xtqrplot

Entity-Specific Marginal Effects from Panel Quantile Regression

Version 1.3.0

Dr Noman Arshed
Senior Lecturer, Department of Business Analytics
Sunway Business School, Sunway University
nouman.arshed@gmail.com

xtqrplot is a Stata package that estimates panel quantile regression across a user-defined grid of quantile windows and produces entity-specific marginal effect visualisations. For each cross-sectional unit (e.g., country, firm, individual), the command identifies that entity's position in the distribution of the dependent variable, extracts the corresponding coefficient, and presents the results as heat-mapped bar charts or matrix plots with 95% confidence intervals. The package supports both balanced and unbalanced panels, four effect types, pre-estimation normality diagnostics, and automatic results tables.

1. Introduction

Panel quantile regression estimates how the relationship between a dependent variable and a set of predictors differs across the conditional distribution of the outcome. After estimation, researchers face a practical challenge: the coefficient at a given quantile applies to all entities, but each entity's economically relevant coefficient is the one at the quantile where that entity's dependent variable actually sits in the distribution.

xtqrplot operationalises this idea. It runs the panel quantile model at each point in a user-defined quantile grid, computes each entity's percentile rank in the dependent variable, assigns that entity to its nearest quantile window, and extracts the corresponding coefficient. The results are presented as colour-coded bar charts (for cross-section or time comparisons) or matrix heat maps (for panel x time comparisons).

1.1 Supported estimators

xtqrplot wraps two panel quantile regression commands available on SSC:

> ***xtqreg*** *(default) — Machado & Santos Silva (2019) method-of-moments estimator. Efficient for large balanced and unbalanced panels. No bootstrap required.*
>
> ***qregpd*** *— Powell (2016) estimator, consistent under non-additive fixed effects. Uses bootstrap standard errors. Slower but more robust to certain heterogeneity patterns.*

1.2 Compatibility

xtqrplot requires Stata 14.0 or later. It has been tested on Stata 17. The package works with both balanced and unbalanced panel data. An optional enhanced heat map is available if heatplot (Jann, 2021) is installed alongside the palettes and colrspace packages.

2. Installation

2.1 Required packages

Install xtqrplot and its required estimator packages from SSC:

```
* Install xtqrplot
copy "https://raw.githubusercontent.com/[repo]/xtqrplot.ado" \
    "C:\Users\[username]\ado\personal\xtqrplot.ado"

* Install required estimators (choose one or both)
ssc install xtqreg
ssc install qregpd
```

2.2 Optional enhanced plotting

For the heatplot-based matrix visualisation (recommended for twoway plots), install the following three packages:

```
ssc install heatplot,  replace
ssc install palettes,  replace
ssc install colrspace, replace
```

Note: If heatplot is not installed, xtqrplot automatically falls back to a 10-bin scatter heat map that requires no additional packages.

2.3 Verify installation

```
which xtqrplot
help xtqrplot
```

3. Syntax

The full syntax of xtqrplot is:

```
xtqrplot depvar indepvars [if] [in] ,
    panelvar(varname) timevar(varname)
    [ method(xtqreg|qregpd)  reps(#)       seed(#)
      nwindows(#)            effect(string) plottype(string)
      saving(string)         replace        noplot  nonormal ]
```

3.1 Required arguments

depvar is the numeric dependent variable. **indepvars** is one or more numeric independent variables. **panelvar(varname)** specifies the cross-sectional identifier (must be numeric). **timevar(varname)** specifies the time identifier (must be numeric).

3.2 All options

Option	Type	Default	Description
`panelvar(varname)`	varname	—	Panel (cross-section) identifier. Required.
`timevar(varname)`	varname	—	Time identifier. Required.
`method(string)`	string	xtqreg	Estimator: xtqreg (default) or qregpd.
`reps(#)`	integer	200	Bootstrap replications for qregpd.
`seed(#)`	integer	—	Random seed for qregpd reproducibility.
`nwindows(#)`	integer	9	Number of interior quantile windows (1–49).
`effect(string)`	string	coef	Effect type: coef, semi, elast, or bp.
`plottype(string)`	string	cross	Plot dimension: cross, time, or twoway.
`saving(string)`	string	—	Base filename for .dta and .gph output files.
`replace`	flag	—	Overwrite existing saved files.
`noplot`	flag	—	Suppress all graphs; tables still produced.
`nonormal`	flag	—	Suppress normality tests and histograms.

4. Effect Types

The effect() option controls the quantity plotted for each entity. Let beta(tau_i) denote the estimated coefficient at entity i's assigned quantile window tau_i. Let X-bar_i and Y-bar_i denote entity i's mean values of a given regressor and the dependent variable (or actual values in the twoway case).

