Comparing models, group differences, and interactions in BRM

Objectives

Comparing coefficients across models in the BRM
Comparing groups in the LRM
Review the BRM
Discuss the problem of interactions in the BRM
Discuss the problem of unobserved heterogeneity
Present a possible solution
Comparing coefficients across models in the BRM

```
. logit ybinary x1, nolog
Logistic regression                             Number of obs     =        500
LR chi2(1)        =     161.77
Prob > chi2       =     0.0000
Pseudo R2         =     0.2335
Log likelihood = -265.54468

------------------------------------------------------------------------------
ybinary |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+--------------------------------------------------
      x1  |   .7388677   .0729611    10.13   0.000     .5958667    .8818688
     _cons |  -.0529777   .1059112   -0.500   0.617    -.2605594    .1546048
------------------------------------------------------------------------------

. estimates store m1

. logit ybinary x2, nolog
Logistic regression                             Number of obs     =        500
LR chi2(1)        =     160.35
Prob > chi2       =     0.0000
Pseudo R2         =     0.2314
Log likelihood = -266.25298

------------------------------------------------------------------------------
ybinary |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+--------------------------------------------------
      x2  |   .4886751   .0482208    10.13   0.000     .394164    .5831861
     _cons |  -.0723833   .1058261   -0.678   0.494    -.2797987    .1350321
------------------------------------------------------------------------------
```
. est store m2

. logit ybinary x1 x2 , nolog

Logistic regression

Number of obs     =        500
LR chi2(2)        =     443.39
Prob > chi2       =     0.0000
Log likelihood = -124.73508
Pseudo R2         =     0.6399

------------------------------------------------------------------------------
ybinary | Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
x1 |  1.78923   .1823053     9.81   0.000     1.431918    2.146541
x2 |  1.173144   .1207745     9.71   0.000     .9364304    1.409858
_cons |  -.2144855   .1626923    -1.32   0.187    -.5333566    .1043856
------------------------------------------------------------------------------

. est store m3
. esttab m1 m2 m3, bic

<table>
<thead>
<tr>
<th></th>
<th>(1) ybinary</th>
<th>(2) ybinary</th>
<th>(3) ybinary</th>
</tr>
</thead>
<tbody>
<tr>
<td>ybinary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>0.739***</td>
<td></td>
<td>1.789***</td>
</tr>
<tr>
<td></td>
<td>(10.13)</td>
<td></td>
<td>(9.81)</td>
</tr>
<tr>
<td>x2</td>
<td></td>
<td>0.489***</td>
<td>1.173***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10.13)</td>
<td>(9.71)</td>
</tr>
<tr>
<td>_cons</td>
<td>-0.0530</td>
<td>-0.0724</td>
<td>-0.214</td>
</tr>
<tr>
<td></td>
<td>(-0.50)</td>
<td>(-0.68)</td>
<td>(-1.32)</td>
</tr>
<tr>
<td>N</td>
<td>500</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>BIC</td>
<td>543.5</td>
<td>544.9</td>
<td>268.1</td>
</tr>
</tbody>
</table>

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

**Question (for you)**

If we found this in the context of the LRM, what would we assume about x1 and x2?
. corr ybinary x1 x2, means
(obs=500)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ybinary</td>
<td>.488</td>
<td>.5003566</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>x1</td>
<td>-2.19e-08</td>
<td>2</td>
<td>-6.32646</td>
<td>6.401608</td>
</tr>
<tr>
<td>x2</td>
<td>3.57e-08</td>
<td>3</td>
<td>-10.56658</td>
<td>9.646875</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>ybinary</th>
<th>x1</th>
<th>x2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ybinary</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>0.5248</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>x2</td>
<td>0.5225</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

BUT...

x1 and x2 are uncorrelated!! What’s going on??
{REFRESHER} Estimated SD of $y^*$

$$SD(\hat{y}^*) = \sqrt{\text{var}(x\hat{\beta}) + \text{var}(\varepsilon)}$$
. est restore m1  
(results m1 are active now)

. listcoef, std

logit (N=500): Unstandardized and standardized estimates

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>x1</td>
<td>0.7389</td>
<td>10.127</td>
<td>0.000</td>
<td>1.478</td>
<td>0.316</td>
<td>0.632</td>
<td>2.000</td>
</tr>
<tr>
<td>constant</td>
<td>-0.0530</td>
<td>-0.500</td>
<td>0.617</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>
est restore m2
(results m2 are active now)

listcoef, std

logit (N=500): Unstandardized and standardized estimates

|         b        z    P>|z|    bStdX    bStdY   bStdXY     SDofX |
|----------|--------|---------|--------|--------|--------|--------|
| constant | -0.0724 -0.684  0.494     .        .        .         .    |
| x2       |  0.4887  10.134  0.000  1.466  0.210  0.629     3.000 |

Observed SD: 0.5004
Latent SD: 2.3322
. est restore m3  
(results m3 are active now)

. listcoef, std

logit (N=500): Unstandardized and standardized estimates

|        b        | z    | P>|z|  | bStdX | bStdY | bStdXY | SDofX |
|---------|-----|-----|-----|-------|-------|--------|-------|
| x1      | 1.7892 | 9.814  | 0.000 | 3.578 | 0.335 | 0.671 | 2.000 |
| x2      | 1.1731 | 9.714  | 0.000 | 3.519 | 0.220 | 0.659 | 3.000 |
| constant| -0.2145 | -1.318 | 0.187 | .     | .     | .     | .     |
What does it all mean?

