Weight Calibration across Packages

Stas Kolenikov

9/23/2019

Weight calibration

- Last step in creating analysis weights in survey data files
- Adjusting the weights so that they sum to known population totals in different subgroups (age, sex, race, ethnicity, geography, etc.)
- Desirable to minimize changes from the input weights (probability of selection, nonresponse adjustments, frame integration, etc.)

Deville & Sarndal (1992)

Contenders

Stata

- `ipfraking` (Kolenikov 2014, 2019)
- `svycal` (official Stata)
- `survwgt` (Winter 2002)
- `sreweight` (Pacifico 2014)

R

- `survey::calibrate()` (Lumley 2010)

SAS

- `rake_and_trim()` (Izrael, Battaglia, Hoaglin, Frankel, Ball, 2017)

Out of scope

- SUDAAN `PROC WGTADJ`, `PROC WGTADJX`
- Stata `ipfweight` (Bergmann 2011)
- R `library(ReGenesees)` (Zardetto 2015)
- R `library(ipfr)` (Ward, Macfarlane 2019)

Expectations

- Produce usable results
- Provide weight diagnostics
- Speed
- Fool proof

Running example

BOLD THINKERS DRIVING REAL-WORLD IMPACT

Running example

CPS 2018 March ASEC data

- estimate control totals based on 13353 adults in CA
- calibrate 8403 adults in TX on
  - sex
  - age (14 categories)
  - race/ethnicity (6 categories)
  - education (5 categories)
  - HH income (9 categories)
  - nativity (3 categories)
  - marital status (6 categories)
  - own vs. rent
  - metro area of TX (23 categories)

Tasks and tests

1. Straight raking
2. Raking with divergent population control totals
3. Raking with bounded weight adjustment ratios [0.3,3]
4. Raking with bounded weight values (2nd and 98th percentile of unrestricted distribution)
5. Linear calibration
6. Linear calibration with trimming
7. (Informative error expected) incorrect specification of control totals

Performance summary

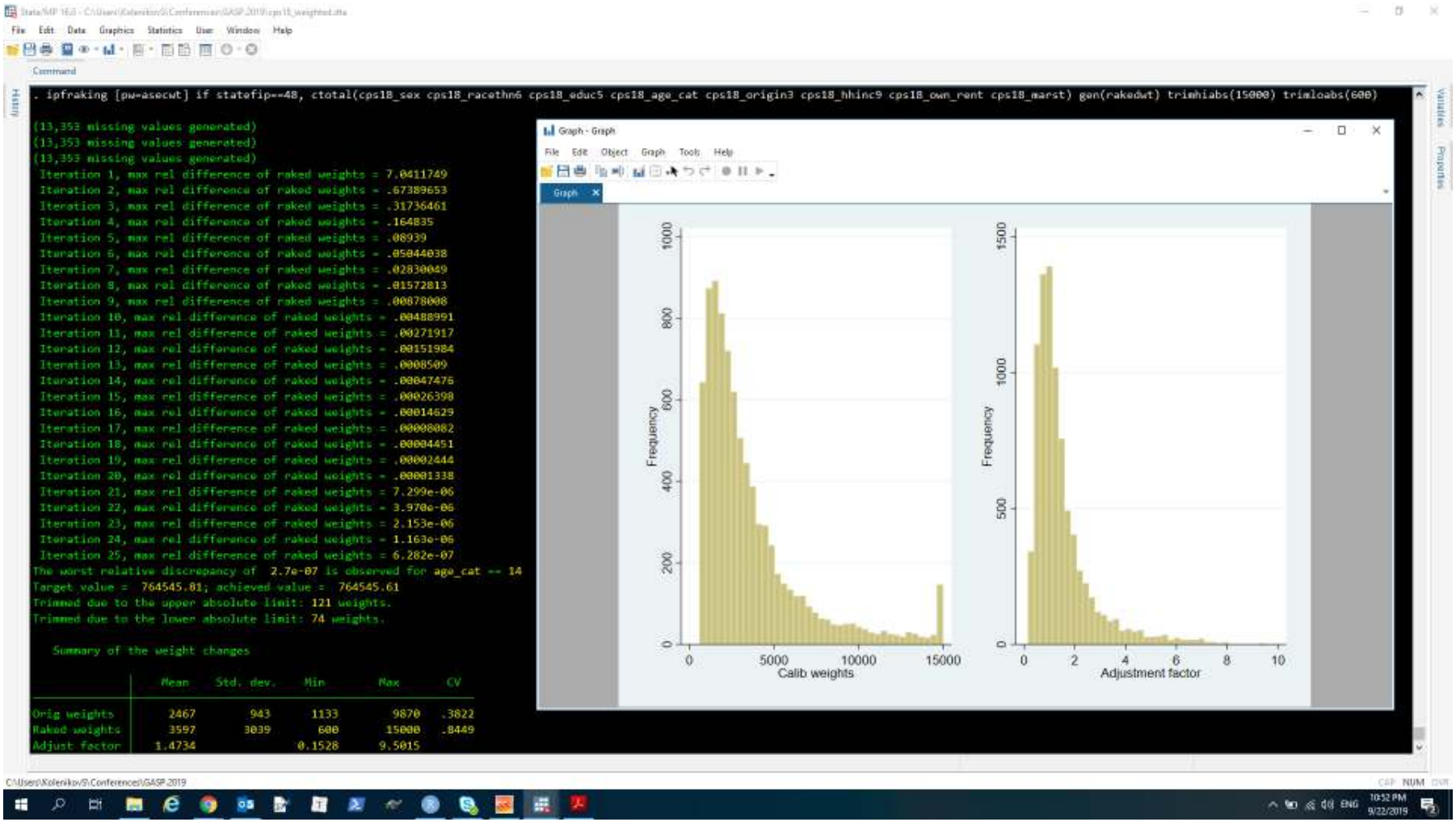
package	TOT	DIV	REL	ABS	LIN	LIN+TR	time
ipfraking	name	+W	+	+	+	N	7.14 sec
svycal	name/=	F	F	N	+	+	0.18 sec
survwgt	order	NW	N	N	N	N	0.80 sec
sreweight	order	F	F	N	+	N	0.19 sec
calibrate	name	-W	..	+	+	+	0.35 sec
rake_and_trim	name+magic	F	-W	+	N	N	61 sec

