Interpreting and Presenting Regression Results

Frederick J. Boehmke

Department of Political Science
University of Iowa

Prepared for presentation at the University of Kentucky.

February 29, 2008

Introduction

- Setup of interaction hypotheses.

Introduction

- Setup of interaction hypotheses.
- Discussion of common claims.

Introduction

- Setup of interaction hypotheses.
- Discussion of common claims.
- Implementing tests.

The Model

$$Y_i = \alpha + \beta_x X_i + \beta_z Z_i + \beta_{xz} X_i Z_i + \epsilon_i.$$

Marginal Effects:

$$\frac{\partial Y_i}{\partial X_i} = \beta_x + \beta_{xz} Z_i;$$

$$\frac{\partial Y_i}{\partial Z_i} = \beta_z + \beta_{xz} X_i.$$

The Hypothesis

Define $\gamma = \beta_x + \beta_{xz} Z_i$.

The Hypothesis

Define $\gamma = \beta_x + \beta_{xz} Z_i$.

$$H_0 : \gamma = 0;$$
$$H_A : \gamma \neq 0.$$

The Test

$$\begin{aligned}\hat{\gamma} &= \hat{\beta}_x + \hat{\beta}_{xz} Z_i, \\ Var[\hat{\gamma}] &= Var(\hat{\beta}_x + \hat{\beta}_{xz} Z_i), \\ &= Var(\hat{\beta}_x) + Z_i^2 Var(\hat{\beta}_{xz}) + 2 Z_i Cov(\hat{\beta}_x, \hat{\beta}_{xz}).\end{aligned}$$

$$\frac{\hat{\gamma}}{\sqrt{Var[\hat{\gamma}]}} \sim t_{n-4}.$$

Frequently Overhead

1. I don't need to include X_i.

Frequently Overhead

1. I don't need to include X_i.
2. I can interpret $\hat{\beta}_x$ directly.

Frequently Overhead

1. I don't need to include X_i.
2. I can interpret $\hat{\beta}_x$ directly.
3. The coefficients are not significant.

I Don't Need to Include X

1. Maybe you have no prediction about its direct effect.

I Don't Need to Include X

1. Maybe you have no prediction about its direct effect.
2. Maybe you have a theory that says its direct effect is zero.

I Don't Need to Include X

1. Maybe you have no prediction about its direct effect.
2. Maybe you have a theory that says its direct effect is zero.
3. Maybe you have a lot of correlation between X and $X \times Z$.

What Happens if you Don't Include X?

$$
\begin{aligned}
\text{True Model:}\quad Y &= \alpha + \beta_x X_i + \beta_z Z_i + \beta_{xz} X_i Z_i + \epsilon_i. \\
\text{Estimate:}\quad Y &= \alpha' + \beta'_z Z_i + \beta'_{xz} X_i Z_i + \epsilon'_i. \\
\text{Implies:}\quad Y &= \alpha' + \beta'_z Z_i + \beta'_{xz} X_i Z_i + (\epsilon_i + \beta_x X_i).
\end{aligned}
$$

What Happens if you Don't Include X?

$$
\begin{aligned}
\text{True Model: } Y &= \alpha + \beta_x X_i + \beta_z Z_i + \beta_{xz} X_i Z_i + \epsilon_i. \\
\text{Estimate: } Y &= \alpha' + \beta'_z Z_i + \beta'_{xz} X_i Z_i + \epsilon'_i. \\
\text{Implies: } Y &= \alpha' + \beta'_z Z_i + \beta'_{xz} X_i Z_i + (\epsilon_i + \beta_x X_i).
\end{aligned}
$$

So we have omitted variable bias!

