Estimation of Quantile Regressions with Multiple Fixed Effects

Fernando Rios-Avila
Levy Economics Institute
Annandale-on-Hudson, NY
friosavi@levy.org

Leonardo Siles
Universidad de Chile
Santiago, Chile
lsiles@fen.uchile.cl

Gustavo Canavire-Bacarreza
The World Bank
Washington, DC
gcanavire@worldbank.org

Abstract.

This is an example of StataJ article made by me

Keywords: st0001, Stata, LaTeX, Quarto, StataJ

1 Introduction

Quantile regression, introduced by Koenker and Bassett (1978), has become an important tool in economic analysis, allowing to examine how the relationship between the dependent and independent variables varies across different points of the conditional distribution of the outcome. While ordinary least squares focuses on analyzing the conditional mean, quantile regression provides a more comprehensive view of how covariates impact the entire conditional distribution of the dependent variable. This can reveal heterogeneous effects that may be otherwise overlooked when analyzing the conditional mean.

A relatively recent development in the literature has focused on extending quantile regression analysis in a panel data setting to account for unobserved, but time fixed heterogeneity. This is particularly important in empirical research, where unobserved heterogeneity can bias estimates of the effects of interest. However, as it is common in the estimation of non-linear models with fixed effects, introducing fixed effects in quantile regression models poses several challenges. On the one hand, the simple inclusion of fixed effects can lead to an incidental parameter problem, which can bias estimates of the quantile coefficients (Neyman and Scott 1948; Lancaster 2000). On the other hand, the computational complexity of estimating quantile regression models with fixed effects can be prohibitive, particularly for large datasets with multiple high-dimensional fixed effects. While many strategies have been proposed for estimating this type of model (see Galvao and Kengo (2017) for a review), none has become standard due to restrictive assumptions regarding the inclusion of fixed effects and the computational complexity.

In spite of the growing interest in estimating quantile regression models with fixed effects in applied research, particularly in the fields of labor economics, health economics, and public policy, among others, there are few commands that allow the estimation of such models. In Stata, there are three main built-in commands available for estimating quantile regression `qreg`, `ivqregress`, and `bayes: qreg`, and none of them allow for

the inclusion of fixed effects, other than using the dummy variable approach. From the community-contributed commands, there is `xtqreg`, which implements a quantile regression model with fixed effects based on the method of moments proposed by Machado and Santos Silva (2019), and more recently `xtmdqr` which implements a minimum distance estimation of quantile regression models with fixed effects described in Melly and Pons (2023). In both cases, these command are constrained to a single set of fixed effects.[1]

To address this, in this paper we introduce two Stata commands for estimating quantile regressions with multiple fixed effects: `mmqreg` and `qregfe`. The first command `mmqreg` is an extension of the method of moments quantile regression estimator proposed by Machado and Santos Silva (2019). The second `qregfe`, implements three other approaches: an implementation of a correlated random effects estimator based on Abrevaya and Dahl (2008), Wooldridge (2019) and Wooldridge (2010, Ch12.10.3); the estimator proposed by Canay (2011), and a proposed modification of this approach. In addition, we also present an auxiliary command `qregplot` for the visualization of the quantile regression models.

Both commands offer the advantage of allowing for the estimation of conditional quantile regressions while controlling for multiple fixed effects. First, they leverage over existing Stata commands, as well as other community-contributed commands, to allow users to estimate quantile regression models and their standard errors under different assumptions. Second, they reduce the impact of the incidental parameters problem depending on the assumptions of the underlying the data generating process. In terms of standard errors, `mmqreg` allows for the estimation of analytical standard errors (see Machado and Santos Silva (2019) and Rios-Avila et al. (2024)), whereas `qregfe` emphasizes the use of bootstrap standard errors. Finally, both commands are designed to be user-friendly, allowing for the estimation of quantile regression models with fixed effects in a single line of code.

The remainder of the paper is organized as follows. Section 2 reviews the methodological framework for quantile regression. Section 3 describes the methods and formulas used by `mmqreg` and `qregfe` commands. Section 4 introduces the commands, along with a brief description of their syntax and options. Section 5 introduces an auxiliary command for the visualization of quantile regression models. Section 6 provides an empirical applications demonstrating their use. Section 7 concludes.