`N`: no documented functionality exists

`W`: issued reasonable warnings

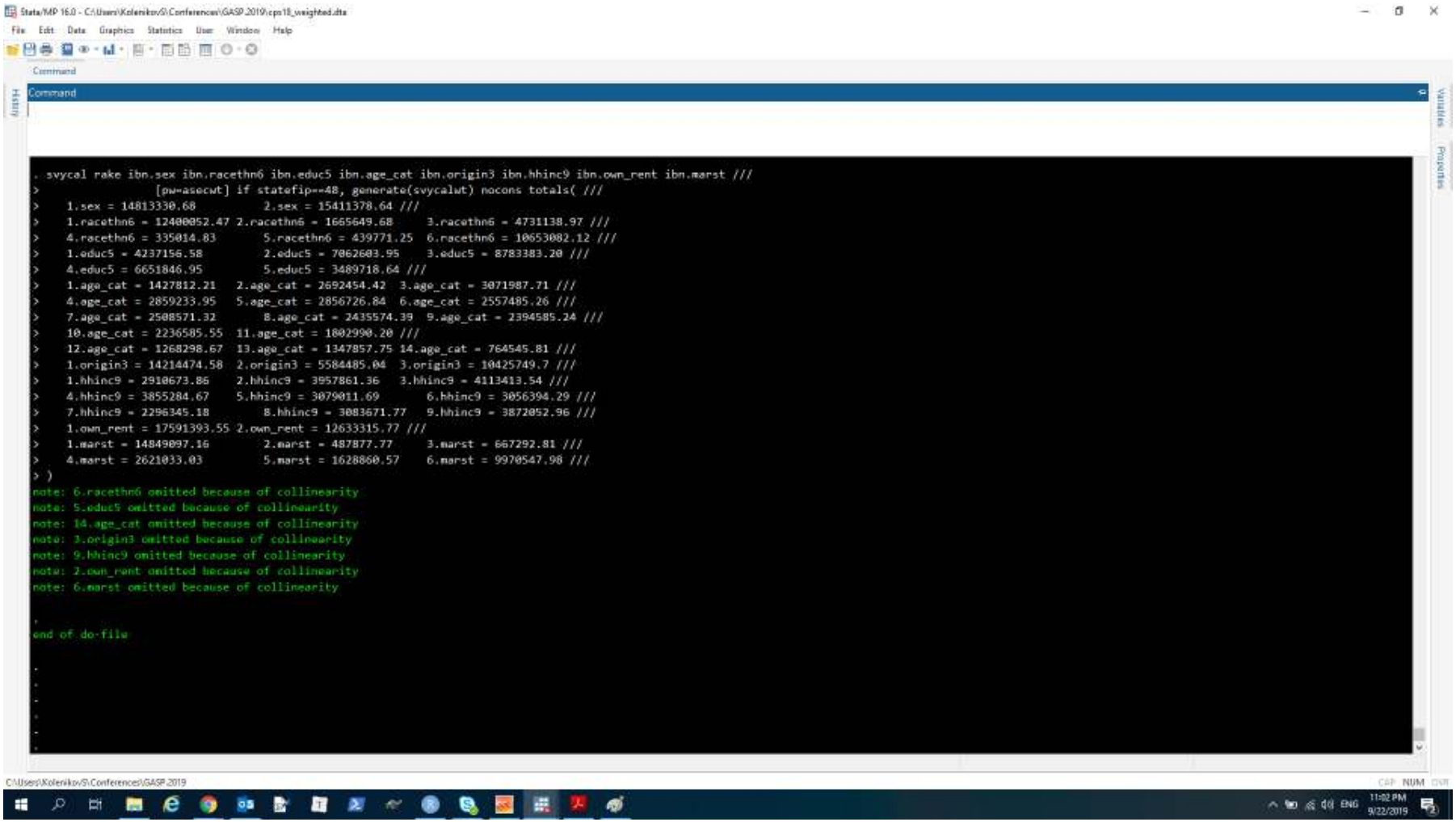
`F`: failed with cryptic error message / no message