Figure: Estimated Marginal Effect when X Included

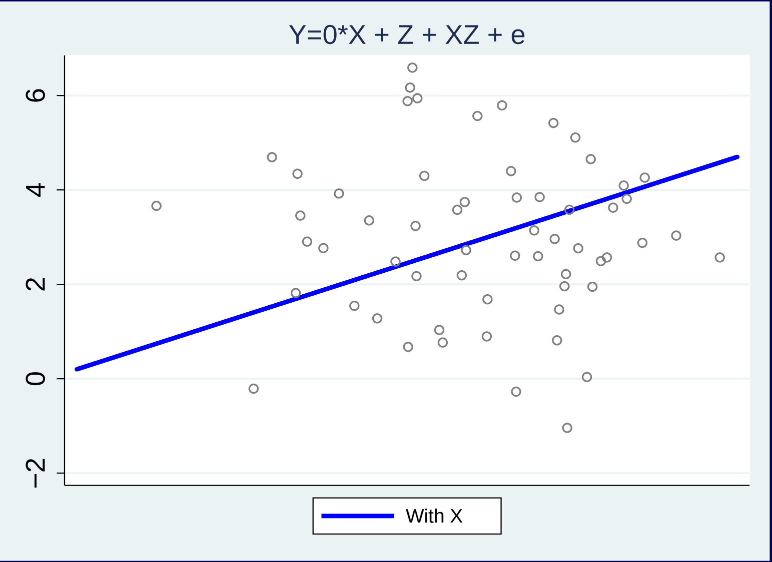


Figure: Marginal Effects from Models with and without X

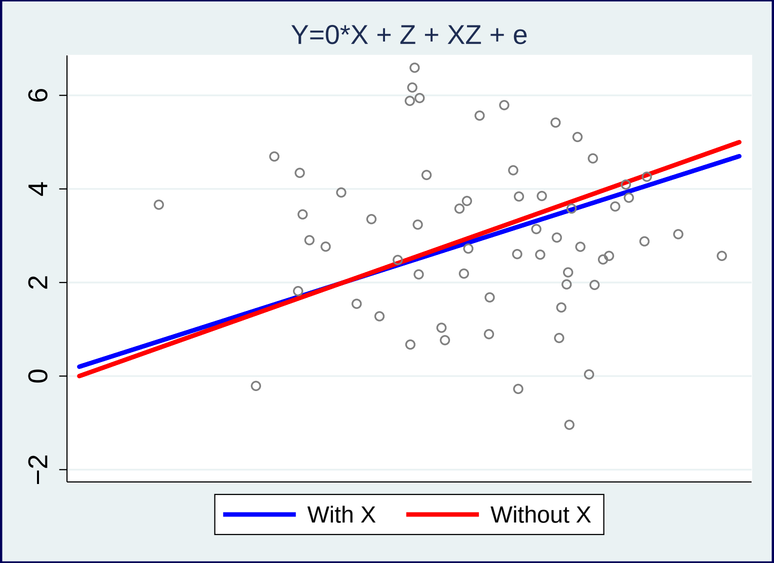


Figure: X Excluded & its Coefficient is not Zero

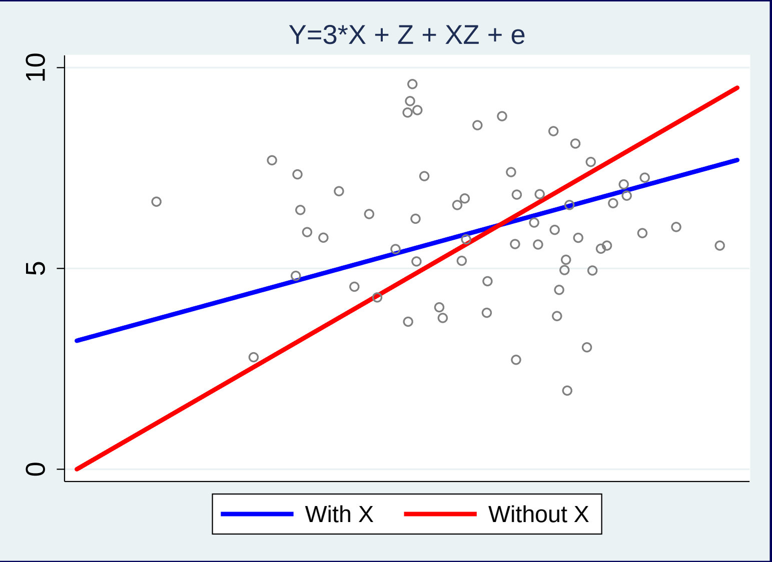


