

Documentation for Quantile Regression for Panel Data (QRPD)*

David Powell[†]
RAND

March 12, 2014

This documentation refers to Stata code for the QRPD and IV-QRPD estimators introduced in Powell [2014]. This code is preliminary and has some limitations, which are discussed below. This code will be updated over time. The do-file is titled `qrpdnew.do`.¹ After downloading, please include the command “do `qrpdnew.do`” in your do-file to load the program. The code limits you to one treatment variable (+ time fixed effects and individual fixed effects). There are two reasons for this. First, the code grid-searches over values for the treatment variables so time considerations are important. Second, obtaining unconditional quantile treatment effects for the treatment variable requires not conditioning on other variables. I will refer you to the paper for more guidance on this issue. I am currently developing code to include more variables.

This new version of the code performs two-step optimal GMM when more than one instrument is included.

I should also note that there are several other quantile estimators for panel data (including Koenker [2004], Harding and Lamarche [2009], Canay [2011], Galvao Jr. [2011], Ponomareva [2011], Graham et al. [2009], and Lamarche [2010]). These estimators include additive terms for the fixed effects and the interpretation of the estimates is very different. Different estimators may be more useful in different applications.

Here is the main information that you need for the code:

1. Type “`ssc install moremata`” into Stata - you will need this file to use this code.
2. Please drop any observations with missing data. The code will not do this for you.
3. The syntax for the command is as follows:

*Special thanks to Travis Smith for helpful suggestions in improving the coding.

[†]`dpowell@rand.org`

¹This version was created November 19, 2013 and replaces an older version of the code.

```
qrpd y, cond_end(x) instr(z) by(person) time(year) min1(-5) max1(5) intvl1(.1) tau(0.5)
```

4. y is your outcome variable.
5. x is the variable of interest (one variable only).
6. z represents the instruments - you can have as many instruments as you'd like. If x is exogenous, then x is also your instrument.
7. year is your time variable (one variable only). The code generates the year fixed effects on its own. Please refer to the paper for a better understanding of what the year fixed effects are doing here. They ensure that the “high quantiles” refer to observations at the top of the distribution in a given year.
8. person - this variable represents the person/household/firm which you want to condition on (these are the individual fixed effects). Only one variable is allowed.
9. min1() - enter the minimum number that you'd like to search over for the coefficient on x
10. max1() - enter the maximum number that you'd like to search over for the coefficient on x
11. intvl1() - the code searches between min1() and max1() at an interval of intvl1()
12. tau() - the quantile of interest. A number between greater than 0 and less than 1 will be read as is. A number greater than or equal to 1 will be divided by 100.

The code will provide coefficients and standards errors for the quantile of interest. It will also save this output in “e(beta1)” and “e(se)”. It is advisable to bootstrap the estimates as well to obtain standard errors instead of the provided standard errors. If you choose to do this, please remember to do a cluster bootstrap (“cluster(person)”). The analytical standard errors adjust for clustering. If bootstrapping, it is advisable to search over a wide range of possible values.

Please email me (dpowell@rand.org) with any concerns or questions.

References

- Ivan A. Canay. A note on quantile regression for panel data models. *The Econometrics Journal*, 14:368–386, 2011.
- Antonio F. Galvao Jr. Quantile regression for dynamic panel data with fixed effects. *Journal of Econometrics*, 164(1):142–157, September 2011.

- Bryan S. Graham, Jinyong Hahn, and James L. Powell. The incidental parameter problem in a non-differentiable panel data model. *Economics Letters*, 105(2):181–182, November 2009.
- Matthew Harding and Carlos Lamarche. A quantile regression approach for estimating panel data models using instrumental variables. *Economics Letters*, 104(3):133 – 135, 2009.
- Roger Koenker. Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74–89, October 2004.
- Carlos Lamarche. Robust penalized quantile regression estimation for panel data. *Journal of Econometrics*, 157(2):396–408, August 2010.
- Maria Ponomareva. Quantile regression for panel data models with fixed effects and small T: Identification and estimation. Working paper, University of Western Ontario, May 2011.
- David Powell. Did the economic stimulus payments of 2008 reduce labor supply? Evidence from quantile panel data estimation. Working papers, RAND Corporation Publications Department, 2014.