



SÉRIE DE NOTAS TÉCNICAS EM ECONOMIA DA UFG

NT N. 05

NOTA TÉCNICA EM ECONOMIA

n. 05

Publicação cujo objetivo é auxiliar na elaboração de aulas e de pesquisas do Curso de Ciências Econômicas da UFG, divulgando demonstrações técnicas e metodológicas ou rotinas computacionais voltadas para a economia.

As opiniões contidas nesta publicação não representam o ponto de vista do NEPEC ou da FACE/UFG, sendo de inteira responsabilidade de seu(s) autor(es).

Reprodução permitida, desde que citada a fonte.

FACE – Faculdade de Administração, Ciências Contábeis e Ciências Econômicas
Curso de Ciências Econômicas
Direção FACE
Maria do Amparo Albuquerque Aguiar
Vice-Direção FACE
Mauro Caetano de Souza
Coordenação do Curso de Ciências Econômicas
Dnilson Carlos Dias

NEPEC – Núcleo de Estudos e Pesquisas Econômicas
Coordenação
Dnilson Carlos Dias

Endereço

Campus Samambaia, Prédio da FACE – Rodovia Goiânia/Nova Veneza, km. 0 – Caixa Postal 131, CEP 74001-970, Goiânia – GO.
Tel. (62) 3521 – 1390

URL

<http://www.face.ufg.br/eco>

Exportando o output de regressões no Stata

FACE/UFG

NEPEC/FACE/UFG
Goiânia – Agosto de 2011
Versão 1.0



Exportando o Output de Regressões no Stata

Sandro Eduardo Monsueto¹

Universidade Federal de Goiás

Resumo:

Esta nota técnica tem por objetivo explicar, a partir de um exemplo, o uso do comando **esttab**, desenvolvido por Jann (2005) para o pacote estatístico Stata. Tal comando permite exportar diretamente para um arquivo em formato RTF, compatível com a maioria dos processadores de texto tradicionais, uma tabela com os resultados das regressões lineares e não lineares, economizando tempo e esforço do usuário.

Palavras chave: Software Stata, esttab, output.

Abstract:

*This technical note aims to explain, from a example, the use of the **esttab** command, developed by Jann (2005) for the Stata statistical package. This command lets you export directly to a RTF file, compatible with most traditional word processors, a table with the results of linear and nonlinear regressions, saving time and user effort.*

Key words: Stata Software, esttab, output.

1. Introdução

Talvez uma das maiores dificuldades que o iniciante, ou incluso o usuário um pouco mais avançado, tenha quando utiliza o software Stata seja a importação dos outputs de regressões para um editor de textos convencional, principalmente os não familiarizados com o sistema LaTeX. Pensando nisso, esta breve nota técnica tem como objetivo introduzir o usuário no uso do comando **esttab**, que pode ser usado tanto para organizar a apresentação dos resultados na tela do Stata, colocando na mesma tabela os diversos modelos estimados, como também permite a importação desta tabela para um arquivo em formato RTF, compatível com a maioria dos editores de texto.

¹ monsueto@face.ufg.br

As informações dispostas aqui são baseadas no próprio arquivo help do comando, elaborado por Jann (2005), e que contém outros exemplos práticos. Aqui é abordada uma combinação de algumas das diversas opções disponíveis com um exemplo específico. Antes de prosseguir, é necessário verificar se os comandos **estout** e **esttab** estão instalados junto com o software. Caso contrário, proceda com a instalação usando a linha de comando:

```
net install estout
ou
net search estout
```

que deve de ser suficiente para instalar os dois comandos ao mesmo tempo. Adicionalmente, são necessários conhecimentos básicos de regressão e arquivos do tipo **do-file**.

2. A sintaxe do comando

De modo geral, os comandos **estout** e **esttab** produzem na tela de output do Stata o mesmo resultado. Os dois possuem linhas de sintaxe e arquivos help independentes, mas a estrutura do primeiro pode ser usada no segundo. Como o **esttab** se mostra mais eficaz para produzir arquivos em formato RTF, ele será o foco desta nota técnica. Contudo, as mesmas funções e explicações são validadas para ambos os comandos.

