STaTa®
Statistics/Data Analysis

_____

help for *taxsim9*
_____

NBER TAXSIM model for federal and state income taxes

Description

>   taxsim9[,full output secondary|wages|interest|long debug plan planval x51
>   replace]

calculates federal and state income tax liability from 21 input variables.
These variables are be defined in your current workspace before calling taxsim.
The program returns values for various tax return concepts, taxes, credits, and
marginal tax rates as new named variables in a new dataset or merged into your
current workspace.

The tax calculator itself is the same FORTRAN program that the NBER has been
updating annually since 1974, but with a simplified input file (no married
filing separate or income averaging, etc). This interface merely packages up
your data and sends it to Cambridge for the actual calculations, which are
returned and merged into your current dataset. No tax calculations are done on
your computer, although there is some range checking of your data.

Input Variables:

year: 4-digit year between 1960 and 2023. Between 1977 and 2015 if state tax is
requested. No default.

state: state identifier. These are numeric codes from 1 to 51 (Alabama to
Wyoming in alphabetical order) with zero indicating "no state tax".

mstat: Marital status 1 for single, 2 for joint, 3 for head of household. No
default.

depx: Number of dependents (usually kids, but can be any age) per tax form.
Must not be less than depchild (below).

agex: The age of the primary taxpayer times 100 plus the age of the secondary
taxpayer (if any). Forexample 6563 would a 65 year old married to a 63 year old.
If zero we assume the taxpayer is of an age eligible for the EITC and the full
AMT exclusion but not for any age exemption or supplemental standard deduction.
It is not an error for a taxpayer to be over 99 years of age but the secondary
taxpayer must be of legal marriagable age. 10301 would be a 102 year old married
to a 101 year old. Prior to July 2016 this field was the number of taxpayers 65
years of age or more, and that usage is still accepted for backward
compatibility, however EITC, and AMT are more accurate with exact ages.

pwages: Wage income of primary taxpayer

swages: Wage income of secondary taxpayer

dividends: Dividend income

otherprop: Interest and other property income. This may be negative. You can put
net alimony here, and subtract any adjustments such as IRAs, Keoghs and self
employment tax, to the extent that yuo may know such items.

pensions: Taxable pension income

gssi: Gross social security benefits. Taxsim will calculate the portion
includable in AGI.

transfers: Non-taxable transfer income, used for calculating low income credits
and property tax credits on state taxes.

rentpaid: Rent paid is used in some state property tax credits.

proptax: Property tax paid is an itemized deduction and is a preference for the
AMT.

otheritem: Taxes paid other than state income taxes. Taxsim will use its own calculated state income tax as an itemized deduction. These are preferences for the AMT.

childcare: Child care expenses are a credit in the federal tax.

ui: Unemployment compensation benefits. Taxsim will calculate the portion included in AGI.

depchild: Number of dependent children under 17 (for child credit), optionally plus 100 times the number of dependent children under 18 plus 10000 times the number of dependent children under 13. e.g. 10302 for 1 child eligible for child care credit, three counted for the EIC and 2 for the child tax credit. Note that to maintain backward compatibility if this value is under 100 then the number of dependents is used for the child care and EIC credits.

mortgage: Mortgage interest paid, possibly plus charitable contibutions, and some minor items that are not preferences for the AMT.

stcg: Short term capital gain or loss (+/-).

ltcg: Long term capital gain or loss (+/-).

The following variables are returned by the procedure:

taxsimid: Record identifer, _n , used to merge taxsim_out.dta back to source workspace.

year: Unchanged from input.

state: Unchanged from input.

fiitax: Federal tax liability, after regular, minimum, and maximum tax, and after refundable credits.

siitax: State tax liability after refundable credits.

fica: FICA tax liability, sum of employer and employee shares.

frate: Marginal federal tax rate wrt primary worker wages. You can change this to the weighted average of the rates on the primary and secondary earners, (equal weights if both are non-workers), secondary earner, property income, or long term gains with the options, **Wages**,**Secondary**, **Interest** or **Long**.

srate: Marginal state tax rate. Same options apply here.

ficar: Marginal FICA rate. The estimate of FICA tax includes both employer and employee portions, and ignores the favorable treatment of self-employment income.

Missing Data:
All variables except year and marital status are optional, and the tax calculation will proceed as if missing variables are zero after printing a warning. A variable is missing if it doesn't show up in a describe command - this is different from missing values for a defined variable.

An individual case missing value for any variable in the list of 21 variables will result in listwise deletion. Zero is an acceptable value for any dollar amount. A zero for the state id, means "no state tax" and in that case zero is returned for the state tax liability. Negative values are allowed only for capital gains and **otherprop**, and will cause listwise deletion if encountered in other variables.