Stata ipfraking



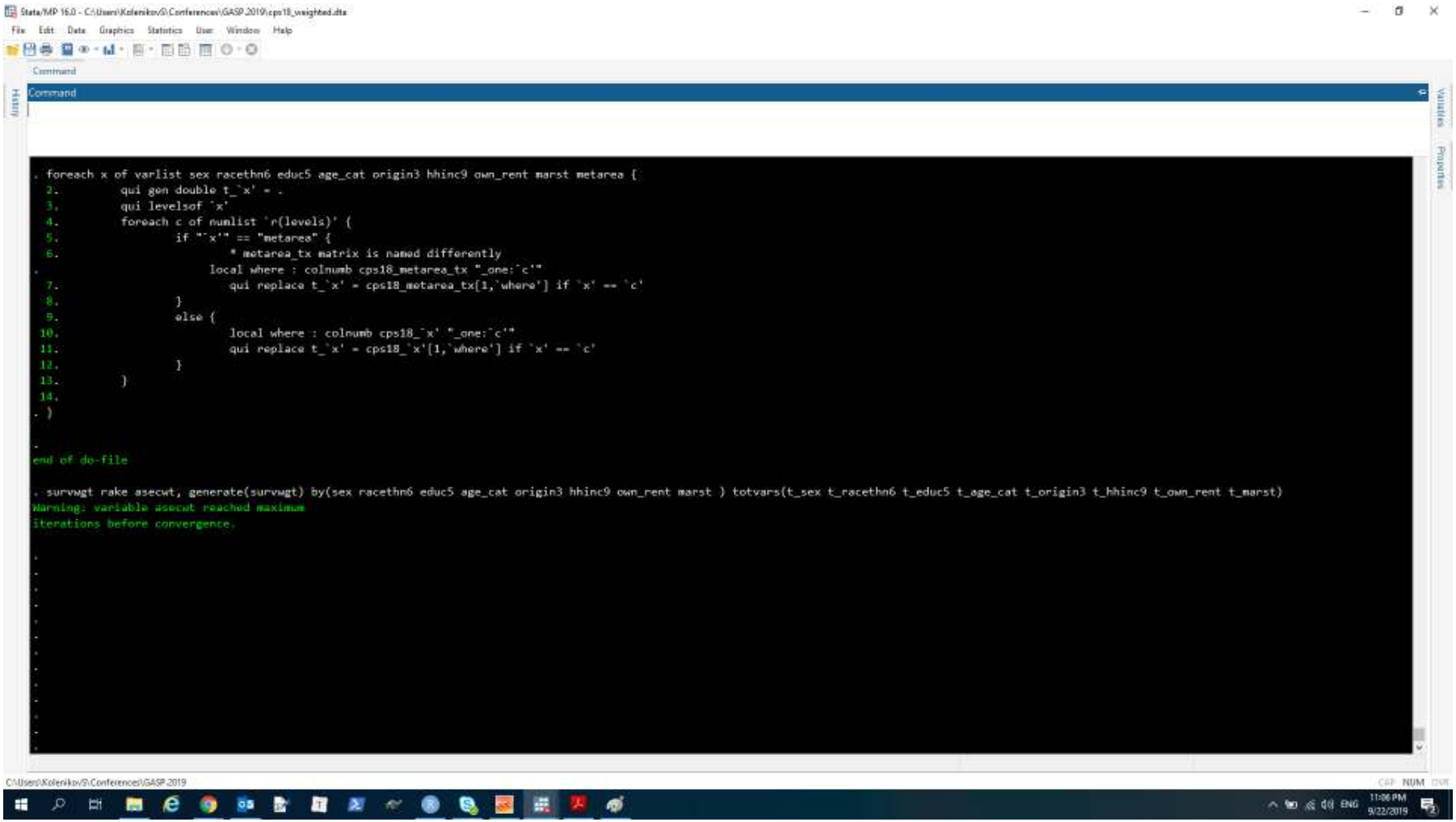
BOLD THINKERS DRIVING REAL-WORLD IMPACT

