

# Testing the Public Goods Theory of Alliances

Joshua Alley\*

January 19, 2019

## 1 Other Estimates of Panel Data Interaction

### 1.1 Alternative Estimators

Robust regression is appropriate for residuals from the growth variable. Several observations of states during war see gigantic increases in spending—the largest value is 140, relative to a median of .063. This generates extremely skewed residuals, so OLS is inefficient.

The skewness of the residuals is readily apparent in Figure 1. Many positive residuals deviate from the normality assumption.

Although OLS is not the best estimator for the spending growth data, it produces similar results. Table 1 summarizes the OLS estimates. The coefficient on changes in allied spending is negative but statistically insignificant. The negative coefficient on GDP indicates the limited validity of these results, but there is still no evidence size modifies the association between changes in allied spending and growth in military expenditures.

The OLS model is a poor fit for the dependent variable. The  $R^2$  is miniscule, and the standard error of the residual is much larger than the robust regression. Despite the lackluster fit, I present these results to show that the conventional estimator generates similar inferences.

---

\*Graduate Student, Department of Political Science, Texas A&M University.

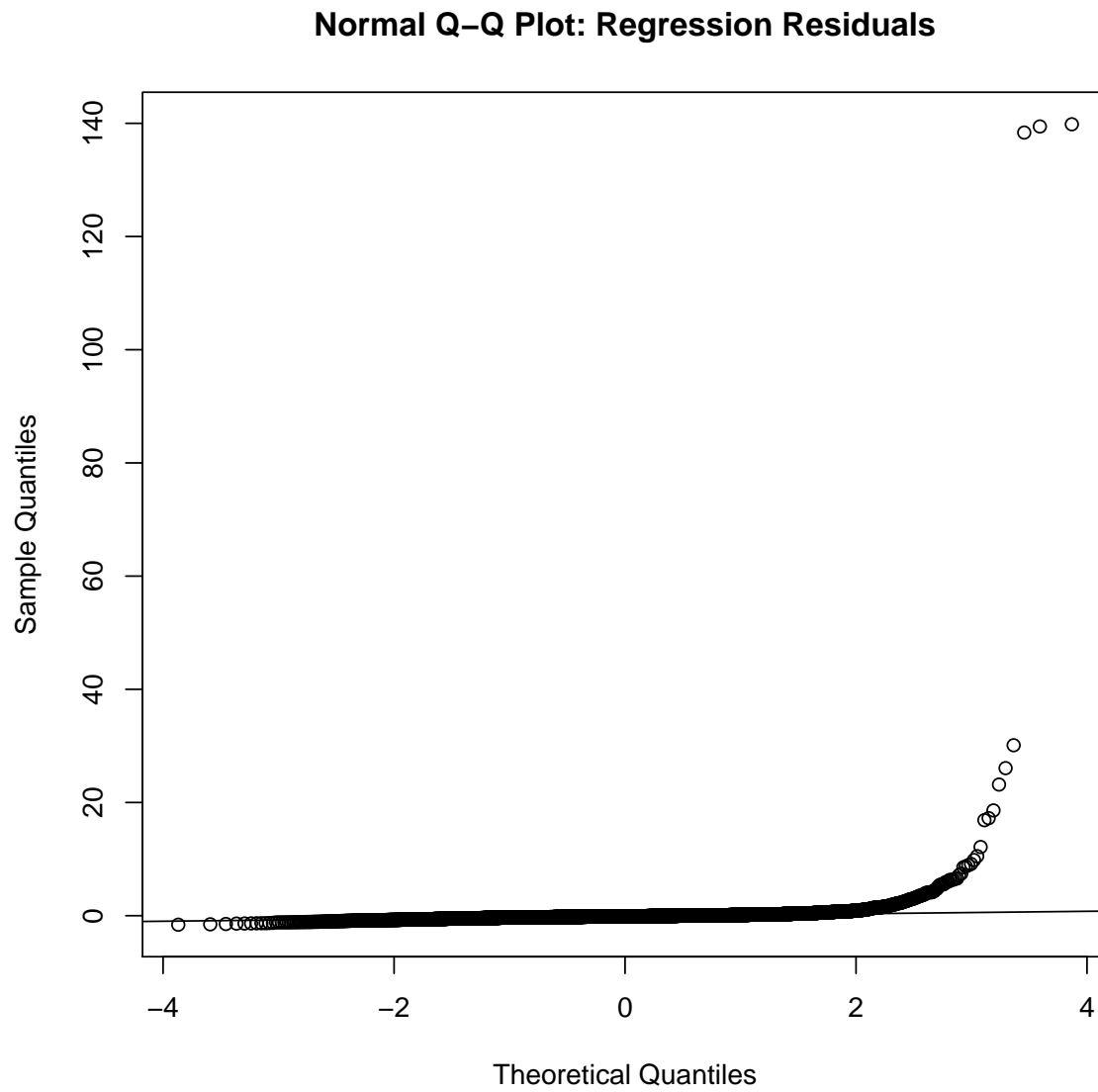


Figure 1: Plot of residuals against normal quantiles. Deviations from the straight line are deviations from the normal distribution.

	Growth Military Spending
Change Allied Spending	−0.235 (0.648)
ln(GDP)	−0.038** (0.017)
Change Allied Spending $\times$ ln(GDP)	0.011 (0.027)
Average Alliance Size	−0.003 (0.002)
Average Alliance Democracy	−0.001 (0.008)
International War	0.476*** (0.141)
Civil War Participant	0.206** (0.104)
POLITY	0.001 (0.005)
External Threat	0.107 (0.152)
Cold War	0.032 (0.058)
Constant	1.074*** (0.413)
Observations	9,139
R <sup>2</sup>	0.003
Residual Std. Error	2.649 (df = 9128)
F Statistic	2.506*** (df = 10; 9128)
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Table 1: OLS estimates of conditional relationship between allied spending, GDP and growth in military spending.

I also checked whether inferences were sensitive to non-random selection into alliances. I employed a Heckman selection model, where the second-stage estimation sample was states in alliances. The second-stage model used the same specification as the robust regression.