Figure: X Excluded by Value of β_x

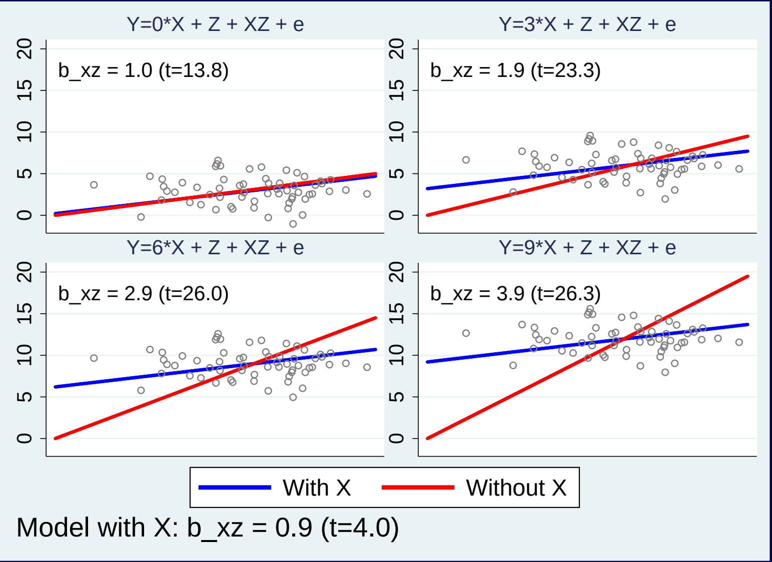
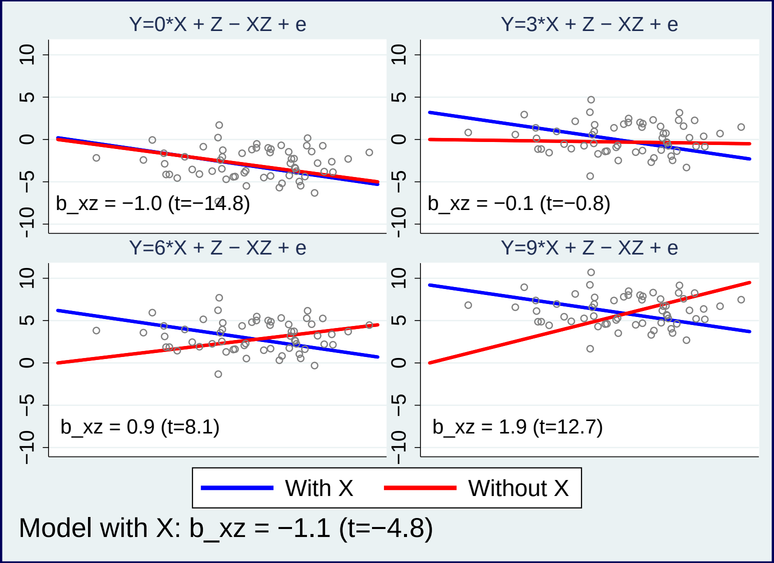


Figure: It Gets Worse!

Figure: It Gets Worse!



Why Interpretation of $\hat{\beta}_x$ is Tricky

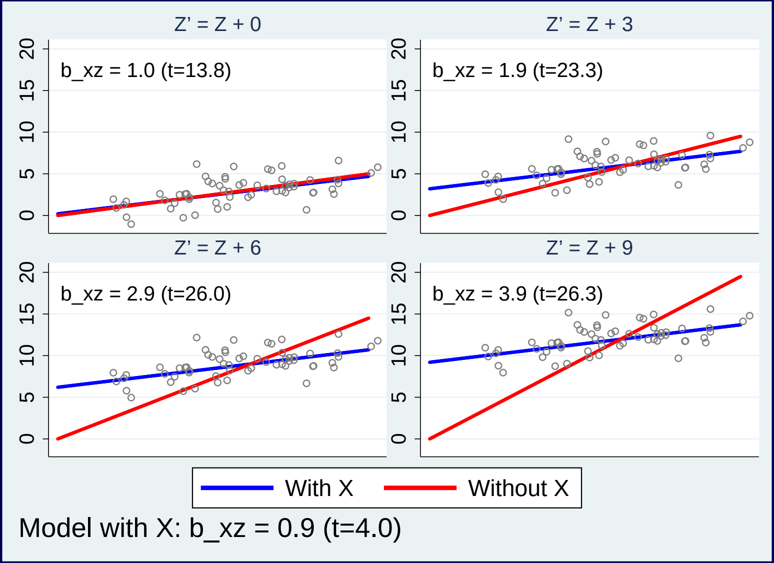
Let $Z_i' = Z_i + c$.