Stata svycal



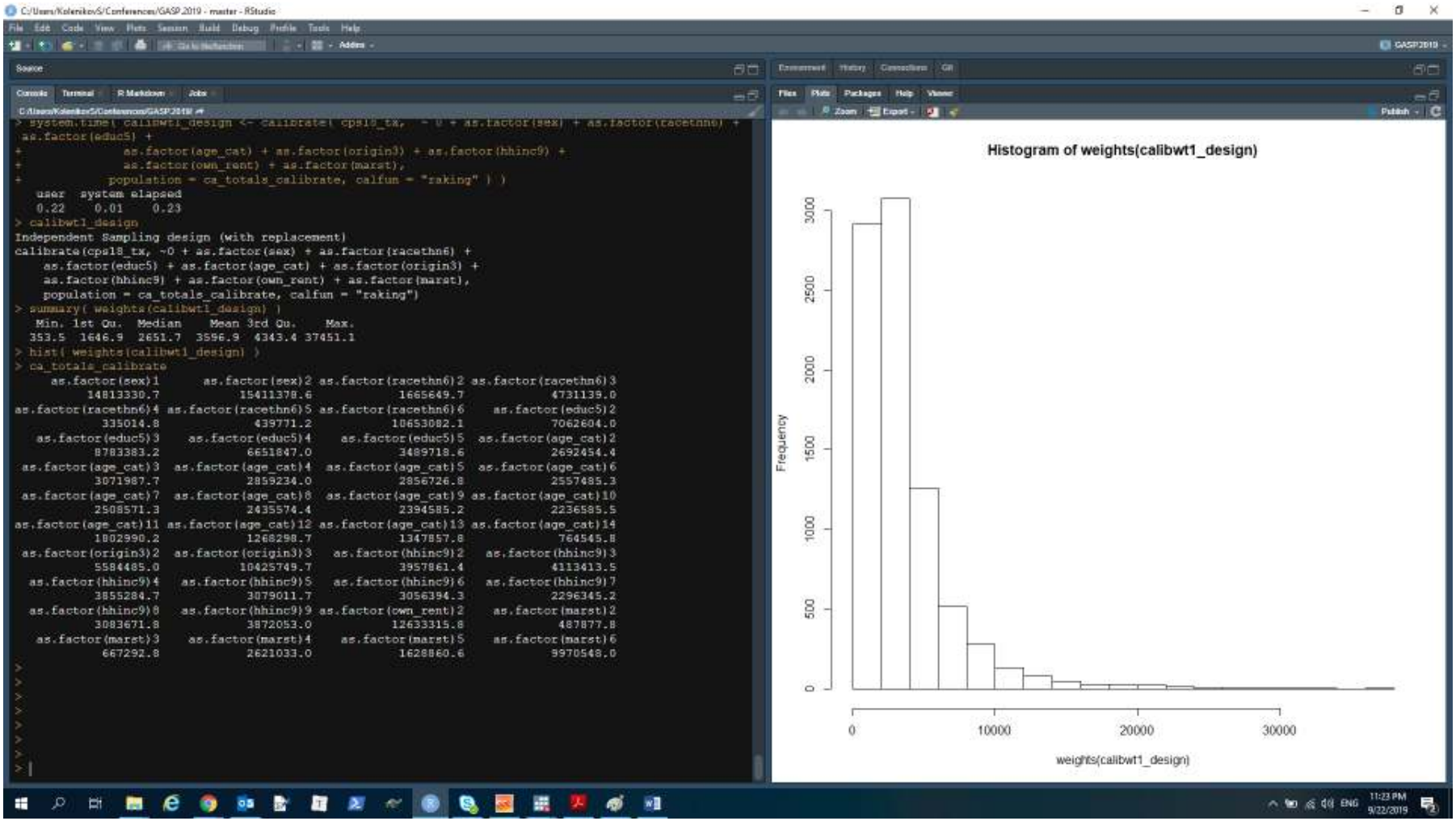
BOLD THINKERS DRIVING REAL-WORLD IMPACT

Stata **survwgt**