Additional Output:

Options

full: Return many intermediate calculations (taxable income, regular tax, individual credits, etc. All the returned variables have labels defined in taxsim9.ado - use "describe" to get the full list or browse the web page below.

output: Specify the name of the output dataset. The default is taxsim_out.dta in the current directory.

secondary: Calculate marginal tax rates with respect to the secondary wage earner. The default is with respect to the primary wage earner.

wages: Calculate marginal tax rates with respect to a weighted average of the primary and secondary wage earner.

interest: Calculate marginal tax rates with respect to interest income.

long: Calculate marginal tax rates with respect to long term gains.

debug: Do not erase intermediate ASCII files written in the current directory. Summarize input variables after cleaning. The intermediate files are essential for debugging. If you are reporting errors to NBER, be sure to create a small (no more than a couple of observations) test file to send to us.

x51: Users of the AHEAD, HRS, and other confidential data are reminded that they are not permitted to send state identifiers to us. If such users need state tax calculations they should set this option which instructs taxsim8.ado (running on your computer) to drop the state identifier from records sent to our server, and instructs the server to return 51 records for each record received, one for each possible state. Then taxsim9.ado (at your computer) will discard the 50 spurious records from the returned data, leaving you with tax calculations for the correct state for each taxpayer, and without sending us anything you aren't allowed to share. Don't laugh - it works and the inefficiency is tolerable.

replace: Merge the taxsim output file into the workspace, possibly overwriting existing variables. Otherwise output is saved to taxsim_out.dta in the current directory. You can run taxsim multiple times with the replace option set and the results will not change unless you change the input data or procedure options.

Plan: Modify the tax calculation according to the list of alternative tax laws listed at "http://www.nber.org/taxsim/taxsim=calc9/plans.html".

PLANVar: Parameter for the **Plan:**.

Notes:

Dollar amounts are rounded to the nearest penny before transmission to the calculator, and calculated amounts are similarly treated.

Extreme marginal tax rates due to notches will become missing values.

The maximum upload file size is vague, but use caution for files larger than 100,000 cases. Start small, and test the server reaction to larger files.

Taxsim9.ado writes out the 21 variables (plus a record id) to a temporary ASCII file and also writes a short ftp script for uploading that file to the NBER taxsim server. Then the same script downloads a file of the same name with the .taxsim suffix added. Our ftp server is modified to run such files through the taxsim program, and return the calculated results rather than the original file. The results are copied to your local hard disk and loaded into the taxsim_out.dta dataset in your home directory. If you wish to examine the process, study the .ado file and ftp.txt files.

Throughput is variable but one user reports 1,500 records/minute with an excellent connection (25 milliseconds ping time). If you experience time-outs, please split your dataset into pieces for sequential processing, if processing is reliable, you may consider larger pieces. I would be interested in hearing reports.

Example:
```
. input state year mstat pwages

. 5 2004 1 50000

. end
```

. taxsim9, full interest replace

The example posits a single taxpayer in California with $50,000 of wage income, and returns the marginal tax rate with respect to interest income, among other values.

Support:

Dan Feenberg would like to speak with anyone using this program - give him a call after you have done some experimenting, but before investing a lot of time and certainly before submitting for publication. If you encounter bugs please prepare a test case with no more than 3 observations and run with the following settings:

. log using foo,text

and also with trace on:

. set trace on

and with the debug option for taxsim:

. taxsim9,debug

set and the files txpydata.raw,results.raw, ftp.txt and msg.txt included with your message. If there is no results.raw file, that probably means that ftp wasn't found or didn't work.

Be sure to provide stata test runs with the smallest possible dataset that demonstrates the problem (preferably one record), run with the debug flag set, and attach the debug files. But always follow the instructions below.

If any tax calculation surprises you, please enter the taxpayer data in the web form mentioned below. A detailed explanation of the calculation will be returned, which may remind you of a code feature or interaction you were unaware of. If you still think the calculation is wrong, follow the instructions given there for sending bug reports. We are very responsive to these communications - they do result in improvements to the calculator, generally within a day or two. But please send us a single record demonstrating the error, not a large file. Most users anticipate we need to be encouraged to respond by the possibility that the error is common and submit large files to demonstrate that. This is not true and we will respond with a request that the file be cut down.

See http://www.nber.org/taxsim/taxsim9/ the TAXSIM web site for additional information about the tax caculator. Although that is the interface for the Web version of taxsim, there is much valuable information there about the operation of the program and the calculation of marginal rates which may surprise you.

A general description of Taxsim is given in http://www.nber.org/taxsim/feenberg-coutts.pdf.

Daniel Feenberg (tax questions)
feenberg@nber.org
617-863-0343

Online: help for taxsim9

Web: http://www.nber.org/taxsim/taxsim-calc9