Why Interpretation of $\hat{\beta}_x$ is Tricky

$$\text{Let } Z_i' = Z_i + c.$$

$$\begin{aligned}
Y_i &= \alpha + \beta_x X_i + \beta_z Z_i' + \beta_{xz} X_i Z_i' + \epsilon_i, \\
&= \alpha + \beta_x X_i + \beta_z (Z_i + c) + \beta_{xz} X_i (Z_i + c) + \epsilon_i, \\
&= \alpha + \beta_x X_i + \beta_z c + \beta_z Z_i + \beta_{xz} X_i Z_i + \beta_{xz} X_i c + \epsilon_i, \\
&= (\alpha + \beta_z c) + (\beta_x + \beta_{xz} c) X_i + \beta_z Z_i + \beta_{xz} X_i Z_i + \epsilon_i, \\
&= \alpha' + \beta_x' X_i + \beta_z Z_i + \beta_{xz} X_i Z_i + \epsilon_i.
\end{aligned}$$

Figure: Adding a Constant to Z Affects Marginal Effect of X



Just Look at the Significance of the Coefficients

- Coefficients possess only limited information.

Just Look at the Significance of the Coefficients

- Coefficients possess only limited information.
- As we've seen, coefficients on constitutive terms are meaningless.

Just Look at the Significance of the Coefficients

- Coefficients possess only limited information.
- As we've seen, coefficients on constitutive terms are meaningless.
- Need to test whether marginal effect is significant at different values of Z.

Table: Number of Citizen Interest Groups per State, 1990

Initiative State	88.50 ∗∗	85.24
	(41.72)	(93.45)
Total Population	17.53 ∗∗	17.56 ∗∗
	(3.92)	(4.04)
Citizen Ideology	1.89	2.02
	(2.88)	(4.34)
Initiative × Ideology		−0.23
		(5.83)
Constant	80.54	82.15
	(56.38)	(70.23)

Table: Assessing Significance of Marginal Effect

Ideology	$\hat{\gamma}$	SE($\hat{\gamma}$)	t	p
-30	92.08	100.66	0.91	0.37
-25	90.94	75.21	1.21	0.23
-20	89.80	53.65	1.67	0.10
-15	88.66	42.39	2.09	0.04
-10	87.52	49.11	1.78	0.08
-5	86.38	68.73	1.26	0.21
0	85.24	93.45	0.91	0.37

Implementing Interaction Tests in Stata

1. Using Stata's `test` command.

Implementing Interaction Tests in Stata

1. Using Stata's `test` command.
2. Using CLARIFY suite.

Implementing Interaction Tests in Stata

1. Using Stata's `test` command.
2. Using CLARIFY suite.
3. Using `grinter`.

Implementing Interaction Tests in Stata

1. Using Stata's `test` command.
2. Using CLARIFY suite.
3. Using `grinter`.
4. Using formulas.

Implementing Interaction Tests in Stata

1. Using Stata's `test` command.
2. Using CLARIFY suite.
3. Using `grinter`.
4. Using formulas.
5. Using simulations.

Using Stata's test Command

```
use boehmke2008-02-29interactions.dta
regress y x z zx
test _b[xz] = 0
test _b[x] + 3*_b[xz]=0
```

Using Stata's test Command

```
use boehmke2008-02-29interactions.dta
regress y x z zx
test _b[xz] = 0
test _b[x] + 3*_b[xz]=0
```

Use loops to automate for many values:

```
forvalues val=1(1)5 {
  test _b[x] + `val'*_b[xz]=0
  }
```

A Slightly More Flexible Version of test

```
summarize x
forvalues val = `r(min)'/`r(max)' {
  local effect = _b[x] + `val'*_b[xz]
  quietly test x + `val'*xz=0
  display `x' , `effect' , r(F) , r(p)
  }
```

Saving Those Values I

```
generate test_val = .
generate test_eff = .
generate test_F = .
generate test_p = .
```

Saving Those Values II

```
summarize x
local i = 1
forvalues val = `r(min)'/`r(max)' {
  local effect = _b[x] + `val'*_b[xz]
  quietly test x + `val'*xz=0
```

Saving Those Values II

```
summarize x
local i = 1
forvalues val = `r(min)'/`r(max)' {
  local effect = _b[x] + `val'*_b[xz]
  quietly test x + `val'*xz=0

  replace test_val = `x' if _n == `i'
  replace test_eff = `effect' if _n == `i'
  replace test_F = `r(F)' if _n == `i'
  replace test_p = `r(p)' if _n == `i'
  local `i' = `i' + 1
  }
```

