The Effect of Unionism on Accidents in U.S. Coal Mining, 1897-1929: Errata

William M. Boal

College of Business and Public Administration
Drake University, Des Moines, IA 50311
Telephone 515-271-3129
Email william.boal@drake.edu

June 2016

In 2009 I published a paper in this journal entitled, “The Effect of Unionism on Accidents in U.S. Coal Mining, 1897-1929.” In that paper I analyzed two panel data sets— a state-level data set where unionism was measured by membership, and a mine-level data set confined to West Virginia where unionism was measured by contract coverage. Applying count-data regression methods to both data sets, I estimated that unionism reduced fatalities in coal mining by “about 40 percent.” However, I recently found errors in the estimates in that paper. In this note I present and discuss corrected estimates.

State-level estimates

Table 3 in my 2009 paper reported estimates of the union effect on fatalities at the state level, computed by estimating NB1 models with state fixed effects. However, I constrained the NB1 dispersion parameter to be constant across states, whereas common practice allows it to vary across states (Cameron and Trivedi 2013, p. 357). In effect, I estimated cross-sectional NB1 models with dummy variables for states rather than true fixed-effects panel NB1 models. I also computed standard errors robust to heteroscedasticity but not to serial correlation. When the NB1 dispersion parameter is allowed to vary across states and standard errors robust to serial correlation are computed (using STATA command *xtnbreg* with cluster-bootstrap standard errors) the results are quite different, as shown in corrected table 3 below. In all columns, the estimated coefficient of unionism is closer to zero (and occasionally positive) and all standard errors are considerably larger. The estimated coefficient of unionism is no longer statistically significant at conventional levels.

In my 2009 paper, I justified the choice of the NB1 model over the simpler Poisson model based on overdispersion: formal tests rejected the Poisson property that the conditional variance equaled the conditional mean. The NB1 model accounts for overdispersion and is more efficient than the Poisson if the true distribution is NB1. However, the Poisson is still consistent provided the conditional mean is correctly specified (Wooldridge 2010, p. 763; Cameron and Trivedi 2013, p. 357).[1] So for comparison, table 3P shows Poisson estimates (computed using STATA command *xtpoisson* with standard errors clustered on states).[2] In all columns, the Poisson estimates of the coefficient of unionism are negative, slightly larger than my 2009 NB1 estimates, and at least twice their standard errors.

Which corrected estimates are preferred—the NB1 estimates in corrected table 3 or the Poisson estimates in table 3P? If the true distribution is NB1, then conditional maximum-likelihood estimates which impose this distribution should have smaller standard errors, but here in general they do not. A comparison of the two tables shows that the standard errors for the union coefficient are about 30 percent *larger* for the NB1 estimates. Moreover, the bootstrap-robust standard errors in table 3 are about twice as large as the default standard errors (not shown in the table) which are computed assuming the NB1 information matrix equality. These anomalous results suggest that the assumptions required for consistency of the NB1 estimates may not be satisfied (Cameron and Trivedi 2010, p. 641). So on balance, I am inclined to favor the Poisson estimates in table 3P. While the point estimates in table 3P are slightly different from those in table 3 of my 2009 paper, and the standard errors are about 50 percent larger, my earlier conclusion of a reduction in fatalities of “about 40 percent” at the state level due to unionism still seems plausible.

Mine-level estimates

Table 6 in my 2009 paper reported estimates of the union effect on fatalities at the mine level, computed by estimating NB1 models with mine fixed effects. Although I did not constrain the NB1 dispersion parameter in the mine-level analysis, I believe I made other errors. To compute the estimates in table 6, I used TSP software, which is highly accurate. However, TSP offers no canned procedures for panel count-data regression, so I was forced to code the NB1 conditional likelihood function[3] myself and to maximize that function using TSP's *ML* procedure. I recently recomputed the estimates using STATA, which does offer canned procedures, with somewhat different results.[4] Corrected table 6 shows point estimates and robust standard errors computed using STATA command *xtnbreg* with cluster-bootstrap standard errors. I have not yet found errors in my 2009 TSP code, but I believe the experts at STATA are better programmers than I am, so I place my bets on these corrected results. The estimates of the coefficient of unionism are slightly smaller than my 2009 estimates but otherwise fairly similar. [5] Based on results in this corrected table, I would conclude that the reduction in fatalities due to unionism was closer to 30 percent than the 40 percent claimed in my 2009 paper.