• The variance of $y^*$ changes from one model to the next, in effect rescaling the coefficients.

• This happens because we fix the residual error variance (how?), so improvements in model fit must result in an improved ratio of explained to residual variance, which results in an increase in total variance!

• As a result, we should not compare our original coefficients across nested models, because the size of the coefficient is directly reflective of the variance of $y^*$, which is reflective of model fit.
So what should I do?

- Present y-standardized coefficients
- Don’t show coefficients for competing models; focus instead on comparison of fit statistics (LRX2, BIC, AIC)
Comparing across groups:

Common research question(s)

Is the effect of a variable the same across groups?

Is the effect of publishing on the probability of gaining tenure the same for men and women?

Do men and women get the same returns for publishing?
LRM

Comparing groups

Estimate separate regression models for two (or more groups) and then compare coefficients across groups

\[ y^m = \alpha^m + \beta^m_{articles} x_{articles} \]
\[ y^w = \alpha^w + \beta^w_{articles} x_{articles} \]

Estimate a single model for all groups combined, with interactions between dummy variables for groups and the variables of interest

\[ y = \alpha + \beta_{female} x_{female} + \beta_{articles} x_{articles} + \beta_{females*articles} x_{females*articles} \]
Do men and women get the same returns for publishing?

\[ H_0 : \beta^m_{\text{articles}} = \beta^w_{\text{articles}} \]

Or, a Chow type test

\[ z = \frac{\hat{\beta}^m_{\text{articles}} - \hat{\beta}^w_{\text{articles}}}{\sqrt{\text{Var}(\hat{\beta}^m_{\text{articles}})} + \text{Var}(\hat{\beta}^w_{\text{articles}})} \]
Comparing groups

Two problems (substantive & statistical)

- The substantive meaning of interactive coefficients are unclear
- The above tests confound:
  - Group differences in the effect of a predictor
  - Group differences in unobserved heterogeneity
BRM redux

Latent variable

Structural model
\[ y^* = \alpha + \beta x + \varepsilon \]

Measurement model
\( y^* \) is linked to the observed \( y \) by the measurement equation:
\[
y = \begin{cases} 
1 & \text{if } y_i^* > 0 \\
0 & \text{if } y_i^* \leq 0 
\end{cases}
\]
Graphically
Probit:

\[
Pr(y = 1 \mid x) = \int_{-\infty}^{x\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = \Phi(x\beta)
\]

Logit:

\[
Pr(y = 1 \mid x) = \frac{\exp(x\beta)}{1 + \exp(x\beta)} = \Lambda(x\beta)
\]
BRM redux

Transformational

How can we transform data to make parameters unbounded, and force $\Pr(y = 1 \mid x) < 1$ and $\Pr(y = 1 \mid x) > 0$?
Probit

$$\Phi^{-1}[\Pr(y = 1 \mid x)] = x\beta$$

$$\Pr(y = 1 \mid x) = \int_{-\infty}^{x\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt = \Phi(x\beta)$$

Logit

$$\text{logit}[\Pr(y = 1 \mid x)] = x\beta$$

$$\Pr(y = 1 \mid x) = \frac{\exp(x\beta)}{1 + \exp(x\beta)} = \Lambda(x\beta)$$
Assumption

In either case, the variance of the error must be assumed

Logit: \( \text{Var}(\varepsilon) = \frac{\pi^2}{3} \)

Probit: \( \text{Var}(\varepsilon) = 1 \)
Impact of assumption

With probit, $\varepsilon$ is rescaled so that:

$$Var(\varepsilon) = Var\left(\frac{\varepsilon}{\sigma}\right) = 1$$

With logit, $\varepsilon$ is rescaled so that:

$$Var(\varepsilon) = \frac{\pi \varepsilon}{\sqrt{3} \sigma} = \frac{\pi^2}{3}$$
Structural model

For men: \( y^* = \alpha^m + \beta^m_{\text{articles}} x_{\text{articles}} + \varepsilon^m \)

For women: \( y^* = \alpha^w + \beta^w_{\text{articles}} x_{\text{articles}} + \varepsilon^w \)

Estimated model (for probit)

For men: \( \frac{y^*}{\sigma^m} = \frac{\alpha^m}{\sigma^m} + \frac{\beta^m_{\text{articles}}}{\sigma^m} x_{\text{articles}} + \frac{\varepsilon^m}{\sigma^m} \)

For women: \( \frac{y^*}{\sigma^w} = \frac{\alpha^w}{\sigma^w} + \frac{\beta^w_{\text{articles}}}{\sigma^w} x_{\text{articles}} + \frac{\varepsilon^w}{\sigma^w} \)
Tests of group differences

We want to test:

\[ H_0 : \beta^m_{\text{articles}} = \beta^w_{\text{articles}} \]

Can only test:

\[ H_0 : \frac{\beta^m_{\text{articles}}}{\sigma^m} = \frac{\beta^w_{\text{articles}}}{\sigma^w} \]

These are only equivalent if:

\[ \sigma^m = \sigma^w \]
But
If in fact:

\[ O^m \neq O^w \]

Then:

"... the standard tests for cross-group differences in the \( \beta \) coefficients tell us nothing about differences in the [true] coefficients" (Allison, 1999).

