Implementing the estimation method from "Estimating Demand Elasticities Using Nonlinear Pricing"

Recall, you are trying to replicate empirically this equation:

$$h_i = G(p_i, \theta_i)$$

You need these variables from your dataset:

- *expend* - Final yearly expenditures for each individual, the h_i from above.
- *deduct* - Level of the deductible, where the marginal price p_i changes
- *dummyside* - dummy identifying the estimation window of interest. Equal to 0 if *expend* is less than the deductible, equal to 1 if *expend* is greater than or equal to the deductible (so 0="L" and 1="R") Missing if the observation is not in the estimation window of interest (though your kernel will take care of this anyway).

Steps:

1. Calculate θ_i
   - The θ_i variable is the percentile of the expenditure within the window.
   - Use the Stata command:
     xtile *theta* = *expend* if *dummyside*!=. , nquantiles(100)
2. Download locpolyslope command
   - Stata 11 and 12 have a command lpoly for local polynomial regression. However this command does not report both the constant and the slope coefficients. I have modified an earlier version of this command to also report the slope coefficients.
   - Download the file called **locpolyslope.ado** to report slope coefficients. Usually this means dropping locpolyslope.ado into your "C:/ado/personal" folder to make Stata recognize this command. locpolyslope uses most of the same options as lpoly.
3. Calculate the slopes $b_L(\theta)$ and $b_R(\theta)$
   Recall, we are estimating a local linear regression for both sides of the nonlinearity:

$$\min_{a_L, b_L} \sum_{\theta_i < \bar{\theta}} \omega(\theta_i) \left(h_i - a_L(\theta) - b_L(\theta)(\theta_i - \theta_0)\right)^2$$

   - Run the local linear regression on both sides of the nonlinearity separately.

- Use the Stata commands:

locpolyslope *expend theta* if *dummyside* == 0, adoonly degree(1) n(#) width(#) generate(*xvar0 yvar0 slopevar0*)

locpolyslope *expend theta* if *dummyside* == 1, adoonly degree(1) n(#) width(#) generate (*xvar1 yvar1 slopevar1*)

where n(#) is the number of points over which you choose to run the llr, and width(#) is the bandwidth you choose. *slopevar0* is $b_L(\theta)$ and *slopevar1* is $b_R(\theta)$. You now have three column variables of size n(#) for each side, the first being the theta smoothing observations *xvar*, the second is the smoothed points *yvar*, and the last is the slope *slopevar*.

4. Plug all estimated values into the elasticity equation.
Recall, the elasticity, η, at the nonlinearity is:

$$\eta = \left[b_L(\bar{\theta}) - b_R(\bar{\theta})\right] \cdot \frac{\bar{\theta}}{\bar{h}}$$

- $\bar{\theta}$ is the value of theta at the nonlinearity - the value in between the last *xvar0* value and the first *xvar1* value.
- $b_L(\bar{\theta})$ is the last value of *slopevar0*.
- $b_R(\bar{\theta})$ is the first value of *slopevar1*
- $\bar{h}$ is the value of the nonlinearity (i.e. the amount of the deductible)