2   The Basics

Quantile regressions allow researchers to identify the heterogenous effect covariates could have over the entire conditional distribution of the dependent variable. Let y_i be the dependent variable, x_i the vector of covariates, and $0 < \tau < 1$ is a parameter such

---

1. There are other community-contributed commands like `xtrifreg`, `rifhdfe`, `qregpd`, `rqr` among others that allow for the estimation of quantile regression models, but do not estimate conditional quantile regressions, but instead focus on unconditional quantile regressions, or quantile treatment effects.

that $q_\tau(y_i|X)$ identifies the τth quantile of the conditional distribution of $y_i|X$. Under the assumption that conditional quantiles are linear functions of the parameters, the quantile regression model can be written as:

$$q_\tau(y_i|X) = x_i\beta(\tau) \tag{1}$$

Where $\beta(\tau)$ is the vector of coefficients that may vary across τ and needs to be estimated, and x_i is a vector of exogenous covariates that may include nonlinear functions of underlying variables. This expression indicates that, conditional on X, the τ-th quantile of Y can be approximated by a linear function of X.

As explained in Wooldridge (2010), the coefficient of quantile regression models can be identified by minimizing the following loss function, with respect to $\beta(\tau)$:

$$\hat{\beta}(\tau) = \min_{\beta(\tau)} \sum_{i=1}^{n} \rho_\tau\big(y_i - x_i\beta(\tau)\big) \tag{2}$$

Where $\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$ is the check function, and $I(\cdot)$ is the indicator function.

Most commands for estimating quantile regression models focus on estimating the above loss function, using linear programming techniques, while others like Kaplan and Sun (2017) (`sivqr`) and Chernozhukov et al. (2022) (`qrprocess`) use other optimization techniques.

When no unobserved heterogeneity is present, quantile regression model can be easily implemented in a panel setting (see Wooldridge (2010)), using a pooled version of the model. However, when unobserved heterogeneity is present explicitly, the estimation of quantile regressions is more challenging. Consider the case of panel data such that the conditional quantile regression model is given by:

$$q_\tau(y_{it}|X_{it}, 1_i) = x_{it}\beta(\tau) + \alpha_i(\tau) \tag{3}$$

This specification is explicitly considering that the unobserve effect is identified for each i_{th} observation, and that this effect varies across quantiles $(\alpha_i(\tau))$. A common approach, yet incorrect due to the incidental parameter problem, is to estimate this model by adding dummy variables for each individual in the quantile regression model (as in Budig and England (2001)), or by demeaning the explanatory variables (as in Budig and Hodges (2010)). In fact, there is no transformation of the data that can eliminate the individual fixed effects, as it happens in standard linear regression models.

In this framework, the problem of the incidental parameter problem occurs because the unobserved factors cannot be differenced out of. In other words, unless specific assumptions are made, the estimation requires the explicit estimation of the unobserved fixed effect. Unfortunately, because the number of available observations per individual fixed effect is limited, they cannot be estimated with precision. In turn, the cumulative

errors in the estimation of the fixed effects will also affect the conditional distribution of the outcome, which quantile regressions leverage on, leading to inconsistent estimates of all parameters. [2]

In the next section, we present a few solutions and implementations for the estimation of quantile regression models with multiple fixed effects.

2.1 Correlated Random Effects: CRE

The first approach we discuss is the use of Correlated Random Effects (CRE) models for the estimation of quantile regression models. The CRE model is an alternative methodology for the estimation of fixed effects models that was proposed by Mundlak (1978) and later generalized by Chamberlain (1982). In contrast with standard fixed effects, the approach allows users to control for time fixed covariates in addition to time-varying covariates. And, in contrast with the random effects model, it does not make the assumption that the unobserved effect is uncorrelated with the observed covariates. Interestingly, in the context of linear models, the CRE model is equivalent to the fixed effects model (Wooldridge 2010).

Consider the following model:

$$y_{it} = x_{it}\beta + z_i\gamma + \alpha_i + u_{it} \tag{4}$$

It is well known that if α_i is correlated with x_{it}, the Random Effects (RE) estimator will be inconsistent. This is similar to the ommited variable bias. The solution proposed by Mundlak (1978) and Chamberlain (1982) was to explicitly account for that correlation in the model, by assuming the unobserved effect α_i is a linear projection of the observed time-varying variables. Specifically:

$$
\begin{aligned}
Mundlack: & \quad \alpha_i = \gamma_0 + \bar{x}_i\gamma + v_i \\
Chamberlain: & \quad \alpha_i = \gamma_0 + x_{i1}\gamma_1 + x_{i2}\gamma_2 + \cdots + x_{iT}\gamma_T + v_i
\end{aligned}
\tag{5}
$$

