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

In the original 2019 presentation,
there was a picture of a truck
with a bumper sticker that said,
"Don't CA my TX", but I was aksed
to remove that picture by a copyright
law firm representing the image holder.
https://www.google.com/search?q=don't+ca+my+tx

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

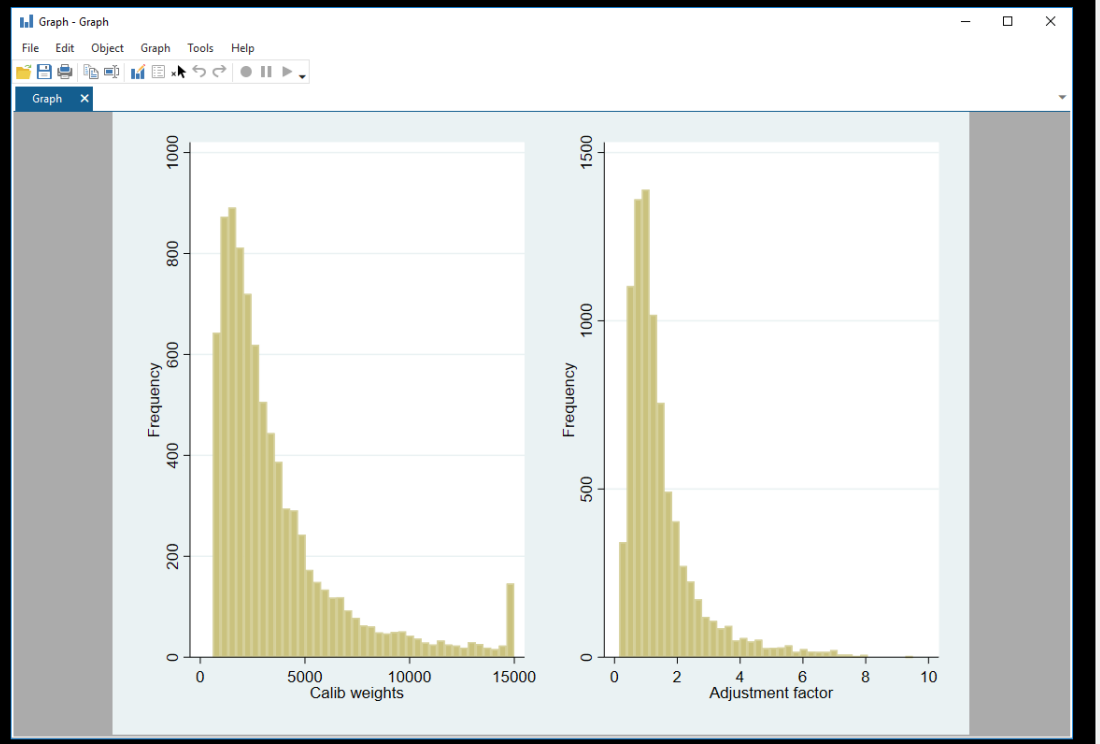
`N`: no documented functionality exists

`W`: issued reasonable warnings

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

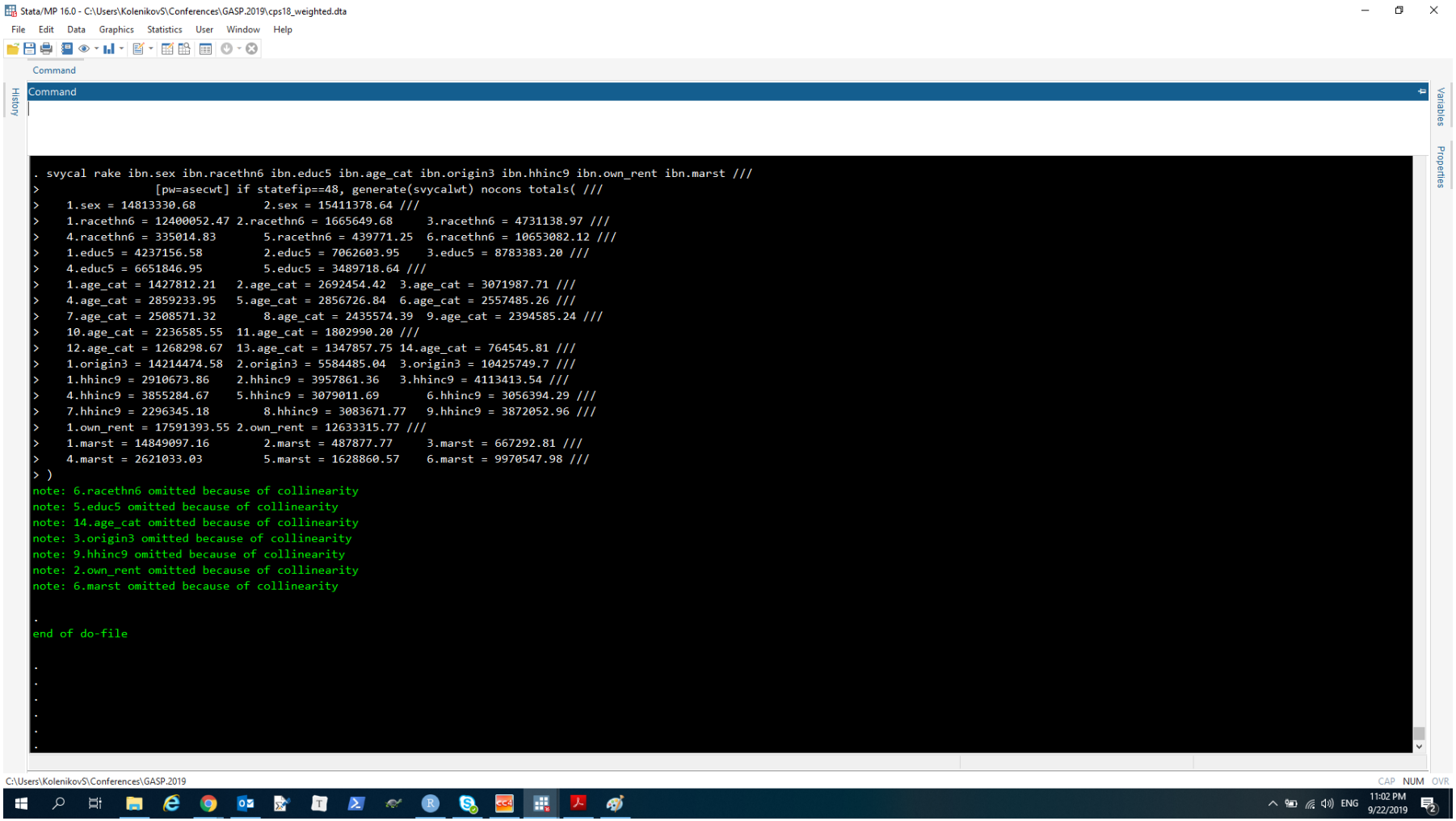
Stata ipfraking

```
. ipfraking [pw=asecwt] if statefip==48, ctotal(cps18_sex cps18_racethn6 cps18_educ5 cps18_age_cat cps18_origin3 cps18_hhinc9 cps18_own_rent cps18_marst) gen(rakedwt) trimhiabs(15000) trimloabs(600)

(13,353 missing values generated)
(13,353 missing values generated)
(13,353 missing values generated)
 Iteration 1, max rel difference of raked weights = 7.0411749
 Iteration 2, max rel difference of raked weights = .67389653
 Iteration 3, max rel difference of raked weights = .31736461
 Iteration 4, max rel difference of raked weights = .164835
 Iteration 5, max rel difference of raked weights = .08939
 Iteration 6, max rel difference of raked weights = .05044038
 Iteration 7, max rel difference of raked weights = .02830049
 Iteration 8, max rel difference of raked weights = .01572813
 Iteration 9, max rel difference of raked weights = .00878008
 Iteration 10, max rel difference of raked weights = .00488991