Are these NB1 estimates trustworthy? For comparison, table 6P shows Poisson estimates (computed using STATA command *xtpoisson* with standard errors clustered on mines). Except for the first column, the Poisson estimates of the union effect in table 6P are much smaller than the NB1 estimates in corrected table 6 (and some of the Poisson estimates are positive). Which estimates are preferred? As noted above, NB1 estimates are more efficient than Poisson if the true distribution is NB1. A comparison of the two mine-level tables shows that the standard errors for the union coefficient are much *smaller* for the NB1 estimates, in contrast to the state results. Moreover, the bootstrap-robust standard errors in corrected table 6 are only about twenty percent larger than the default standard errors (not shown in the table) which are computed assuming the NB1 information matrix equality. While they do not constitute a formal test, these results (in contrast with those for the state-level data) are compatible with the true distribution being NB1. So on balance, I am inclined to favor the NB1 estimates in corrected table 6.

Summary

In this note, I have presented corrected estimates of the effect of unionism on fatalities in coal mining with cluster-robust standard errors. The mine-level estimates are smaller and many of the standard errors (now allowing for serial correlation) are larger than those reported in my 2009 paper. The corrected estimates still support my 2009 claim that unionism reduced fatalities in coal mining, but that support is admittedly weaker and less robust across specifications.

References

Boal, William M. 2009. "The Effect of Unionism on Accidents in U.S. Coal Mining, 1897-1929." Industrial Relations 48 (1): 97-120.

Cameron, A. Colin and Pravin K. Trivedi. 2010. Microeconometrics Using Stata. College Station, Texas: Stata Press.

Cameron, A. Colin and Pravin K. Trivedi. 2013. Regression Analysis of Count Data. Cambridge: Cambridge University Press.

Wooldridge, Jeffrey M. 2010. Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts: MIT Press.

Corrected Table 3:
NB1 ML estimates of coal mining fatalities for state-level data, 1902-1929

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of workers who are members of the UMWA	0.015 (0.225)	0.008 (0.242)	-0.069 (0.244)	-0.078 (0.345)	-0.182 (0.340)	-0.512 (0.376)
Log total number of workers	0.814 (0.165)	0.878 (0.152)	0.849 (0.151)	0.830 (0.167)	0.858 (0.185)	1.052 (0.205)
Log average number of days worked	0.662 (0.191)	0.672 (0.183)	0.702 (0.180)	0.683 (0.205)	0.799 (0.243)	0.917 (0.255)
Fraction of output mined by machine	0.282 (0.262)	0.308 (0.293)	0.292 (0.323)	0.351 (0.405)	0.176 (0.430)	1.036 (0.716)
Number of important mine safety regulations		-0.041 (0.027)	-0.038 (0.027)	-0.042 (0.033)	-0.037 (0.032)	-0.054 (0.045)
Law increasing employer liability (dummy variable)			0.244 (0.199)	0.273 (0.227)	0.127 (0.237)	0.164 (0.242)
Workers' Compensation law (dummy variable)			0.186 (0.205)	0.206 (0.237)	0.123 (0.233)	0.180 (0.237)
Fraction union during Lewis era [a]			0.253 (0.176)	0.187 (0.194)	0.298 (0.234)	0.378 (0.215)
Fraction of output strip-mined				0.808 (0.837)	0.667 (0.808)	1.318 (1.075)
Fraction of workers on 8-hour day					-0.135 (0.396)	-0.032 (0.342)
Fraction of workers on 9-hour day					-0.411 (0.435)	-0.074 (0.494)
Fraction of workers on 10-hour day					-0.340 (0.407)	-0.007 (0.356)
Fraction of output "shot off the solid"						1.254 (0.807)
Number of states	25	24	24	24	24	24
Number of (state x year) observations	394	383	383	333	318	206
Log-likelihood function	-1504.0	-1478.3	-1472.9	-1270.4	-1188.7	-714.3

[a] Fraction of workers who are members of the UMWA times a dummy variable for year 1919 or later.

