

Does size matter?
(to biosecurity risk)

Andrew Robinson and Tony Arthur

CEBRA, University of Melbourne
ABARES, DAWR

November 22, 2017



What presents biosecurity risk?

Propagule pressure.

What drives propagule pressure?

That depends on the underlying system.

Two competing models

1. constant rate, or
2. constant quantum.

(Probably, the truth lies within.)

Biosecurity risk is not necessarily just the rate.

We've spent a lot of time comparing pathways by rate.

What if there's more to the story?

Case Study: cut flowers pathway



Pathway Characteristics

```
nrow(cf.dat) # Number of Lines
```

```
## [1] 67504
```

```
length(unique(cf.dat$quarantine.entry)) # Number of entries
```

```
## [1] 34244
```

```
length(unique(cf.dat$country)) # Number of countries
```

```
## [1] 54
```

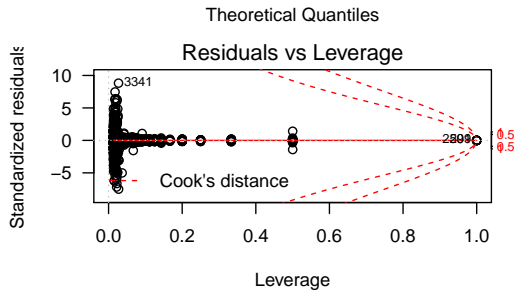
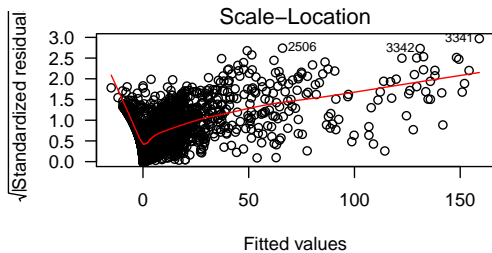
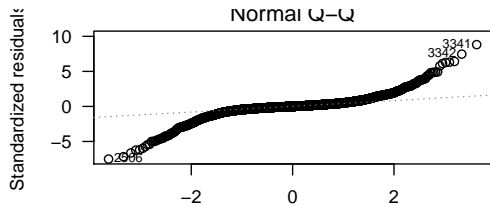
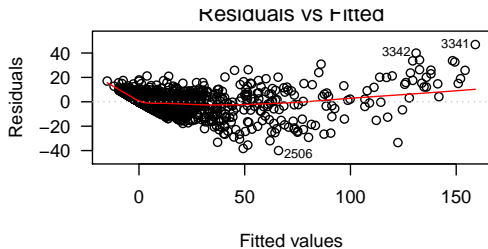
```
length(unique(cf.dat$stariff.num)) # Number of products
```

```
## [1] 8
```

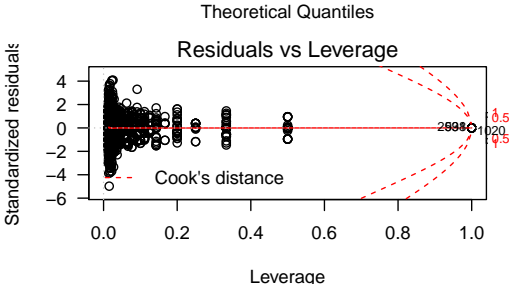
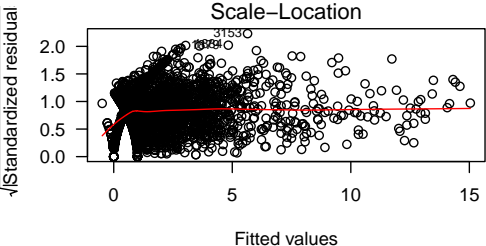
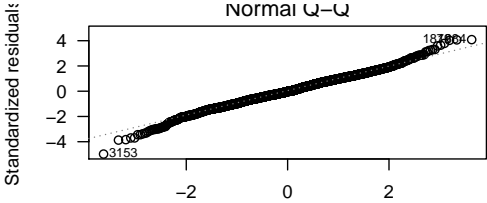
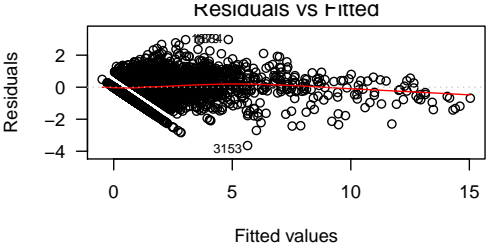
Regression on Monthly Counts

```
cf.m.lm.0 <- lm(fails ~ lines + country * tariff, data = cf.month.all)
```

Diagnostics



Sqrt Transformation Diagnostics



ANOVA

```
anova(cf.m.lm.0)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: fails
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
## lines	1	1074075	1074075	36831.939	< 2.2e-16	***
## country	53	59088	1115	38.231	< 2.2e-16	***
## tariff	7	5869	838	28.750	< 2.2e-16	***
## country:tariff	107	42390	396	13.585	< 2.2e-16	***
## Residuals	3420	99732	29			

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Switch term order to be sure.