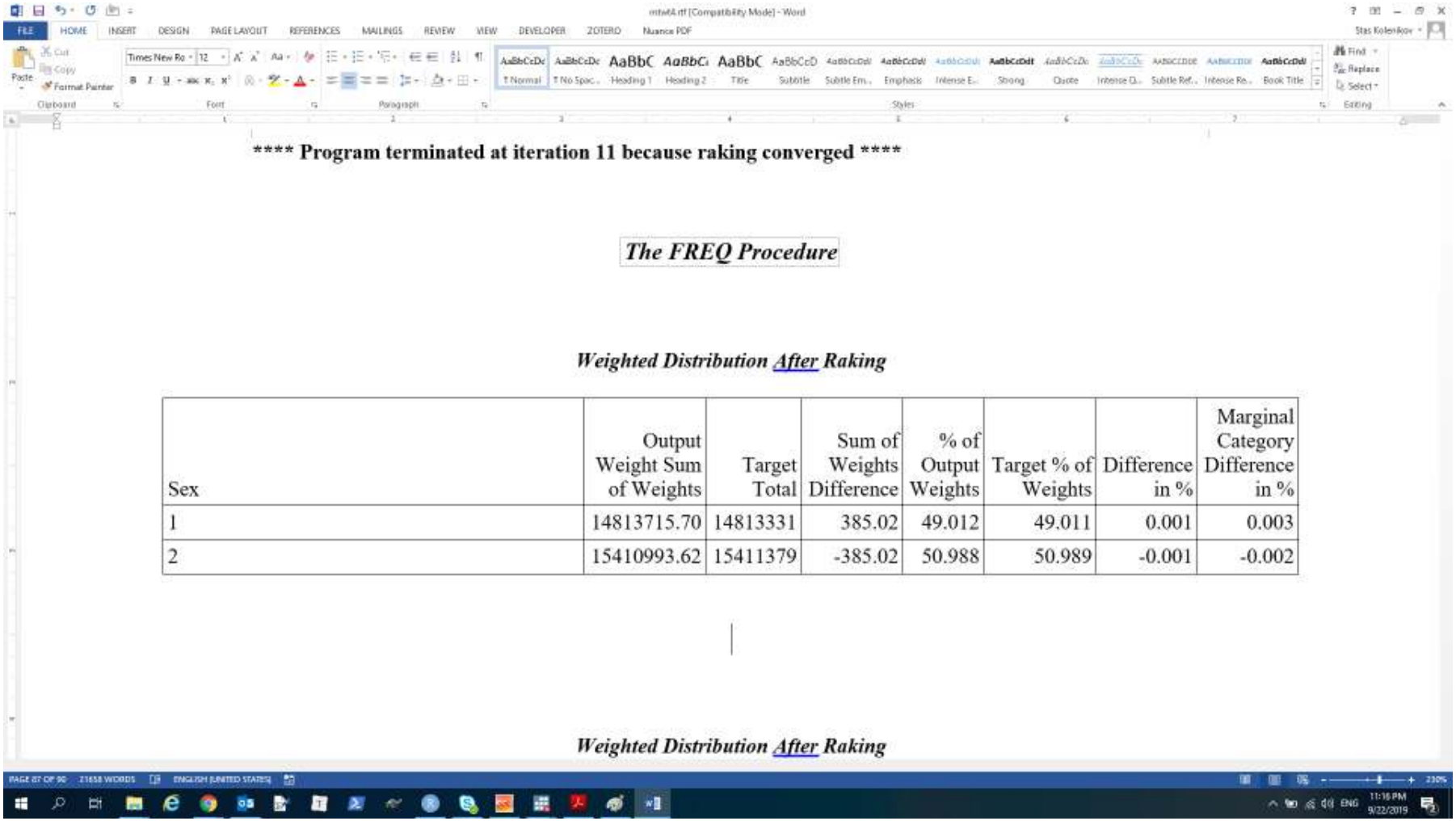
BOLD THINKERS DRIVING REAL-WORLD IMPACT

R survey::calibrate()



