIBLINGS LIBRARY

. estimates store EARNIV  
. reg LGEARN S EXP ASVABC MALE ETHBLACK ETHHISP

Source	SS	df	MS				
Model	65.490707	6	10.9151178				
Residual	121.216936	533	.227423895				
Total	186.707643	539	.34639637				
LGEARN	Coef.	Std. Err.	t	P >  t	[95% Conf.		
S	.0883257	.0109987	8.03	0.000	.0667196	Inter	
EXP	.0227131	.0050095	4.53	0.000	.0128724	Inter	
ASVABC	.0129274	.0028834	4.48	0.000	.0072633	Inter	
MALE	.2652878	.042235	6.28	0.000	.1823203	Inter	
ETHBLACK	.0077265	.0715863	0.11	0.914	-.1328994	Inter	
ETHHISP	.0536544	.0937966	0.57	0.568	-.1306019	Inter	
_cons	.4002952	.1663149	2.41	0.016	.0735821	Inter	

Number of obs = 540  
F(6,533) = 47.99  
Prob > F = 0.000  
R-squared = 0.353  
Adj R-squared = 0.345  
Root MSE = .478



Окончание табл. 8.4

estimates store EARNOLS  
hausman EARNIV EARNOLS, constant  
— Coefficients —

	(b) EARNIV	(B) EARNOLS	(b - B) Difference	sqrt(diag (V_b - V_B)) S.E.
S	.111379	.0883257	.0230533	.0464029
EXP	.0258798	.0227131	.0031667	.0063889
ASVABC	.0092263	.0129274	-.0037011	.0074527
MALE	.2619787	.2652878	-.0033091	.0076842
ETHBLACK	-.0121846	.0077265	-.019911	.0405924
ETHHISP	.0457639	.0536544	-.0078904	.018018
cons	.2258512	.4002952	-.174444	.3513736

b = consistent under Ho and Ha; obtained from ivreg  
B = inconsistent under Ha, efficient under Ho;  
obtained from regress

Test: Ho: difference in coefficients not systematic  
 $\chi^2(7) = (b - B)'[(V_b - V_B)^{-1}](b - B) = 0.25$   
 Prob >  $\chi^2 = 0.9999$

Воспользуйтесь данными файла Dougherty.dta

## 1) Переменные

Тип переменной обозначен С для непрерывных и D для фиктивных (dummy) переменных.

MALE (D) – 1 для мужчин и 0 для женщин,

ETHBLACK (D) – 1 для чернокожих и 0 для белых и испаноязычных,

ETHHISP (D) – 1 для испаноязычных и 0 для всех остальных,

ETHWHITE (D) – 1 для белых и 0 для всех остальных,

AGE (C) – возраст индивидуума,

S (C) - количество лет обучения,

ASVAB2 (C) – результаты теста по арифметике,

ASVAB3 (C) – результаты теста по правописанию,

ASVAB4 (C) – результаты теста по пониманию прочитанного материала,

ASVAB5 (C) – результаты теста на скорость выполнения численных операций,

ASVAB6 (C) – результаты теста на скорость восприятия прочитанного материала,

ASVABC = 0.5 ASVAB2 + 0.25 ASVAB3 + 0.25 ASVAB4 (C) – обобщенная переменная, характеризующая способности школьника,

HEIGHT (C) – рост индивидуума в 1985 г.,

WEIGHT85 (C) – вес индивидуума в 1985 г.,

WEIGHT94 (C) – вес индивидуума в 1994 г.,

SM (C) - количество лет обучения матери индивидуума,

SF (C) - количество лет обучения отца индивидуума,

SIBLINGS (C) – число родных братьев или сестер индивидуума,

EARNINGS (C) - почасовая заработная плата в 1994 г.

Возможно, переменная ASVABC неадекватно характеризует способности индивидуума. Ошибки измерения могут привести к смещению оценки МНК соответствующего коэффициента.

а) Оцените регрессию LGARN на S и ASVABC

1) Используя метод наименьших квадратов,

- 2) Используя метод инструментальных переменных, выбрав переменную SM в качестве инструментальной для ASVABC.

Проведите тест Хаусмана для определения, есть ли значимое различие между оценками метода инструментальных переменных и метода наименьших квадратов.

- б) Повторить проделанные в п. а) вычисления, используя для переменной ASVABC переменные SM, SF, SIBLINGS, LIBRARY в качестве инструментальных.

*Методические рекомендации по выполнению упражнения 1 в пакете STATA*

*а) Последовательно наберите в командном окне*

*gen LG EARN = ln(EARNINGS)*

*ivreg LG EARN S EXP MALE ETHBLACK ETHHISP (ASVABC = SM)*

*estimates store IV*

*reg LG EARN S EXP MALE ETHBLACK ETHHISP ASVABC*

*estimates store OLS*

*hausman IV OLS, constant*

- б) То же самое, только в 2 сверху строчке в скобках набираете ASVABC = SM SF SIBLINGS LIBRARY*