State fixed effects and year fixed effects included in all equations.
Cluster-bootstrap standard errors based on 400 replications in parentheses.

Table 3P:
Poisson ML estimates of coal mining fatalities for state-level data, 1902-1929

Regressor	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of workers who are members of the UMWA	-0.421	-0.381	-0.458	-0.673	-0.830	-0.992
	(0.170)	(0.170)	(0.179)	(0.242)	(0.270)	(0.332)
Log total number of workers	1.006	1.022	1.089	1.177	1.044	1.052
	(0.159)	(0.166)	(0.183)	(0.141)	(0.126)	(0.163)
Log average number of days worked	0.838	0.851	0.909	0.803	0.940	0.988
	(0.094)	(0.100)	(0.098)	(0.157)	(0.222)	(0.241)
Fraction of output mined by machine	0.003	0.288	0.372	0.274	-0.034	0.918
	(0.286)	(0.260)	(0.295)	(0.384)	(0.403)	(0.550)
Number of important mine safety regulations		-0.083	-0.082	-0.089	-0.065	-0.056
		(0.031)	(0.035)	(0.040)	(0.031)	(0.037)
Law increasing employer liability (dummy variable)			-0.047	-0.005	0.007	0.146
			(0.130)	(0.117)	(0.104)	(0.130)
Workers' Compensation law (dummy variable)			-0.076	-0.038	0.030	0.084
			(0.156)	(0.158)	(0.162)	(0.159)
Fraction union during Lewis era [a]			0.233	0.220	0.425	0.468
			(0.163)	(0.174)	(0.228)	(0.233)
Fraction of output strip-mined				0.807	0.380	0.871
				(1.073)	(0.989)	(0.982)
Fraction of workers on 8-hour day					-0.140	0.081
					(0.189)	(0.134)
Fraction of workers on 9-hour day					-0.376	0.148
					(0.263)	(0.213)
Fraction of workers on 10-hour day					-0.410	0.019
					(0.276)	(0.212)
Fraction of output "shot off the solid"						1.230
						(0.831)
Number of states	25	24	24	24	24	24
Number of (state x year) observations	394	383	383	333	318	206
Log-likelihood function	-2856.4	-2732.3	-2716.4	-2426.2	-2135.1	-1296.2

[a] Fraction of workers who are members of the UMWA times a dummy variable for year 1919 or later.

State fixed effects and year fixed effects included in all equations.
Cluster-robust standard errors in parentheses.

Corrected Table 6:
NB1 ML estimates of coal mining fatalities for West Virginia mine-level data, 1897-1928

Regressor	(1)	(2)	(3)	(4)	(5)
Mine under union operation (dummy variable)	-0.312 (0.062)	-0.315 (0.064)	-0.320 (0.065)	-0.206 (0.098)	-0.217 (0.098)
Log total number of workers	0.410 (0.065)	0.411 (0.065)	0.412 (0.066)	0.410 (0.065)	0.412 (0.065)
Log number of days worked	0.527 (0.070)	0.527 (0.070)	0.527 (0.071)	0.480 (0.072)	0.490 (0.073)
Annual time trend	0.005 (0.004)	0.004 (0.006)	0.003 (0.007)	0.012 (0.008)	0.017 (0.009)
Fraction of output mined by machine	-0.037 (0.113)	-0.038 (0.113)	-0.039 (0.113)	-0.041 (0.113)	-0.048 (0.113)
Number of important mine safety regulations		0.005 (0.026)	0.002 (0.027)	0.043 (0.031)	0.029 (0.033)
Workers' Compensation law			0.030 (0.076)	-0.043 (0.076)	-0.082 (0.076)
World War I dummy variable [a]				-0.137 (0.097)	-0.138 (0.097)
Postwar (Lewis era) dummy variable [b]				-0.228 (0.104)	-0.227 (0.104)
Union dummy during Lewis era [c]				-0.185 (0.133)	-0.208 (0.137)
Statewide tonnage under union contract (fraction)					0.184 (0.171)
Number of mines	443	443	443	443	443
Number of (mine x year) observations	6870	6870	6870	6870	6870
Log-likelihood function	-5841.7	-5841.7	-5841.6	-5835.4	-5834.8

