

Эконометрика, 2017-2018, 2 модуль

Семинар 5.1

27.11.17 для

Группы Э_Б2015_Э_3

Семинарист О.А.Демидова

Выбор функциональной формы модели

- I. По данным файла `dougherty.dta`, используя тест Бокса-Кокса, с помощью статистического пакета STATA, оцените параметры модели

$$EARNINGS^{(\theta)} = \beta_0 + \beta_1 S^{(\lambda)} + \beta_2 ASVABC^{(\lambda)} + \varepsilon$$

- 1) Когда переменные левой и правой части преобразуются не одинаково:

Это можно сделать с помощью команды:

```
boxcox EARNINGS S ASVABC, model (theta)
```

- 2) Когда переменные в обеих частях модели преобразуются одинаково.

Это можно сделать с помощью команды:

```
boxcox EARNINGS S ASVABC, model (lambda),
```

- 3) Когда преобразуется только зависимая переменная

Это можно сделать с помощью команды

```
boxcox EARNINGS S ASVABC, model (lhsonly),
```

- 4) Когда преобразуются только независимые переменные

Это можно сделать с помощью команды

```
boxcox EARNINGS S ASVABC, model (rhsonly),
```

- 5) поэкспериментируйте с проведением теста Бокса-Кокса в модели с другим набором переменных (если Вы не хотите преобразовывать какие-то переменные, например, `dummy`, то это можно сделать с помощью команды, аналогичной следующей:

```
boxcox EARNINGS S ASVABC, notrans(MALE) model (theta)
```

```
boxcox EARNINGS S ASVABC, notrans(MALE) model (lambda)
```

- II. По данным файла `dougherty.dta` выберите между линейной, полулогарифмической и линейной в логарифмах моделями с помощью теста

А) РЕ теста Дэвидсона, Уайта и МакКиннона

б) Бера и МакАлера.

а) РЕ теста Дэвидсона, Уайта и МакКиннона

Необходимые команды

Выбор между линейной и линейной в логарифмах моделью

Сначала оценим линейную модель с помощью команды:

```
reg EARNINGS S ASVABC,
```

Сохранить предсказанные значения зависимой переменной можно с помощью команды:

```
predict y_hat
```

Предварительно создав необходимые дополнительные переменные,

аналогично оценим линейную в логарифмах модель с помощью команды,

```
gen lnEARNINGS = ln(EARNINGS)
```

```
gen lnS = ln(S)
```

```
gen lnASVABC = ln(ASVABC)
```

```
reg lnEARNINGS lnS lnASVABC
```

и сохраним предсказанные значения зависимой переменной с помощью команды:

```
predict ln_y_hat.
```

Теперь переходим к шагу 2, оценим дополнительные модели.

Сначала создадим дополнительную разность для линейной модели:

```
gen lin_add= ln(y_hat)- ln_y_hat
```

и оценим эту модель:

```
reg EARNINGS S ASVABC lin_add
```

Создадим также дополнительную разность для линейной в логарифмах модели:

```
gen log_add=y_hat-exp(ln_y_hat)
```

оценим эту модель:

```
reg lnEARNINGS lnS lnASVABC log_add
```

Если

- 1) оба коэффициента при дополнительных разностях значимы или оба незначимы, то выбрать посредством теста Дэвидсона, Уайта и МакКиннона невозможно,
- 2) если незначим только коэффициент при дополнительной разности в линейной модели, то лучше линейная модель,
- 3) если незначим только коэффициент при дополнительной разности в линейной в логарифмах модели, то лучше линейная в логарифмах модель.

Выбор между линейной и полулогарифмической моделями

Сначала оценим линейную модель с помощью команды:

```
reg EARNINGS S ASVABC
```

Сохранить предсказанные значения зависимой переменной можно с помощью команды:

```
predict y_hat
```

Аналогично оценим полулогарифмическую модель с помощью команды:

```
reg lnEARNINGS S ASVABC
```

и сохраним предсказанные значения зависимой переменной с помощью команды:

```
predict semiln_y_hat.
```

Теперь переходим к шагу 2, оценим дополнительные модели.

