The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$ 98	1-year subscription	$138
2-year subscription	$165	2-year subscription	$245
3-year subscription	$225	3-year subscription	$345
1-year student subscription	$ 75	1-year student subscription	$ 99
1-year university library subscription	$125	1-year university library subscription	$165
2-year university library subscription	$215	2-year university library subscription	$295
3-year university library subscription	$315	3-year university library subscription	$435
1-year institutional subscription	$245	1-year institutional subscription	$285
2-year institutional subscription	$445	2-year institutional subscription	$525
3-year institutional subscription	$645	3-year institutional subscription	$765
Electronic only		**Electronic only**	
1-year subscription	$ 75	1-year subscription	$ 75
2-year subscription	$125	2-year subscription	$125
3-year subscription	$165	3-year subscription	$165
1-year student subscription	$ 45	1-year student subscription	$ 45

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Stata, Stata Press, Mata, Mata, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2013)
13, Number 2, pp. 315–322

Importing U.S. exchange rate data from the Federal Reserve and standardizing country names across datasets

Betul Dicle
New Orleans, LA
bkdicle@gmail.com

John Levendis
Loyola University New Orleans
New Orleans, LA
jlevendi@loyno.edu

Mehmet F. Dicle
Loyola University New Orleans
New Orleans, LA
mfdicle@loyno.edu

Abstract. `fxrates` is a command to import historical U.S. exchange rate data from the Federal Reserve and to calculate the daily change of the exchange rates. Because many cross-country datasets use different spellings and conventions for country names, we also introduce a second command, `countrynames`, to convert country names to a common naming standard.

Keywords: dm0069, fxrates, countrynames, exchange rates, country names, standardization, data management, historical data

1 Introduction

Economic and financial researchers must often convert between currencies to facilitate cross-country comparisons. We provide a command, `fxrates`, that downloads daily foreign exchange rates relative to the U.S. dollar from the Federal Reserve's database.

Working with multiple cross-country datasets, such as international foreign exchange rates, introduces a unique problem: variations in country names. They are often spelled differently or follow different grammatical conventions across datasets. For example, North Korea is often different among datasets; it could be "North Korea", "Korea, North", "Korea, Democratic People's Republic", or even "Korea, DPR". Likewise, "United States of America" is often "United States", "USA", "U.S.A.", "U.S.", or "US". A dataset may have country names in all caps. Country names could also have inadvertent leading or trailing spaces. Thus we provide a second command, `countrynames`, that renames many country names to follow a standard convention. The command is, of course, editable, so researchers may opt to use their own naming preferences.

2 The fxrates command

2.1 Syntax

`fxrates` [*namelist*] [`, period(2000|1999|1989) chg(ln|per|sper)`
`save(`*filename*`)`]

2.2 Options

namelist is a list of country abbreviations for the countries whose foreign exchange data you wish to download from the Federal Reserve's website. Exchange rates for all available countries will be downloaded if *namelist* is omitted. The list of countries includes the following:

`al`	Australia	`ma`	Malaysia
`au`	Austria	`mx`	Mexico
`be`	Belgium	`ne`	Netherlands
`bz`	Brazil	`nz`	New Zealand
`ca`	Canada	`no`	Norway
`ch`	China, P.R.	`po`	Portugal
`dn`	Denmark	`si`	Singapore
`eu`	Economic and Monetary Union member countries	`sf`	South Africa
`ec`	European Union	`ko`	South Korea
`fn`	Finland	`sp`	Spain
`fr`	France	`sl`	Sri Lanka
`ge`	Germany	`sd`	Sweden
`gr`	Greece	`sz`	Switzerland
`hk`	Hong Kong	`ta`	Taiwan
`in`	India	`th`	Thailand
`ir`	Ireland	`uk`	United Kingdom
`it`	Italy	`ve`	Venezuela
`ja`	Japan		

`period(2000|1999|1989)` specifies which block of dates to download. The Federal Reserve foreign exchange database is separated into three blocks: one ending in 1989, a second for 1990–1999, and a third for 2000 through the present. The default (obtained by omitting `period()`) is to download the three separate files and merge them automatically so that the user has all foreign exchange market data available. You can specify one or more periods. If you know which data range you wish to download, however, you can save time by specifying which of the three blocks to download. Specifying all three periods is equivalent to the default of downloading all the data.