Para facilitar, os exemplos serão feitos usando uma base de dados disponível no próprio Stata, sobre automóveis (*auto.dta*) – a mesma usada no arquivo help do comando, com algumas alterações. Para carregá-la e realizar as mudanças necessárias, use a seguinte seqüência de comandos:

```
sysuse auto
replace price = price / 1000
replace weight = weight / 1000
```

Basicamente, a seqüência para usar o comando **esttab** (ou **estout**) é a seguinte:

- 1º. Estimar o modelo (MQO, probit, VI, etc).
- 2º. Armazenar o resultado na memória e fornecer uma etiqueta para o modelo com o comando **estimates store**.
- 3º. Executar o comando **esttab**.

Os modelos podem ser tanto lineares de M.Q.O. (**regress**) ou de variáveis instrumentais (**ivreg**) como também os não lineares de probabilidade (**probit**, **dprobit**, **logit**) e outros. Recomenda-se cuidado com modelos com mais de uma equação (**mlogit**, **sqreg**, etc) e não exagerar na quantidade de modelos na mesma tabela. Um primeiro exemplo de uso desta seqüência de comandos pode ser obtido desta forma:

```

reg price weight mpg, r
estimates store Modelo1
esttab Modelo1

```

Obviamente, podem ser produzidos e comparados diversos modelos alternativos na mesma tabela de resultados:

```

reg price weight mpg foreign, r
estimates store Modelo1
reg price weight, r
estimates store Modelo2
reg price mpg, r
estimates store Modelo3
esttab Modelo1 Modelo2 Modelo3

```

O resultado desta última sequência aparece na Figura 1 abaixo:

. esttab Modelo1 Modelo2 Modelo3			
	(1)	(2)	(3)
	price	price	price
weight	3.465*** (4.46)	2.044*** (5.24)	
mpg	0.0219 (0.27)		-0.239*** (-4.16)
foreign	3.673*** (5.52)		
_cons	-5.854 (-1.51)	-0.00671 (-0.01)	11.25*** (8.18)
N	74	74	74
t statistics in parentheses			
* p<0.05, ** p<0.01, *** p<0.001			

Figura 1. Exemplo simples de resultado do comando esttab

Resultados mais complexos podem ser obtidos com o uso das diversas opções do comando e também existe a alternativa de montar um arquivo no formato RTF ao invés de exibir os resultados na tela do programa. Nesta nota técnica não são abordadas todas as opções disponíveis. Porém, o usuário com um pouco mais de familiaridade com os comandos do Stata não deve de encontrar dificuldade para utilizar as opções mais complexas ao seu caso específico, utilizando as informações do arquivo de help do comando.

O exemplo que vamos explorar gera como resultado na tela do Stata o que aparece na Figura 2, que apresenta uma tabela com os coeficientes estimados, erros padrão entre

parênteses e estatísticas adicionais, como o R^2 ajustado e o p-valor F, além dos tradicionais asteriscos que simbolizam o nível de significância de cada coeficiente.

	(1) Price	(2) Price	(3) Price
Peso do Carro	3.465 * (0.78)	2.044 * (0.39)	
Km Rodados	0.022 (0.08)		-0.239 * (0.06)
Tipo de Carro	3.673 * (0.66)		
Constante	-5.854 (3.87)	-0.007 (1.03)	11.253 * (1.38)
R2 Ajustado	0.4781	0.2802	0.2087
Número de obs.	74	74	74
F	15.23	27.51	17.28
Prob>F	0.00	0.00	0.00
Erros padrão entre parênteses *** p<0.10, ** p<0.05, * p<0.01			

Figura 2. Exemplo de resultado do comando estab

Além disso, pode ser gerado um arquivo no formato RTF que contém a seguinte tabela:

	(1) Price		(2) Price		(3) Price	
Peso do Carro	3.465	*	2.044	*		
	(0.78)		(0.39)			
Km Rodados	0.022				-0.239	*
	(0.08)				(0.06)	
Tipo de Carro	3.673	*				
	(0.66)					
Constante	-5.854		-0.007		11.253	*
	(3.87)		(1.03)		(1.38)	
R2 Ajustado	0.4781		0.2802		0.2087	
Número de obs.	74		74		74	
F	15.23		27.51		17.28	
Prob>F	0.00		0.00		0.00	
Erros padrão entre parênteses *** p<0.10, ** p<0.05, * p<0.01						

Para gerar estes resultados, o comando **esttab** foi usado da seguinte forma:

```
#d;
esttab Modelo1 Modelo2 Modelo3 using resultado.rtf,
replace

cells(
    b (star fmt(3))
    se(par fmt(2))
)

varlabels(
    _cons Constante
    weight "Peso do Carro"
    mpg "Km Rodados"
    foreign "Tipo de Carro"
)

label

stats(
    r2_a N F p,
    fmt(4 0 2 2)
    labels("R2 Ajustado" "Número de obs." F "Prob>F")
)

collabels(none)
starlevels(*** 0.10 ** 0.05 * 0.01)
stardetach
legend
note ("Erros padrão entre parênteses")
;
```

Abaixo, são explicadas as opções utilizadas no exemplo.

```
esttab Modelo1 Modelo2 Modelo3 using resultado.rtf,  
replace
```

Esta linha define os modelos armazenados que entrarão na tabela e o arquivo onde será gravado o resultado final². A opção **replace** indica que se já existir um arquivo com o mesmo nome o mesmo será substituído por um novo. Adicionalmente, pode ser usada a opção **append** no lugar de **replace**, para adicionar um novo resultado ao arquivo RTF ao invés de substituí-lo. A opção **using nome_do_arquivo** deve ser retirada se o usuário desejar exibir os resultados apenas na tela do próprio Stata, sem gerar o arquivo RTF (ou seja, para gerar apenas o exibido na Figura 2). Isso é recomendável principalmente se são os primeiros modelos estimados ou quando se está interessado apenas em comparar as estimativas antes de enviá-los ao relatório final, por exemplo. Neste caso, a linha de comando correta seria apenas **esttab Modelo1 Modelo2 Modelo3**, junto com as demais opções disponíveis.

² Outras opções de formatos também podem ser usadas, como HTML ou txt.

```
cells(
  b (star fmt(3))
  se(par fmt(2))
)
```

A opção **cells** define o que vai aparecer na parte principal da tabela de resultados, como coeficientes, erros padrão ou a estatística t, além de definir o seu formato. É possível exibir tantas informações disponíveis quanto queira o usuário. Algumas informações disponíveis são:

- b: coeficientes estimados (ou efeitos marginais, se a opção **margin** for usada)
- se: erros padrão
- t: valor da estatística t
- p: p-valor
- ci: intervalo de confiança

Note que ainda existem sub-opções de formatação para cada uma destas informações:

star: ordena que apareçam os asteriscos (ou outro símbolo definido pelo usuário) de significância do lado da informação desejada. No caso do exemplo, os símbolos são exibidos ao lado dos coeficientes (**b**). Os símbolos podem ser definidos com a opção **starlevels**, comentada mais abaixo.

par: a informação será exibida entre parênteses. No exemplo, serão os erros padrões (se).

fmt(): define o número de casas decimais.

```
varlabels(
  _cons Constante
  weight "Peso do Carro"
  mpg "Km Rodados"
  foreign "Tipo de Carro"
)
```

Esta opção troca os nomes das variáveis por etiquetas definidas pelo usuário. Com isso, a variável *weight*, por exemplo, pode aparecer na tabela final com a etiqueta “*Peso do Carro*”. Se a etiqueta não for composta, com apenas uma palavra, as aspas são opcionais. A constante tem como nome original *_cons*. Uma alternativa a esta configuração é o uso da opção **label**, caso o usuário tenha etiquetado anteriormente as variáveis usando o comando **label variable**.

```
stats(
  r2_a N F p,
  fmt(4 0 2 2)
  labels("R2 Ajustado" "Número de obs." F "Prob>F")
)
```