The main difference between both approaches was that Chamberlain's is more flexible by allowing all realizations of the time-varying variables to explain the unobserved effect. In contrast, Mundlak's approach only considers the average of the time-varying variables, which is a more restrictive specification. Using either model specification, if we substitute Equation 5 into Equation 4, the final model can be written as:

$$y_{it} = x_{it}\beta + z_i\gamma + \gamma_0 + f(x_{it})\Gamma + v_i + u_{it} \tag{6}$$

where $f(x_{it})$ can be the full set of time-varying variables or just the average of them. Interestingly, either method provides the same results if the panel data is balanced, and

---

2. This is similar to the measuring error problem of dependent variables in quantile regression models discussed in Hausman et al. (2021).

all covariates are strictly exogenous. However, this identity breaks down in other cases (see Abrevaya (2013)).

The strategy proposed by Abrevaya and Dahl (2008) was to extend the CRE model (Chamberlain (1982) style) for the estimation of quantile regression models. This, however, has some limitations. First, when the number of periods is large, the number of additional regressors grows quickly, which can lead to other problems during estimation. Second, while the application of Chamberlain (1982) projection approach for unbalance data is possible (see Abrevaya (2013)), it is not straightforward to implement in practice, specially for the framework of quantile regression. Instead, we follow Wooldridge (2010) and Wooldridge (2019), and use the Mundlak representation of the CRE model for the estimation of quantile regression models. Wooldridge (2019) has shown that this can be easily applied for cases with unbalanced panels, and the estimation of non-linear models.

Specifically, Wooldridge (2010) suggests that we could estimate the quantile regression model using the Mundlak representation of the CRE model:

$$q_\tau(y_{it}|x_{it}, \bar{x}_i) = x_{it}\beta(\tau) + \bar{x}_i\gamma(\tau)$$

One of the benefits from this approach is that it will not only allow to easily use unbalanced panels, but may also provide an approach to control for multiple fixed effects, as discussed in Baltagi (2023). For example, for a case with two dimensions of fixed effects (say individual and time), we could model each time varying variable as follows:

$$x_{it}^k - E(x_{it}^k) = \lambda_i^{xk} + \lambda_t^{xk} + \varepsilon_{it}$$

Here, we use the centered transformation of the explanatory variable, that is $x_{it}^k - E(x_{it}^k)$, so that all $\lambda's$ have an expected value of zero. Also, λ_i^{xk} and λ_t^{xk} are the equivalent to $\bar{x}_i$ in the Mundlak one-way fixed effects model, and they can be obtained using an iterative process similar to Rios-Avila (2015) or Correia (2016).[3] The final model can be written as:

$$q_\tau(y_{it}|x_{it}, \lambda_i^x, \lambda_t^x) = x_{it}\beta(\tau) + \lambda_i^x\gamma_i(\tau) + \lambda_t^x\gamma_t(\tau)$$

Which could be extended to any number of fixed effects.

2.2 Canay (2011) Estimator

The second approach under consideration is the estimator proposed by Canay (2011). This paper argues that the estimator proposed by Abrevaya and Dahl (2008), and thus the implementation described above, may not provide consistent estimates of the quantile regression coefficients, as long as there is a disturbance ε_{it} left after modeling

---

3. Internally, we use Correia (2016) `reghdfe` to obtain the predicted fixed effects

Equation 5. However, under the assumption that the unobserved effect is a pure location shift, that is α_i does not vary across quantiles, they propose that a two-step estimator can be used to consistently estimate the quantile regression coefficients. The first one eliminating the unobserved fixed effect, and the second one estimating the quantile coefficients.

To better understand how this estimator works, let us consider the following data generating process (DGP):

$$y_i = \beta_0(\rho_i) + \beta_1(\rho_i)x_{1i} + \beta_2(\rho_i)x_{2i}$$

This represents the random coefficient approach to quantile regression models, where the coefficients of the model are allowed to vary as a function of a random variable ρ_i, which follows a uniform distribution. ρ_i is unobserved, but it is assumed to be independent of the covariates. Given this DGP, the conditional quantile function can be written as:

$$q_\tau(y_i|x_{1i}, x_{2i}) = \beta_0(\tau) + \beta_1(\tau)x_{1i} + \beta_2(\tau)x_{2i}$$

In this specification, the impact of x_k on y is allowed to vary across quantiles. However, we could also impose the assumption that some of the coefficients are constant across quantiles. For example, if $\beta_2(\tau) = \beta_2$ for all τ, we would be assuming that x_2 only has a location effect on y. In fact, if all coefficients (but β_0) are constant across quantiles, the model could just as well be estimated using OLS. And under this scenario, it may be convenient to estimate the model imposing this restriction in the model. This is the main idea behind the Canay estimator.