Using the CLARIFY Suite of Commands

```
estisimp regress y x z xz
setx (x z xz) mean
simqi, fd(ev) changex(x 0 1 xz 0 1)
```

Using the CLARIFY Suite of Commands

```
estisimp regress y x z xz
setx (x z xz) mean
simqi, fd(ev) changex(x 0 1 xz 0 1)

summarize z
forvalues val = `r(min)'/`r(max)' {
   simqi, fd(ev) changex(x 0 1 xz 0 `val')
  }
```

Using `grinter` to Graph Marginal Effect

- `grinter` automates graphing marginal effect for simple interaction.

Using `grinter` to Graph Marginal Effect

- `grinter` automates graphing marginal effect for simple interaction.
- Graphs $\partial Y / \partial X = \hat{\beta}_x + \hat{\beta}_{xz} Z$ against values of Z.

Using `grinter` to Graph Marginal Effect

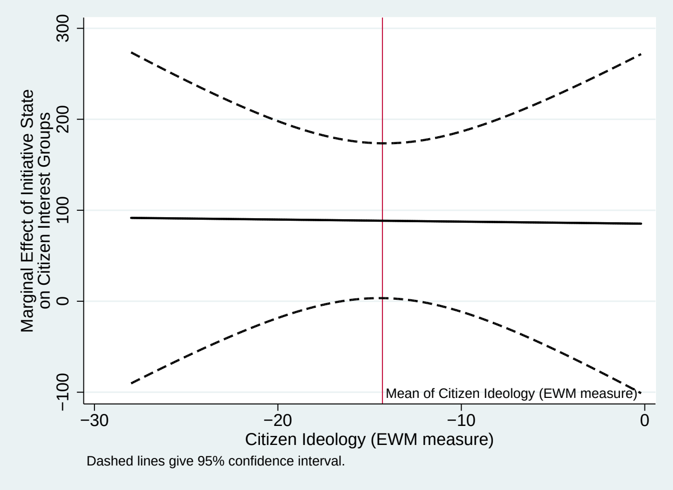
- `grinter` automates graphing marginal effect for simple interaction.
- Graphs $\partial Y / \partial X = \hat{\beta}_x + \hat{\beta}_{xz} Z$ against values of Z.
- Adds confidence interval to asses whether it includes zero.

Using `grinter` to Graph Marginal Effect

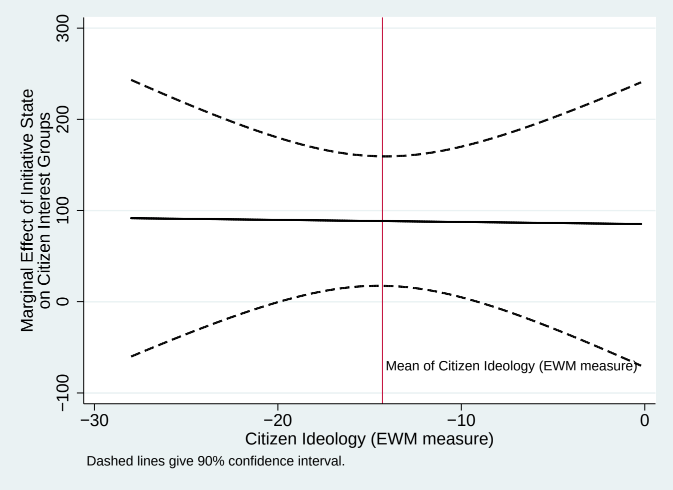
- `grinter` automates graphing marginal effect for simple interaction.
- Graphs $\partial Y / \partial X = \hat{\beta}_x + \hat{\beta}_{xz} Z$ against values of Z.
- Adds confidence interval to asses whether it includes zero.
- Basic syntax:

```
grinter x, inter(xz) const02(z).
```