 Iteration 11, max rel difference of raked weights = .00271917
 Iteration 12, max rel difference of raked weights = .00151984
 Iteration 13, max rel difference of raked weights = .0008509
 Iteration 14, max rel difference of raked weights = .00047476
 Iteration 15, max rel difference of raked weights = .00026398
 Iteration 16, max rel difference of raked weights = .00014629
 Iteration 17, max rel difference of raked weights = .00008082
 Iteration 18, max rel difference of raked weights = .00004451
 Iteration 19, max rel difference of raked weights = .00002444
 Iteration 20, max rel difference of raked weights = .00001338
 Iteration 21, max rel difference of raked weights = 7.299e-06
 Iteration 22, max rel difference of raked weights = 3.970e-06
 Iteration 23, max rel difference of raked weights = 2.153e-06
 Iteration 24, max rel difference of raked weights = 1.163e-06
 Iteration 25, max rel difference of raked weights = 6.282e-07
The worst relative discrepancy of  2.7e-07 is observed for age_cat == 14
Target value =  764545.81; achieved value =  764545.61
Trimmed due to the upper absolute limit: 121 weights.
Trimmed due to the lower absolute limit: 74 weights.

   Summary of the weight changes

                |      Mean   Std. dev.       Min        Max         CV
----------------+--------------------------------------------------------
Orig weights    |      2467        943       1133       9870       .3822
Raked weights   |      3597       3039        600      15000       .8449
Adjust factor   |    1.4734                 0.1528     9.5015
```

