Supplementary Materials:
Bivariate Left-Censored Bayesian Model for Predicting Exposure: Preliminary Analysis of Worker Exposure during the *Deepwater Horizon* Oil Spill

Caroline Groth[1], Sudipto Banerjee[2*], Gurumurthy Ramachandran[3], Mark R. Stenzel[4], Dale P. Sandler[5], Aaron Blair[6], Lawrence S. Engel[7], Richard K. Kwok[5] and Patricia A. Stewart[8]

[1] Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota 55455.

[2] Department of Biostatistics, University of California-Los Angeles, Los Angeles, California 90095.

[3] Department of Environmental Health and Engineering, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205.

[4] Exposure Assessments Applications, LLC, Arlington, Virginia 22207

[5] National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

[6] National Cancer Institute, Bethesda, Maryland 20892

[7] Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

[8] Stewart Exposure Assessments, LLC, Arlington, Virginia 22207

[*] *email:* sudipto@ucla.edu

1 Oil Weathering

A complicating factor related to oil spills is that the composition of released crude oil does not remain constant. The components in the oil can evaporate, dissolve into water, or be broken down by sunlight or bacteria. These processes are referred to as oil weathering and will result in a change in crude oil composition.

When the *Deepwater Horizon* sank it caused the oil to be released 5000 feet below the surface. Some of the weathering occurred to the oil as it rose to the surface, but once oil reached the surface, significant oil weathering occurred within a few days. During the period when the well was leaking, fresh leaking oil replaced the weathered oil on the surface, leading to a pseudo steady state that should result in an approximately constant crude oil composition and should lead to observable and strong correlations over the period the oil was being released. In cases where the integrity of the surface barrier of the oil or tar (arising from the weathered oil) was maintained, such as is the case with undissolved submerged oil, the volatile components of the crude should remain within the crude oil plume or tar and only be released when the surface barrier is broken. Thus, release of crude oil volatile components could occur after the well was capped at significant distances in space or time from the original source of the spill. This phenomenon is also why the THC concentration can vary over time and space. From the perspective of this paper, it is not necessary to definitively identify the reason for the correlations between THC and the chemicals of interest, but rather, if, by empirical observation, correlations are constant over a defined time period, then these correlations can be used to estimate exposures to the THC components for the workers involved in the response.

2 Linear Relationship Background

Since the relationship is known to be linear in nature, we further assessed whether an intercept was necessary using a cross-validation approach. From this analysis (not shown), we concluded that including an intercept allowed us to predict exposure levels while minimizing the influence of outlying observations. Additionally, chemical concentrations are often depicted on a ln scale to better meet normality assumptions. This proved to be appropriate for our data. Therefore, the final model uses an intercept and the natural log scale for both the predictor and response.

GM, AM, and GSD Calculations

To obtain GM estimates for Y_i from our model, using the posterior samples of each parameter, we calculated the mean based on the model $(\beta_{0i} + \beta_{1i}\mu_i)$ for each EG i using the estimated mean of X (μ_i). Then, we exponentiated this distribution to obtain the distribution of the GM of Y_i for each EG. Similarly, for the GM estimates for X_i, we exponentiated the posterior samples of μ_i for each EG i. To obtain the GSD distributions of X_i and $Y \mid X_i$, we exponentiated the posterior samples for the standard deviation of X_i and $Y \mid X_i$ respectively. To obtain, the variance of Y_i, we calculated the covariance of X_i and Y_i for each EG i (using $cov_i = \sigma^2_{X_i}\beta_{1i}$). Then, using formulas for conditional variances, we solved for the distribution of the variance of Y (where $\sigma^2_{Y_i} = \sigma^2_{Y|X_i} + cov_i^T \left(\sigma^2_{X_i}\right)^{-1} cov_i$). Then, finally using the posterior distributions of the GSD and GM of Y_i previously calculated, we calculated the distribution of the arithmetic mean (AM) of Y_i for EG i using $AM = GM \times e^{\left(\frac{1}{2}(ln(GSD))^2\right)}$. Likewise, for the AM of X_i we used the same AM formula with the GM and GSD of X_i for each EG i for the GM and GSD respectively. These equations allowed us to obtain posterior distributions of all parameters of interest. In this paper, our primary interest is AM exposure estimates.

3 RJAGS Model

n.adapt=1000 (Standard)

```
model { for (i in 1:N){
is.notcensoredx[i] ~ dinterval(X[i],cx[i])
X[i] ~ dnorm(mux[SEG[i]],tausqx[SEG[i]])
is.notcensoredy[i] ~ dinterval(Y[i],cy[i])
Y[i] ~ dnorm(mu[i], sigmasq[SEG[i]])
mu[i]<- beta[SEG[i],1] + beta[SEG[i],2]*X[i]
}

mu.beta[1:2] ~dmnorm(mean[],prec[,])
Omega[1:2,1:2] ~ dwish(W[,],2)

for(k in 1:NSEG) {
mux[k] ~ dnorm(0,0.00001)
beta[k,1:2] ~ dmnorm(mu.beta[],Omega[,])
sigmasq[k] ~ dgamma(0.01,0.01)
tausqx[k] ~ dgamma(0.01,0.01)
mu.seg[k] <- beta[k,1] + beta[k,2]*mux[k]
sigmaxsq[k] <- 1/tausqx[k]
sigmaysq[k] <- 1/sigmasq[k]
cov[k]<-sigmaxsq[k]*beta[k,2]
variancey[k]<-sigmaysq[k]+cov[k]*tausqx[k]*cov[k]
corr[k]<-cov[k]/(sqrt(variancey[k])*sqrt(sigmaxsq[k]))
}
LSigma[1:2,1:2]<-inverse(Omega[,])

}
```

4 Openbugs Model

```
model{

for (i in 1:N) {

X[i] ~ dnorm(mux[SEG[i]], tausqx[SEG[i]]) I(,cx[i])

Y[i] ~ dnorm(mu[i], sigmasq[SEG[i]]) I(,cy[i])

mu[i] <- beta[SEG[i],1] + beta[SEG[i],2]*X[i]

}


mu.beta[1:2] ~dmnorm(mean[],prec[,])

Omega[1:2,1:2] ~ dwish(W[,],2)


for(k in 1:NSEG) {

mux[k] ~ dnorm(0,0.00001)

beta[k,1:2] ~ dmnorm(mu.beta[],Omega[,])

sigmasq[k] ~ dgamma(0.01,0.01)

tausqx[k] ~ dgamma(0.01,0.01)

mu.seg[k] <- beta[k,1] + beta[k,2]*mux[k]

sigmaxsq[k] <- 1/tausqx[k]

sigmaysq[k] <- 1/sigmasq[k]

cov[k]<-sigmaxsq[k]*beta[k,2]

variancey[k]<-sigmaysq[k]+cov[k]*tausqx[k]*cov[k]

corr[k]<-cov[k]/(sqrt(variancey[k])*sqrt(sigmaxsq[k]))

}

}
```

5 Simulation Study

5.1 Methods

We performed three simulation experiments under different levels of LOD censoring. Table S1 describes the parameters we set for all three scenarios. We started off by dividing a set of 300 observations into 10 groups of various sizes ranging from 9 to 92, which is similar to what we see in the GuLF STUDY measurement data.

We set the true parameters using common characteristics of the data (not shown). Since the model uses a natural logged response and predictor, the parameters listed reflect what the parameters would be on the natural log scale for both X and Y. Specifically, we selected for each group a slope parameter between 0.6 and 1, which, for example, corresponds to the slopes generally found between ln(THC) (X) and ln(xylene) (Y). Based on previous regression models, we set intercept values between -2.5 and -0.5 (in ln(ppb) units), where an intercept can be interpreted as the mean estimate of ln(xylene) when ln(THC) is 0.

Then, also using common characteristics of the data, we set the mean of X to be between 5 and 7.25 ln(ppb) (148.4 ppb and 1408.1 ppb). Next, we set the variances to be between 1.44 and 5.29 for X and 0.49 and 3.24 for $Y \mid X$, corresponding to GSDs ranging from 3.3 to 10 for X, and from 2 to 6 for a second chemical Y. We then generated Xs from $N(\mu_x, \sigma^2_X)$ and, for each generated X, we drew a Y from $N(\beta_{0i} + \beta_{1i}X, \sigma^2_{Y|X_i})$, where μ_i, $\sigma^2_{X_i}$, β_{0i}, β_{1i}, and $\sigma^2_{Y|X_i}$ are as defined earlier.