A first-stage probit model predicts the presence of at least one alliance. The first stage model includes lagged levels of military spending, the log of GDP, regime type, international conflict, and external threat as covariates. Table 2 summarizes the output of the second-stage equation.

In the selection model, change in allied spending has no association with growth in spending. GDP is positively associated with military expenditures, but the interaction term remains statistically insignificant. Even after accounting for non-random selection into alliances, there is little evidence of a conditional relationship between alliance spending and state size.

## **1.2 Continuous modifying Variable**

Hainmueller, Mummolo and Xu (2019) show linearity assumptions and a lack of support in the range of the modifying variable can generate misleading inferences in interactive models. They suggest estimating associations with binning and kernel estimators to check for non-linearity and adequate support. Figure 2 plots the results of the binning estimator.

The binning estimator suggests that the linearity assumptions of the robust regression are acceptable. There is still no evidence state size modifies the impact of allied spending on growth in military spending, however. The marginal effect of allied spending cannot be distinguished from zero across the range of GDP.

The kernel estimator also generates little evidence of a conditional relationship. As Figure 3 demonstrates, the marginal effect of spending is mostly positive. Negative point estimates of the marginal effect are indistinguishable from zero. Relaxing the linearity assumptions of the robust regression does little to change substantive inferences about the presence of a conditional relationship.

	Growth Military Spending
Change Allied Spending	−0.304 (0.266)
ln(GDP)	0.201*** (0.017)
Change Allied Spending $\times$ ln(GDP)	0.013 (0.011)
International Conflict	0.362*** (0.140)
Civil War Participant	0.005 (0.049)
POLITY	−0.022*** (0.004)
External Threat	−0.098 (0.151)
Cold War	0.102*** (0.026)
Constant	−5.443*** (0.416)
Observations	9,235

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 2: Heckman Sample Selection Regression Estimates: Allied Spending, GDP, and Growth in military spending 1816-2007.

