Example of NLP approach to calibration

Don Boyd to me May 21

Hi Yimeng,

One reason I wanted to do the calibration for the 150 plans was because I wasn't sure I could explain clearly what I meant by the nonlinear program (NLP) approach - an enhanced version of what NMR did (they didn't have information on duration and they do not appear to have used an optimization approach - it was a cruder adjustment).

The attached should be enough information for you to see what I mean.

I applied the approach to the simplest check-able problem we have: UCRP closed-group benefit payments to year1 retirees. Here's what I did:

The idea was to take model-produced benefit payments and push them around so that they more closely match the (ordinarily unobserved) true benefit payments. For true benefits, I used the UCRP closed-group Segal benefits. For the model-produced benefits, I used your model-produced benefits. These data are in the attached Excel file.

Here are the steps I used. They can be improved upon, and there will be some real-world complications with other plans that I did not have to contend with here, but you'll get the basic idea from this approach.

Step 1: Adjust the model-produced benefits proportionately, so that year 1 matches the known year 1 exactly (e.g., from an actuarial valuation)

Step 2: Do a crude adjustment to step 1, multiplying the year2-year n values by a constant ratio, so that the PV of step 2 matches the PV of the true benefits (which might be known from the AV)

Step 3: Adjust each year by calculating new values, based on an NLP, where:
a) It has 4 constraints:
(1) the new (step3) values must have pv = known pv of the benefits (per AV)
(2) the pv of the step3 values at discount rate minus 1% must match what we believe they would be (we would calculate this using a duration estimate from a CAFR - here I was able to calculate it directly since I know the Segal values)
(3) pv must match at discount rate plus 1%
(4) year 1 value for step3 must equal known year 1 value

b) the NLP solver satisfies these constraints by choosing the step 3 values while minimizing a penalty or distortion function that is larger when a step 3 value is further away from the step2 values (on the theory that those values, which were largely based on a model - mortality rates, etc. - are pretty good and we don't want to move too far from them). I used a simple penalty function in which we minimize the sum of the squared deviations of step 3 from step 2. Other penalty functions could be useful - for example, we could weight each squared difference to make years far in the future less important than years that are closer, or something like that, but I thought this would be simple and good to illustrate the approach.

c) I used the step2 values as the starting point for the NLP

You can see the results below (retired is the annual benefit payment to retirees per Segal, in $ millions; retired.yy is your estimates; and rstep1-rstep3 are as I desribed above). To my eye, the step3 values look much closer to the (ordinarily unobserved) "true" values than the unadjusted model-based benefit payments and certainly are much better than the crude step 1 and step 2 adjusted values.

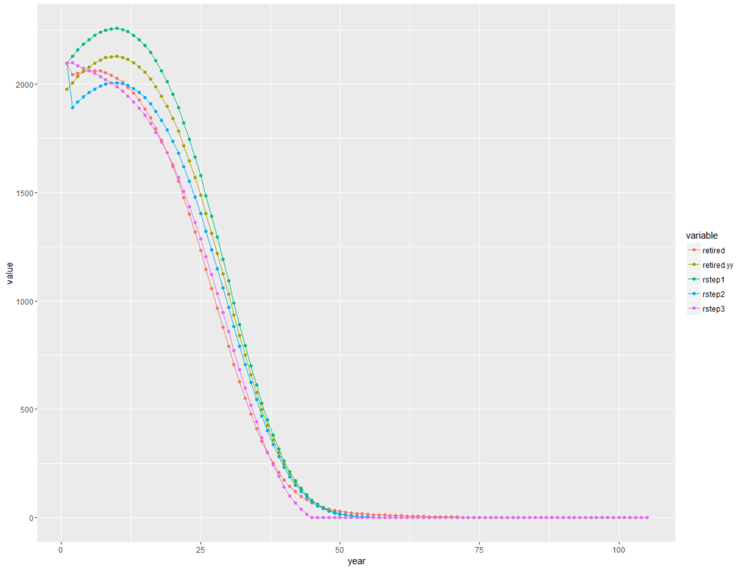
I used the ipoptr package which provides an interface to ipopt, an open-source interior point optimizer that is very flexible -- it can solve very large scale continuous nonlinear problems, with equality and inequality nonlinear constraints and with bounds on the solution variables (we obviously don't want them to be negative). You may recall that I put this in the dropbox a long time ago (..\Pension simulation project\Packages). It can be difficult to build from source so I recommend using the compiled version I put there - copy entire folder to the directory with your R libraries.

There are many potential variants on this approach. (e.g., if for some reason we cannot produce model-based values to use as starting values, we probably could estimate a generic benefit structure for actives, term vested, or retirees - each appears to have pretty distinctive polynomial shape, reflecting mortality/separation and nominal growth). Oftentimes we will not have the luxury I had here of targeting retirees (and then actives and term vested) separately - we may only have AAL in total (all outflows) and we may have to target total outflows. But again, the basic idea would be the same: (1) develop a good set of starting point values using a benefit model or perhaps polynomial estimation approach, and then (2) push it around to hit four constraints - the 3 pvs based on duration plus the year 1 value - while minimizing a measure of distortion relative to the starting point.

(BTW, I have also used a variant on this NLP approach to push a workforce matrix around so that the adjusted matrix has an average age consistent with summary data from an AV.)

Let's talk about this after you have a chance to work through it.

Don



ucrp noodling(2)

r

ucrp_djb

Spreadsheet