The parameters described above were kept for all simulation studies. After assigning the parameters, three scenarios were defined. For the first scenario, the censoring levels were below 31% in both X and Y, corresponding to lower levels of LOD censoring. In the second scenario, the censoring on X remained the same as scenario 1, but we increased the censoring in Y to 25-50%. Finally, in the third scenario, censoring on X remained as in scenario 1, but the censoring ranged from 25-70% in Y, to demonstrate a scenario with highly censoring (censoring $> 50\%$) in the predicted variable of some groups. Censoring levels among the groups varied within a scenario, allowing for similar sample sizes to have different censoring levels. To be consistent, the percent censored in Y was always greater than or equal to the percent censored in X (as is generally seen in our GuLF STUDY data).

In order to create censoring, we determined the quantiles in each scenario corresponding to above or below each percentage censoring. All values below the quantile were censored or became missing. Following this, a set of LODs was assigned for each group in a uniform distribution just below the quantile chosen. This allowed for multiple LODs for each group, due to, in our data, different durations of sampling (i.e. 4-18 hr).

We implemented our Bayesian models by running an additional 10,000 MCMC iterations after 5,000 initial iterations for burn-in. In our model, we used inverse-gamma priors on the variance components. We also conducted simulation studies using informative uniform priors on the standard deviations with GSDs ranging from 1.01 to 12. Estimates of the intercept and slope parameters were similar, but the variance estimates varied more under the inverse-gamma priors as expected.

In our model, we used an inverse-Wishart prior on $\mathbf{V}_\beta$ with 2 degrees of freedom. The 2 by 2 scale matrix of this prior had upper left element 200, lower right element 0.2, and 0 otherwise. A normal prior was placed on μ_β with a mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{V}_\mu$. The variance-covariance matrix $\mathbf{V}_\mu$ had variances of 1,000,000 and covariances of 0. We used a normal prior on each μ_i with mean 0 and variance 100,000 for all 10 groups. Then, finally, we used an inverse-gamma distribution on the $\sigma^2_{X_i}$ and $\sigma^2_{Y|X_i}$ for each group with shape parameter 0.01 and rate (1/scale) parameter 0.01.

We also compared our hierarchical Bayesian EG model in (3) with three simpler models for each of the three scenarios described earlier. For model comparisons, we replicated the observed Xs and observed Ys from the respective models. In the first model, only an intercept was included for prediction of X and Y; X was not used in the estimation of Y, and each group was modeled separately. This assumed different variances for each group where we simply modeled means, not accounting for additional information. The second model had a common intercept and common slope, where groups were not modeled separately but as one group. The third model used varying intercepts for groups but assumed that all groups had the same slope estimate. In all of the above models, we account for censoring in X and Y. D-statistics were used to compare models.

5.2 Results

The results of the model comparison for all three scenarios indicate that the hierarchical Bayesian EG model was preferred according to the D-statistic (Table S2). The D-statistic, in all scenarios, was lower for the hierarchical Bayesian EG model than for the other three model types. This finding demonstrates that if groups really did have their own intercept and slope estimates, the hierarchical Bayesian EG model would be preferred over the simpler models. Across all three scenarios, modeling the groups separately was meaningful. The D-statistics should not be compared across scenarios since the datasets between the three scenarios were fundamentally different due to the different levels of censoring. G-statistics were consistently higher for the common intercept and common slope model, indicating that there were great deviations between the replicates based on this model and the real values. The real values were not generated based on a single regression line, and were, as described above, based on individual regression lines per group. Thus, this finding was expected.

The credible intervals (CI) are provided for the intercept (β_0), slope (β_1), variance of X (σ^2_X), and variance of $Y \mid X$ ($\sigma^2_{Y|X}$) to see if they contained the true value of the parameter (Table S3). In all scenarios, all parameters were contained within the credible intervals. In scenario 1, all slopes were significantly positive, although group 6 was barely so. This particular group had a small number of non-censored samples below 10 that likely led to the wide credible interval. Thus, slope estimation was reasonable in this scenario and followed what we would expect based on the values we provided. The upper bounds on the variance of X in groups 3 and 7 were quite high. However, in both of these cases, we had set the highest variances for these parameters of the groups, so this result was expected.

In scenario 2 with moderate censoring in Y (25-50% censoring), the credible intervals tended to be slightly wider for the slopes compared to scenario 1. With increased censoring, there was less certainty and smaller non-censored sample sizes to estimate the true parameters. For group 6, the slope was insignificant as seen from the 95% credible interval, which marginally includes 0. We note that the 90% credible interval (not shown) did not include 0, indicating significance at this level. For most groups, the upper bounds of the variance of $Y \mid X$ increased from scenario 1 to scenario 2. As censoring increased in Y, there may have been more variability that went into estimation of Y at lower values of X, increasing the variance of $Y \mid X$ in some cases.

In scenario 3 containing some high levels of censoring in Y, the group with the highest censoring, group 2, had an increased median posterior intercept and a decreased median posterior slope compared to scenarios 1 and 2 (medians not shown). Since censoring was relatively high in this scenario, our model began to use inference from other groups to model this group. Overall, the slopes for the other groups were lower than this group. Therefore, this group's slope estimate at very high levels of censoring closely reflected the slopes of other groups.

To summarize, these results highlight that the model performed well under a variety of levels of censoring and that the 95% credible intervals contained the true parameters. It is expected that as censoring increases the relationships will change, but the model clearly was able to generate reasonable estimates and model the data adequately at levels $< 70\%$ censoring.

		True Parameter Values					All Scenarios	Scenario 1	Scenario 2	Scenario 3
Group	N	β_0	β_1	σ^2_X	$\sigma^2_{Y\mid X}$	μ_x	$X < LOD$ (%)	$Y < LOD$ (%)	$Y < LOD$ (%)	$Y < LOD$ (%)
1	9	-1.50	0.70	2.25	1.00	5.50	22.2	22.2	33.3	33.3
2	50	-2.50	1.00	2.25	1.44	6.75	10.0	10.0	50.0	70.0
3	10	-2.00	0.80	4.00	3.24	6.50	20.0	30.0	30.0	40.0
4	92	-2.00	0.80	2.25	1.00	6.50	15.2	25.0	41.3	59.8
5	20	-1.50	0.70	1.44	0.49	6.80	5.0	10.0	35.0	35.0
6	12	-1.20	0.65	2.56	1.44	5.50	25.0	25.0	25.0	25.0
7	15	-2.00	0.80	5.29	3.24	6.20	13.3	13.3	26.7	33.3
8	14	-2.50	1.00	1.44	1.00	7.25	14.3	14.3	35.7	50.0
9	16	-0.50	0.60	2.25	1.69	5.00	18.8	25.0	43.8	43.8
10	62	-1.50	0.70	2.89	1.96	6.25	21.0	30.6	30.6	59.7

Table S1: Simulation study scenarios for assessing oil-related chemical exposure. β_0 is the intercept, β_1 is the slope, σ^2_X is the variance of X, $\sigma^2_{Y\mid X}$ is the variance of $Y \mid X$, and μ_x is the mean of X.

Scenario	Model	D-Statistic	P	G
1	Intercept Only Model	2710.0	1681.7	1028.4
	Common Intercept and Common Slope	2775.0	1591.8	1183.3
	Common Slope and Varying Intercepts	2748.5	1706.6	1041.8
	Hierarchical Bayesian EG Model	2690.8	1664.1	1026.6
2	Intercept Only Model	2656.6	1632.0	1024.6
	Common Intercept and Common Slope	2731.8	1535.3	1196.5
	Common Slope and Varying Intercepts	2635.5	1617.9	1017.6
	Hierarchical Bayesian EG Model	2602.9	1595.9	1007.0
3	Intercept Only Model	2517.0	1476.5	1040.5
	Common Intercept and Common Slope	2555.1	1362.1	1193.0
	Common Slope and Varying Intercepts	2415.1	1431.5	983.6
	Hierarchical Bayesian EG Model	2404.2	1417.8	986.4

Table S2: Model Comparison for our simulation study of models for assessing exposure to oil-related chemicals

