How Much Should We Trust Estimates from Multiplicative Interaction Models?
Simple Tools to Improve Empirical Practice

Jens Hainmueller (Stanford)  Jonathan Mummolo (Princeton)  Yiqing Xu (UCSD)
Motivation

• Political scientists often build theories based on conditionality: the effect of D on Y depends on X

• Most widely used way to explore conditional relationships is to estimate multiplicative interaction models

\[ Y = \mu + \alpha D + \eta X + \beta (D \cdot X) + \epsilon \]

• Brambor, Clark and Golder (2006) provides a useful checklist; cited over 3,000 times
Motivation

- We argue that two important problems remain with empirical practice:
  
  - **Linear interaction effect assumption (LIE)**

\[
Y = \mu + \alpha D + \eta X + \beta (D \cdot X) + \epsilon \\
ME_D = \frac{\partial Y}{\partial D} = \alpha + \beta X
\]
Hellwig and Samuels (2007) CPS
Motivation

• We argue that two important problems remain with empirical practice:

• Linear interaction effect assumption (LIE)

• **Lack of common support**
  
  • We need variation on the treatment throughout the range of the moderator, otherwise…
Motivation

• We argue that two important problems remain with empirical practice:

• Linear interaction effect assumption (LIE)

• Lack of common support
  • interpolation /extrapolation
  • model dependence
  • fragile results (King and Zeng 2006)
How to diagnose?

Look at the data...
Simulated Samples

\[ Y_i = 5 - 4X_i - 9D_i + 3D_i X_i + \epsilon_i, \quad i = 1, 2, \ldots, 200. \]

\[ X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(3, 1) \]

\[ D_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(0.5) \]

\[ \epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 4) \]

\[ ME_D = -9 + 3X_i \]
Simulated Samples

\[ Y_i = 5 - 4X_i - 9D_i + 3D_i X_i + \epsilon_i, \quad i = 1, 2, \cdots, 200. \]

\[ X_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(3, 1) \]

\[ D_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(3, 1) \]

\[ \epsilon_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 4) \]

\[ ME_D = -9 + 3X_i \]
Diagnostic Tools

- Scatterplot
- Binning Estimator
- Kernel Estimator
Scatterplots

Treatment = 0

Treatment = 1
Scatterplots
Generalized Additive Model (GAM)
Diagnostic Tools

- Scatterplots
- Binning Estimator
- Kernel Estimator
Binning Estimator

1. Cut X into three (or more) bins

\[ G_1 = \begin{cases} 
1 & X < \delta_{1/3} \\
0 & \text{otherwise} 
\end{cases}, \quad G_2 = \begin{cases} 
1 & X \in [\delta_{1/3}, \delta_{2/3}) \\
0 & \text{otherwise} 
\end{cases}, \quad G_3 = \begin{cases} 
1 & X \geq \delta_{2/3} \\
0 & \text{otherwise} 
\end{cases} \]

2. Estimate key coefficients within each bin

\[ Y = \sum_{j=1}^{3} \left\{ \mu_j + \alpha_j D_i + \eta_j (X - x_j) + \beta_j (X - x_j) D \right\} G_j + \gamma Z + \epsilon \]

3. Obtain marginal effects at specified points

\[ ME(x_j) = \alpha_j = \alpha + \beta x_j \]
Binning Estimator

Marginal effect of D on Y

Moderator: X
Binning Estimator

Moderator: X

Marginal effect of D on Y

0.0 2.5 5.0 7.5

0 0

-10
Binning Estimator

Marginal effect of D on Y

Moderator: X

L M H
Binning Estimator

Marginal effect of D on Y
Binning Estimator

Dichotomous Treatment

Continuous Treatment

Marginal effect of D on Y

Moderator: X

Dichotomous Treatment

Continuous Treatment

Moderator: X
Binning Estimator

Dichotomous Treatment

Continuous Treatment

Marginal effect of D on Y
Diagnostic Tools

- Scatterplots
- Binning Estimator
- Kernel Estimator
Kernel Estimator

Assume the true model is:

\[ Y = f(X) + g(X)D + \gamma(X)Z + \epsilon. \]

1. Imagine each bin is now represented by a “kernel”

\[ K \left( \frac{X_i - x_0}{h} \right) \]

2. Estimate key coefficients within each “bin”

\[ L = \sum_{i=1}^{N} \left\{ \left[ Y_i - \bar{\mu} - \bar{\alpha}D_i - \bar{\eta}(X_i - x_0) - \bar{\beta}D_i(X_i - x_0) - \bar{\gamma}Z_i \right]^2 K \left( \frac{X_i - x_0}{h} \right) \right\} \]