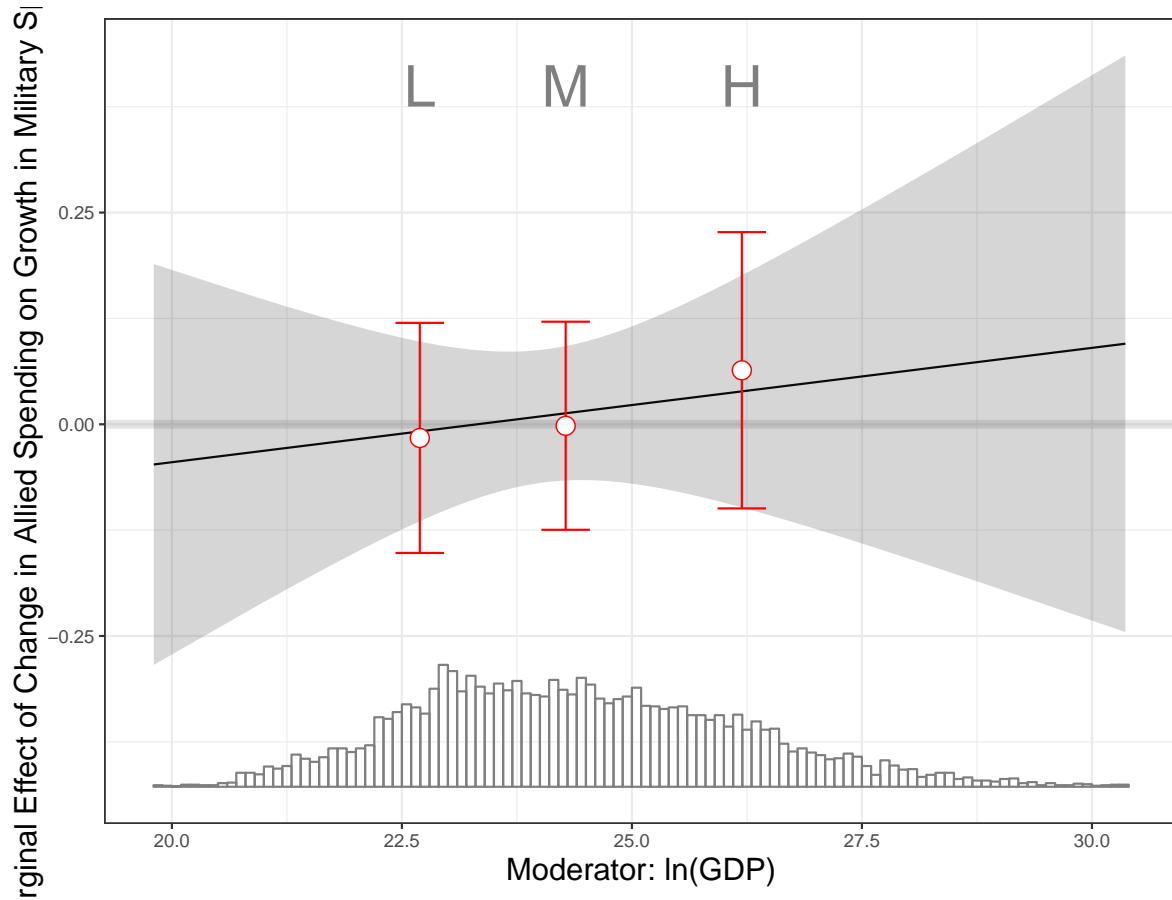


Figure 2: Binning estimates of interaction between changes in allied spending and GDP.