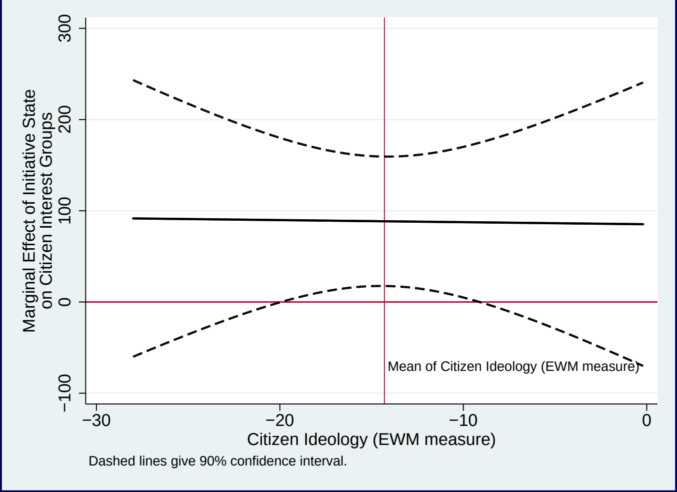
```
grinter init, inter(initideo) const02(ideology)
```



```
grinter init, inter(initideo) const02(ideology)
clevel(90)
```



```
grinter init, inter(initideo) const02(ideology)
clevel(90) yline(0)
```



```
grinter init, inter(initideo) const02(ideology)
clevel(90) yline(0) kdensity
```

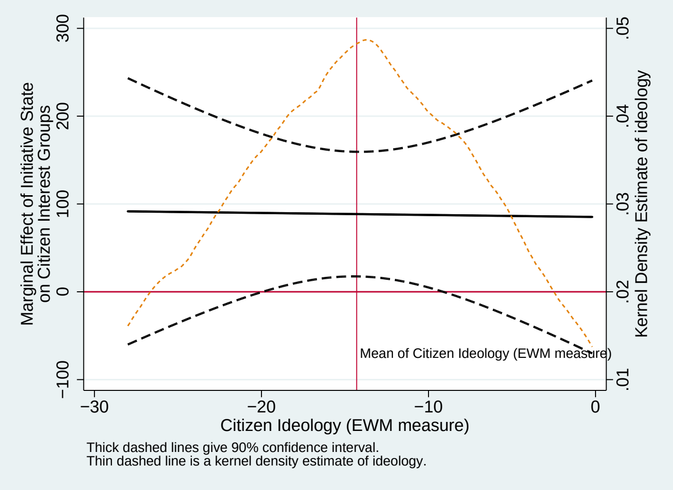
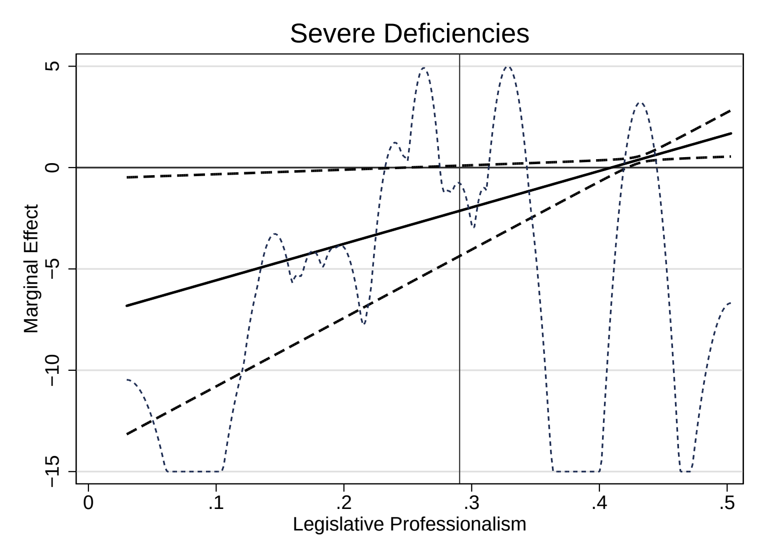


Figure: A More Complicated Example



Marginal Effects in Non-Linear Models

- More difficult than in OLS since marginal effect depends on all covariates.

Marginal Effects in Non-Linear Models

- More difficult than in OLS since marginal effect depends on all covariates.
- But the basic principle remains the same: determine

$$\partial Y / \partial X.$$

Marginal Effects in Non-Linear Models

- More difficult than in OLS since marginal effect depends on all covariates.
- But the basic principle remains the same: determine

$$\partial Y / \partial X.$$

- Getting confidence intervals is more difficult, but simulation helps.