“[I]n the presence of \textit{even fairly small differences} in residual variation, naive comparisons of coefficients [across groups] can \textit{indicate differences where none exist, hide differences that do exist, and even show differences in the opposite direction} of what actually exists” (Hoetker, 2004)
If the estimated relationship is:

$$\beta^m_{\text{articles}} > \beta^w_{\text{articles}}$$

and, in fact:

$$\sigma^m < \sigma^w$$

then...?
In a nutshell

We could have this problem:

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True model</strong></td>
<td>$y^* = \alpha + \beta_{\text{articles}} x_{\text{articles}} + \epsilon$</td>
<td>$y^* = \alpha + 2\beta_{\text{articles}} x_{\text{articles}} + 2\epsilon$</td>
</tr>
<tr>
<td><strong>Estimated model</strong></td>
<td>$y^* = \alpha + \beta_{\text{articles}} x_{\text{articles}} + \epsilon$</td>
<td>$y^* = \alpha + \beta_{\text{articles}} x_{\text{articles}} + \epsilon$</td>
</tr>
</tbody>
</table>

Hide differences that do exist!

Or, alternatively, this problem:

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True model</strong></td>
<td>$y^* = \alpha + \beta_{\text{articles}} x_{\text{articles}} + \epsilon$</td>
<td>$y^* = \alpha + \beta_{\text{articles}} x_{\text{articles}} + 2\epsilon$</td>
</tr>
<tr>
<td><strong>Estimated model</strong></td>
<td>$y^* = \alpha + \beta_{\text{articles}} x_{\text{articles}} + \epsilon$</td>
<td>$y^* = \alpha + .5\beta_{\text{articles}} x_{\text{articles}} + \epsilon$</td>
</tr>
</tbody>
</table>

Indicate differences where none exist!
Unobserved heterogeneity

It is reasonable to believe that women have more heterogeneous career patterns than men?

• More unobserved heterogeneity...

• Unmeasured variables affecting the chances of promotion may be more important for women than for men.

• That difference could explain why the coefficient for men is larger than for women
Allison’s (1999) solution

Assume that the effects of one variable are the same across groups.

Ratio of rescaled $\beta$ coefficients for that variable = ratio of standard deviation of the errors.
In other words

If we believe our coefficients for z should be equal:

Men: \( y^* = \alpha^m + \beta^m_{articles} x_{articles} + \beta^m_z z + \varepsilon^m \)

Women: \( y^* = \alpha^w + \beta^w_{articles} x_{articles} + \beta^w_z z + \varepsilon^m \)

Then we can assume that:

\[
\frac{\beta^m_z}{\beta^w_z} = \frac{\beta^m_z / \sigma^m}{\beta^w_z / \sigma^w} = \frac{\beta^m_z / \sigma^m}{\beta^w_z / \sigma^w} = \frac{\sigma^m}{\sigma^w}
\]

This provides the information we need to test:

\( H_0: \beta^m_{articles} = \beta^w_{articles} \)
Question (for you)

Under what conditions would you feel comfortable asserting that the effect of a variable should be equivalent across groups?
Williams’ (2009, 2010) solution

“Heterogeneous choice model” (a.k.a. location-scale model)

- Allows modeling of a more complex variance term
- Can be used for both binary and ordinal logit models
- \texttt{oglm} commands in Stata

But:

“[O]ne cannot distinguish empirically between the hypothesis of uniform proportionality of effects across transitions and the hypothesis that group differences between parameters of binary regressions are artifacts of heterogeneity between groups in residual variation.” (Hauser & Andrew 2006)
Long’s solution: Comparing groups using predicted probabilities

Predicted probabilities are invariant to the arbitrary assumption about the distribution of the error.