Com esta opção são definidas as informações adicionais da tabela de resultados, que aparecem na parte inferior, como o número de observações ou o R^2 . Na primeira parte da opção devem ser informados os dados que se deseja exibir. Estas informações são baseadas no grupo de macros `e()` gerado por cada comando de regressão, que armazenam uma série de constantes sobre os resultados estimados. Para ver a lista completa de macros armazenadas, basta conferir no help de cada comando, geralmente na parte final, no tópico *command saves the following in e()*. No exemplo, são usadas as macros armazenadas pelo comando `reg` referentes ao R^2 ajustado (`r2_a`), o número de observações (`N`), o valor da estatística F (`F`) e o p-valor do teste F (`p`).

As sub-opções formatam a apresentação destas estatísticas:

`fmt()`: define o número de casas decimais, na mesma seqüência, dos dados apresentados. No exemplo, o R^2 ajustado fica com duas casas, o número de observações sem nenhuma, o valor F com duas e o p-valor F também com duas (4 0 2 2).

`labels()`: fornece etiquetas para as informações. Devem ser usadas aspas se a etiqueta tiver mais de uma palavra.

`collabels(none)`

Evita que uma terceira linha de títulos seja adicionada na tabela.

`starlevels(*** 0.10 ** 0.05 * 0.01)`

Define os símbolos que serão usados para cada nível de significância dos coeficientes individuais.

`stardetach`

Esta opção separa os asterísticos (ou símbolos definidos pelo usuário) em uma coluna distinta dos coeficientes.

`legend`

`note ("Erros padrão entre parênteses")`

Estas duas opções definem as informações adicionais, já fora da tabela de resultados, como comentários adicionais e a legenda de símbolos de significância.

Abaixo, segue toda a seqüência de comandos usada no exemplo, com alguns comentários. O usuário pode copiá-la para um arquivo **do-file** e editar da forma que desejar.

```

sysuse auto
replace price = price / 1000
replace weight = weight / 1000

* Estimando os modelos de exemplo
reg price weight mpg foreign, r
estimates store Modelo1 /*armazena na memória o modelo estimado com a */
reg price weight, r /*etiqueta Modelo1 */
estimates store Modelo2
reg price mpg, r
estimates store Modelo3

* Comando esttab com comentários
#d;
esttab Modelo1 Modelo2 Modelo3 using resultado.rtf,
replace /* ou append, caso deseja acrescentar e não substituir*/

cells( /* Define o que vai aparecer no output principal*/
    b (star fmt(3))
    se(par fmt(2))
)

varlabels( /* Define etiquetas para as variáveis*/
    _cons Constante
    weight "Peso do Carro"
    mpg "Km Rodados"
    foreign "Tipo de Carro"
)

label /* Define que o comando usará as etiquetas pré definidas
para as variáveis */

stats( /* Define as informações que apareceram na parte de baixo da
tabela */
    r2_a N F p,
    fmt(4 0 2 )
    labels("R2 Ajustado" "Número de obs." F "Prob>F")
)

collabels(none) /*Dá um nome para as colunas de modelos - none significa
sem nome*/

starlevels(*** 0.10 ** 0.05 * 0.01) /*define os níveis de sig e os
símbolos */

stardetach /* separa os asteriscos do coeficiente em outra coluna */

legend
note ("Erros padrão entre parênteses")
;

* OUTRAS OPÇÕES:
* style(): define um estilo de apresentação. Se a tabela ficar fora de
formato, use style(fixed)
* delimiter (";") /*facilita a exportação para o excel, colocando o
delimitador */
* margin: exhibe os efeitos marginais. Só é válido se foi usado o comando
mxf depois de cada modelo estimado.

```

Referências bibliográficas

Jann, B. (2004): "ESTOUT: Stata module to make regression tables," Statistical Software Components S439301, Boston College Department of Economics, revised 12 Oct 2009.