`chg(ln|per|sper)` is the periodic return. Three different percent changes can be calculated for the adjusted closing price: natural log difference, percentage change, and symmetrical percentage change. Whenever one of these is specified, a new variable is created with the appropriate prefix: `ln` for the first-difference of logs method, `per` for the percent change, and `sper` for the symmetric percent change.

`save(`*filename*`)` is the output filename. *filename* is created under the current working directory.

2.3 Using fxrates to import historical exchange rate data

▷ Example

In this example, we use `fxrates` to import the entire daily exchange rate dataset from the Federal Reserve. Because we did not specify the countries, `fxrates` downloads data from all countries. Because we did not specify the period, `fxrates` defaults to downloading data for all available dates.

```
. fxrates
au does not have 00
be does not have 00
  (output omitted)
ve does not have 89
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
date	10551	11403.8	4264.338	4018	18788
_al	10145	.8764938	.2391958	.4828	1.4885
_au	7013	15.21975	3.999031	9.5381	26.0752
_be	7021	38.61327	8.036983	27.12	69.6
_bz	4134	1.944	.6851041	.832	3.945
_ca	10158	1.227691	.1689277	.9168	1.6128
_ch	7592	6.110467	2.380834	1.5264	8.7409
_dn	10151	6.702978	1.32041	4.6605	12.3725
_eu	3131	1.197019	.1959648	.827	1.601
_ec	4902	1.137739	.1776148	.6476	1.4557
(*output omitted*)					
_sd	10151	6.633961	1.631	3.867	11.027
_sz	10152	1.796176	.7191334	.8352	4.318
_ta	6665	31.36318	4.002814	24.507	40.6
_th	7571	30.99084	7.293934	20.36	56.1
_uk	10152	1.779829	.3176284	1.052	2.644
_ve	4127	1.528076	1.111678	.1697	4.3

Output such as `au does not have 00` indicates that there were no observations in a particular block of years (in this case, the 2000–present block) for the particular country. When this appears, it is most often the case that the currency has been discontinued, as when Austria started using the euro.

◁

▷ Example

In this second example, we download the exchange rates of the U.S. dollar versus the French franc, the German deutschmark, and the Hong Kong dollar for all the available dates.

```
. fxrates fr ge hk
fr does not have 00
ge does not have 00
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
date	10550	11404.5	4263.934	4021	18788
_fr	7021	5.673864	1.227564	3.8462	10.56
_ge	7021	2.143872	.5509681	1.3565	3.645
_hk	7652	7.63056	.5069678	5.127	8.7

◁

▷ Example

In this example, we download the exchange rate data for United States versus France, Germany, and Hong Kong. Because no period was specified, `fxrates` downloads the data from all available dates. We also specified that `fxrates` calculate the daily percent change, calculated in two different ways: as the log first-difference and as the arithmetic daily percent change. The log-difference percent change for each country is prefixed by `ln`; the arithmetic percent change for each country is prefixed by `per`.

```
. fxrates fr ge hk, chg(ln per)
fr does not have 00
ge does not have 00
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
date	10550	11404.5	4263.934	4021	18788
_fr	7021	5.673864	1.227564	3.8462	10.56
_ge	7021	2.143872	.5509681	1.3565	3.645
_hk	7652	7.63056	.5069678	5.127	8.7
ln_fr	6743	-.0000122	.0061762	-.0416059	.0587457
per_fr	6743	6.85e-06	.0061803	-.0407522	.0605055
ln_ge	6743	-.0001283	.0064045	-.0414075	.0586776
per_ge	6743	-.0001078	.0064049	-.0405619	.0604333
ln_hk	7363	.000054	.0023756	-.0410614	.0653051
per_hk	7363	.0000568	.0023914	-.0402298	.0674847

◁

Example

In this final example, we download the U.S. dollar exchange rate versus the Japanese yen and the Mexican peso. We calculate the daily percent change by calculating the first-differences of natural logs for the data ending in 1999 (that is, for the data ending in 1989 plus the data from 1990 through 1999).

```
. fxrates ja mx, period(1999 1989) chg(ln)
mx does not have 89
. summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
date	7565	9315	3057.56	4021	14609
_ja	7267	195.5763	74.42725	81.12	358.44
_mx	1541	7.258319	2.154108	3.1	10.63
ln_ja	6980	-.0001587	.0063255	-.056302	.0625558
ln_mx	1477	.0005535	.0132652	-.1796934	.1926843

3 The countrynames command

3.1 Syntax

`countrynames` *countryvar*

3.2 Description

The command `countrynames` changes the name of a country in a dataset to correspond to a more standard set of names. By default, `countrynames` creates a new variable, `_changed`, containing numeric codes that indicate which country names have been changed. A code of `0` indicates no change; a code of `1` indicates that the country's name has been changed. We recommend you run `countrynames` on both datasets whenever two different cross-country datasets are being merged. This minimizes the chance that a difference in names between datasets will prevent a proper merge from occurring. However, if you wish to keep a variable with the original names, you need to copy the variable to another variable. For example, before running `countrynames` *country*, you would need to type `generate` *origcountry* `=` *country*.