3. Obtain marginal effects at each evaluation point

\[ \hat{g}(x_0) = \hat{\alpha}(x_0) + \hat{\beta}(x_0)x_0. \]
Kernel Estimator

Marginal effect of D on Y

Moderator: X

25
Kernel Estimator

Marginal effect of D on Y

Moderator: X
Kernel Estimator

Marginal effect of D on Y
Kernel Estimator

Marginal effect of D on Y

Moderator: X
Kernel Estimator

Marginal effect of D on Y

Moderator: X
Kernel Estimator

Dichotomous Treatment

Marginal effect of D on Y

Moderator: X

Continuous Treatment

Marginal effect of D on Y

Moderator: X
How prevalent are these issues in published work?
Case Selection

• Time span: 2006-2015

• APSR and AJPS: All papers citing BCG (2006) or having a marginal effects plot

• JOP, IO and CPS: All papers citing BCG (2006)

• Criteria
  • Conditional effect central to main argument
  • Linear models (OLS and Fixed Effects)
  • Moderator has more than 5 values (continuous)
  • Replicable (40—>22 papers and 46 interaction effects)
Huddy et al. (2015) APSR

FIGURE 2. The Marginal Effect of Experimental Party Threat and Reassurance on Anger and Enthusiasm by Partisan Identity and Ideological Issue Intensity

A. Blog Study: Anger

B. Blog Study: Enthusiasm

C. Blog Study: Anger

D. Blog Study: Enthusiasm
Huddy et al. (2015) APSR
Huddy et al. (2015) APSR

![Graph showing the marginal effect of threat on anger as a function of partisan identity. The graph includes a line of best fit and error bars for different levels of partisan identity (L, M, H).]
Marginal effect of lag CBI on FDI
Marginal effect of regime type on overhang
Marginal effect of change in valence on change in vote share
Marginal effect of anger on policy opinion

$-30$
$-20$
$-10$
$-0.8$
$-0.4$
$0.25$
$0.0$
$0.5$
$Banks & Valentino (2012)

Moderator: change in party dispersion
Moderator: imports
Moderator: Polity

$5$
$-0.5$

$10$
$0.0$
$0.4$
$0.50$

$Banks & Valentino (2012)$

Marginal effect of CBI on M2 change
Marginal effect of NPC membership on return on assets
Marginal effect of economy on election

$-2$
$1$
$-1$
$-0.05$
$20$
$-10$
$0.25$
$0.00$
$2$
$15$
$0.5$
$0.75$

$Hellwig & Samuels (2007)$

Marginal effect of CBI on inflation
Marginal effect of Share Non-Citizens on Change in Ed. Services
Marginal effect of treatment on question count (a)

$-0.0$
$-5$
$-2$
$-1$
$0$
$1$
$2$
$3$
$4$
$5$
$6$
$7$
$8$
$9$
$10$
$15$
$60000$
$80000$
$20000$

$Malesky et al. (2012)$

Marginal effect of lag CBI on 10
Marginal effect of Share Non-Citizens on Change in Social Services
Marginal effect of treatment on critical questions % (a)

$-2$
$-1$
$0$
$1$
$2$
$3$
$4$
$5$
$6$

$Bodea and Hicks (2015b)$

Marginal effect of women's only petition % on total names/1000
Marginal effect of presidential elections on effective no. of parties
Marginal effect of UN authorization on rallies

$-2$
$-1$
$0$
$1$
$2$
$3$
$4$
$5$
$6$
$7$
$8$
$9$
$10$
$20$
$40$
$60$
$80$

$Bodea and Hicks (2015b)$

Marginal effect of economy on election
Marginal effect of lag positive reinforcement on renewable energy share
Marginal effect of UN authorization on rallies

$-2$
$-1$
$0$
$1$
$2$
$3$
$4$
$5$
$6$
$7$
$8$
$9$
$10$
$20$
$40$
$60$
$80$

$Chapman (2009)$

Marginal effect of lag positive reinforcement on renewable energy share
Marginal effect of UN authorization on rallies

$-2$
$-1$
$0$
$1$
$2$
$3$
$4$
$5$
$6$
$7$
$8$
$9$
$10$
$20$
$40$
$60$
$80$

$CPS$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$5$
$-0.5$

$10$
$0.0$
$0.4$
$0.50$

$Pelc (2011)$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$IO$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$Clark & Leiter (2014)$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$CPS$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$Truex (2014)$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$APSR$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$Vernby (2013)$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$APSR$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$Malesky et al. (2012)$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$APSR$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$Malesky et al. (2012)$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$APSR$

Marginal effect of change in party dispersion
Marginal effect of imports
Marginal effect of Polity

$0.0$
$0.5$

$H$
Common Problems

1. Lack of common support
2. Severe interpolation
3. Nonlinearity
Common Problems

1. Lack of common support
2. Severe interpolation
3. Nonlinearity
FIGURE 2. Marginal effect of UN authorization by affinity with the Security Council.

