Changing the graphical presentation of multiplicative interaction effects: Improving the communication of empirical social research

Dag Tanneberg

General considerations

Consider the standard linear model with a single multiplicative interaction term. Let x denote the predictor of substantive interest and let z be the moderator. The fully specified regression fully

$$y = f(x, z, \epsilon)y = \beta_0 + \beta_1 x + \beta_2 z + \beta_3 xz + \epsilon$$

β_1	β_2	β_3	Joint Effect
> 0	> 0	> 0	Mutual augmentation
< 0	> 0	> 0	β_1 attenuated, β_2 augmented
> 0	< 0	> 0	β_1 augmented, β_2 attenuated
> 0	> 0	< 0	Mutual attenuation
< 0	> 0	< 0	β_1 augmented, β_2 attenuated
> 0	< 0	< 0	β_1 attenuated, β_2 augmented
< 0	< 0	< 0	Mutual augmentation
< 0	< 0	> 0	Mutual attenuation

Assuming $x, z \geq 0$.

Toy data setup

Original: Kam & Franzese, Multiplicative Interaction Terms Download: Government duration data

```
##   id govdur   PS  NP PD
## 1  2   17.1 52.0 1.0  0
## 2  3   32.2 58.9 2.2  1
## 3  4   11.9 59.0 4.3  0
## 4  5   12.3 52.1 2.7  0
## 5  6   32.8 54.3 1.0  1
## 6  7   30.0 55.0 1.0  0
```

Fit model & generate predicted values

```r
# fit model & generate predictions ------------------------
fit <- lm(govdur ~ NP*PS + PD, data = dta)
summary(fit)
```

```
##
## Call:
## lm(formula = govdur ~ NP * PS + PD, data = dta)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.2900 -4.4670 -0.1628  4.2164 14.3963
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  59.2735    26.4553   2.241  0.03870 *
## NP          -31.3703    11.3451  -2.765  0.01324 *
## PS           -0.5861     0.4544  -1.290  0.21435
## PD            9.8470     3.2039   3.073  0.00689 **
## NP:PS         0.4686     0.1863   2.516  0.02222 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.869 on 17 degrees of freedom
## Multiple R-squared:  0.6044, Adjusted R-squared:  0.5113
## F-statistic: 6.494 on 4 and 17 DF,  p-value: 0.002315
```