And, they provide a clear substantive interpretation.
Example

```
.setac cda_tnure01
.setcbo tenue female articles prestige , compact
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Unique</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>tenure</td>
<td>2797</td>
<td>2</td>
<td>.122989</td>
<td>0</td>
<td>1</td>
<td>Is tenured?</td>
</tr>
<tr>
<td>female</td>
<td>2797</td>
<td>2</td>
<td>.377547</td>
<td>0</td>
<td>1</td>
<td>Scientist is female?</td>
</tr>
<tr>
<td>articles</td>
<td>2797</td>
<td>48</td>
<td>7.05041</td>
<td>0</td>
<td>73</td>
<td>Total number of articles.</td>
</tr>
<tr>
<td>prestige</td>
<td>2797</td>
<td>98</td>
<td>2.64659</td>
<td>.65</td>
<td>4.8</td>
<td>Prestige of department.</td>
</tr>
</tbody>
</table>

```
.setac sum tenure female articles prestige
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>tenure</td>
<td>2797</td>
<td>.122989</td>
<td>.3284832</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>female</td>
<td>2797</td>
<td>.377547</td>
<td>.4848602</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>articles</td>
<td>2797</td>
<td>7.05041</td>
<td>6.575682</td>
<td>0</td>
<td>73</td>
</tr>
<tr>
<td>prestige</td>
<td>2797</td>
<td>2.64659</td>
<td>.7769724</td>
<td>.65</td>
<td>4.8</td>
</tr>
</tbody>
</table>
. logit tenure i.female articles prestige, nolog

Logistic regression                               Number of obs   =       2945
               LR chi2(3)      =     132.04
               Prob > chi2     =     0.0000
               Log likelihood = -1033.5545                       Pseudo R2       =     0.0600

------------------------------------------------------------------------------
tenure |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------|--------------|------------------|--------|---------|------------------|------------------|
female  |             |                  |        |         |                 |                 |
Female  | -.285549    | .1226154         | -2.33  | 0.020   | -.5258707       | -.0452273        |
articles| .0811034    | .0074894         | 10.83  | 0.000   | .0664245        | .0957824         |
prestige | -.3895368   | .0790089         | -4.93  | 0.000   | -.5443914       | -.2346822        |
_cons   | -1.530872   | .203357          | -7.53  | 0.000   | -1.929444       | -1.132299        
------------------------------------------------------------------------------
.// mgen for women and men as articles change and prestige is at mean
.mgen, at(articles=(0(5)40) female=0) stub(M)

Predictions from: margins, at(articles=(0(5)40) female=0) predict(pr)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Unique</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mpr1</td>
<td>9</td>
<td>9</td>
<td>.3229053</td>
<td>.0745109</td>
<td>.6622114</td>
<td>pr(y=Tenure) from margins</td>
</tr>
<tr>
<td>Mll1</td>
<td>9</td>
<td>9</td>
<td>.2719532</td>
<td>.0615271</td>
<td>.5591988</td>
<td>95% lower limit</td>
</tr>
<tr>
<td>Mul1</td>
<td>9</td>
<td>9</td>
<td>.3738575</td>
<td>.0874947</td>
<td>.7652241</td>
<td>95% upper limit</td>
</tr>
<tr>
<td>Marticles</td>
<td>9</td>
<td>9</td>
<td>20</td>
<td>0</td>
<td>40</td>
<td>Total number of articles.</td>
</tr>
</tbody>
</table>

.mgen, at(articles=(0(5)40) female=1) stub(F)

Predictions from: margins, at(articles=(0(5)40) female=1) predict(pr)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Unique</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fpr1</td>
<td>9</td>
<td>9</td>
<td>.2742684</td>
<td>.0571391</td>
<td>.5971256</td>
<td>pr(y=Tenure) from margins</td>
</tr>
<tr>
<td>Fll1</td>
<td>9</td>
<td>9</td>
<td>.2206695</td>
<td>.044239</td>
<td>.4840426</td>
<td>95% lower limit</td>
</tr>
<tr>
<td>Full1</td>
<td>9</td>
<td>9</td>
<td>.3278673</td>
<td>.0700392</td>
<td>.7102085</td>
<td>95% upper limit</td>
</tr>
<tr>
<td>Farticles</td>
<td>9</td>
<td>9</td>
<td>20</td>
<td>0</td>
<td>40</td>
<td>Total number of articles.</td>
</tr>
</tbody>
</table>

.label var Fpr1 "Women"
.label var Mpr1 "Men"
.label var Marticles "Number of articles"
. twoway (connected w1pl w1x, msym(`sf') clcol(`cf') mcol(`cf')) ///
  (connected m1pl m1x, msym(`sm') clcol(`cm') mcol(`cm')), ///
  subtitle("Scientist at departments with average prestige") ///
  xlabel(0(5)40) ylabel(0(.2)1.0) xtitle("Number of articles") ///
  legend(pos(11) ring(0) cols(1)) ///
  ytitle("Pr(tenure)") caption(job'-01-noX.png, size(tiny))

. graph export `job'-01-noX.png, width(1200) replace
(file groups01-spost-trm-01-noX.png written in PNG format)
{Aside: Interaction in the BRM}

If we believed that men and women benefitted differently from publishing, we might also consider adding in an interaction effect, via a product term, to our model.

\[ y = \alpha + \beta_{\text{articles}} x_{\text{articles}} + \beta_{\text{female}} x_{\text{female}} + \beta_{\text{articles*female}} x_{\text{articles*female}} \]
\{Aside: Interaction in the BRM\}

But, of course, the BRM already includes an interaction between female & publications by way of its functional form.