Effect	Formula	Interpretation
`coef`	beta(tau_i)	Raw coefficient at entity's assigned quantile window. Measures change in the conditional quantile of Y per unit change in X.
`semi`	beta(tau_i) x X-bar_i	Semi-elasticity. Scales the coefficient by the entity's own mean level of X, capturing the effect in the same units as Y.
`elast`	beta(tau_i) x X-bar_i / Y-bar_i	Elasticity. Expresses the effect as a percentage change in Y relative to the entity's own Y level. Requires Y > 0.
`bp`	beta(tau_i) x 10,000	Basis points. Multiplies the coefficient by 10,000. Useful for financial rate variables where raw coefficients are very small.

Note: For effect(elast), entities with Y-bar_i <= 0 are assigned missing effects and a warning is printed. Use effect(coef) or effect(semi) for data with non-positive dependent variables.

Note: All effect types include 95% confidence intervals computed as beta +/- 1.96 x SE, shown as error bars on cross and time plots.

5. Plot Types

5.1 cross (default)

Produces a horizontal bar chart with entities on the y-axis. One chart is produced per independent variable. For each entity, the bar shows the marginal effect at that entity's assigned quantile window. Red bars indicate positive effects; blue bars indicate negative effects. Error bars show the 95% confidence interval. A secondary x-axis (0 to 100) shows a green circle at each entity's exact percentile rank of the mean dependent variable, so the assignment context is visible alongside the effect.

5.2 time

Identical to cross but the y-axis shows time periods rather than entities. The data are collapsed to period means before percentile ranking, so the plot answers: for each year, what is the marginal effect at the quantile where that year's average outcome sits?

5.3 twoway

Produces a matrix heat map with panel units on the y-axis and time periods on the x-axis. Each cell shows the effect for that specific panel unit x time period combination, using each observation's actual (not mean) value of the dependent variable to determine quantile assignment. Coefficient values are printed inside each cell.

If heatplot is installed, a continuous diverging colour scale (blue for negative, red for positive) with a proper legend bar is produced. Otherwise a 10-bin RdBu scatter fallback is used automatically.

> **Warning: The twoway plot is best suited for panels with limited dimensions (up to ~15 entities and ~20 time periods). For larger panels, the printed matrix table is the more informative output.**

In addition to the plot, xtqrplot always prints a formatted matrix table (rows = entities, columns = years, cells = effect value) regardless of whether the plot succeeds.

6. Quantile Windows

xtqrplot runs the panel quantile model at nwindows evenly-spaced interior quantile points. The step size is 100/(nwindows+1) and the grid runs from step to nwindows x step:

nwindows	Step size	Quantile grid
1	50	50 (median only)
3	25	25, 50, 75
4	20	20, 40, 60, 80
9	10	10, 20, 30, 40, 50, 60, 70, 80, 90 (default)
19	5	5, 10, 15, ..., 90, 95
24	4	4, 8, 12, ..., 92, 96
49	2	2, 4, 6, ..., 94, 96, 98

Note: Boundary quantiles (0 and 100) are never included because both xtqreg and qregpd are numerically unstable near the distribution extremes.

Note: Each entity is assigned to its nearest quantile window based on the percentile rank of its mean (or actual) dependent variable. Ties are broken by nearest-neighbour rounding.

7. Output Produced

7.1 Screen output

• ***Normality diagnostics*** *— Skewness-Kurtosis (sktest), Shapiro-Wilk (swilk), and Shapiro-Francia (sfrancia) tests for the dependent variable and every independent variable. Histograms with kernel density overlays are drawn with bin count equal to nwindows+1 so visual categories match the quantile grid. Suppress with nonormal.*

• ***Quantile-by-quantile progress*** *— Each estimated quantile is reported as it completes, with [ok] or an error message.*

• ***Median-level results table*** *— Coefficient, standard error, z-statistic, and two-tailed p-value at the quantile nearest q50. A replication command is printed so the exact model can be reproduced.*

• ***Entity results table*** *— One row per entity showing: entity name, mean of the dependent variable, exact percentile rank, assigned quantile window, and computed marginal effect for each independent variable.*

• ***Matrix table (twoway only)*** *— Full rows=entities, cols=years matrix of coefficient values printed to the screen regardless of whether the heat map plot succeeds.*

7.2 Saved dataset

When saving(filename) is specified, a .dta dataset is saved containing:

Variable	Description
`_my`	Mean (or actual) value of the dependent variable for the entity
`_mx#`	Mean (or actual) value of the #th independent variable
`_pctrank`	Exact percentile rank of _my in the distribution (0-100)
`_q_win`	Assigned quantile window (nearest grid point, in percent)
`_eff#`	Computed marginal effect for the #th independent variable
`_eff#_lo`	Lower bound of 95% confidence interval for _eff#
`_eff#_hi`	Upper bound of 95% confidence interval for _eff#

8. Examples

8.1 Simulate example data

The following code generates a 10-country, 10-year panel with a normally-distributed x1 and a right-skewed chi-squared x2, producing genuine heterogeneity in quantile effects:

```
clear
set seed 2024
set obs 100
gen country_id = ceil(_n/10)
gen year        = mod(_n-1,10) + 2010
gen x1 = rnormal(2,1)
gen u1 = runiform() \ gen u2 = runiform()
gen x2 = -ln(u1) - ln(u2) - 2   // chi2(2)-2, right-skewed
drop u1 u2
gen fe  = rnormal(0,0.8)
bysort country_id: replace fe = fe[1]
gen eps = rnormal(0,1)*(1 + 0.5*abs(x2))
gen y   = 1 + 0.4*x1 + 0.3*x2 + fe + eps
label define clab 1 "Argentina" 2 "Brazil" 3 "Chile" 4 "Colombia"
    5 "Ecuador" 6 "Mexico" 7 "Panama" 8 "Peru"
    9 "Uruguay" 10 "Venezuela"
label values country_id clab
label variable y  "GDP Growth"
label variable x1 "Trade Openness"
label variable x2 "Investment Shock"
xtset country_id year
```

8.2 Basic cross-section plot (all defaults)

```
xtqrplot y x1 x2, panelvar(country_id) timevar(year)
```

Note: Runs xtqreg at quantiles 10, 20, ..., 90. Produces a horizontal bar chart per regressor. Normality tests are run first for all variables.

8.3 Time-dimension plot, semi-elasticity

```
xtqrplot y x1 x2, panelvar(country_id) timevar(year)  \
    plottype(time) effect(semi) nonormal
```

8.4 Twoway heat map (requires heatplot)

```
ssc install heatplot, replace
ssc install palettes, replace
ssc install colrspace, replace

xtqrplot y x1 x2, panelvar(country_id) timevar(year)  \
    plottype(twoway) nonormal
```

8.5 Basis points, finer grid, save all output

```
capture mkdir results

xtqrplot y x1 x2, panelvar(country_id) timevar(year)  \
    effect(bp) nwindows(19) nonormal                   \
    saving(results/xtqrplot_out) replace
```

8.6 Powell qregpd with seed

```
xtqrplot y x1, panelvar(country_id) timevar(year)  \
    method(qregpd) reps(300) seed(42) nwindows(4)
```

8.7 Table only, no graph

```
xtqrplot y x1 x2, panelvar(country_id) timevar(year)  \
    noplot nonormal saving(data/results) replace

use data/results_cross.dta, clear
list country_id _pctrank _q_win _eff1 _eff2 _eff1_lo _eff1_hi
```

9. Quality Checks

xtqrplot performs the following quality and safety checks automatically. These run without any user action and are designed to prevent silent errors, protect against data loss, and guide the user when estimation fails.

#	Check	When	Action if failed
1	Option validation	Before estimation	Exit rc=198 with descriptive message listing valid values.
2	Package availability	Before estimation	Exit rc=199 with ssc install command to fix.
3	Panel structure (xtset)	Before estimation	Exit rc=459 with guidance on numeric ID requirements.
4	Sample size	After marksample	Print warning if N < 30. Estimation continues.
5	Normality diagnostics	Before regression	Run sktest, swilk, sfrancia + histograms for all variables. Suppress with nonormal.
6	Per-quantile convergence	During estimation	Stop at first failure; print quantile number, rc, and 4 likely causes.
7	Missing effect count	After effect computation	Print count of missing effects per variable so assignment failures are visible.
8	CI consistency	After effect computation	Missing SE produces missing CI rather than incorrect bounds.
9	File overwrite protection	Before saving	Exit rc=602 if file exists and replace not specified.
10	Plot failure isolation	During plotting	Each variable's plot captured separately; failure of one does not block others.

A formal pre-submission test suite (xtqrplot_test_suite.do) covering 37 tests across 11 test blocks is distributed with the package. Run it after installation with:

```
do xtqrplot_test_suite.do
```

Each test prints [PASS] or [FAIL]. Zero FAIL lines confirms the package is working correctly in your Stata environment. A log file (xtqrplot_test_log.txt) is saved automatically.