	Scenario1				Scenario 2			
	β_0	β_1	σ_X^2	$\sigma_{Y\|X}^2$	β_0	β_1	σ_X^2	$\sigma_{Y\|X}^2$
Group	CI	CI	CI	CI	CI	CI	CI	CI
1	(-4.89, -0.56)	(0.46, 1.23)	(0.72, 12.29)	(0.20, 2.30)	(-5.30, -0.50)	(0.44, 1.28)	(0.67, 9.96)	(0.21, 3.23)
2	(-3.19, -0.80)	(0.74, 1.11)	(1.93, 4.51)	(0.73, 1.71)	(-4.29, -0.96)	(0.75, 1.22)	(1.95, 4.64)	(0.84, 2.80)
3	(-4.63, 1.96)	(0.20, 1.00)	(2.55, 27.37)	(0.77, 9.88)	(-4.98, 1.94)	(0.20, 1.04)	(2.59, 28.67)	(0.77, 10.81)
4	(-3.13, -1.05)	(0.68, 0.96)	(1.78, 3.46)	(0.63, 1.27)	(-3.12, -0.83)	(0.66, 0.96)	(1.77, 3.43)	(0.51, 1.13)
5	(-2.59, 2.00)	(0.19, 0.87)	(0.62, 2.42)	(0.43, 1.85)	(-3.63, 1.42)	(0.28, 1.00)	(0.64, 2.47)	(0.32, 2.00)
6	(-3.33, 2.12)	(0.01, 0.89)	(1.38, 10.95)	(0.69, 8.14)	(-3.60, 2.16)	(-0.01, 0.92)	(1.38, 11.93)	(0.68, 8.40)
7	(-5.40, -1.36)	(0.57, 1.15)	(3.04, 15.76)	(0.83, 4.87)	(-6.43, -1.77)	(0.61, 1.27)	(3.17, 19.30)	(0.83, 5.82)
8	(-4.31, 1.98)	(0.39, 1.21)	(0.60, 3.69)	(0.55, 3.37)	(-5.22, 2.42)	(0.33, 1.30)	(0.61, 3.78)	(0.60, 5.20)
9	(-2.86, 1.30)	(0.22, 0.91)	(1.86, 10.70)	(0.78, 5.07)	(-4.12, 1.25)	(0.22, 1.04)	(1.90, 11.26)	(0.83, 8.20)
10	(-3.03, -0.40)	(0.50, 0.87)	(2.63, 6.07)	(0.99, 2.47)	(-3.11, -0.46)	(0.50, 0.88)	(2.65, 6.02)	(0.97, 2.53)

	Scenario 3			
	β_0	β_1	σ_X^2	$\sigma_{Y\|X}^2$
Group	CI	CI	CI	CI
1	(-5.17, -0.43)	(0.44, 1.25)	(0.67, 9.02)	(0.21, 3.40)
2	(-3.65, 0.14)	(0.66, 1.14)	(1.93, 4.54)	(0.48, 2.26)
3	(-5.64, 2.00)	(0.20, 1.09)	(2.57, 26.46)	(0.87, 14.75)
4	(-3.38, -0.48)	(0.62, 0.98)	(1.76, 3.41)	(0.48, 1.31)
5	(-3.59, 1.37)	(0.29, 0.99)	(0.64, 2.49)	(0.32, 1.97)
6	(-3.60, 2.03)	(0.01, 0.92)	(1.34, 10.49)	(0.69, 8.02)
7	(-6.63, -1.71)	(0.62, 1.30)	(3.05, 16.65)	(0.62, 4.55)
8	(-5.48, 2.54)	(0.28, 1.28)	(0.60, 3.83)	(0.77, 11.13)
9	(-4.06, 1.22)	(0.22, 1.05)	(1.85, 11.14)	(0.81, 7.42)
10	(-4.26, -0.66)	(0.53, 0.99)	(2.68, 6.16)	(0.92, 3.21)

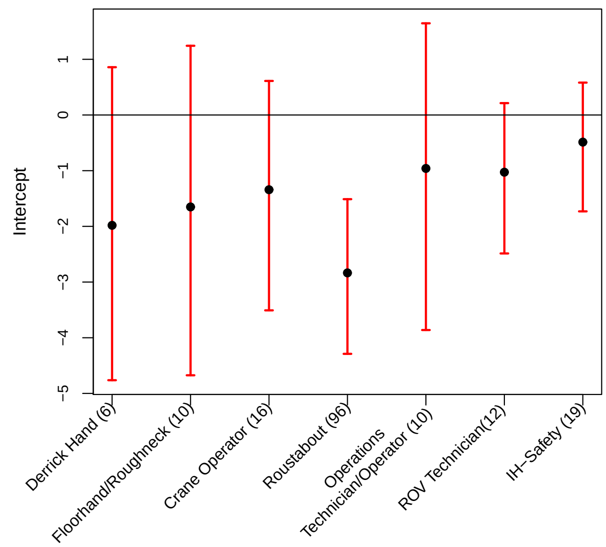
Table S3: Simulation study credible intervals for parameters in our hierarchical Bayesian EG model. β_0 is the intercept, β_1 is the slope, σ_X^2 is the variance of X, and $\sigma_{Y|X}^2$ is the variance of $Y \mid X$. The median and 95% credible intervals (CI) are reported for each parameter.

6 Illustrative Example from EGs on the *DDIII*: Results with Inverse-gamma

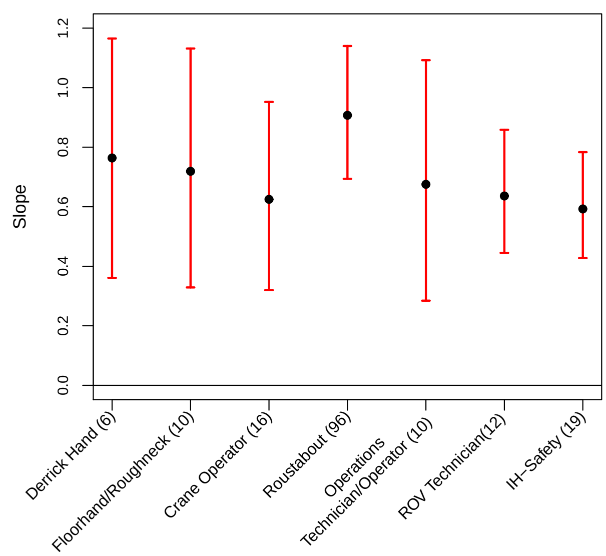
6.1 Prior and Model Specification

In our model, we used an inverse-Wishart prior on $\mathbf{V}_\beta$ with 2 degrees of freedom. The 2 by 2 scale matrix of this prior had upper left element 200, lower right element 0.2, and 0 otherwise. A normal prior was placed on μ_β with a mean vector $\mathbf{0}$ and variance-covariance matrix $\mathbf{V}_\mu$. The variance-covariance matrix $\mathbf{V}_\mu$ had variances of 1,000,000 and covariances of 0. We used a normal prior on each μ_i with mean 0 and variance 100,000 for all 10 groups. Then, finally, we used an inverse-gamma distribution on the $\sigma^2_{X_i}$ and $\sigma^2_{Y|X_i}$ for each group with shape parameter 0.01 and rate (1/scale) parameter 0.01.

6.2 Figures



(a) Intercepts



(b) Slopes

Figure S1: Intercepts and slopes for the EGs on the *DDIII* from May 15-July 15, 2010. The upper panel displays the intercepts from the regression model. The lower panel displays the slopes.The dots in each bar represent the median posterior samples.

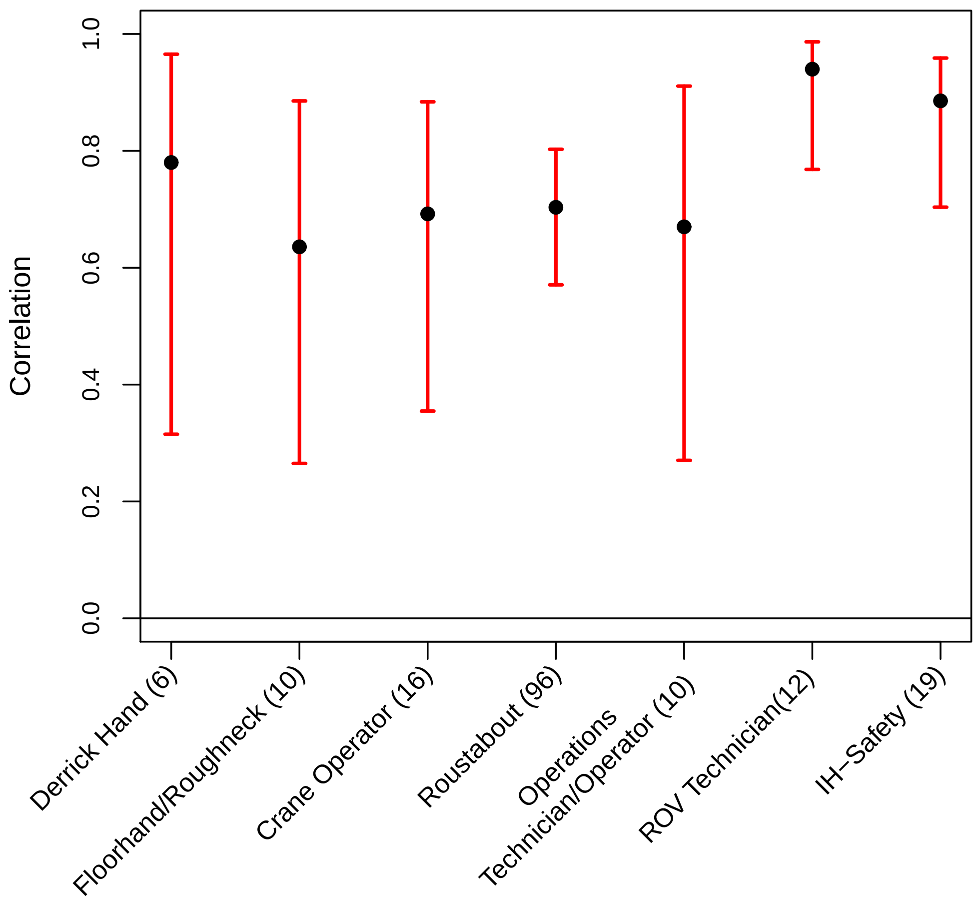


Figure S2: 95% Credible intervals of the posterior correlation estimates by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