BOLD THINKERS DRIVING REAL-WORLD IMPACT

SAS rake_and_trim()



****** Program terminated at iteration 11 because raking converged ******

The FREQ Procedure

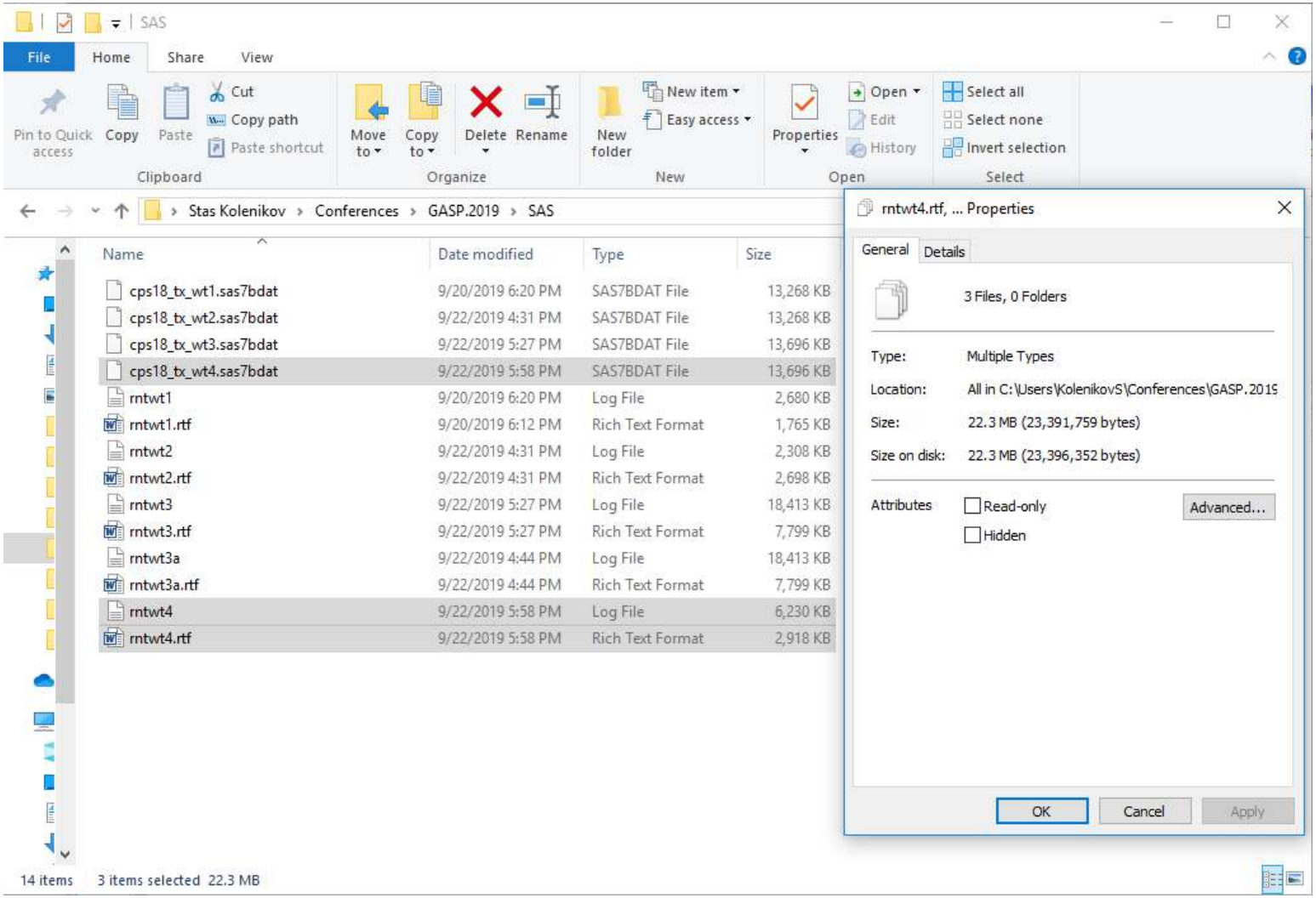
Weighted Distribution After Raking

Sex	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
1	14813715.70	14813331	385.02	49.012	49.011	0.001	0.003
2	15410993.62	15411379	-385.02	50.988	50.989	-0.001	-0.002