\textbf{Review:}

Consider the model:
\[ \Pr(y = 1 \mid x, z) = \Phi(1 + 1x + .75z) \]

Fix \( z = -8 \):
\[ \Pr(y = 1 \mid x, z = -8) = \Phi(1 + 1x + [.75 \times -8]) = \Phi(-5 + 1x) \]

Increase \( z \) to \(-4\):
\[ \Pr(y = 1 \mid x, z = -4) = \Phi(1 + 1x + [.75 \times -4]) = \Phi(-2 + 1x) \]
{Aside: Interaction in the BRM}
Aside: Interaction in the BRM

Because of this necessary feature, scholars have distinguished between two types of interaction effects in the BRM:

- Interaction effects due to *compression*, or due solely to the functional form of the BRM
  - As we move away from the linear portion of the curve towards 0 or 1, groups must converge → thus differing slopes/second derivatives for groups

- Interactions modeled via *product terms*, as traditionally done in the LRM
Figure 2  Logit Models Illustrating How Marginal Effects on Pr(Y) Vary with the Values of Independent Variables

Panel A depicts the model Pr(Y) = G(-4 + X_1 + X_2). Panel B depicts the model Pr(Y) = G(-4 + X_1 + X_2 - 0.30X_1X_2). In both cases, G() is the logit link function. Arrows indicate the marginal effect (m.e.) of the curve, i.e., ∂Pr(Y)/∂X_1, at the indicated point.
Question (for you)

Nagler (1991, 1994) argues that, in fact, interaction effects due to compression are solely a methodological artifact and thus are not substantively meaningful.

Do you agree?


{Aside: Interaction in the BRM}

So, unlike in the LRM, a significant product term is neither necessary nor sufficient for determining whether certain coefficients differ by group.

**Necessary:**

Model may be interactive without a product term.

**Sufficient:**

- A significant product term may actually make your model more linear/less interactive than the model WITHOUT a product term.
- A non-significant product term may fail to represent significant second-differences at some points of the data space.
Figure 2  Logit Models Illustrating How Marginal Effects on \( \text{Pr}(Y) \) Vary with the Values of Independent Variables

Panel A depicts the model \( \text{Pr}(Y) = G(-4 + X_1 + X_2) \). Panel B depicts the model \( \text{Pr}(Y) = G(-4 + X_1 + X_2 - 0.30X_1X_2) \). In both cases, \( G() \) is the logit link function. Arrows indicate the marginal effect (m.e.) of the curve, i.e., \( \partial \text{Pr}(Y)/\partial X_1 \), at the indicated point.
{Aside: Interactions in the BRM}

As a further complication:

The sign on the coefficient may not always be representative of the change that is occurring.
B: Illustrative Logit Model
With Product Term

\[ \hat{\beta}_{X_1X_2} = -0.30 \]
So how do we decide when to include a product term?

Appeal to theory:

Berry et al. (2010): “We believe that this decision must be based on an explicit theory about the effects on the unbounded latent variable y* [...] This is because even though the researcher’s hypothesis is about the effects of variables on Pr(Y) [...] it is y* [...] that is specified as a linear function of the independent variables.”

- Remember: compression interactions are still limited by the functional form of the logit/probit

In the BRM, all interactions are not theorized equal[ly].
Verify via predicted probabilities:

Full extent of interactivity only revealed by examining differences in predicted probabilities.
**{Aside: Interactions in the BRM}**

**A simple example**

```stata
. logit tenure i.female##c.articles prestige , nolog
```

Logistic regression

|                | Coef.     | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|----------------|-----------|-----------|------|-----|----------------------|
| female         |           |           |      |     |                      |
| Female         | .1785884  | .1789557  | 1.00 | 0.318| -1.721584  to  0.5293352 |
| articles       | .1038231  | .0100759  | 10.30| 0.000| .0840748  to  0.1235714 |
| female#c.articles | -0.0496277 | .0141527  | -3.51| 0.000| -0.0773666  to -0.0218889 |
| prestige       | -0.3610535 | .0792061  | -4.56| 0.000| -0.5162946  to  -0.2058125 |
| _cons          | -1.809395  | .2203708  | -8.21| 0.000| -2.241314  to  -1.377476 |

```. est sto simpleX
```
```
. esttab simple simpleX, mti bic aic

--------------------------------------------
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>simple</td>
<td>simpleX</td>
</tr>
<tr>
<td>tenure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.female</td>
<td>-0.286*</td>
<td>0.179</td>
</tr>
<tr>
<td></td>
<td>(-2.33)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>articles</td>
<td>0.0811***</td>
<td>0.104***</td>
</tr>
<tr>
<td></td>
<td>(10.83)</td>
<td>(10.30)</td>
</tr>
<tr>
<td>prestige</td>
<td>-0.390***</td>
<td>-0.361***</td>
</tr>
<tr>
<td></td>
<td>(-4.93)</td>
<td>(-4.56)</td>
</tr>
<tr>
<td>0.female#c~s</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(. )</td>
<td>(. )</td>
</tr>
<tr>
<td>1.female#c~s</td>
<td>0</td>
<td>-0.0496***</td>
</tr>
<tr>
<td></td>
<td>(-3.51)</td>
<td>(-3.51)</td>
</tr>
<tr>
<td>_cons</td>
<td>-1.531***</td>
<td>-1.809***</td>
</tr>
<tr>
<td></td>
<td>(-7.53)</td>
<td>(-8.21)</td>
</tr>
</tbody>
</table>
|--------------------------------------------
| N    | 2945 | 2945 |
| AIC  | 2075.1 | 2065.0 |
| BIC  | 2099.1 | 2094.9 |
|--------------------------------------------
t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001
```
{Aside: Interactions in the BRM}