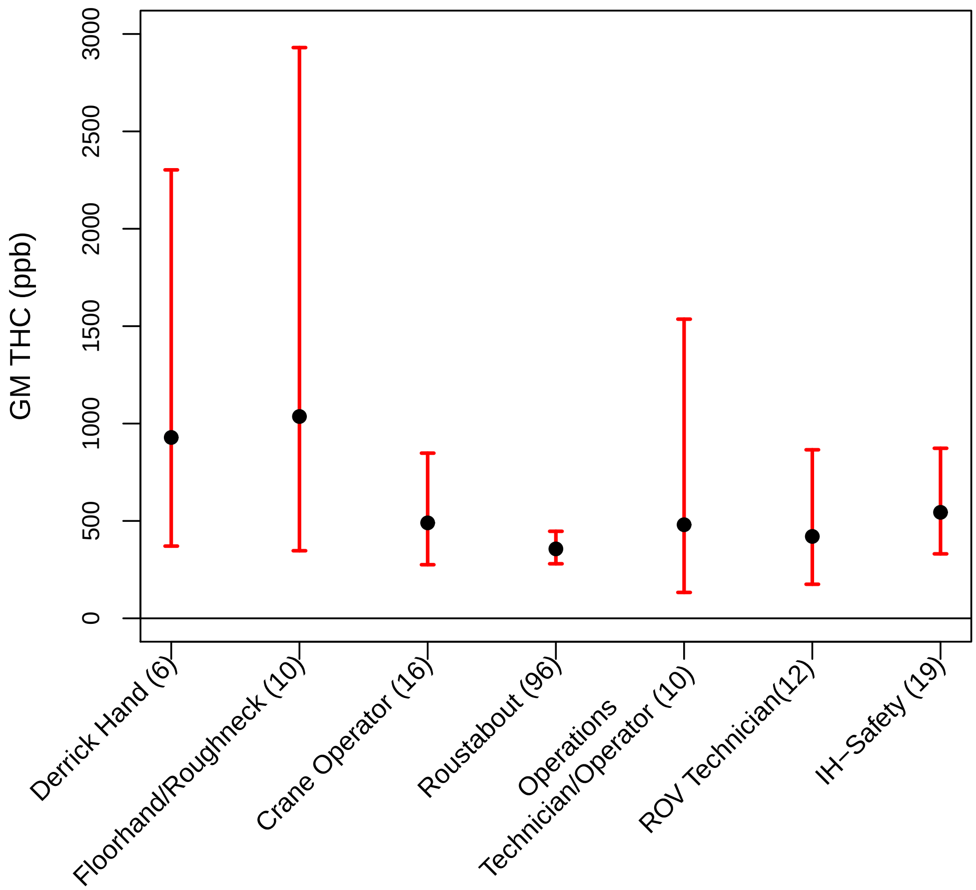


Figure S3: 95% Credible intervals of the posterior GMs estimates for THC by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

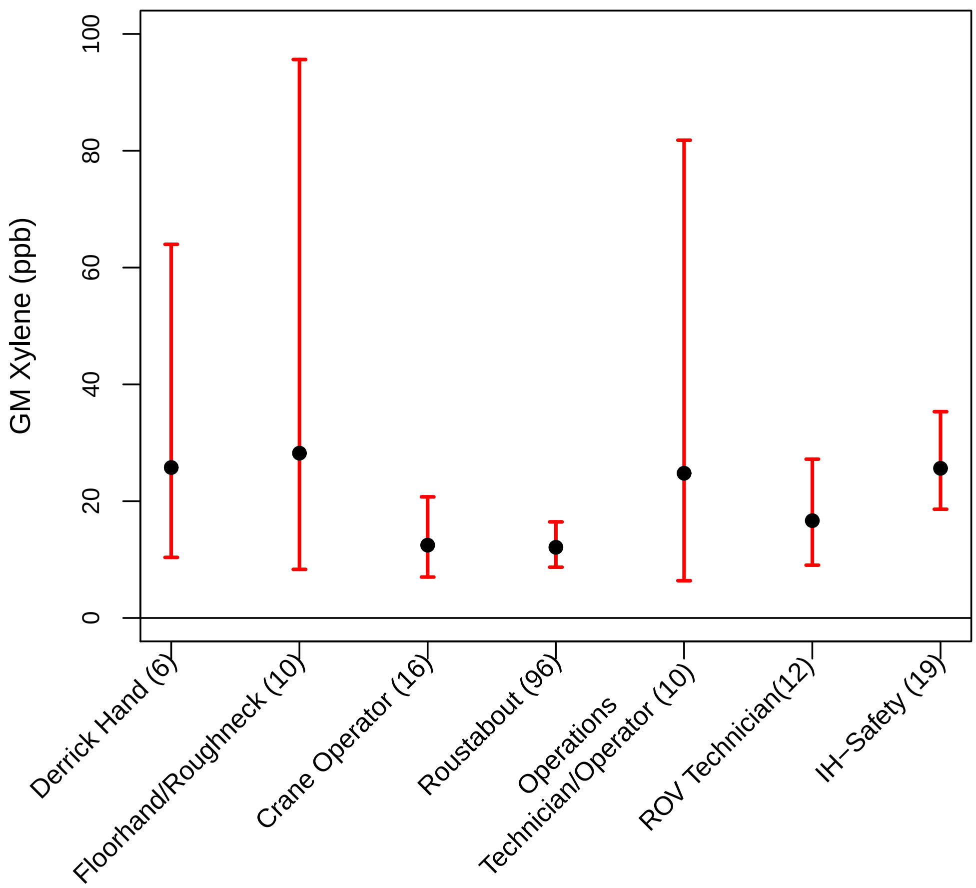
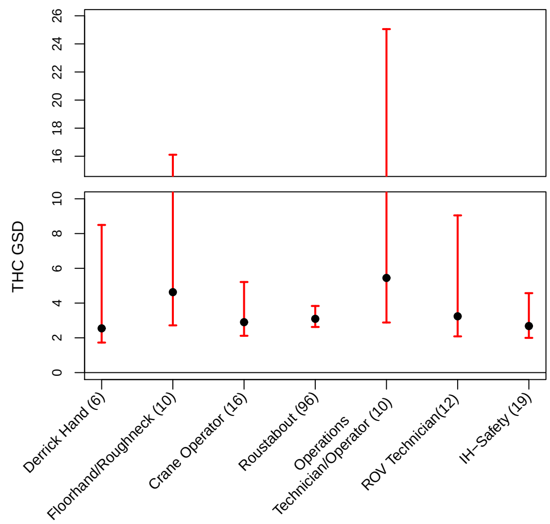
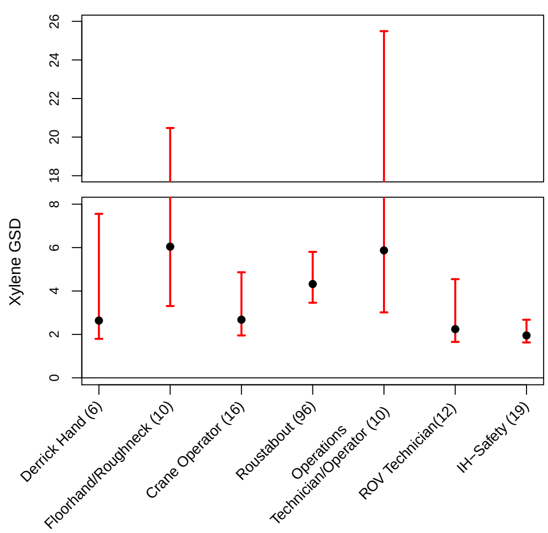


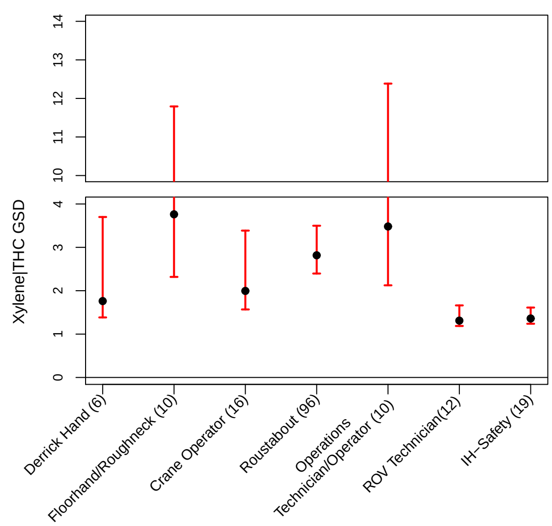
Figure S4: 95% Credible intervals of the posterior GMs estimates for xylene by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.