The estimator proposed by Canay (2011) imposes the assumption that the unobserved effect is a pure location shift. More explicit, assumes that the data generating process in a panel data setting, can be written as follows:

$$y_{it} = \beta_0(\rho_{it}) + \beta_1(\rho_{it})x_{it} + \alpha_i$$

According to Canay (2011), under the assumption that α_i is constant across quantiles, we could consistently estimate the quantile regression coefficients $\beta\tau$, by simply transforming the dependent variable as follows:

$$y_{it} - \alpha_i = \tilde{y}_{it} = \beta_0(\rho_{it}) + \beta_1(\rho_{it})x_{it}$$

And then use standard quantile methods on $\tilde{y}_{it}$. More formally, he suggests the following two-step estimator:

1. Using an OLS regression, run the following model:

$$y_{it} = x_{it}\beta + \alpha_i + \varepsilon_{it}$$

and esitmate $\hat{\alpha}_i$.[4]

2. Transform the dependent variable as $\tilde{y}_{it} = y_{it} - \hat{\alpha}_i$, and estimate the quantile regression model:

$$q_\tau(y_{it} - \hat{\alpha}_i|x_{it}) = q_\tau(\tilde{y}_{it}|x_{it}) = x_{it}\beta(\tau)$$

This simple approach allows for the identification of the quantile coefficients, by imposing the assumption that the unobserved effect is a pure location shift. And like the CRE model, it can be extended to multiple fixed effects, as long as one is willing to assume that the unobserved effects are pure location shifts. For example, consider a case with two fixed effects dimensions (individual and time), the model could be written as:

$$y_{it} = x_{it}\beta + \alpha_i + \alpha_t + \varepsilon_{it}$$

As before α_i and α_t, assumed constant across quantiles, could be estimated using OLS, and the quantile regression model estimated using the transformed dependent variable.

$$q_\tau(y_{it} - \hat{\alpha}_i - \hat{\alpha}_t|x_{it}) = q_\tau(\tilde{y}_{it}|x_{it}) = x_{it}\beta(\tau)$$

Which again can be easily estimated using standard quantile regression methods.

2.3 Modified Canay(2011) Estimator

Perhaps one of the main limitations of the Canay estimator is that it assumes that the unobserved effect is a pure location shift. If this assumption is violated, it may lead to inconsistent estimates of the quantile coefficients. To address this limitation, we propose a small modification to the Canay estimator.

We start by assuming that the unobserved effect represents some characteristics of the individual that are constant across quantiles, and that can be compared across individuals. Under this consideration, the data generating process can be written as:

$$y_{it} = \beta_0(\rho_{it}) + \beta_1(\rho_{it})x_{it} + \gamma(\rho_{it})\alpha_i$$

Next, similar to Canay (2011), we propose to estimate the unobserved effect α_i using OLS, and assume that, in average the $E[\gamma(\rho_{it})] = 1$. Up to this point, the estimator is the same as the Canay estimator. However, we propose that instead of imposing

---

4. Empirically, this can be done using `reghdfe` command.

$\gamma(\rho_{it})\alpha_i$ to be constant across quantiles, we allow for the prediction of the unobserved effect α_i to vary across quantiles. This can be done by estimating the following model:

$$q_\tau(y_{it}|x_{it}, \hat{\alpha}_i) = x_{it}\beta(\tau) + \gamma(\tau)\hat{\alpha}_i$$

As before, this model can be extended to multiple fixed effects, by simply estimating the unobserved effects using OLS, and then estimating the quantile regression model using the predicted unobserved effects. The main advantage over Canay (2011) is that this estimator allows for the unobserved effect to have a different impact on the dependent variable across quantiles, which may be more realistic in many applications. However, it assumes the OLS estimator does allow for the consistent estimation of an unobserved effect that is comprarable across individuals, which may not always be the case.

2.4   Method of Moments Quantile Regression Machado and Santos Silva (2019)

This approach of estimating regression quantiles with multiple fixed effects distinguishes itself from the other estimators reviewed above in the sense that (Machado and Santos Silva 2019) introduce location-scale effects of fixed effects upon the distribution of interest. Compared to Canay's estimator, the Method of Moments Quantile Regression (MMQREG) not only allows the α_i's to affect Y_{it} through location shifts, but rather MMQREG is able to identify the scale shifts that alter different points of the distribution belonging to Y.