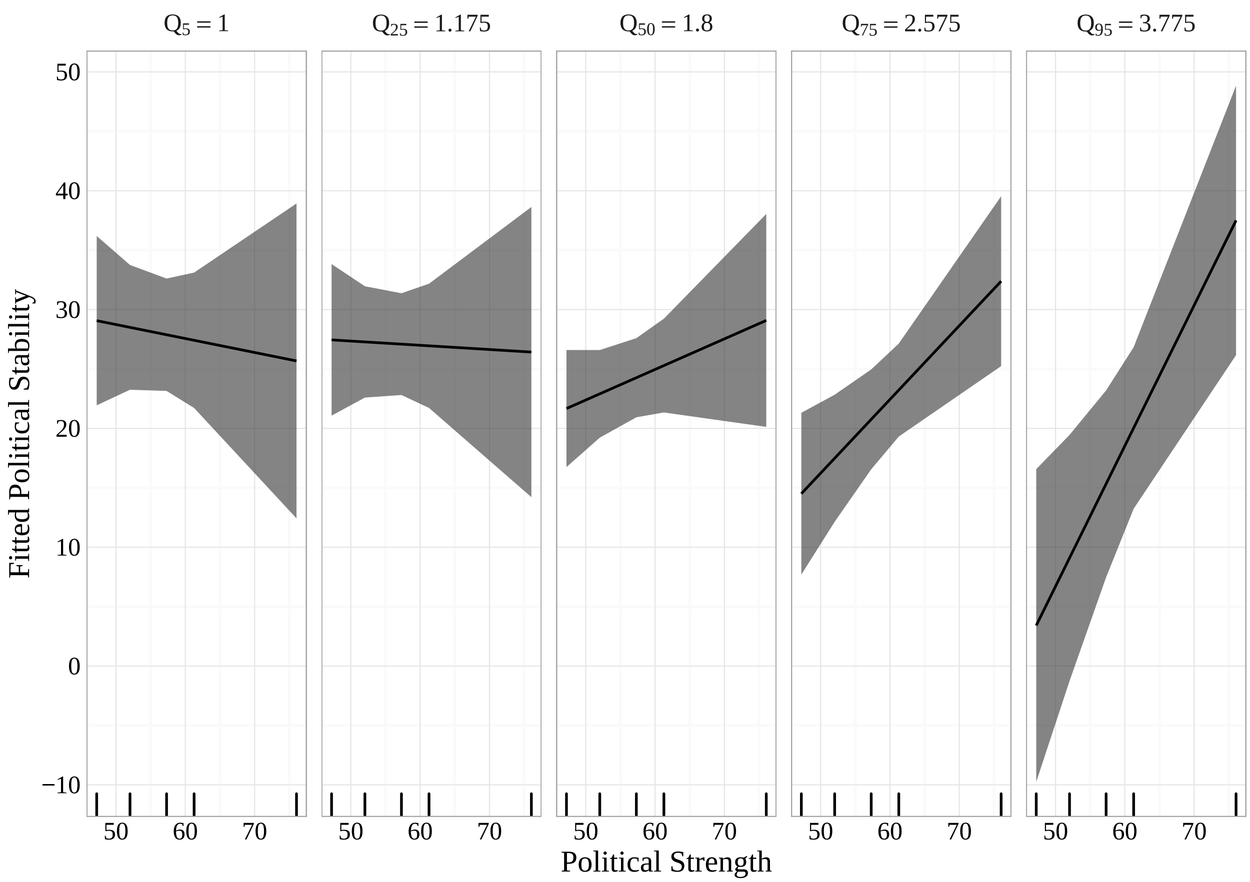
```r
quantiles <- c(.05, .25, .5, .75, .95)
pred.dta <- expand.grid(
  # expand.grid covers all quantile combinations
  NP = quantile(dta[, 'NP'], quantiles),
  PS = quantile(dta[, 'PS'], quantiles),
  PD = mean(dta[, 'PD'])
)
pred.y <- predict(
  fit, newdata = pred.dta, type = 'response',
  se.fit = TRUE, interval = 'confidence'
)
pdta <- data.frame(
  pred.dta, pred.y[['fit']], se.fit = pred.y[['se.fit']]
)
rm(quantiles, pred.y)
pdta <- within(pdta,
  NP.cat <- factor(NP, levels = sort(unique(NP)),
    labels = c(
      'Q[5] == 1', 'Q[25] == 1.175', 'Q[50] == 1.8',
      'Q[75] == 2.575', 'Q[95] == 3.775'
    )
  )
)
```