10. Error Catalogue

The table below lists all anticipated errors, their Stata return codes, causes, and recommended remedies.

Code	Cause	Remedy
198	Invalid option value (method, effect, plottype, nwindows)	Check spelling and allowed values. See Options section.
199	Required package not installed (xtqreg or qregpd)	Run: ssc install xtqreg or ssc install qregpd
459	Cannot xtset the data with given panelvar/timevar	Ensure both ID variables are numeric. Run xtset manually to diagnose.
602	Output file already exists	Add the replace option to overwrite existing files.
2000	Empty estimation sample	Check if/in conditions. Ensure no variable has all-missing values.
Other	Estimation failure at quantile tau	Too few obs (reduce nwindows), collinear regressor, or try method(xtqreg).

9.1 Estimation failure at a specific quantile

When estimation fails at one quantile window, xtqrplot stops and reports the quantile percentage and return code. Common causes:

- *Too few observations relative to nwindows — reduce nwindows() or use a larger sample.*
- *A regressor with no within-entity variation (time-invariant variable) — remove it.*
- *Convergence failure in qregpd — increase reps(), add seed(), or switch to method(xtqreg).*
- *Extreme collinearity between regressors — check VIF before running xtqrplot.*

11. Methodological Notes

10.1 Quantile window assignment

For cross and time plots, the dependent variable is first collapsed to entity (or period) means. The percentile rank of each mean in the overall distribution determines the quantile window. Assignment uses nearest-neighbour rounding: each entity is assigned to the window whose quantile value is closest to its percentile rank.

For twoway plots, no collapse is performed. Each individual observation Y_it is ranked in the full distribution of all Y_it values, so the quantile assignment is observation-specific.

10.2 Confidence intervals

95% confidence intervals are computed as beta +/- 1.96 x SE at each quantile window. These are asymptotic CIs based on the standard errors from the underlying estimator. For qregpd they are bootstrap-based; for xtqreg they are analytic.

10.3 Unbalanced panels

xtqrplot supports unbalanced panels. The collapse step (cross/time plots) computes means over whatever observations are available for each entity or period. The twoway plot assigns effects to cells where data exist; missing cells are shown as dots in the matrix table.

10.4 Interpreting the heat map

Colour intensity in the twoway plot reflects the magnitude of the marginal effect at each cell's assigned quantile. Blue cells indicate negative effects; red cells indicate positive effects. Cells with similar colours can be compared directly. However, because quantile windows are discrete, nearby cells may be assigned to different windows, producing apparent discontinuities that reflect the coarseness of the grid rather than genuine economic discontinuities.

12. References

Jann, B. (2021). *heatplot: Stata module to create heat plots and hexagon plots. Statistical Software Components S458598, Boston College Department of Economics.*

Koenker, R. and Bassett, G. (1978). *Regression quantiles. Econometrica, 46(1), 33-50.*

Machado, J.A.F. and Santos Silva, J.M.C. (2019). *Quantiles via moments. Journal of Econometrics, 213(1), 145-173. https://doi.org/10.1016/j.jeconom.2019.04.009*

Powell, D. (2016). *Quantile regression with nonadditive fixed effects. RAND Working Paper WR-1088. https://www.rand.org/pubs/working_papers/WR1088.html*

Royston, P. (1991). *Shapiro-Francia test for normality. Stata Technical Bulletin, 2, 16-17.*

13. Model Performance Disclaimer

xtqrplot is a visualisation and post-estimation tool. Model performance and the validity of results depend entirely on the underlying panel quantile regression and the data provided.

Convergence: *Results are only meaningful when the panel quantile regression converges successfully at each quantile window. Always inspect the quantile-by-quantile output and verify that coefficients move sensibly across quantiles.*

Window assignment sensitivity: *The quantile window assignment is a nearest-neighbour approximation. Entities near the boundary between two windows may be assigned to either; small samples amplify this sensitivity. Increasing nwindows provides finer resolution but requires more computation.*

Twoway heat map: *The heat map assigns each observation its individual quantile rank. For panels with many missing cells, colour patterns may reflect the distribution of available data rather than genuine economic variation.*

Effect interpretation: *The marginal effects produced by xtqrplot are at-the-quantile effects, not average treatment effects. They describe how the conditional quantile of Y responds to X at the quantile where each entity sits in the distribution.*

Active development notice: xtqrplot is under active development. If you observe any irregularity — unexpected output, errors not listed in the error catalogue, or results that appear economically implausible — please report them to the author at **nouman.arshed@gmail.com** with a minimal reproducible example. Your feedback directly improves the package for the research community.