We begin by defining the DGP of the location scale model:

$$Y_{it} = \alpha_i + X'_{it}\beta + (\delta_i + X'_{it}\gamma)u_{it} \tag{7}$$

Where parameters α_i and δ_i capture the individual fixed effects. Note that, compared to equation [eq:dgp_canay], the fixed effects not only enter the model in an additive fashion, instead they also have a multiplicative effect upon the error term. In addition, the U_{it} are i.i.d. across i and t, statistically independent of X_{it} and satisfy $E(U) = 0$ and $E(|U|) = 1$ both of which normalize the random variable.

Our location scale model in equation [eq:dgp_mmqreg] implies that:

$$Q_\tau(Y_{it}|X_i) = [\alpha_i + \delta_i q(\tau)] + X'_{it}\beta + X'_{it}\gamma q(\tau)$$

Where the scalar coefficient $\alpha_i(\tau) \equiv \alpha_i + \delta_i q(\tau)$ is the quantile-τ fixed effect for individual i, which represent how time-invariant variables have *different impacts on different regions* of the conditional distribution of Y_{it}. However, our real interest is in the regression quantile coefficients:

$$\beta_\tau = \beta + q(\tau)\gamma$$

Which are simply a linear combination of the location coefficients (β) and the scale coefficients (γ), where the second vector of coefficients is weighted by the value of the τ-th quantile of the variable of interest Y_{it}. (Machado and Santos Silva 2019) develop the following algorithm for implementing the MMQREG estimator:

1. Obtain $\hat{\beta}_k$ by regressing time-demeaned Y_{it} on time-demeaned controls X_{it}, i.e. obtain $\hat{\beta}$ by the within estimator.

2. Estimate the $\hat{\alpha}_i$'s and calculate the residuals $\hat{R}_{it} = Y_{it} - \hat{\alpha}_i - X'_{it}\hat{beta}$.

3. Obtain $\hat{\gamma}_k$ by the within estimator using $|\hat{R}_{it}|$ as the dependent variable.

4. Estimate $\hat{\delta}_i$ by taking the time average of $|\hat{R}_{it}| - X'_{it}\hat{\gamma}$.

5. Estimate $q(\tau)$ by $\hat{q}$, which corresponds to the regression quantile of standardized residuals $[\hat{R}_{it}/(\hat{\gamma}_i + X'_{it}\hat{\gamma})]$ upon an intercept term.

Note that Steps 1 and 2 from the MMQREG algorithm are the same as those performed in Canay's estimator. However, the location-scale model used in (Machado and Santos Silva 2019) adds three additional steps that are required to estimate $\hat{\gamma}$ and $\hat{q}$, so that regression quantile coefficients β_τ are allowed to affect not only the location of the distribution, but also its shape.

Add notes on the assumptions of MMQREG here *if* necessary.

Statistical inference can be performed using the asymptotic distribution of the estimator derived in (Machado and Santos Silva 2019). Expanding on this literature, (Rios-Avila et al. 2024) propose methods for computing alternative standard errors using the empirical influence functions of the estimators. Robust and clustered standard errors can be estimated following this approach, and are readily available as options in the `mmqreg` command in Stata. Another extension of the MMQREG estimator due to (Rios-Avila et al. 2024) is to allow for the inclusion of multiple fixed effects using an application of the Frisch-Waugh-Lovell (FWL) theorem to partial out the effect of variables capturing unobserved heterogeneity from both dependent and explanatory variables. Likewise, the command `mmqreg` allows for multiple fixed effects, as will be shown next.

Table 1: Government spending and coalition government - MMQREG Estimator

Column 1	Column 2	Column 3
Data 1	Data 2	Data 3
Data 4	Data 5	Data 6

As we allow for both location and scale shifts in our estimation method, the monotonic pattern observed in the first row of table 3 is reversed. Now, the farther we move from the median of the distribution in the direction of greater government spending, the lower is the effect of a permanent change from single-party to coalition upon spending. The median coefficient, once again, remains close to the mean coefficient (2.36) under MMQREG estimation. Results from table 4 support the conclusions from (Persson et al. 2007) regarding types of government and its spending during the last year of legislature, expanding the results to the response of the distribution of spending upon changes in explanatory variables.