(a) GSD of THC



(b) GSD of Xylene



(c) GSD of Xylene|THC

Figure S5: Variance components of the model by EG on the *DDIII* from May 15-July 15, 2010. The top panel displays the GSD for THC. The bottom left panel displays the GSD for xylene. The bottom right panel displays the GSD for xylene given THC. The dots in each bar represent the median posterior samples.

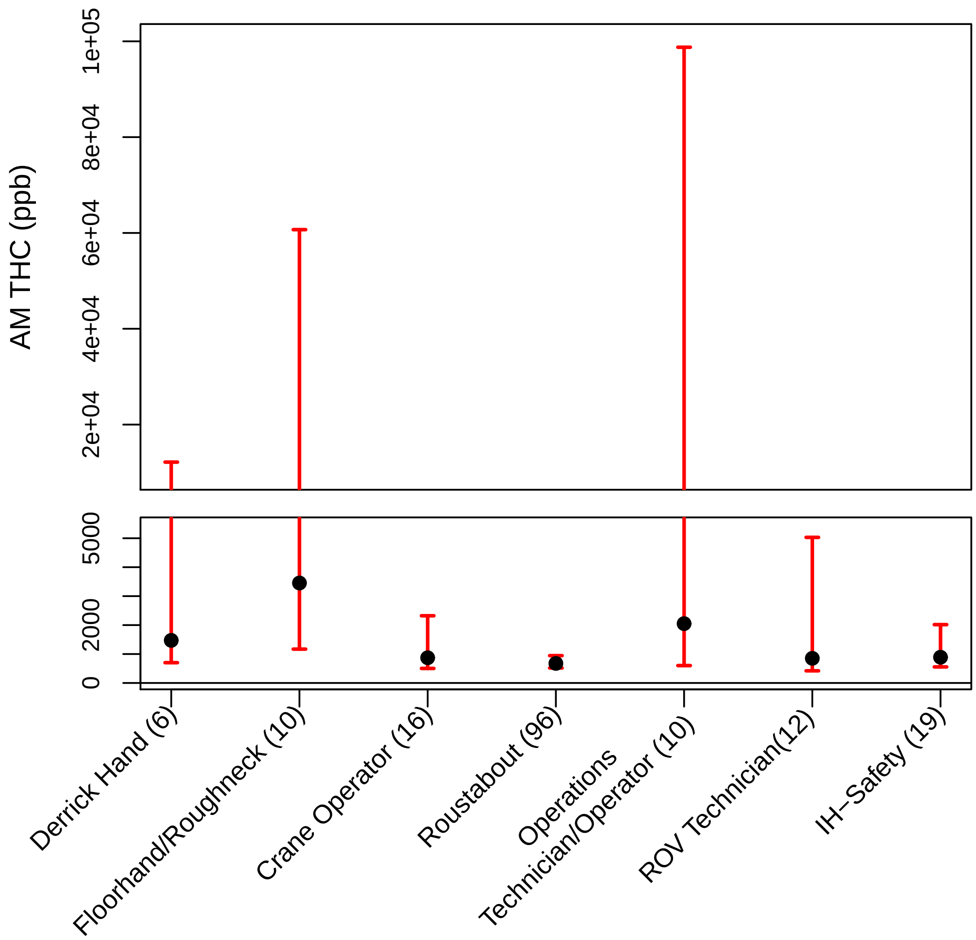


Figure S6: 95% Credible intervals of the posterior AMs estimates for THC by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

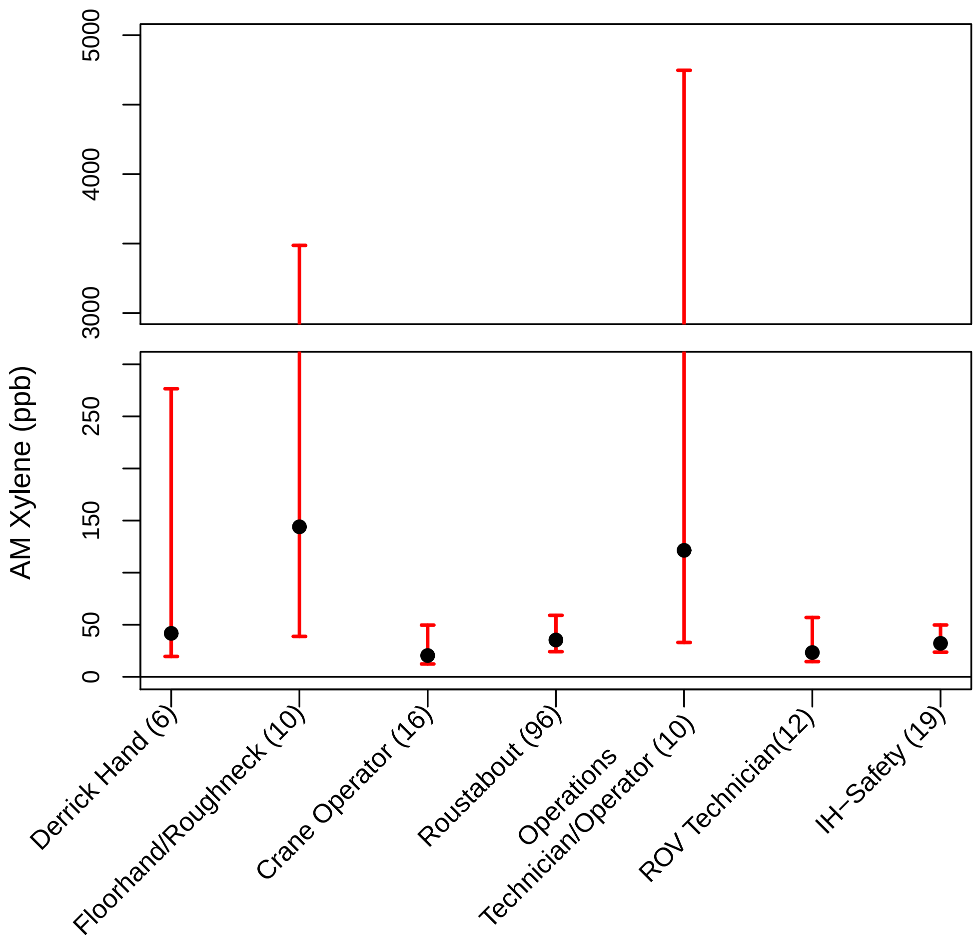


Figure S7: 95% Credible intervals of the posterior AMs estimates for xylene by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

Uniform Prior Supplementary Materials

Simulation Study: Results with Uniform Priors on Standard Deviations

	Scenario1				Scenario 2			
	β_0	β_1	σ^2_X	$\sigma^2_{Y\|X}$	β_0	β_1	σ^2_X	$\sigma^2_{Y\|X}$
Group	CI	CI	CI	CI	CI	CI	CI	CI
1	(-4.93, -0.48)	(0.44, 1.23)	(0.77, 5.78)	(0.22, 2.97)	(-5.41, -0.48)	(0.44, 1.30)	(0.74, 5.71)	(0.24, 3.73)
2	(-3.15, -0.78)	(0.74, 1.10)	(1.98, 4.61)	(0.73, 1.76)	(-4.43, -0.94)	(0.75, 1.24)	(1.99, 4.82)	(0.87, 2.97)
3	(-4.55, 1.86)	(0.21, 1.01)	(2.36, 6.09)	(0.86, 5.75)	(-4.87, 1.77)	(0.22, 1.05)	(2.38, 6.10)	(0.88, 5.77)
4	(-3.14, -1.04)	(0.67, 0.96)	(1.82, 3.50)	(0.64, 1.27)	(-3.19, -0.85)	(0.66, 0.97)	(1.79, 3.52)	(0.52, 1.18)
5	(-2.69, 2.05)	(0.19, 0.88)	(0.65, 2.66)	(0.45, 2.04)	(-3.89, 1.39)	(0.28, 1.03)	(0.68, 2.77)	(0.35, 2.22)
6	(-3.33, 2.04)	(0.01, 0.89)	(1.44, 5.94)	(0.78, 5.67)	(-3.53, 2.01)	(0.02, 0.93)	(1.44, 5.95)	(0.78, 5.62)
7	(-5.47, -1.27)	(0.56, 1.16)	(2.86, 6.11)	(0.88, 4.80)	(-6.57, -1.84)	(0.62, 1.29)	(2.86, 6.10)	(0.91, 5.24)
8	(-4.42, 2.01)	(0.38, 1.23)	(0.67, 4.34)	(0.59, 3.82)	(-5.25, 2.31)	(0.34, 1.31)	(0.66, 4.33)	(0.68, 5.02)
9	(-3.03, 1.26)	(0.22, 0.93)	(1.97, 6.00)	(0.84, 5.12)	(-4.13, 1.12)	(0.24, 1.07)	(1.93, 6.01)	(0.90, 5.61)
10	(-3.12, -0.41)	(0.50, 0.88)	(2.71, 5.78)	(1.01, 2.55)	(-3.12, -0.46)	(0.50, 0.88)	(2.70, 5.70)	(1.01, 2.53)

