METHODS AND NAMING

We ran 96\* models. All models assessed the relationship between violations and injuries.

All models included an exposure term (employment hours), as well as the following covariates:

* Number of inspection hours,
* Mine time (number of quarters/years, depending on specification, for which the mine has been active),
* State fixed effects, and
* Time fixed effects (quarters/years, depending on specification)

In all models, standard errors were clustered by mine.

For models with a binary outcome variable type, we ran logit and probit models.

For models with a continuous outcome variable type, we ran poisson and negative binomial models.

We name the models using the form: **A.B.C.D.E.F**, where:

**A**: Injury type[[1]](#footnote-1)

* Maintenance and repair (**MR**)
* Pinning and striking (**PS**)

**B**: Unit of analysis

* Mine-year (**Y**)
* Mine-quarter (**Q**)

**C**: Outcome variable type

* Binary: any injuries (**B**)
* Continuous: how many injuries (**C**)

**D**: Predictor variable level

* Part violations (**P**)
* Subpart violations (**SP**)

**E**: Predictor variable type

* Count of violations (**V**)
* Count of significant and substantial violations (**SSV**)
* Number of penalty points (**PP**)

**F**: Predictor variable lag

* No lag: violations today predict injuries today (**1**)
* Lag 1: violations yesterday predict injuries today (**2**)
* Cumulative lag 4: cumulative violations over the last 4 active quarters predict injuries today (**3**)
* Cumulative lag all: cumulative violations from the first active quarter predict injuries today (**4**)

*Example*:

MR.Y.P.B.SSV.1 = model assessing relationship between the number of non-lagged, part-level, significant and substantial violations and the occurrence of any injury – controlling for covariates.

MR-BINARY-PART (**MR**.X.**B**.**P**.X.X) MODELS – MODEL FITNESS

*maintenance and repair injuries, binary injuries, part-level violations*

***All*** *specifications* fail to account for overdispersion of data.

Comparing the SD of true and predicted values, we see considerable dissimilarity:

For **Y** specifications:

SD

True 0.50

Predicted ~0.29

For **Q** specifications:

SD

True 0.43

Predicted ~0.23

Note: **Y.2**, **Y.4**, and **Q.4** specifications (across **V**, **SSV**, and **PP** specifications) exhibit considerably higher/lower SDs of the (predicted) outcome variable than listed above.

***All*** *specifications* approximate the mean of the outcome variable well.

Note: **Q.2**, **Q.3**, and **Q.4** specifications (across **V**, **SSV**, and **PP** specifications), as well as **Q.PP.1** and **Y.PP.2** specifications, exhibit biased means of the (predicted) outcome variable.

***Q*** *vs.* ***Y*** *specifications*

* **Q** specifications violate model assumptions; **Y** specifications satisfy model assumptions
* **Q** specifications have inferior log pseudo-likelihoods compared to **Y** specifications
* **Q** specifications have higher correct classifications (~81%) than do **Y** specifications (~77%)
* **Q** specifications have lower false positive rates (~5%) than do **Y** specifications (~25%)
* **Q** specifications have higher false negative rates (~55%) than do **Y** specifications (~20%)
* **Q** specifications have inferior correct classifications compared to the null model; **Y** specifications have superior correct classifications compared to the null model
* **Q** specifications approximate the mean of the (predicted) outcome variable worse than do **Y** specifications – see above

***V*** *vs.* ***SSV*** *vs.* ***PP*** *specifications*

* **PP** specifications have superior log pseudo-likelihoods compared to **V** and **SSV** specifications
  + Exceptions: in **Y.1**, **Y.3**, **Q.3**, **Y.4**, and **Q.4** specifications, **PP** specifications have similar log pseudo-likelihoods as **V** and **SSV** specifications

***1*** *vs.* ***2*** *vs.* ***3*** *vs.* ***4*** *specifications*

* For **Q** specifications, the specifications are ordered such that **1** < **2** < **3** < **4** with regards to performance of log pseudo-likelihood
* For **Q** specifications, the specifications are ordered such that **1** > **2** > **3** > **4** with regards to correct classification, false positive rate, and false negative rate.
* For **Y** specifications, all lag constructions produce similar log pseudo-likelihoods, correct classifications, false positive rates, and false negative rates (though patterns along other dimensions of model specification hold)

MR-COUNT-PART (**MR**.X.**C**.**P**.X.X) MODELS – MODEL FITNESS

*maintenance and repair injuries, count injuries, part-level violations*

Across***all*** *specifications*, the data is overdispersed, with alpha estimated around 0.2.

***All*** *specifications* approximate the mean, standard deviation, and range of the outcome variable poorly.

Comparing the mean, SD, and range of true and predicted values, we see considerable dissimilarity:

For **Y** specifications:

Mean SD Min Max

True 1.88 3.27 0 37

Predicted ~1.92 ~2.90 ~0 ~32

For **Q** specifications:

Mean SD Min Max

True 0.41 0.96 0 14

Predicted ~0.46 ~0.70 ~0 ~9

Note: There is considerable variation in predicted mean, SD, min, and max across other dimensions of specification (i.e., **V**, **SSV**, **PP**, and **1**, **2**, **3**, **4**) – see below.