We must also note that statistical inference —with robust standard errors displayed in table 4— for the (Persson et al. 2007) dataset show that many coefficients that with the previous estimators were not statistically different from zero, are now significant even at the 99% confidence level. This is due to the ability to compute robust standard errors for the MMQREG estimator, contrasting with the CRE and Canay estimators, for which we estimated the covariance matrix using the bootstrap. Although not displayed here, our findings are robust to clustering standard errors for country so that the coefficients of 25th and 50th quantiles corresponding to the indicator variable for coalition government remain statistically significant.

3 References

Abrevaya, J. 2013. The projection approach for unbalanced panel data. *The Econometrics Journal* 16(2): 161–178. http://dx.doi.org/10.1111/j.1368-423X.2012.00389.x.

Abrevaya, J., and C. M. Dahl. 2008. The Effects of Birth Inputs on Birthweight. *Journal of Business & Economic Statistics* 26(4): 379–397.

Baltagi, B. H. 2023. The two-way Mundlak estimator. *Econometric Reviews* 42(2): 240–246. https://www.tandfonline.com/doi/full/10.1080/07474938.2023.2178139.

Budig, M. J., and P. England. 2001. The Wage Penalty for Motherhood. *American Sociological Review* 66(2): 204–225. http://dx.doi.org/10.1177/000312240106600203.

Budig, M. J., and M. J. Hodges. 2010. Differences in Disadvantage. *American Sociological Review* 75(5): 705–728. http://dx.doi.org/10.1177/0003122410381593.

Canay, I. A. 2011. A simple approach to quantile regression for panel data. *The Econometrics Journal* 14(3): 368–386.

Chamberlain, G. 1982. Multivariate regression models for panel data. *Journal of Econometrics* 18(1): 5–46.

Chernozhukov, V., I. Fernández-Val, and B. Melly. 2022. Fast algorithms for the quantile regression process. *Empirical Economics* 62(1): 7–33. https://doi.org/10.1007/s00181-020-01898-0.

Correia, S. 2016. A Feasible Estimator for Linear Models with Multi-Way Fixed Effects. *Unpublished Manuscript* .

Galvao, A. F., and K. Kengo. 2017. Quantile regression methods for longitudinal data. In *Handbook of quantile regression*, 363–380. Chapman and Hall/CRC.

Hausman, J., H. Liu, Y. Luo, and C. Palmer. 2021. Errors in the Dependent Variable of Quantile Regression Models. *Econometrica* 89(2): 849–873. http://dx.doi.org/10.3982/ECTA14667.

Kaplan, D. M., and Y. Sun. 2017. Smoothed estimating equations for instrumental variables quantile regression. *Econometric Theory* 33(1): 105–157. Publisher: Cambridge University Press.

Koenker, R., and G. Bassett. 1978. Regression Quantiles. *Econometrica* 46(1): 33–50.

Lancaster, T. 2000. The incidental parameter problem since 1948. *Journal of Econometrics* 95(2): 391–413.

Machado, J. A., and J. Santos Silva. 2019. Quantiles via moments. *Journal of Econometrics* 213(1): 145–173.

Melly, B., and M. Pons. 2023. Minimum Distance Estimation of Quantile Panel Data Models. Unpublished working paper.

Mundlak, Y. 1978. On the Pooling of Time Series and Cross Section Data. *Econometrica* 46(1): 69. http://dx.doi.org/10.2307/1913646.

Neyman, J., and E. L. Scott. 1948. Consistent Estimates Based on Partially Consistent Observations. *Econometrica* 16(1): 1–32.

Persson, T., G. Roland, G. Tabellini, et al.. 2007. Electoral rules and government spending in parliamentary democracies. *Quarterly Journal of Political Science* 2(2): 155–188.

Rios-Avila, F. 2015. Feasible Fitting of Linear Models with N Fixed Effects. *The Stata Journal* 15(3): 881–898. Publisher: SAGE Publications.

Rios-Avila, F., L. Siles, and G. Canavire-Bacarreza. 2024. Estimating Quantile Regressions with Multiple Fixed Effects through Method of Moments. *Working Paper* .

Wooldridge, J. M. 2010. *Econometric analysis of cross section and panel data*. MIT press.

———. 2019. Correlated Random Effects Models with Unbalanced Panels. *Journal of Econometrics* 211(1): 137–150.

About the authors

Fernando Rios-Avila is a Research Scholar at the Levy Economics Institute of Bard College.

Gustavo Canavire-Bacarreza is a Senior Economist at the World Bank.