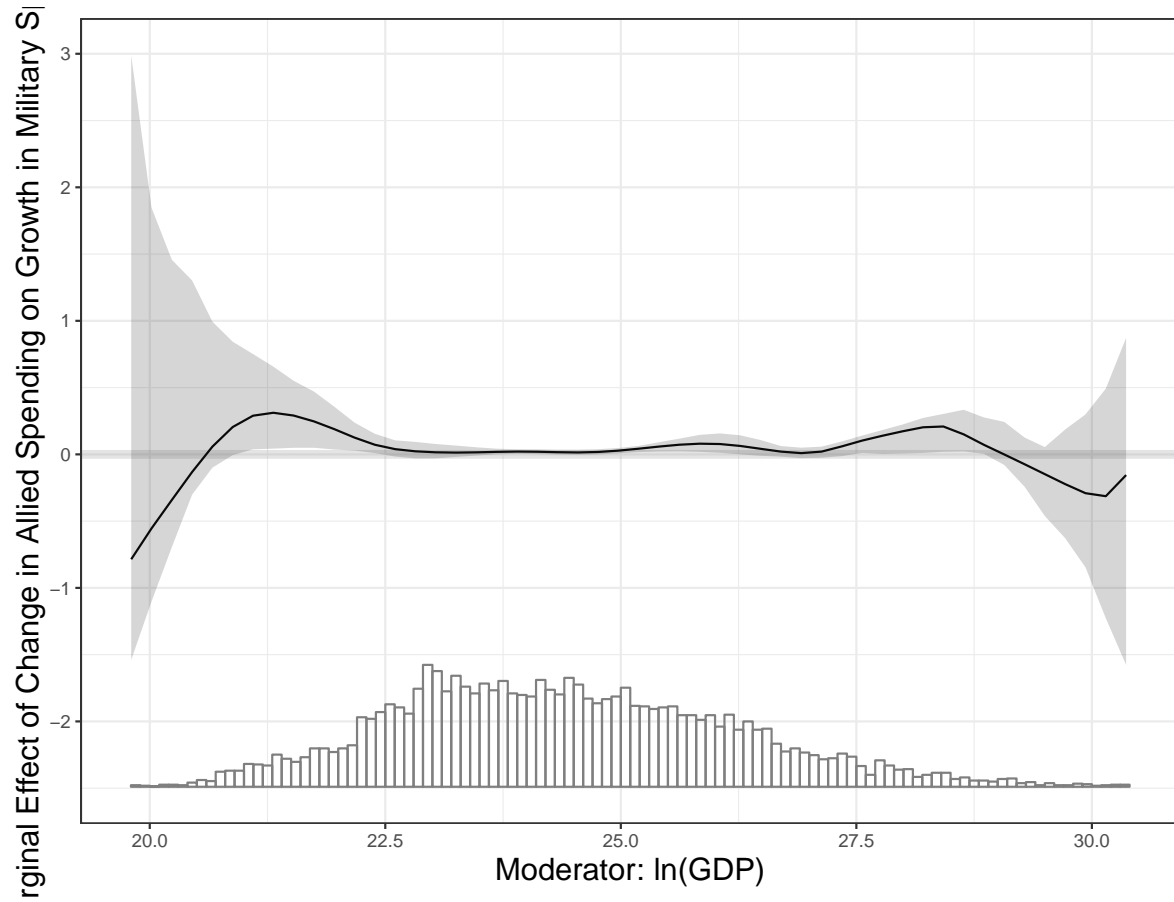


Figure 3: Kernel estimate of interaction between changes in allied spending and GDP.