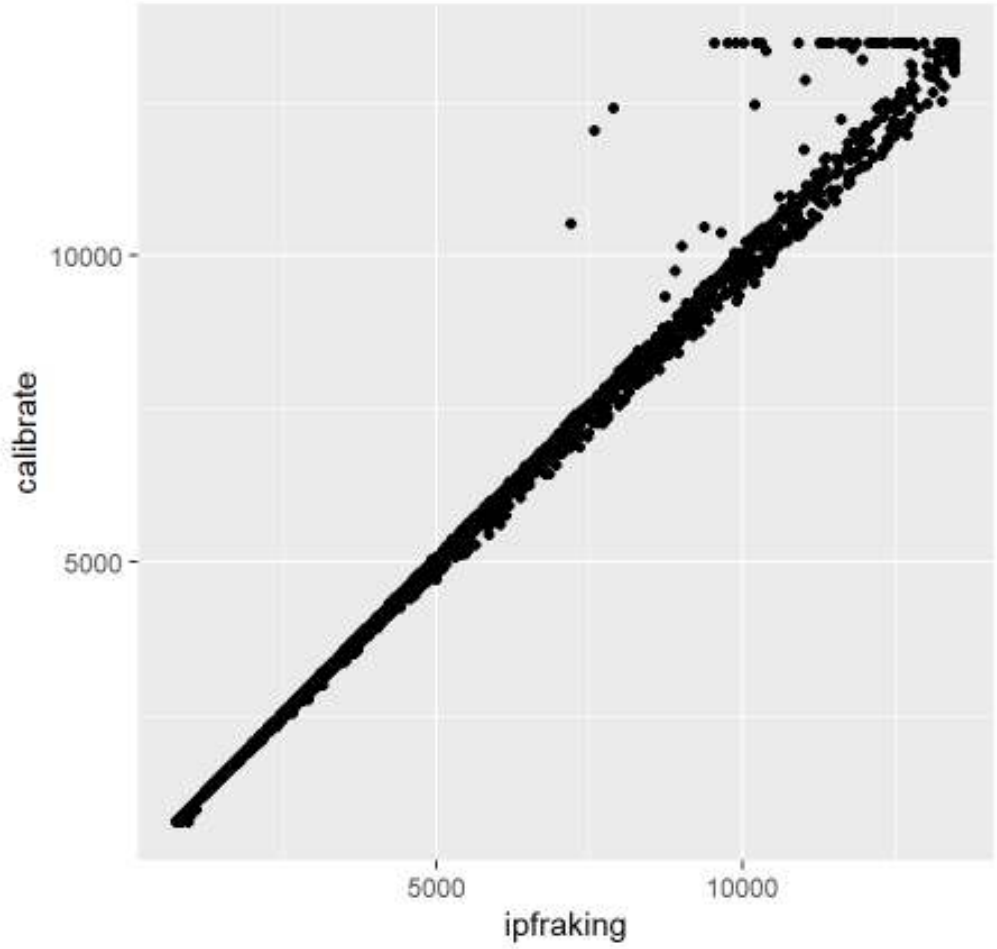
Weighted Distribution After Raking

BOLD THINKERS DRIVING REAL-WORLD IMPACT

SAS rake_and_trim()



Weight trimming ≠ methodology



BOLD THINKERS DRIVING REAL-WORLD IMPACT

Misspecified control totals

package	Extra in pop	Extra in data	Wrong order
ipfraking	E	E	
svycal	E	!!!	
survwgt	N/A	N/A	!!!
sreweight	E	E	!!!
calibrate	E	E	
rake_and_trim	!?!?	!!!	

E: issued an error and stopped

!!!: did not issue an error – results highly suspect!

Thanks and out

Questions?

- stas_kolenikov@abtassoc.com
- @StatStas on Twitter