Presentation using Predicted Values

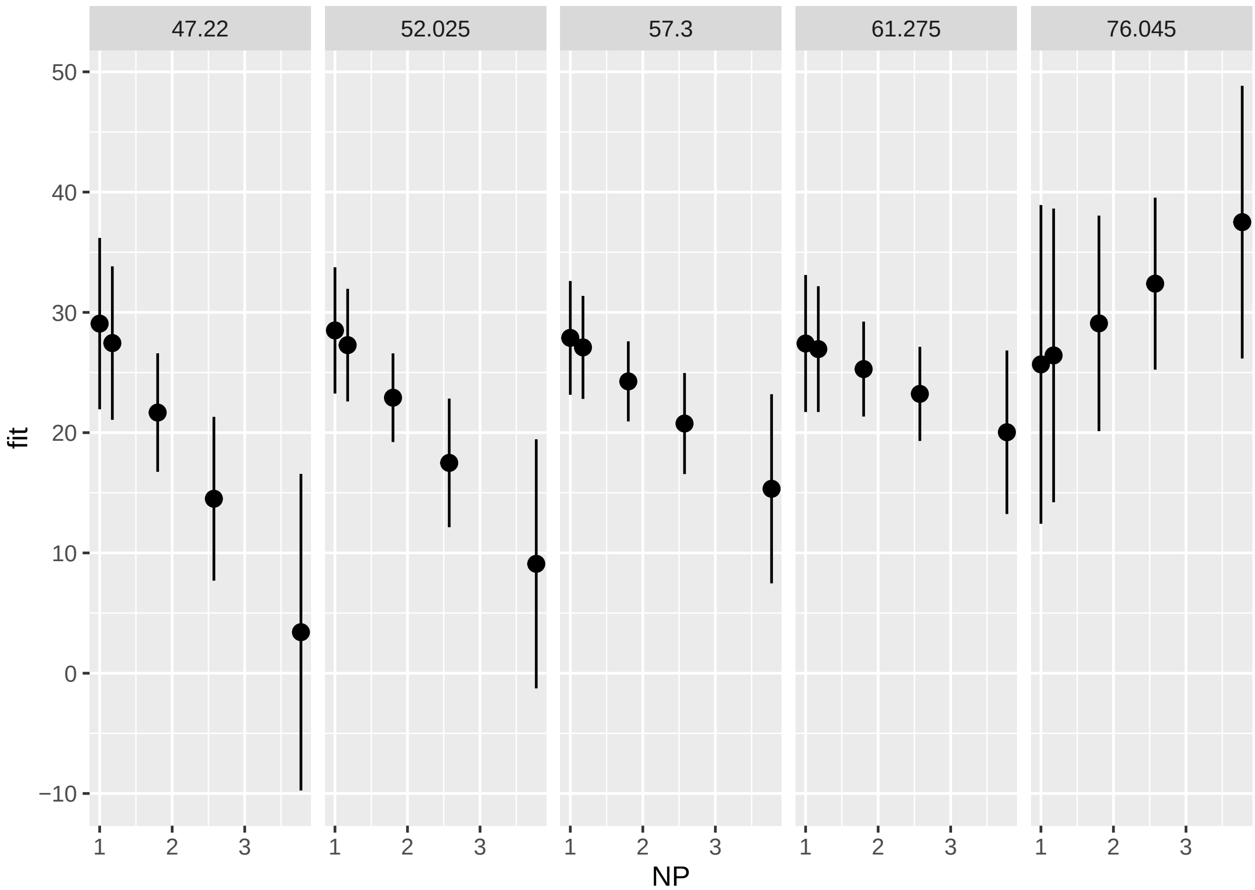
An equally powerful presentation of interaction effects focuses on the relationship between the response and the focal predictor Such effect displays plot the predicted values of the response against the focal predictor, while fixing the moderator at select values. In effect, the marginal effect on the focal predictor is replaced with the expected level of the dependent variable itself, and the conditional relationship can be interpreted in the metric of the depedent variable itself. Using these graphs makes it fairly easy to translate formal statistical models into the substantive quantities that led to the formulation of the conditional hypothesis in the first place.

This variant of effect displays becomes a particularly effective means for communicating statistical results if at least the the focal predictor is a continuous variable. In the past many scholars resorted to either discretization or focused on

```
ggplot(pdta, aes(x = PS, y = fit)) +
  geom_ribbon(aes(ymin = lwr, ymax = upr), alpha = .6) +
  geom_line() +
  geom_rug(sides = 'b') +
  labs(
    x = 'Political Strength', y = 'Fitted Political Stability'
  ) +
  facet_grid(~NP.cat, labeller = label_parsed) +
  theme_minimal(base_size = 12, base_family = 'serif') +
  theme(panel.border = element_rect(size = .3, colour = 'grey65', fill = 'transparent'))
```



```
ggplot(pdta, aes(x = NP, y = fit)) +
  geom_pointrange(aes(ymin = lwr, ymax = upr)) +
  facet_grid(~PS)
```