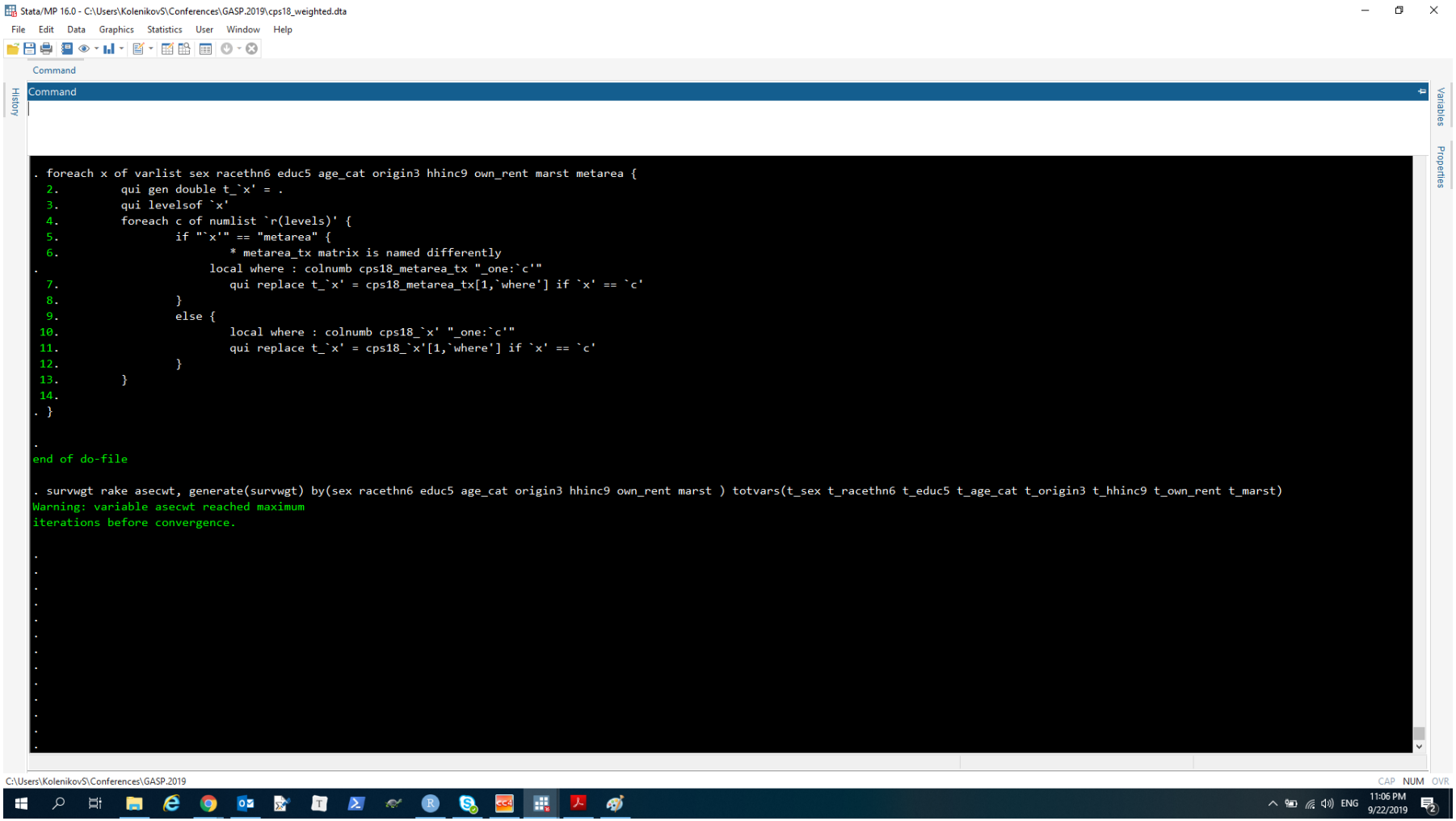


BOLD THINKERS DRIVING REAL-WORLD IMPACT

Stata svycal