```
cf.m.lm.2 <- lm(terms(sqrt(fails) ~ country * tariff + lines,  
                keep.order = TRUE), data = cf.month.all)
```

```
anova(cf.m.lm.2)
```

```
## Analysis of Variance Table
```

```
##
```

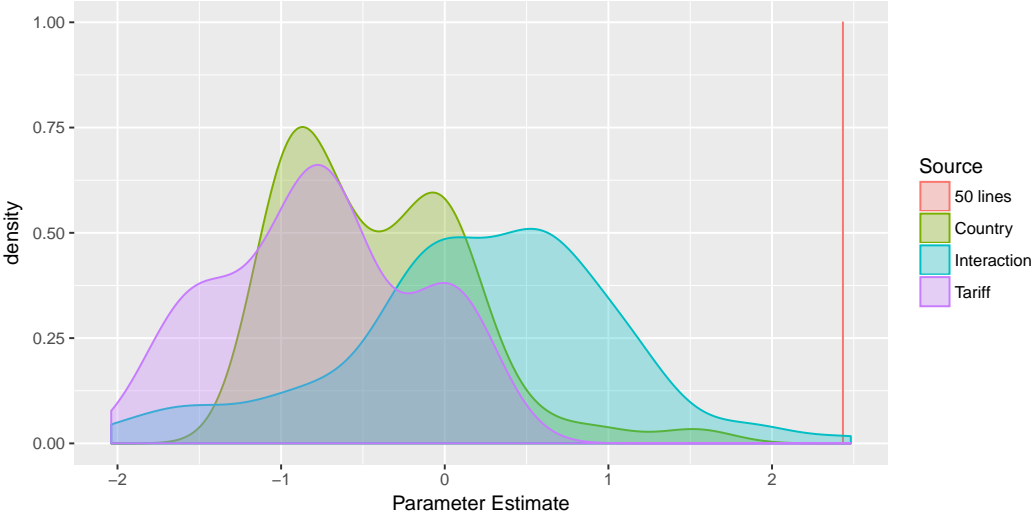
```
## Response: sqrt(fails)
```

```
##           Df Sum Sq Mean Sq  F value    Pr(>F)  
## country      53 5587.1   105.42   193.377 < 2.2e-16 ***  
## tariff        7 2259.0    322.72   591.998 < 2.2e-16 ***  
## country:tariff 107 3737.6     34.93    64.078 < 2.2e-16 ***  
## lines         1 2361.6  2361.65 4332.222 < 2.2e-16 ***  
## Residuals   3420 1864.4     0.55
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Coefficient Comparison



Relative Importance

```
library(relaimpo)

cf.m.lm.2.ri <-
  calc.relimp(lm(sqrt(fails) ~ country + tariff + lines,
                 data = cf.month.all))
```

```
cf.m.lm.2.ri@lmg

##   country   tariff    lines
## 0.2226866 0.0769322 0.5522962
```

55.2% of variability correlated to number of lines. A lot!

Robustness: Use Weeks instead of Months

```
cf.m.lm.3.ri <-  
  calc.relimp(lm(sqrt(fails) ~ country + tariff + lines,  
                data = cf.week.all))
```

```
cf.m.lm.3.ri@lmg
```

```
##      country      tariff      lines  
## 0.21917477 0.06712094 0.45578590
```

45.6% of variability correlated to number of lines. Still a lot!

Conclusion

- ▶ Volume matters a lot in this instance. If you want to predict future biosecurity threats, then you really need to predict the volume.
- ▶ A high-volume, low-rate pathway will present a greater biosecurity risk than a low-volume high-rate pathway, *ceteris paribus*.
- ▶ Of course, risk is easier to ameliorate when it is easier to find. But if we want to get all cost-benefit-y, then we should do so explicitly.

Thanks to ...

The Department of Agriculture and Water Resources for provision of data and funding, and New Zealand's Ministry for Primary Industries for support.