***Q*** *vs.* ***Y*** *specifications*

* **Q** specifications violate model assumptions; **Y** specifications satisfy model assumptions
* **Q** specifications have inferior log pseudo-likelihoods compared to **Y** specifications

***V*** *vs.* ***SSV*** *vs.* ***PP*** *specifications*

* **PP** specifications have superior log pseudo-likelihoods compared to **V** and **SSV** specifications
* **PP** specifications approximate the mean, standard deviation, and range of the (predicted) outcome variable more poorly than do **V** and **SSV** specifications
  + Exceptions: in **Y.2**, **Q.2**, **Y.3**, and **Q.3** specifications, **V** specifications predict a lower maximum value of the outcome variable than do **PP** specifications (both are under-estimates)
* **SSV** specifications approximate the range of the outcome variable closest to the true values compared to **V** and **PP** specifications

***1*** *vs.* ***2*** *vs.* ***3*** *vs.* ***4*** *specifications*

* For **Q** specifications, the specifications are ordered such that **1** < **2** < **3** < **4** with regards to performance of log pseudo-likelihood
* For **Y** specifications, all lag constructions produce similar log pseudo-likelihoods (though patterns along other dimensions of model specification hold)
* **4** specifications underestimate the range of the outcome variable more than other lag constructions (i.e., **1**, **2**, and **3** specifications)

MR-BINARY-SUBPART (**MR**.X.**B**.**SP**.X.X) MODELS – MODEL FITNESS

*maintenance and repair injuries, binary injuries, subpart-level violations*

***All*** *specifications* fail to account for overdispersion of data.

Comparing the SD of true and predicted values, we see considerable dissimilarity:

For **Y** specifications:

SD

True 0.50

Predicted ~0.30

For **Q** specifications:

SD

True 0.43

Predicted ~0.23

Note: **Q.4** specifications (across **V**, **SSV**, and **PP** specifications) exhibit considerably higher SDs than listed above.

***All*** *specifications* approximate the mean of the outcome variable poorly.

Comparing the mean of true and predicted values, we see considerable dissimilarity:

For **Y** specifications:

Mean

True 0.55

Predicted ~0.53

For **Q** specifications:

Mean

True 0.24

Predicted ~0.26

Note: For **Y** specifications, **PP** specifications produce even more biased means of the (predicted) outcome variable (across all lag specifications). For **Q** specifications, **4** specifications also produce more biased means of the (predicted) outcome variable (across **V**, **SSV**, and **PP** specifications), as do **2** and **3** specifications for **V** and **SSV** specifications.

***Q*** *vs.* ***Y*** *specifications*

* **Q** specifications violate model assumptions; **Y** specifications satisfy model assumptions (exceptions: **Y.SSV.4**, **Y.PP.4**)
* **Q** specifications have inferior log pseudo-likelihoods compared to **Y** specifications
* **Q** specifications have higher correct classifications (~81%) than do **Y** specifications (~77%)
* **Q** specifications have lower false positive rates (~7%) than do **Y** specifications (~25%)
* **Q** specifications have higher false negative rates (~55%) than do **Y** specifications (~20%)
* **Q** specifications have inferior correct classifications compared to the null model; **Y** specifications have superior correct classifications compared to the null model

***V*** *vs.* ***SSV*** *vs.* ***PP*** *specifications*

* **PP** specifications have superior log pseudo-likelihoods compared to **V** and **SSV** specifications
* **PP** specifications have higher correct classifications than do **V** and **SSV** specifications

***1*** *vs.* ***2*** *vs.* ***3*** *vs.* ***4*** *specifications*

* For **Q** specifications, the specifications are ordered such that **1** ~ **2** < **3** < **4** with regards to performance of log pseudo-likelihood
* For **Q** specifications, the specifications are ordered such that **1** ~ **2** > **3** ~ **4** with regards to correct classification
* For **Q** specifications, the specifications are ordered such that **1** < **2** < **3** < **4** with regards to false positive rate
* For **Q** specifications, the specifications are ordered such that **1** > **2** > **3** > **4** with regards to false negative rate
* For **Y** specifications, all lag constructions produce similar log pseudo-likelihoods, correct classifications, false positive rates, and false negative rates (though patterns along other dimensions of model specification hold)

MR-COUNT-SUBPART (**MR**.X.**C**.**SP**.X.X) MODELS – MODEL FITNESS

*maintenance and repair injuries, count injuries, subpart-level violations*

Across***most*** *specifications*, the data is overdispersed, with alpha estimated around 0.1 – 0.2.

Exceptions: **Y.PP.1**, **Y.SSV.2**, **Y.PP.2**, **Q.PP.2**, **Y.PP.3**, **Q.PP.3**, **Q.PP.4**

***All*** *specifications* approximate the mean, standard deviation, and range of the outcome variable poorly.

***Q*** *vs.* ***Y*** *specifications*

* **Q** specifications violate model assumptions; **Y** specifications satisfy model assumptions
* **Q** specifications have inferior log pseudo-likelihoods compared to **Y** specifications

***V*** *vs.* ***SSV*** *vs.* ***PP*** *specifications*

* **PP** specifications have superior log pseudo-likelihoods compared to **V** and **SSV** specifications
* For **Y** specifications, **PP** specifications approximate the range of the (predicted) outcome variable better than do **V** and **SSV** specifications
  + Exceptions: in **Y.4**, the **V** specification predicts a maximum value of the outcome variable closer to the true value than does the **PP** specification
* For **Y** specifications, **V** specifications generally (dramatically) over-estimate the maximum value of the outcome variable

***1*** *vs.* ***2*** *vs.* ***3*** *vs.* ***4*** *specifications*

* For **Q** specifications, the specifications are ordered such that **1** < **2** < **3** < **4** with regards to performance of log pseudo-likelihood
* For **Q** specifications, **3** and **4** specifications underestimate the range of the (predicted) outcome variable more than other lag constructions (i.e., **1** and **2** specifications)
* For **Y** specifications, all lag constructions produce similar log pseudo-likelihoods (though patterns along other dimensions of model specification hold)

1. \* At the time of this summary, only models using the MR injury type have been run [↑](#footnote-ref-1)