Interpreting odds ratios [?] in the BLM

No straightforward way to interpret unexponentiated interaction coefficients

Some (Williams, 2011) argue that there is no straightforward way to interpret odds ratios for interaction effects

But we can make some conceptual sense of them.
. listcoef, help

logit (N=2945): Factor change in odds

Odds of: Tenure vs NoTenure

|              | b   | z     | P>|z| | e^b | e^bStdX | SDofX |
|--------------|-----|-------|-----|-----|---------|-------|
| female       |     |       |     |     |         |       |
| Female       | 0.1786 | 0.998 | 0.318 | 1.196 | 1.091   | 0.486 |
| articles     | 0.1038 | 10.304 | 0.000 | **1.109** | 2.014   | 6.745 |
| female#c.c~s |     |       |     |     |         |       |
| Female       | -0.0496 | -3.507 | 0.000 | **0.952** | 0.744   | 5.959 |
| prestige     | -0.3611 | -4.558 | 0.000 | 0.697 | 0.757   | 0.771 |
| constant     | -1.8094 | -8.211 | 0.000 | .     | .       | .     |

b = raw coefficient  
z = z-score for test of b=0  
P>|z| = p-value for z-test  
e^b = exp(b) = factor change in odds for unit increase in X  
e^bStdX = exp(b*SD of X) = change in odds for SD increase in X  
SDofX = standard deviation of X
Moving from probabilities to odds may help

. **Articles=1
. margins, at(female=(0 1) art=1) atmeans post

Adjusted predictions

Model VCE : OIM

Expression : Pr(tenure), predict()

1._at : female = 0
       articles = 1
       prestige = 2.633975 (mean)

2._at : female = 1
       articles = 1
       prestige = 2.633975 (mean)

<table>
<thead>
<tr>
<th>Delta-method</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Margin</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
</tr>
<tr>
<td>_at</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.0655847</td>
<td>.0068408</td>
<td>9.59</td>
<td>0.000</td>
<td>.0521771</td>
</tr>
<tr>
<td>2</td>
<td>.0739447</td>
<td>.0089751</td>
<td>8.24</td>
<td>0.000</td>
<td>.0563538</td>
</tr>
</tbody>
</table>
. **Compute odds ratio
. di (0.0739/(1-0.0739))/(0.0656/(1-0.0656))
1.1366207

. di exp(0.179)*exp(-0.0496)
1.1381453
. **Articles=7
. margins, at(female=(0 1) art=7) atmeans post

Adjusted predictions                                      Number of obs = 2945
Model VCE : OIM

Expression : Pr(tenure), predict()

1._at        : female          = 0
               articles        = 7
               prestige        = 2.633975 (mean)

2._at        : female          = 1
               articles        = 7
               prestige        = 2.633975 (mean)

|          | Delta-method
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Margin</td>
<td>Std. Err.</td>
<td>z</td>
<td>P&gt;</td>
<td>z</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
<td>--------</td>
<td>--------</td>
<td>------------------</td>
<td>--------</td>
</tr>
<tr>
<td>_at</td>
<td>1</td>
<td>.115716</td>
<td>.0078353</td>
<td>14.77</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>.0995312</td>
<td>.009047</td>
<td>11.00</td>
<td>0.000</td>
</tr>
</tbody>
</table>

. **Compute odds ratio
. di (0.1000/(1-0.1000))/(0.1157/(1-0.1157))
. 84922693

. di exp(0.179)*exp(7*(-0.0496))
. 84518478
\{Aside: Interactions in the BRM\}

Plotting predicted probabilities

\[. \text{mgen}, \text{at(female}=0 \text{ articles}=(0(5)40)) \text{ stub(Mx)}\]

Predictions from: \text{margins, at(female}=0 \text{ articles}=(0(5)40)) \text{ predict(pr)}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Unique</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mxpr1</td>
<td>9</td>
<td>9</td>
<td>.3814966</td>
<td>.0613916</td>
<td>.7973321</td>
<td>pr(y=Tenure) from margins</td>
</tr>
<tr>
<td>Mxll1</td>
<td>9</td>
<td>9</td>
<td>.3230883</td>
<td>.0481615</td>
<td>.6986448</td>
<td>95% lower limit</td>
</tr>
<tr>
<td>Mxul1</td>
<td>9</td>
<td>9</td>
<td>.4399049</td>
<td>.0746218</td>
<td>.8960194</td>
<td>95% upper limit</td>
</tr>
<tr>
<td>Mxarticles</td>
<td>9</td>
<td>9</td>
<td>20</td>
<td>0</td>
<td>40</td>
<td>Total number of articles.</td>
</tr>
</tbody>
</table>

\[. \text{mgen}, \text{at(female}=1 \text{ articles}=(0(5)40)) \text{ stub(Fx)}\]