Сначала создадим дополнительную разность для линейной модели:

```
gen lin_adds= ln(y_hat)- semiln_y_hat
```

и оценим эту модель:

```
reg EARNINGS S ASVABC lin_adds
```

Создадим также дополнительную разность для полулогарифмической модели:

```
gen semilog_add=y_hat-exp(ln_y_hat)
```

оценим эту модель:

```
reg lnEARNINGS S ASVABC semilog_add
```

Если

- 1) оба коэффициента при дополнительных разностях значимы или оба незначимы, то выбрать посредством теста Дэвидсона, Уайта и МакКиннона невозможно,
- 2) если незначим только коэффициент при дополнительной разности в линейной модели, то лучше линейная модель,
- 3) если незначим только коэффициент при дополнительной разности в полулогарифмической модели, то лучше полулогарифмическая модель.

б) Тест Бера и МакАлера

Необходимые команды

Выбор между линейной и полулогарифмической моделями

Сначала оценим полулогарифмическую модель с помощью команды:

```
reg lnEARNINGS S ASVABC
```

и сохраним предсказанные значения зависимой переменной с помощью команды:

```
predict semiln_y_hat.
```

Аналогично оценим линейную модель с помощью команды:

```
reg EARNINGS S ASVABC
```

Сохранить предсказанные значения зависимой переменной можно с помощью команды:

```
predict y_hat
```

Теперь переходим к шагу 2, оценим дополнительные модели.

```
1) gen y1= exp(semiln_y_hat)
```

```
reg y1 S ASVABC
```

Сохранить остатки регрессии можно с помощью команды

```
predict res1, resid
```

```
2) gen y2= ln(y_hat)
```

```
reg y2 S ASVABC
```

Сохранить остатки регрессии можно с помощью команды

```
predict res2, resid
```

Теперь переходим к шагу 3, оценив еще 2 дополнительные модели.

```
reg lnEARNINGS S ASVABC res1
```

```
reg EARNINGS S ASVABC res2
```

Если

- 1) оба коэффициента при res1 и res2 незначимы или оба значимы, то выбрать посредством теста Бера и МакАлера невозможно,

- 2) если незначим только коэффициент при $res1$, то лучше полулогарифмическая модель,
- 3) если незначим только коэффициент при $res2$, то лучше линейная модель.

III. По данным файла `dougherty.dta` выберите между линейной и полулогарифмической моделями с помощью теста Зарембки.

Необходимые команды

Выбор между линейной и полулогарифмической моделями

```
reg EARNINGS S ASVABC
gen lnEARNINGS = ln(EARNINGS)
reg lnEARNINGS S ASVABC
```

means EARNINGS

Variable	Type	Obs	Mean	[95% Conf. Interval]	
EARNINGS	Arithmetic	540	19.71924	18.48493	20.95355
	Geometric	540	16.3442	15.54379	17.18584
	Harmonic	540	13.77391	13.05586	14.57555

```
gen EARNINGSstar= EARNINGS/16.3442
gen lnEARNINGSstar = ln(EARNINGSstar)
reg EARNINGSstar S ASVABC
сохраните RSS с помощью команды scalar rss3=e(rss)
reg lnEARNINGSstar S ASVABC
сохраните RSS с помощью команды scalar rss4=e(rss)
```

Используя RSS из оцененных регрессий, следует рассчитать тестовую F – статистику

$$scalar \chi^2 = 0.5 * 540 * abs(ln(rss4/rss3))$$

$$display \chi^2$$

и p-value для этой статистики:

$$display \chi^2_{tail}(1, \chi^2)$$

Если рассчитанное p-value не превышает выбранного уровня значимости, то основная гипотеза отвергается, есть разница между исходными линейной и полулогарифмическими моделями.