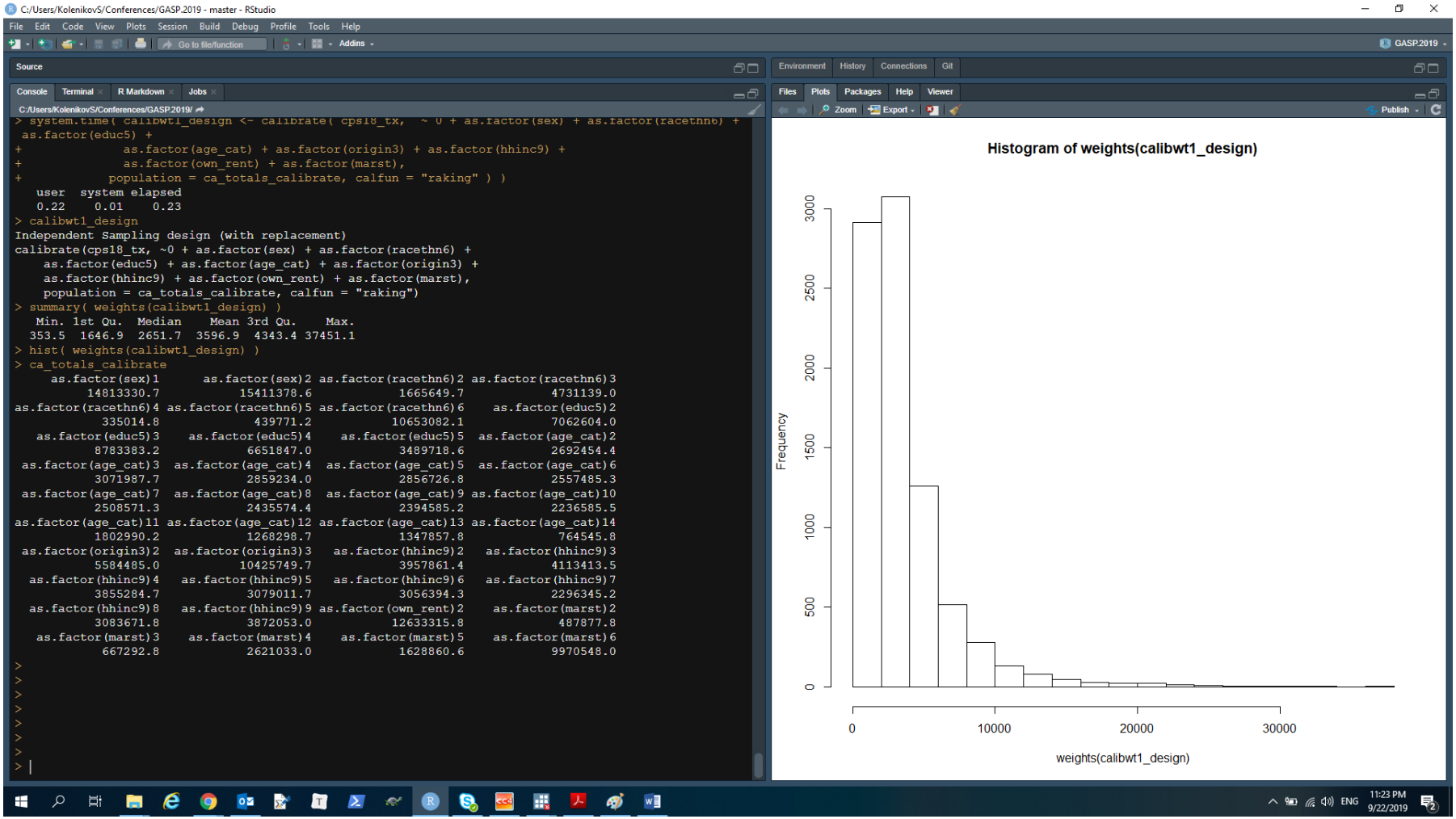


Stata survwgt



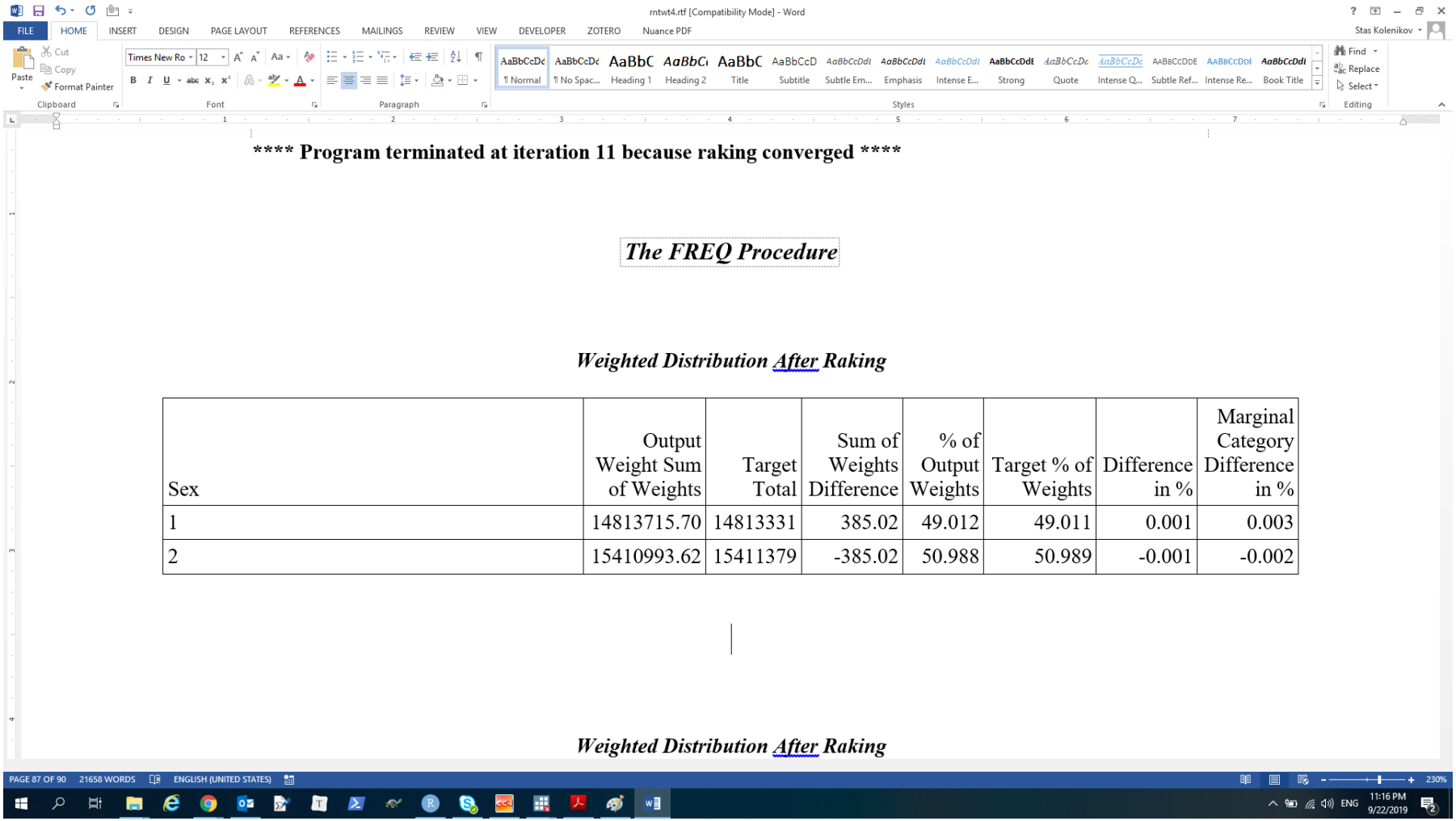
BOLD THINKERS DRIVING REAL-WORLD IMPACT

R survey::calibrate()



BOLD THINKERS DRIVING REAL-WORLD IMPACT

SAS rake_and_trim()