	Scenario 3			
	β_0	β_1	σ^2_X	$\sigma^2_{Y\|X}$
Group	CI	CI	CI	CI
1	(-5.40, -0.41)	(0.44, 1.28)	(0.75, 5.68)	(0.24, 4.02)
2	(-3.88, 0.15)	(0.65, 1.16)	(1.99, 4.66)	(0.49, 2.67)
3	(-5.11, 1.80)	(0.21, 1.06)	(2.38, 6.10)	(0.98, 5.90)
4	(-3.27, -0.50)	(0.62, 0.97)	(1.76, 3.46)	(0.48, 1.31)
5	(-3.85, 1.32)	(0.30, 1.02)	(0.67, 2.66)	(0.36, 2.28)
6	(-3.45, 2.01)	(0.02, 0.92)	(1.46, 5.94)	(0.77, 5.67)
7	(-6.73, -1.73)	(0.63, 1.30)	(2.90, 6.12)	(0.70, 4.53)
8	(-5.44, 2.47)	(0.29, 1.30)	(0.65, 4.12)	(0.89, 5.78)
9	(-4.04, 1.09)	(0.24, 1.05)	(1.92, 5.99)	(0.92, 5.62)
10	(-4.56, -0.68)	(0.53, 1.02)	(2.75, 5.82)	(0.96, 3.47)

Table S4: Simulation study credible intervals for parameters in our hierarchical Bayesian EG model using uniform priors. β_0 is the intercept, β_1 is the slope, σ^2_X is the variance of X, and $\sigma^2_{Y|X}$ is the variance of $Y \mid X$. The median and 95% credible intervals (CI) are reported for each parameter.

Illustrative Example from EGs on the *DDIII*: Results with Uniform Priors on Standard Deviation

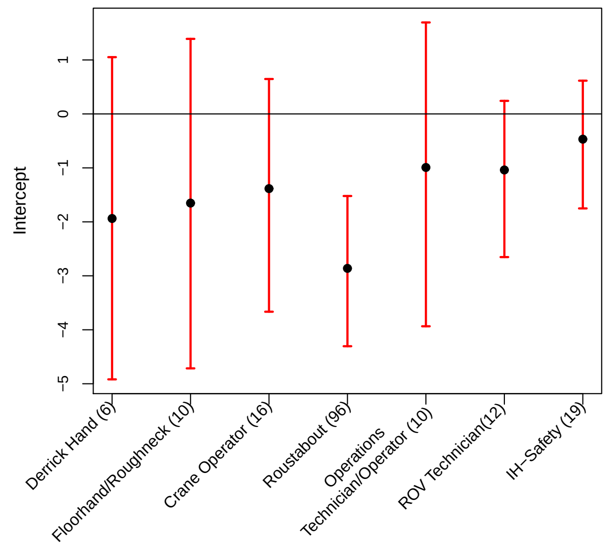
Model Parameter	Median	95% Credible Interval
μ_{β_0}	-1.47	(-6.52, 3.47)
μ_{β_1}	0.71	(0.44, 0.97)
$\mathbf{V}_{11}$	32.93	(13.07, 126.43)
$\mathbf{V}_{22}$	0.06	(0.02, 0.27)
$\rho(\beta_0, \beta_1)$	-0.16	(-0.75, 0.57)

Table S5: Posterior inference for the hyperparameter (global parameter) estimates in (**??**) on the *DDIII* between May 15-July 15, 2010. The parameters V_{11} and V_{22} are the diagonal elements of $\mathbf{V}_{\beta}$. The correlation between the intercepts and slopes is reported as $\rho(\beta_0, \beta_1)$. It was not significant and does not feature in the substantive inference.

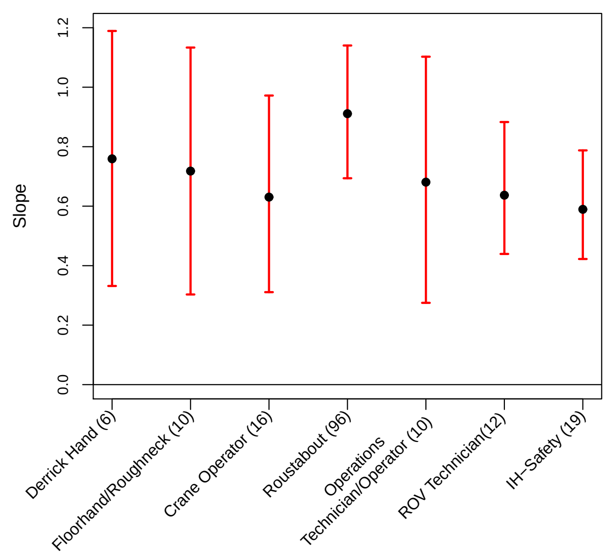
Parameter	Derrick Hand Median	Derrick Hand 95% CI	Floorhand/Roughneck Median	Floorhand/Roughneck 95% CI	Crane Operator Median	Crane Operator 95% CI	Roustabout Median	Roustabout 95% CI
Intercept	-1.94	(-4.92, 1.05)	-1.65	(-4.72,1.39)	-1.38	(-3.67, 0.65)	-2.86	(-4.30, -1.52)
Slope	0.76	(0.33, 1.19)	0.72	(0.30, 1.13)	0.63	(0.31, 0.97)	0.91	(0.69, 1.14)
Correlation	0.77	(0.28, 0.96)	0.63	(0.25, 0.87)	0.69	(0.34, 0.89)	0.70	(0.57, 0.80)
GSD of THC	2.84	(1.79, 8.75)	4.98	(2.81, 10.82)	3.07	(2.18, 5.93)	3.12	(2.64, 3.88)
GSD of Xylene | THC	1.90	(1.41, 4.68)	4.09	(2.40, 10.08)	2.09	(1.60, 4.06)	2.84	(2.41, 3.53)
GSD of Xylene	2.99	(1.88, 8.36)	6.55	(3.51, 15.95)	2.87	(2.02, 5.74)	4.38	(3.50, 5.89)
AM of THC (ppb)	1630	(710, 13786)	3783	(1221, 26941)	919.96	(508, 2798)	680.49	(519, 961)
AM of Xylene (ppb)	47	(20, 368)	165	(40, 2085)	22	(12, 61)	36	(24, 60)

Parameter	Operations Technician or Operator Median	Operations Technician or Operator 95% CI	ROV Technician Median	ROV Technician 95% CI	IH-Safety Median	IH-Safety 95% CI
Intercept	-0.99	(-3.93, 1.69)	-1.04	(-2.65, 0.24)	-0.47	(-1.75, 0.62)
Slope	0.68	(0.27, 1.10)	0.64	(0.44, 0.88)	0.59	(0.42, 0.79)
Correlation	0.65	(0.27, 0.88)	0.94	(0.75, 0.99)	0.89	(0.70, 0.96)
GSD of THC	5.61	(3.01,11.25)	3.62	(2.18, 9.57)	2.77	(2.03, 5.06)
GSD of Xylene | THC	3.83	(2.22, 10.07)	1.33	(1.19, 1.80)	1.37	(1.24, 1.65)
GSD of Xylene	6.29	(3.17, 16.84)	2.41	(1.70, 4.96)	1.99	(1.64, 2.86)
AM of THC (ppb)	2078	(610, 14731)	936	(433, 5601)	922	(566, 2306)
AM of Xylene (ppb)	130	(34, 1754)	24	(15, 65)	33	(24, 52)

Table S6: Preliminary Results:*DDIII* May 15-July 15, 2010 hierarchical Bayesian EG model parameter estimates using uniform priors



(a) Intercepts



(b) Slopes

Figure S8: Intercepts and slopes for the EGs on the *DDIII* from May 15-July 15, 2010. The upper panel displays the intercepts from the regression model. The lower panel displays the slopes. The dots in each bar represent the median posterior samples.