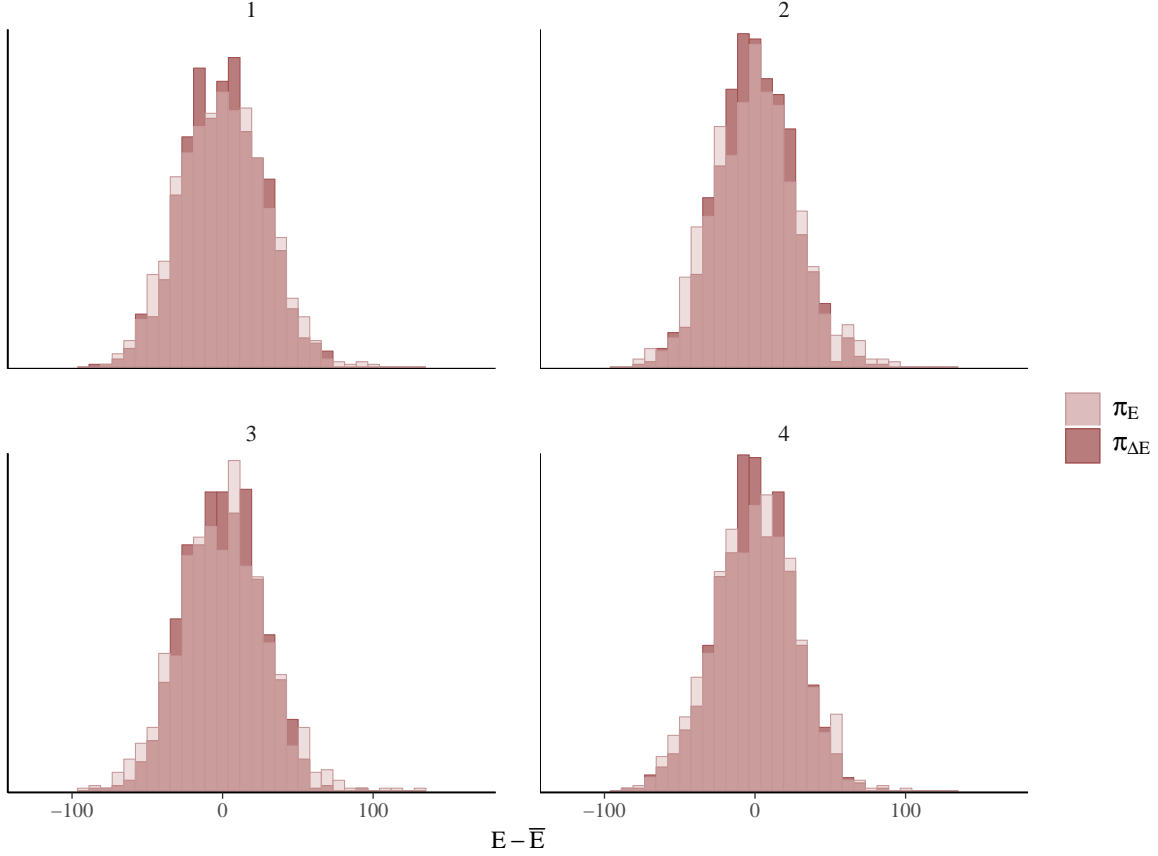


Figure 4: Energy plot of multilevel model results. Greater overlap in the two histograms indicates adequate exploration of the posterior distribution.

## 2 Convergence of Multilevel Model

There were no divergent iterations in sampling, which would have invalidated the results. However, there are other threats to inference from the posterior samples. Given heavy tails in military spending growth, STAN might have struggled to explore the posterior distribution.

Energy plots can diagnose this problem. Figure 4 plots the marginal energy distribution and the first differenced distribution. If the two histograms do not overlap, sampling was impeded by heavy tails. The substantial overlap in the distributions for all four chains in Figure 4 indicates this was not a problem.

The split  $\hat{R}$  statistic is another way to assess convergence.  $\hat{R}$  compares the behavior of each



chain by measuring the ratio of the average variance of draws within each chain to the variance of the pooled draws across chains. When  $\hat{R}$  is close to 1, all the chains have similar variance, and are therefore in equilibrium.

The standard heuristic is that an  $\hat{R}$  greater than 1.1 is problematic. Figure 5 plots the  $\hat{R}$  statistic for every parameter in the model. No parameters generate concern, even at a more conservative threshold of 1.05.