[a] Equals one for 1917 and 1918, and equals zero otherwise.
[b] Equals zero through 1918, and equals one thereafter.
[c] Equals zero through 1918, and equals union dummy variable thereafter.

Mine fixed effects included in all equations.
Cluster-bootstrap standard errors based on 400 replications in parentheses.

Table 6P:
Poisson ML estimates of coal mining fatalities for West Virginia mine-level data, 1897-1928

Regressor	(1)	(2)	(3)	(4)	(5)
Mine under union operation	-0.332	-0.207	-0.190	0.432	0.387
(dummy variable)	(0.315)	(0.368)	(0.326)	(0.400)	(0.380)
Log total number of workers	0.348	0.299	0.298	0.286	0.288
	(0.273)	(0.256)	(0.256)	(0.253)	(0.254)
Log number of days worked	0.425	0.447	0.446	0.293	0.327
	(0.223)	(0.223)	(0.222)	(0.230)	(0.232)
Annual time trend	-0.017	0.021	0.023	0.019	0.030
	(0.012)	(0.022)	(0.016)	(0.014)	(0.016)
Fraction of output mined by machine	0.000	0.010	0.014	0.013	-0.007
	(0.271)	(0.263)	(0.261)	(0.262)	(0.264)
Number of important mine safety		-0.195	-0.191	-0.046	-0.078
regulations		(0.119)	(0.133)	(0.124)	(0.134)
Workers' Compensation law			-0.072	-0.220	-0.325
			(0.281)	(0.273)	(0.248)
World War I dummy variable [a]				-0.545	-0.544
				(0.269)	(0.270)
Postwar (Lewis era) dummy				-0.194	-0.176
variable [b]				(0.227)	(0.228)
Union dummy during Lewis era [c]				-1.260	-1.328
				(0.439)	(0.458)
Statewide tonnage under union					0.553
contract (fraction)					(0.303)
Number of mines	443	443	443	443	443
Number of (mine x year) observations	6,870	6,870	6,870	6,870	6,870
Log-likelihood function	-8507.7	-8460.5	-8459.7	-8322.0	-8313.6

[a] Equals one for 1917 and 1918, and equals zero otherwise.
[b] Equals zero through 1918, and equals one thereafter.
[c] Equals zero through 1918, and equals union dummy variable thereafter.

Mine fixed effects included in all equations.
Cluster-robust standard errors in parentheses.

Endnotes

[1] Consistency here refers to the probability limit of the estimator as the number of states approaches infinity. With only 25 states in this sample, the property of consistency is perhaps not so helpful here.

[2] These same coefficient estimates were reported in a working paper version of my 2009 paper as "table B2," but the standard errors were not robust to serial correlation. Cameron and Trivedi (2010, p. 635) advise that "the importance of using cluster-robust standard errors [in panel data] cannot be overemphasized."

[3] The NB1 conditional likelihood function is given in Cameron and Trivedi 2013, p. 257, eq. 9.43.

[4] I also corrected data errors in about a dozen observations, but this made almost no difference in the estimates.

[5] Also, the number of mines and the number of observations reported in the corrected table are smaller than in my 2009 paper, but this is not due to a reduction in the sample size. STATA automatically drops mines with only one observation or with no fatalities in any years, because such mines contribute nothing to the conditional likelihood function.