Predictions from: \text{margins, at(female}=1 \text{ articles}=(0(5)40)) \text{ predict(pr)}

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Unique</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fxpr1</td>
<td>9</td>
<td>9</td>
<td>.2075626</td>
<td>.0724697</td>
<td>.3998722</td>
<td>pr(y=Tenure) from margins</td>
</tr>
<tr>
<td>Fxll1</td>
<td>9</td>
<td>9</td>
<td>.145391</td>
<td>.0546684</td>
<td>.2467551</td>
<td>95% lower limit</td>
</tr>
<tr>
<td>Fxul1</td>
<td>9</td>
<td>9</td>
<td>.2697342</td>
<td>.0902709</td>
<td>.5529892</td>
<td>95% upper limit</td>
</tr>
<tr>
<td>Fxarticles</td>
<td>9</td>
<td>9</td>
<td>20</td>
<td>0</td>
<td>40</td>
<td>Total number of articles.</td>
</tr>
</tbody>
</table>
Without interaction:

Scientists at departments with average prestige

- Women
- Men

Pr(tenure) vs Number of articles
With interaction:

Scientist at departments with average prestige

- **Women**
- **Men**

Number of articles vs. \( Pr(\text{tenure}) \)
{</Aside>}
. logit tenure i.female##c.articles i.female##c.prestige, nolog

Logistic regression

|          | Coef.   | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|----------|---------|-----------|------|-------|---------------------|
| tenure   |         |           |      |       |                     |
| female   |         |           |      |       |                     |
| Female   | 0.4611594 | 0.4315736 | 1.07 | 0.285 | -0.3847094 - 1.307028 |
| articles | 0.1031505 | 0.0100957 | 10.22| 0.000 | 0.0833633 - 0.1229377 |
| female#c.articles | | | | | |
| Female   | -0.0458669 | 0.0151276 | -3.03| 0.002 | -0.0755165 - 0.0162173 |
| prestige | -0.3239936 | 0.0942152 | -3.44| 0.001 | -0.5086521 - 0.1393352 |
| female#c.prestige | | | | | |
| Female   | -0.1243385 | 0.1735868 | -0.72| 0.474 | -0.4645624 - 0.2158854 |
| _cons    | -1.897    | 0.252495  | -7.51| 0.000 | -2.391881 - 1.402119 |

Number of obs = 2945
LR chi2(5) = 144.71
Prob > chi2 = 0.0000
Log likelihood = -1027.2213
Pseudo R2 = 0.0658

---

. eststo simple

2X
. margins, at(female=(0 1)) atmeans post

Adjusted predictions

Model VCE : OIM

Expression : Pr(tenure), predict()

1._at : female = 0
articles = 7.209168 (mean)
prestige = 2.633975 (mean)

2._at : female = 1
articles = 7.209168 (mean)
prestige = 2.633975 (mean)

|            Delta-method                      |
|            | Margin | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|-------------|--------|-----------|------|------|---------------------|
| _at         |        |           |      |      |                     |
| 1           | .1184948 | .0079391  | 14.93 | 0.000 | .1029344 - .1340553 |
| 2           | .0994126 | .009196   | 10.81 | 0.000 | .0813887 - .1174365 |

. mlincom(2-1)

<table>
<thead>
<tr>
<th></th>
<th>lincom</th>
<th>pvalue</th>
<th>ll</th>
<th>ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.019</td>
<td>0.116</td>
<td>-0.043</td>
<td>0.005</td>
</tr>
</tbody>
</table>
. **females
. logit tenure articles prestige if female==1, nolog

Logistic regression

|                         | Coef.  | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------------------------|--------|-----------|-------|-------|----------------------|
| articles                | .0572837 | .011266   | 5.08  | 0.000 | .0352028 .0793646    |
| prestige                | -.4483321 | .1457939  | -3.08 | 0.002 | -.734083 -.1625812  |
| _cons                   | -1.435841 | .350003   | -4.10 | 0.000 | -2.121834 -.7498473 |

. eststo bygroupF
. margins, at(articles=7.209168 prestige=2.633975)

Expression : Pr(tenure), predict()
at : articles = 7.209168
prestige = 2.633975

| | Delta-method
| Margin | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------------------------------|---------|---------|-------|-------|----------------------|
| _cons                        | .0994126 | .009196 | 10.81 | 0.000 | .0813887 .1174365   |
**males**

.logit tenure articles prestige if female==0, nolog

Logistic regression

Number of obs = 1824
LR chi2(2) = 114.10
Prob > chi2 = 0.0000

Log likelihood = -656.93699
Pseudo R2 = 0.0799

| tenure | Coef.  | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|--------|--------|-----------|-----|-----|---------------------|
| articles | 0.1031505 | 0.0100957 | 10.22 | 0.000 | 0.0833633 - 0.1229377 |
| prestige  | -0.3239937 | 0.0942152 | -3.44 | 0.001 | -0.5086521 - -0.1393352 |
| _cons    | -1.897 | 0.252495 | -7.51 | 0.000 | -2.391881 - -1.402119 |