3.3 Using the countrynames command to convert country names to a common naming standard

▷ Example

In this example, we use two macroeconomic datasets that have countries named slightly differently. The first dataset is native to and shipped with Stata.

```
. sysuse educ99gdp, clear
(Education and GDP)
```

Though the dataset is very small, it suffices for our purposes. Notice the spelling of `United States` in this dataset.

```
. list
```

	country	public	private
1.	Australia	.7	.7
2.	Britain	.7	.4
3.	Canada	1.5	.9
4.	Denmark	1.5	.1
5.	France	.9	.4
6.	Germany	.9	.2
7.	Ireland	1.1	.3
8.	Netherlands	1	.4
9.	Sweden	1.5	.2
10.	United States	1.1	1.2

```
. save temp1.dta, replace
(note: file temp1.dta not found)
file temp1.dta saved
```

In fact, all the spellings in this dataset correspond with the preferred names listed in `countrynames`, so nothing is required of us here. We could run `countrynames` just to be on the safe side, but it would not have any effect. It is, however, good practice to run `countrynames` whenever merging datasets to maximize the chances that the two datasets use the same country names.

The second dataset, using World Health Organization data, is from Kohler and Kreuter (2005). The data are available from the Stata website.

```
. net from http://www.stata-press.com/data/kk/
  (output omitted)
. net get data
  (output omitted)
. use who2001.dta, clear
```

Notice how the United States is called `United States of America` in this dataset.

```
. list country
```

	country
1.	Afghanistan
2.	Albania
	(*output omitted*)
180.	United States of America
	(*output omitted*)
187.	Zambia
188.	Zimbabwe

We now run `countrynames` on this dataset to standardize the names of the countries. This will rename `United States of America` to `United States`, as it was in the first dataset.

```
. countrynames country
. list country _changed
```

	country	_changed
1.	Afghanistan	0
2.	Albania	0
	(*output omitted*)	
180.	United States	1
	(*output omitted*)	
187.	Zambia	0
188.	Zimbabwe	0

Notice that the generated variable, `_changed`, is equal to `1` for the `United States` entry; this indicates that its name was once something different.

Having run `countrynames` on both datasets, we have increased the chances that countries in both datasets follow the same naming convention. We are now safe to merge the datasets:

```
. drop _changed
. sort country
. merge 1:1 country using temp1.dta
```

Result	# of obs.	
not matched	180	
from master	179	(_merge==1)
from using	1	(_merge==2)
matched	9	(_merge==3)

The merge results table above is important: It is the result of merging a dataset that used the `countrynames` command (master: `who2001.dta`) with a dataset that did not use the `countrynames` command (using: `temp1.dta`). If the dataset using the command includes a country name that is not renamed with `countrynames`, then it will appear in the merge results table.

```
. sort country
. list country
```

	country
1.	Afghanistan
2.	Albania
	(*output omitted*)
180.	United Arab Emirates
	(*output omitted*)
188.	Zambia
189.	Zimbabwe

◁

3.4 How to edit preferred country names within the countrynames command

It is possible to add, remove, or change country name entries within the `countrynames` command. After opening the `countrynames.ado` file with a do-file editor (any text editor), you can delete country name entries, add new entries, or change spellings according to your preferences. Any changes made to the `countrynames.ado` file should be saved. The `discard` command will refresh the `countrynames` command with your updates along with all the ado installations to Stata. We recommend that you confirm updates to the `countrynames` command with a merge table.[1]

4 Reference

Kohler, U., and F. Kreuter. 2005. *Data Analysis Using Stata*. College Station, TX: Stata Press.

About the authors

Betul Dicle recently earned her PhD from the political science department of Louisiana State University.

John Levendis is an assistant professor of economics at Loyola University New Orleans.

Mehmet F. Dicle is an assistant professor of finance at Loyola University New Orleans.

1. Please note that if we ever do an update to our program, the user edits to the ado-file will be lost when users grab the updated ado-file.