rmtwt4.rtf [Compatibility Mode] - Word

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

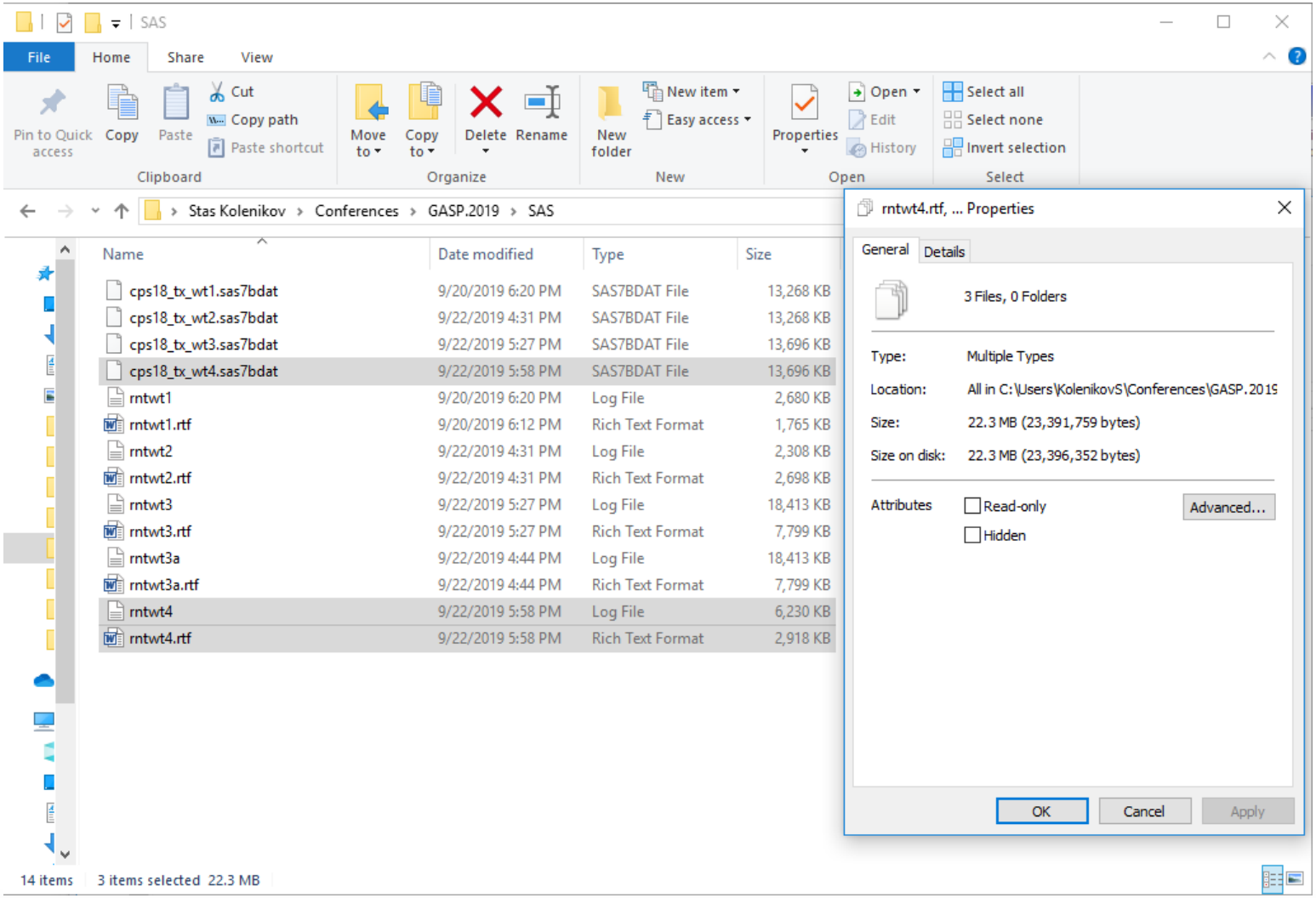
The FREQ Procedure

Weighted Distribution After Raking

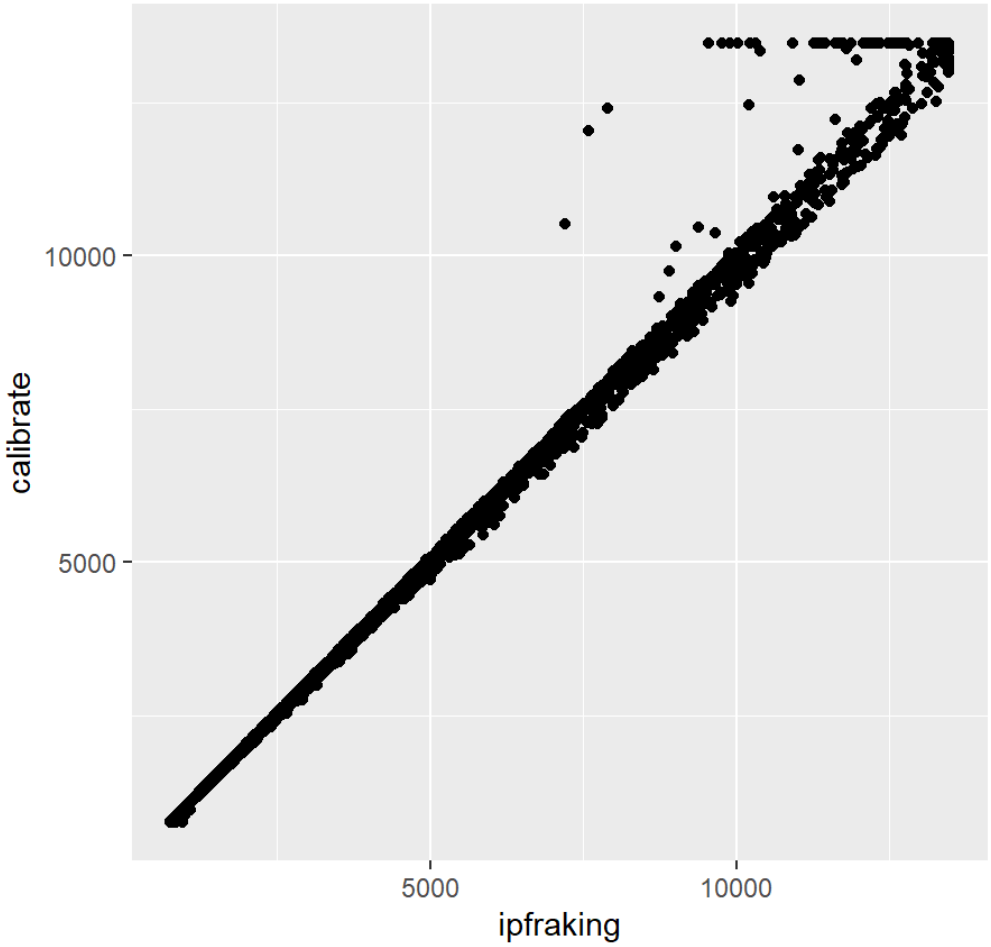
Sex	Output Weight Sum of Weights	Target Total	Sum of Weights Difference	% of Output Weights	Target % of Weights	Difference in %	Marginal Category Difference in %
1	14813715.70	14813331	385.02	49.012	49.011	0.001	0.003
2	15410993.62	15411379	-385.02	50.988	50.989	-0.001	-0.002

Weighted Distribution After Raking

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