Note: Dashed lines give 95 percent confidence interval.
Chapman (2009) IO

![Graph showing the relationship between UN authorization and US affinity with UN Security Council.]
Chapman (2009) IO

Moderator: US affinity with UN Security Council

Marginal effect of UN authorization on rallies
Chapman (2009) IO
Common Problems

1. Lack of common support
2. Severe interpolation
3. Nonlinearity
Malesky, Schuler and Tran (2012) APSR

FIGURE 1. Intensity of Treatment Effect

Change in Questions Asked

6th vs. 5th Session

6th Session vs. Average

Change in Criticism

6th vs. 5th Session

6th Session vs. Average

Note: Displays the marginal effect of treatment on number of critical questions asked and percentage of critical questions, based on internet penetration, which impacts the intensity experienced by delegates. The panels are derived from the fully-specified models (4, 8, 9, and 10) in Table 5. Triangles demonstrate marginal effects, with range bars representing 90% Confidence Intervals.
Malesky, Schuler and Tran (2012) APSR

Marginal effect of treatment on question count (a)

Moderator: internet penetration
Malesky, Schuler and Tran (2012) APSR

Dropping 4 extreme values (< 0.9% of data)
Malesky, Schuler and Tran (2012) APSR

Marginal effect of treatment on question count (a)
Malesky, Schuler and Tran (2012) APSR

Using block bootstrap
Common Problems

1. Lack of common support
2. Severe interpolation
3. Nonlinearity
Clark and Golder (2006) CPS

Figure 2
The Marginal Effect of Temporally Proximate Presidential Elections on the Effective Number of Electoral Parties

Marginal Effect of Presidential Elections

Effective Number of Presidential Candidates

--- 90% Confidence Intervals
Marginal effect of presidential elections on effective no. of parties

Moderator: effective no. of pres. candidates

Marginal Effect of presidential elections on effective no. of parties

Moderator: effective no. of pres. candidates
Scoring Cases

• Reject equal effects at low vs. high bin (+1)

• *No* severe extrapolation or interpolation (+1)
  — L-kurtosis < 0.16

• Monotonicity (+1)

• Fail to reject linear model (+1)
  — (Wald test of binned model vs. linear model)

• Our coding rules are conservative.
# Replication Results by Journal

<table>
<thead>
<tr>
<th>Journal</th>
<th>Papers</th>
<th>Cases</th>
<th>Low v. High</th>
<th>No Severe Extrap.</th>
<th>Monotonic</th>
<th>Linear</th>
<th>Score [0-4]</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJPS</td>
<td>5</td>
<td>9</td>
<td>0.22</td>
<td>0.22</td>
<td>0.44</td>
<td>0.56</td>
<td>1.4</td>
</tr>
<tr>
<td>APSR</td>
<td>7</td>
<td>17</td>
<td>0.29</td>
<td>0.59</td>
<td>0.41</td>
<td>0.29</td>
<td>1.6</td>
</tr>
<tr>
<td>CPS</td>
<td>5</td>
<td>10</td>
<td>0.30</td>
<td>0.70</td>
<td>0.40</td>
<td>0.50</td>
<td>1.9</td>
</tr>
<tr>
<td>IO</td>
<td>3</td>
<td>7</td>
<td>0.14</td>
<td>0.29</td>
<td>0.14</td>
<td>0.00</td>
<td>0.6</td>
</tr>
<tr>
<td>JOP</td>
<td>2</td>
<td>3</td>
<td>0.67</td>
<td>0.67</td>
<td>0.00</td>
<td>0.33</td>
<td>1.7</td>
</tr>
<tr>
<td><strong>Total/ Mean</strong></td>
<td><strong>22</strong></td>
<td><strong>46</strong></td>
<td><strong>0.28</strong></td>
<td><strong>0.50</strong></td>
<td><strong>0.35</strong></td>
<td><strong>0.35</strong></td>
<td><strong>1.1</strong></td>
</tr>
</tbody>
</table>
Recommendations

• Draw diagnostic *scatterplot*
  — check linearity of conditional expectation functions
  — check common support and that moderator is evenly distributed

• Compute the conditional marginal effects using the binning estimator and plot the estimates with a histogram of the moderator

• Apply the kernel estimator

• Check other model assumptions to the extent possible!
Thanks.

• R and Stata Package available:
  • SSC (STATA): ssc install interflex, replace all
  • CRAN (R): install.packages(“interflex”)
• More info on: yiqing.org/software