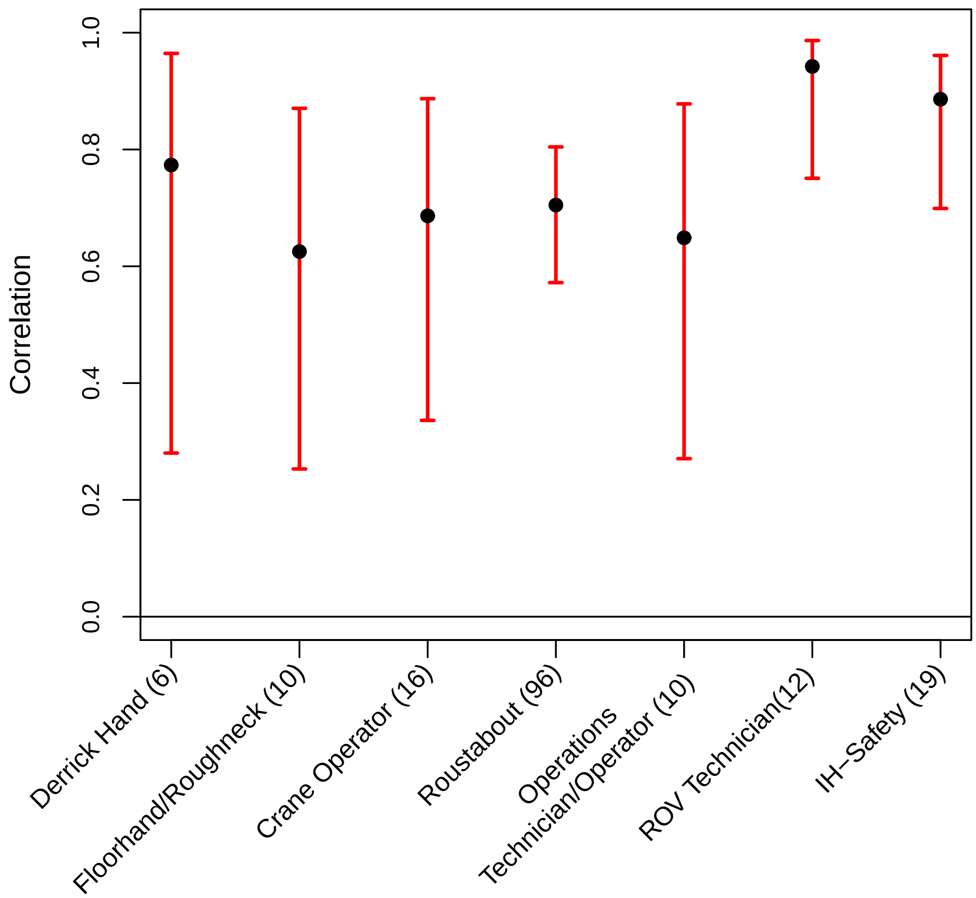


Figure S9: 95% Credible intervals of the posterior correlation estimates by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

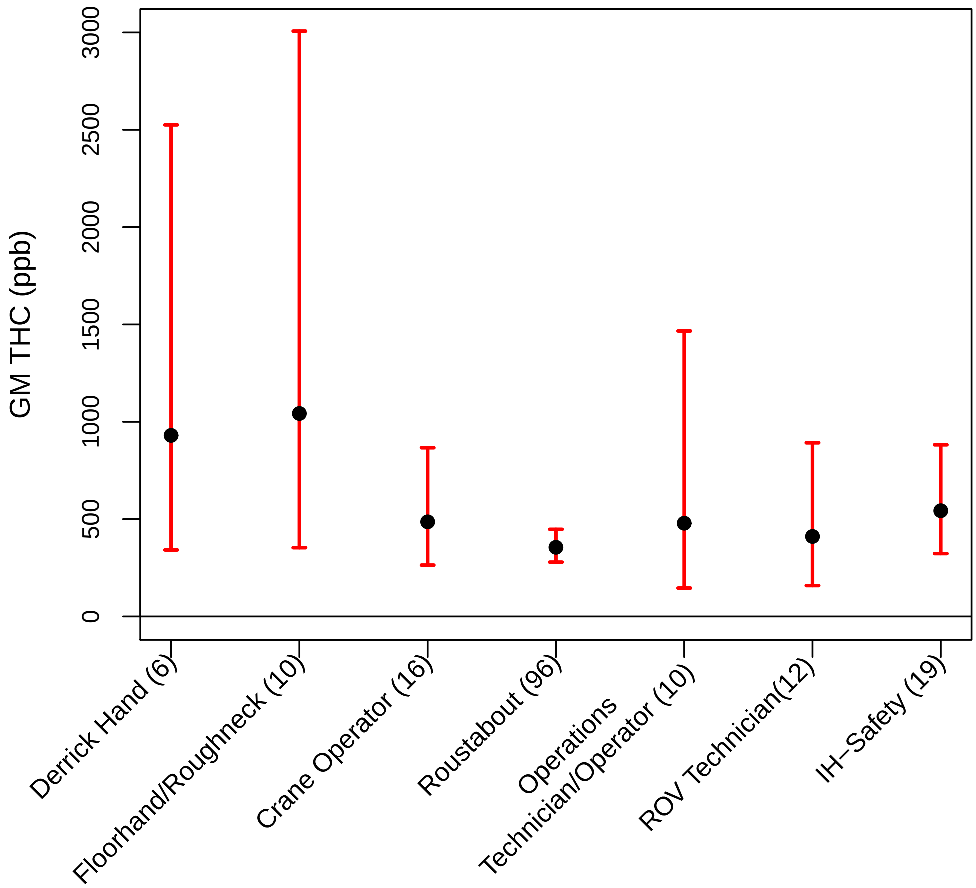


Figure S10: 95% Credible intervals of the posterior GMs estimates for THC by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

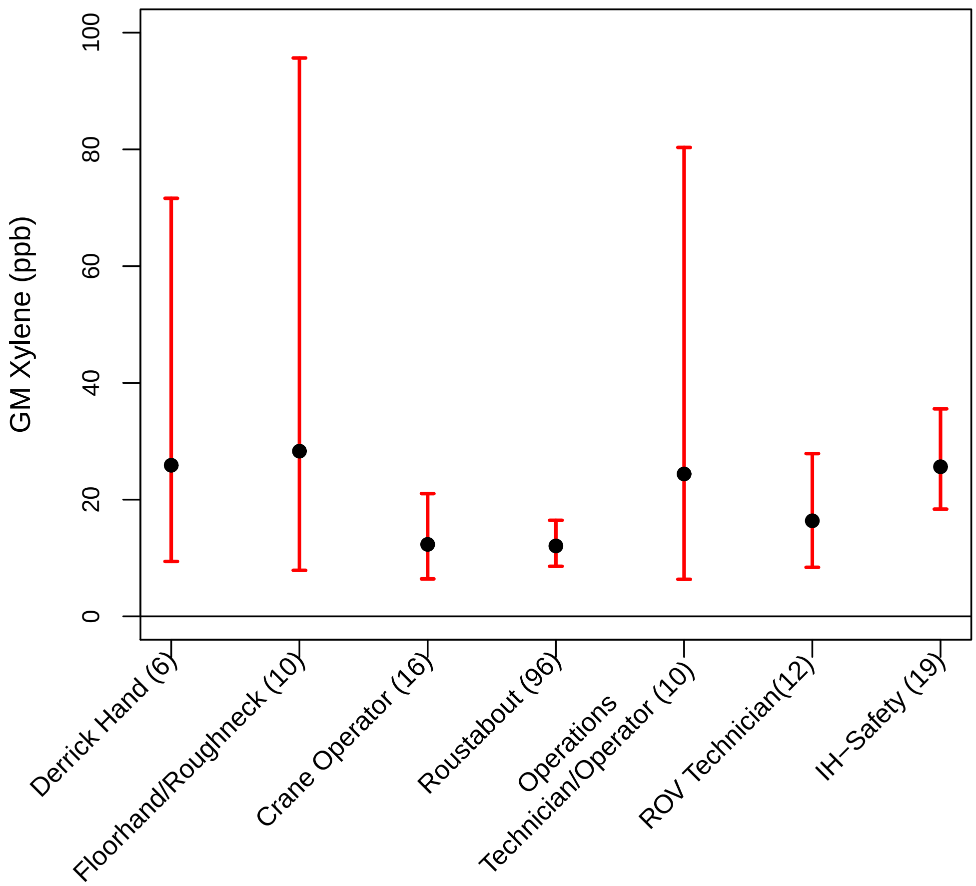
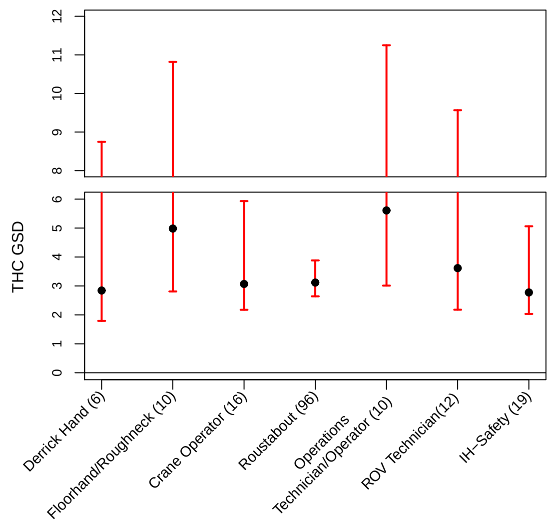
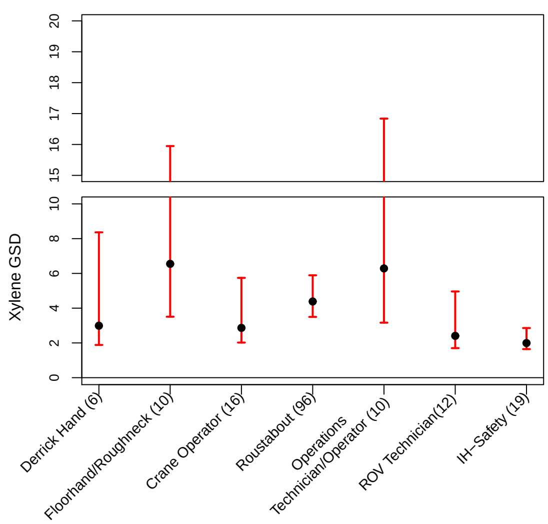