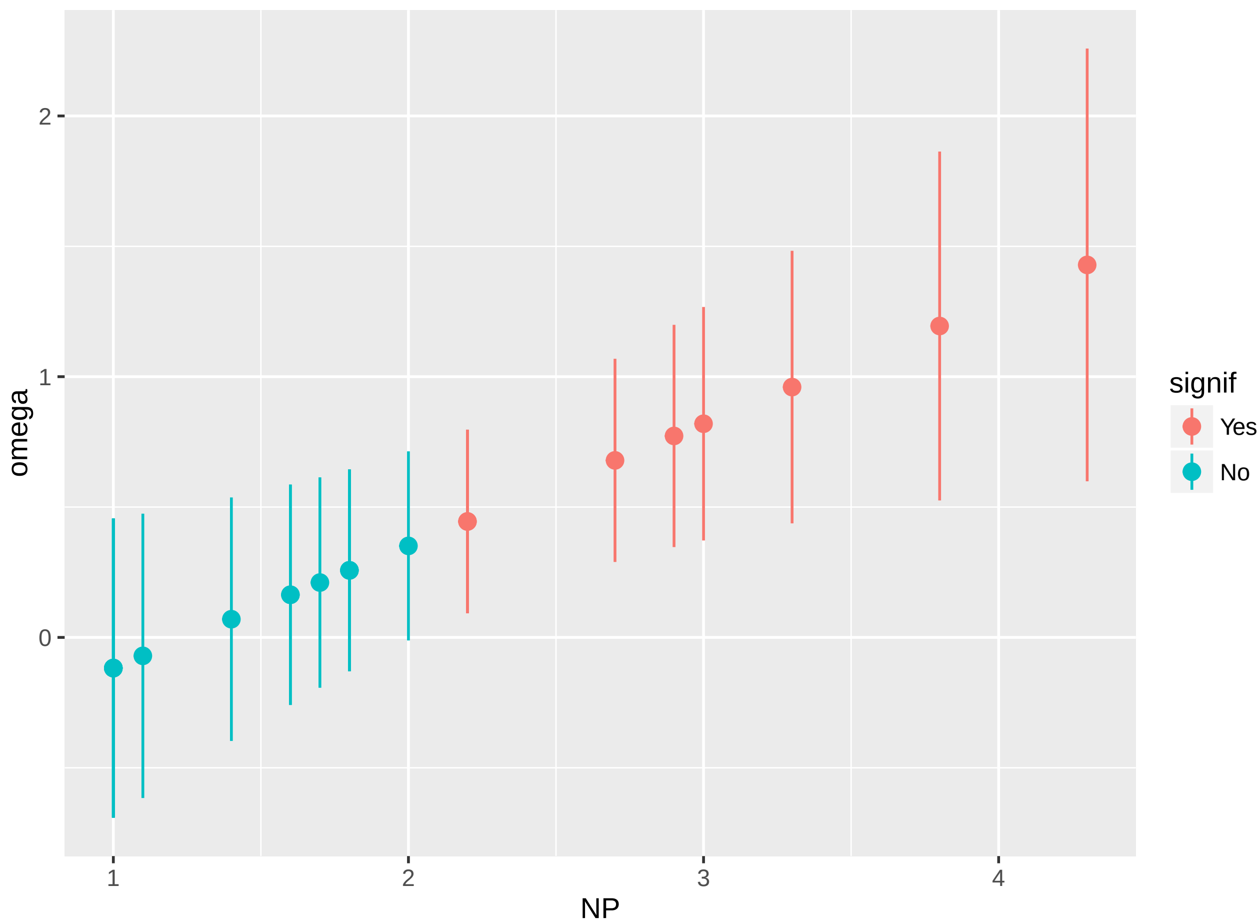
Version 2: Simple slopes approach

The conditional marginal effect of the focal predictor is plotted against the moderator. This plots emphasizes areas of statistical significance.

```
beta <- coef(fit)
omega <- beta['PS'] + beta['NP:PS'] * dta[, 'NP']
se.fit <- sqrt(
  diag(vcov(fit))['PS'] +
  2 * dta[, 'NP'] * vcov(fit)['PS', 'NP:PS'] +
  (dta[, 'NP']^2) * diag(vcov(fit))['NP:PS']
)

upr <- omega + 1.96 * se.fit
lwr <- omega - 1.96 * se.fit
pdta.bauer <- data.frame(
  omega = omega,
  upr = upr,
  lwr = lwr,
```

```r
  se.fit = se.fit,
  NP = dta[, 'NP']
)
pdta.bauer <- within(pdta.bauer, {
  signif <- ifelse(lwr < 0 & upr > 0, 1, 0)
  signif <- factor(signif, 0:1, c('Yes', 'No'))
  }
)
rm(beta, omega, se.fit, upr, lwr)
ggplot(
  data = pdta.bauer,
  aes(x = NP, y = omega, ymin = lwr, ymax = upr, colour = signif)
) +
  geom_pointrange()
```



Version 3: Chloropleth map

This map plots both independent variables against each other. The expected value of the response is used to as fill color. This plots ignores uncertainty.

```r
# simple map ----------------------------------------------
ggplot(data = pdta,
```

```r
  aes(
    x = factor(format(NP, digits = 2)),
    y = factor(format(PS, digits = 3)),
    fill = fit #, alpha = 1/se.fit
  )
) +
  geom_tile(colour = 'white', size = .3) +
  scale_fill_gradient(low = '#2166ac', high = '#b2182b') +
  labs(
    x = 'Number of Cabinet Parties',
    y = 'Political Strength',
    fill = 'Fitted Government Duration'
  ) +
  theme_minimal() +
  theme(
    text = element_text(family = 'serif'),
    legend.key.height = grid::unit(.5, 'lines'),
    legend.position = 'top',
    legend.direction = 'horizontal',
    axis.ticks = element_blank(),
    panel.grid = element_blank(),
    plot.background = element_blank()
  )
```