Pretty pictures we need for the paper, *publication quality*

RBLandau 20180824

Notes on pictures:

- Aspect ratio 16:10, raw size 1600x1000 or greater for publication, 800x500 for easier viewing while editing. [Do journals have standards for sizes and aspects? *Science* does, but I've never looked at others.]
- All log-log.
- Almost all need decade lines for loss rates: 1%, 0.1%, 0.01%, I think. Horizontal dotted lines with small labels. Would be good if the lines were easily distinguished, e.g., by weight, but not clear that's possible. What would Tufte say?
- X half-life scales 2-1000 or 2-100.
- Legends upper right (if I can force legends at all; may have to be manual annotations, yuck).
- Attempt consistent color labeling for number of copies, but not sure this is reasonable in R.
- Describe data plots: #copies, losses, error injection, log scales x and y, percent losses, trimmed means (25%=midmean, maybe only 10%) of sample size=21, colors=black or blue usually best case recommendation.
- Describe why low half-lives and why 3,4 copies: otherwise numbers too small to make clear pictures, best to look at where the differences are exaggerated but parallel.
- Note that one would never actually use disks with such low half-lives, but they are needed to show the phenomena.
- I will have to regerate the data for most of these, rats.
- Comments inside the graph area will be done with annotation(), also rats.

(Note: the numbers are just for reference during discussions and editing.)

	idea / assertion	basics	details	wayzit	needs data & work

	idea / assertion	basics	details	wayzit	needs data & work
CALIBRATIONPIC: Copies1 Calibration	for copies=1 losses are unacceptable and auditing can't help	cop=1, lifem=2-1000	line at 0.1% or so, maybe decades below 1%, many annotations	pictures/noauditcalibrationcopies1	How do I draw looong arrows in R? All the Unicode arrows are short. There is supposed to be a way to place an image, so I could draw an arrow. Is this a plausible aspect ratio? Need better title and captions?

	idea / assertion	basics	details	wayzit	needs data & work
NOAUDITCOPIES: No audit needs many copies	without auditing, need many copies to minimize losses	cop=1,2,3,5,10 lifem=2-1000	1% line and 0.1%, clear legend for n-copies		
COPIES5SUFFICES: Is Copies=5 sufficient for all non-shock conditions?	with annual auditing, need only a few copies	cop=3,4,5 lifem=2-1000	1% line and 0.1% line, maybe shorten lifem=2-100		
COPIES4MAYBE: Is Copies=4 sufficient for calm periods?	is 4 sufficient in calm periods?	cop=4,5 lifem=2-100, various audit methods	decade lines for all graphs, I think		
COPIES5LONGTERM: Is Copies=5 okay for long periods?	is 5 sufficient for longer periods, 30-50 years?	lifem=2-1000	lines 1% and decades below		
COPIES6FORSHOCKS: Need Copies=6 for shock conditions?	are 6 necessary in shock periods?	cop=5,6 lifem=2-100 shocks freq=2yr dur=1yr, span=2,3, impact=50,67,75?,100%	see which is most striking visually		

	idea / assertion	basics	details	wayzit	needs data & work
RANDOMCONTRAST: Random auditing direct contrast	even generous random auditing still worse than total auditing	WITH vs WITHOUT replacement, cop=3 maybe also 5	4 or 10 segments for direct comparison		
FIVEYEARRANDOM: Five year random audit cycle	long-term lazy random auditing vs total	WITH replacement, cop=3,4,5	5 year cycle in 5 segments, hence 20% per year	again want direct comparison with total audit WITH vs WITHOUT replacement	
SHOCKSSEGMENTS: Shocks frequent segments better	frequent segments improve survival slightly in no-shock	1 segment vs 4 and maybe 10	cop=4,5 lifem2-100		
LARGEDOCTARGETS: Large docs are bigger targets	larger docs are larger targets for random errors	docsize=5,50,500,5000MB, usual lifem, losses are 100% for large docs at low half-life	legend must be clear		

	idea / assertion	basics	details	wayzit	needs data & work
DOCSIZETRADEOFF: Docsize vs error rate table	TABLE: tradeoff of docsize vs error rate	docsize5-5000 lifem=2-10000	nicely shaded png or pdf of spreadsheet excerpt; graph overlays too much to make the point		
THREATSTABLE: Threats table	TABLE: threats to content, from google doc	make it look like the other tables, bad word wraps	redo in md so it matches if possible to change text color emphasis		
GLITCHERRORRATES: Glitch just increases error rate	glitch just like higher error rate	cop=5 show line shifted to lower half-life with significant glitch	probably one month per quarter 50% or 67% increase in error rate, cherry-pick for clearest appearance		

	idea / assertion	basics	details	wayzit	needs data & work
SHOCKSMORETHAN4: Medium shocks need more than 4 copies	shocks: moderate not tolerated with 3 or 4 copies, the current pic is not too bad		redo with larger samples to smooth out the numbers		
SEVERESHOCKS5OR6: Severe shocks require 5 or 6 copies	severe shocks may require 6 copies	copies=4,5,6	shocks freq2yr dur1yr span2,3 imp50,80,100		
Less than 1% loss no audit	how many copies for < 1% long loss? no audit	long=30yr			
ONEPCTLONGTERM: Less than 1% loss annual audit	how many copies for < 1% long loss? annual audit	long=30yr			

	idea / assertion	basics	details	wayzit	needs data & work
CALIBRATIONTABLE: Calibration table	TABLE: calibration test of simulation results	theoretical vs simulated results	copies=1, 100 samples, vs straight Poisson, in ppm; have in spreadsheet, get nice png or pdf	tables/calibration	
20					
21					

Notes from 20180821

- Need convincing pic for five audited copies in calm days. Add 1% and 0.1% lines and see what it looks like. Check also for much longer periods, e.g., 30 and 50 years.
- Random: compare same number of segments, just with vs without replacement. How many copies? Year, quarter, month. Maybe also 5 year cycle with 5 segments, done with replacement, just for comparison.
- Use names not numbers for the figures to allow us to reorder them. Numbers assigned at the last minute.
- Segments: ignore monthly, but leave the 2 year line.
- Compression: may do this in a table rather than figure. Can we do a dramatic picture/graph?
- Is there a function that relates one extra copy to a reduction in loss? Need to look across a lot of empirical data.