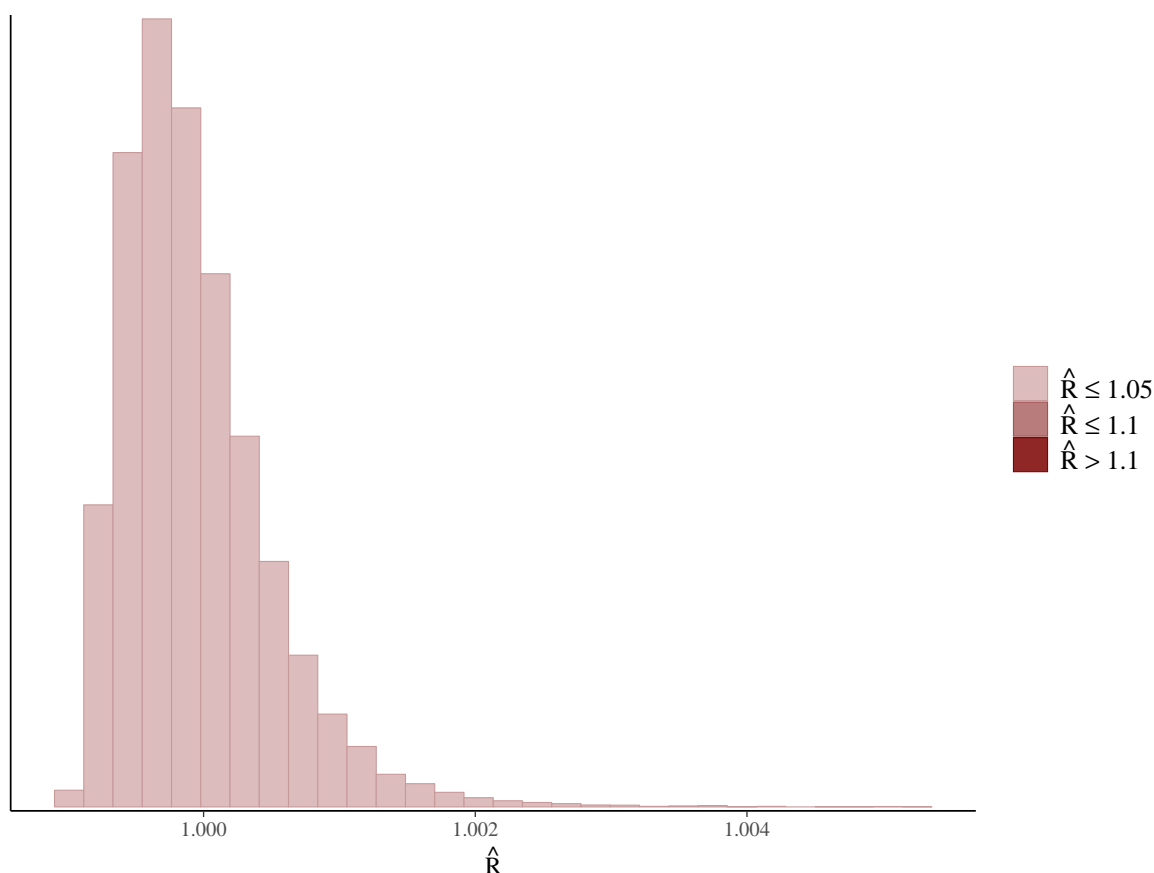


Figure 5: Histogram of split  $\hat{R}$  statistic for all parameters in the multilevel model.

## 2.1 Multilevel Model Priors

All priors are specified to be weakly informative relative to the scale of the data (Gelman, Simpson and Betancourt, 2017). I summarize the prior distributions for each set of parameters in

$$\begin{aligned}
p(\alpha) &\sim N(0, 1) \\
p(\sigma) &\sim \text{half-}N(0, 1) \\
p(\alpha^{yr}) &\sim N(0, \sigma^{yr}) \\
p(\sigma^{yr}) &\sim N(0, 1) \\
p(\alpha^{st}) &\sim N(0, \sigma^{st}) \\
p(\sigma^{st}) &\sim \text{half-}N(0, 1) \\
p(\sigma^{all}) &\sim \text{half-}N(0, 1) \\
p(\beta) &\sim N(0, 1) \\
p(\gamma) &\sim N(0, 1) \\
p(\nu) &\sim \text{gamma}(2, 0.1)
\end{aligned}$$

Table 3: Summary of Priors in Multilevel Model

Table 3.  $p(\nu)$  a well-behaved prior for the degrees of freedom in a t-distribution (Juárez and Steel, 2010). Given that median growth in military expenditures is 0.06, the standard normal priors on the  $\beta$  and  $\gamma$  parameters are quite diffuse.

To facilitate estimation, I use a non-centered parameterization for the state and year varying intercepts (Betancourt and Girolami, 2015). The prior is equivalent, but a non-centered parameterization decouples the mean and variance, which makes sampling easier. I also employ a sparse matrix representation of the alliance membership matrix  $\mathbf{Z}$  to speed up estimation.

## References

- Betancourt, Michael and Mark Girolami. 2015. Hamiltonian Monte Carlo for hierarchical models. In *Current Trends in Bayesian Methodology with Applications*, ed. Satyanshu K. Upadhyay, Umesh Singh, Dipak K. Dey and Appaia Loganathan. Chapman and Hall/CRC Press pp. 79–102.
- Gelman, Andrew, Daniel Simpson and Michael Betancourt. 2017. “The prior can generally only be understood in the context of the likelihood.” *arXiv preprint arXiv:1708.07487*.
- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. “How Much Should We Trust Estimates from Multiplicative Interaction Models?: Simple Tools to Improve Empirical Practice.” *Political Analysis*.
- Juárez, Miguel A and Mark FJ Steel. 2010. “Model-Based Clustering of Non-Gaussian Panel Data Based on Skew-t Distributions.” *Journal of Business & Economic Statistics* 28(1):52–66.