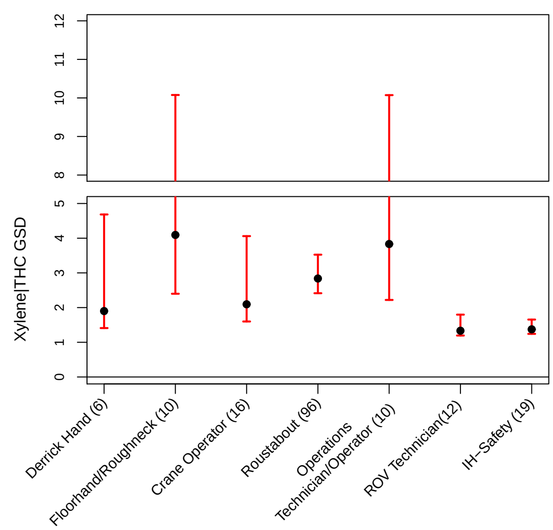
Figure S11: 95% Credible intervals of the posterior GMs estimates for xylene by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.



(a) GSD of THC



(b) GSD of Xylene



(c) GSD of Xylene|THC

Figure S12: Variance components of the model by EG on the *DDIII* from May 15-July 15, 2010. The top panel displays the GSD for THC. The bottom left panel displays the GSD for xylene. The bottom right panel displays the GSD for xylene given THC. The dots in each bar represent the median posterior samples.

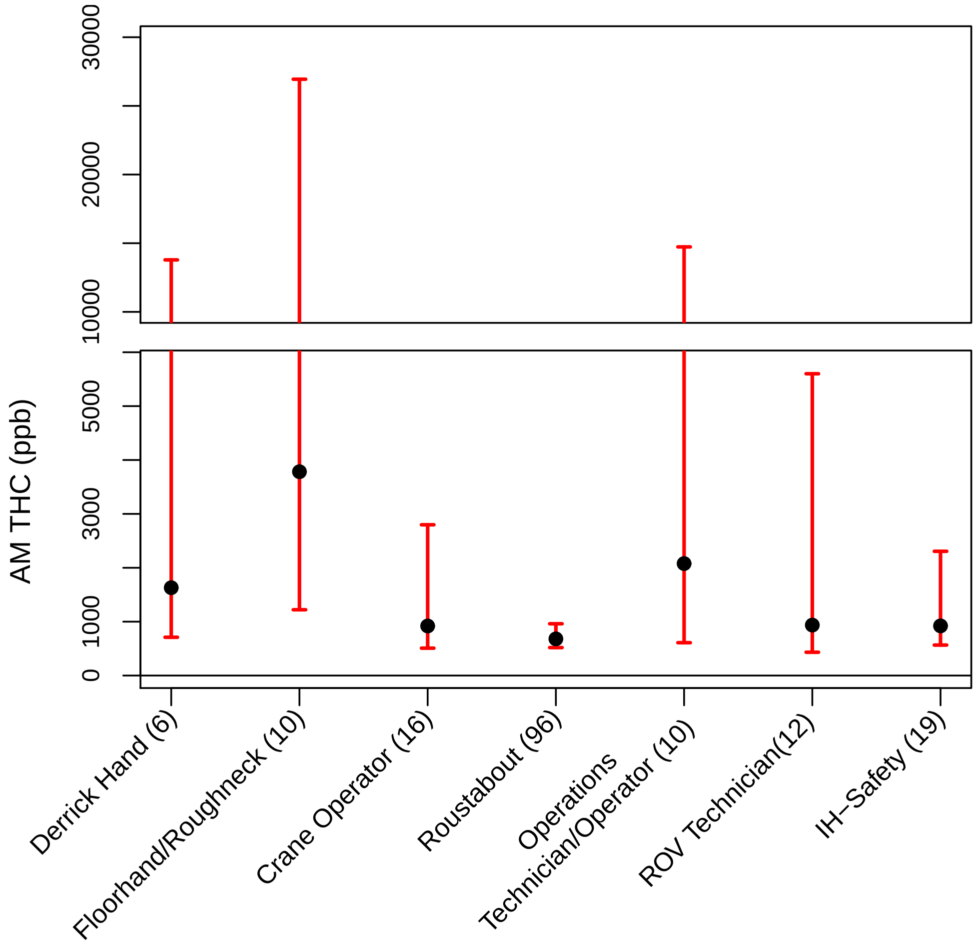


Figure S13: 95% Credible intervals of the posterior AMs estimates for THC by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.

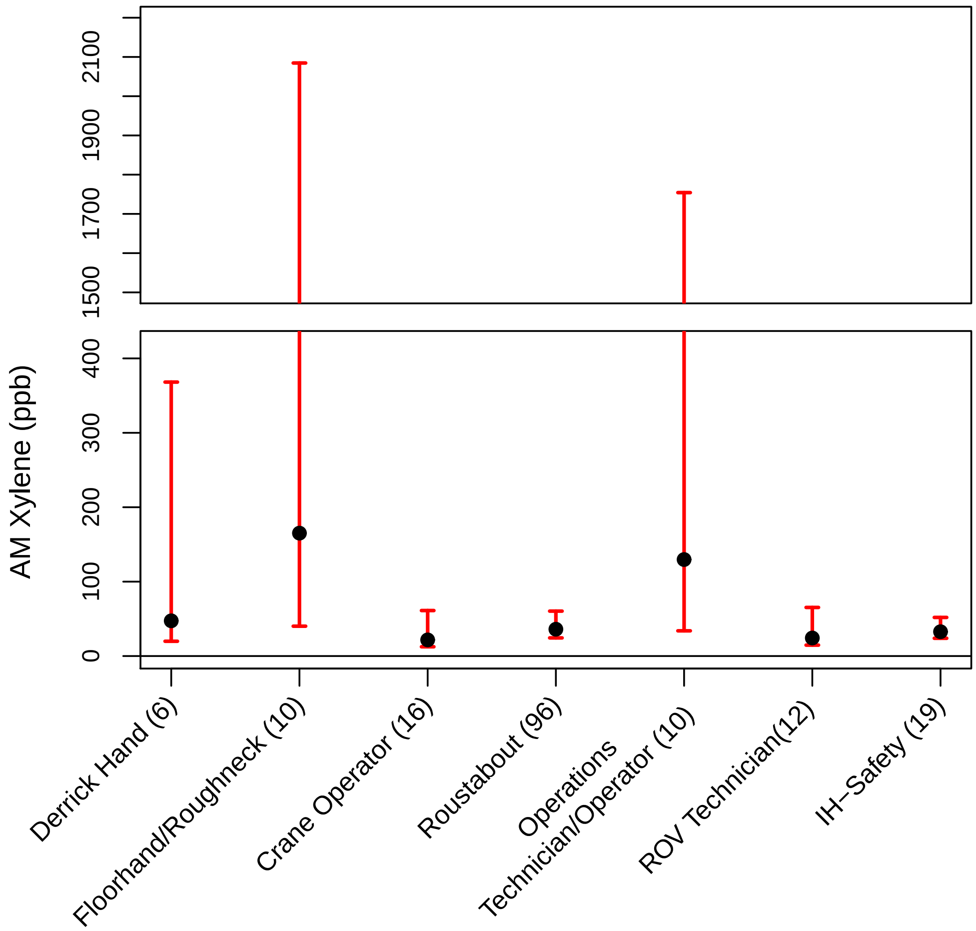


Figure S14: 95% Credible intervals of the posterior AMs estimates for xylene by EG on the *DDIII* from May 15-July 15, 2010. The dots in each bar represent the median posterior samples.