.eststo bygroupM

.margins, at(articles=7.209168 prestige=2.633975)

Expression : Pr(tenure), predict()
at : articles = 7.209168
prestige = 2.633975

| Delta-method |
|-------------|-----------|-----|-----|---------------------|
| Margin | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
| _cons  | 0.1184948 | 0.0079391 | 14.93 | 0.000 | 0.1029344 - 0.1340552 |

Group Differences \ 67
. logit tenure **ibn.female** female#c.articles female#c.prestige, **nocon nolog**

Logistic regression                               Number of obs   =       2945
Wald chi2(6)    =    1168.63
Log likelihood = -1027.2213                       Prob > chi2     =     0.0000

------------------------------------------------------------------------------
tenure |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------------
female  |                                                 
   Male  |   -1.897    .252495   -7.51  0.000   -2.391881   -1.402119   
 Female |   -1.435841    .350003   -4.10  0.000   -2.121834   -.7498473

female#c. |                                                 
c.articles |                                                 
   Male  |   .1031505   .0100957   10.22  0.000    .0833633    .1229377
 Female |   .0572837   .0112666    5.08  0.000    .0352028    .0793646

c. |                                                 
c.prestige |                                                 
   Male  |   -.3239937   .0942152   -3.44  0.001   -.5086521   -.1393352
 Female |   -.4483321   .1457939   -3.08  0.002   -.734083    -.1625812

------------------------------------------------------------------------------

. est sto fullX


```
.margins, at(female=(0 1)) atmeans

Adjusted predictions  Number of obs   =       2945
Model VCE    : OIM
Expression   : Pr(tenure), predict()

1._at        : female          =           0
               articles        =    7.209168 (mean)
               prestige        =    2.633975 (mean)

2._at        : female          =           1
               articles        =    7.209168 (mean)
               prestige        =    2.633975 (mean)

+----------------------------------------------------------------
|            Delta-method
|     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
|----------------------------------------------------------------
|            _at  |
|     1       |   .1184948   .0079391   14.93   0.000     .1029344    .1340552
|     2       |   .0994126   .0091960   10.81   0.000     .0813887    .1174365
+----------------------------------------------------------------
```

.
. esttab simpleX2 bygroupF bygroupM fullX //
> , mtitles(simpleX2 bygroupF bygroupM fullX) ///</
> nostar nogaps nonumbers nodepvars

<table>
<thead>
<tr>
<th></th>
<th>simpleX2</th>
<th>bygroupF</th>
<th>bygroupM</th>
<th>fullX</th>
</tr>
</thead>
<tbody>
<tr>
<td>tenure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.female</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.female</td>
<td>0.425</td>
<td>0.0637</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(5.24)</td>
<td>(10.12)</td>
<td></td>
</tr>
<tr>
<td>articles</td>
<td>0.110</td>
<td>0.0637</td>
<td>0.110</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.12)</td>
<td>(5.24)</td>
<td>(10.12)</td>
<td></td>
</tr>
<tr>
<td>0.female#c~s</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.female#c~s</td>
<td>-0.0459</td>
<td></td>
<td>0.0637</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-2.82)</td>
<td></td>
<td>(5.24)</td>
<td></td>
</tr>
<tr>
<td>prestige</td>
<td>-0.360</td>
<td>-0.466</td>
<td>-0.360</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-3.72)</td>
<td>(-3.13)</td>
<td>(-3.72)</td>
<td></td>
</tr>
<tr>
<td>0.female#c~e</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.female#c~e</td>
<td>-0.107</td>
<td></td>
<td>-0.466</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.60)</td>
<td></td>
<td>(-3.13)</td>
<td></td>
</tr>
<tr>
<td>_cons</td>
<td>-1.847</td>
<td>-1.421</td>
<td>-1.847</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-7.20)</td>
<td>(-3.97)</td>
<td>(-7.20)</td>
<td></td>
</tr>
</tbody>
</table>

---

N       2797   1056   1741   2797

t statistics in parentheses
Plotted at prestige = 4

Pr(tenure) vs. Number of articles for Men and Women.
Plotted at prestige = 2

- 95% confidence interval
- Male-Female difference
Plotted at prestige = 1

- The graph shows the difference in the probability of men and women as a function of the number of articles.
- The x-axis represents the number of articles, ranging from 0 to 50.
- The y-axis represents the difference in probability, ranging from -1 to 1.
- The graph includes a solid red line and a dashed purple line, indicating the trend for men and women, respectively.
- At prestige = 1, the probability difference increases as the number of articles increases.

Group Differences
Plotted at prestige = 2
Plotted at prestige = 3
Plotted at prestige = 5

Number of articles

Pr(men) - Pr(women)
Prestige + publications: In three dimensions
Resources


End Group Differences