Logit

$$
\begin{aligned}
\Pr(Y_i = 1|X) &= \Pr(X_i\beta > 0|X), \\
&= 1 - \Pr(-X_i\beta < 0|X), \\
&= 1 - F(-X_i\beta), \\
&= 1 - \frac{\exp(-X_i\beta)}{1 + \exp(-X_i\beta)}, \\
&= \frac{1}{1 + \exp(-X_i\beta)}, \\
&= (1 + \exp(-X_i\beta))^{-1}.
\end{aligned}
$$

Marginal Effect in Logit I

$$X_i\beta = \alpha + \beta_x X_i + \beta_z Z_i + \beta_{xz} X_i Z_i.$$

Marginal Effect in Logit I

$$X_i\beta = \alpha + \beta_x X_i + \beta_z Z_i + \beta_{xz} X_i Z_i.$$

$$\begin{aligned}
\frac{\partial \Pr(Y_i = 1|W_i)}{\partial X} &= \frac{\partial(1+\exp(-X_i\beta))^{-1}}{\partial X}, \\
&= -\frac{\partial(1+\exp(-X_i\beta))}{\partial X}(1+\exp(-X_i\beta))^{-2}, \\
&= -\frac{\partial(-X_i\beta)}{\partial X}\exp(-X_i\beta)(1+\exp(-X_i\beta))^{-2}, \\
&= (\beta_x + \beta_{xz} Z_i)\exp(-X_i\beta)(1+\exp(-X_i\beta))^{-2};
\end{aligned}$$

Marginal Effect in Logit II

$$\begin{aligned}
&= (\beta_x + \beta_{xz} Z_i) \exp(-X_i\beta)(1 + \exp(-X_i\beta))^{-2}, \\
&= (\beta_x + \beta_{xz} Z_i) \left(\frac{\exp(-X_i\beta)}{1 + \exp(-X_i\beta)} \right) (1 + \exp(-X_i\beta))^{-1}, \\
&= (\beta_x + \beta_{xz} Z_i) \left(\frac{\exp(-X_i\beta)}{1 + \exp(-X_i\beta)} \right) \Pr(Y_i = 1|X_i), \\
&= (\beta_x + \beta_{xz} Z_i) \left(1 - \frac{1}{1 + \exp(-X_i\beta)} \right) \Pr(Y_i = 1|X_i), \\
&= (\beta_x + \beta_{xz} Z_i)(1 - \Pr(Y_i = 1|X_i)) \Pr(Y_i = 1|X_i), \\
&= (\beta_x + \beta_{xz} Z_i) \Pr(Y_i = 0|X_i) \Pr(Y_i = 1|X_i).
\end{aligned}$$

Estimating Marginal Effect in Logit

- Trying to estimate:

$$(\beta_x + \beta_{xz} Z_i) \Pr(Y_i = 0 | X_i) \Pr(Y_i = 1 | X_i).$$

- Use:

$$(\hat{\beta}_x + \hat{\beta}_{xz} Z_i) \widehat{\Pr(Y_i = 0} | X_i) \widehat{\Pr(Y_i = 1} | X_i).$$

- Generate confidence interval by sampling J times from distribution of estimated coefficients:

$$\hat{\beta}^j \sim N(\hat{\beta}, Var(\hat{\beta})).$$

Logit Marginal Effects in Stata

```
use boehmke2008-02-29interactions-logit.dta
logit y x z xz
generat x_val = 1
generat z_val = 1
drawnorm beta_x beta_z beta_xz alpha,
  means(e(b)) cov(e(V))
```

Logit Marginal Effects in Stata

```
use boehmke2008-02-29interactions-logit.dta
logit y x z xz
generat x_val = 1
generat z_val = 1
drawnorm beta_x beta_z beta_xz alpha,
  means(e(b)) cov(e(V))

generat xb_hat = alpha + beta_x*x_val + beta_z*z_val
   + beta_xz*x_val*z_val
generat gamma_hat = beta_x + beta_xz*z_val
generat pi1_hat = 1/(1+exp(-xb_hat))
```

Logit Marginal Effects in Stata

```
use boehmke2008-02-29interactions-logit.dta
logit y x z xz
generat x_val = 1
generat z_val = 1
drawnorm beta_x beta_z beta_xz alpha,
  means(e(b)) cov(e(V))

generat xb_hat = alpha + beta_x*x_val + beta_z*z_val
   + beta_xz*x_val*z_val
generat gamma_hat = beta_x + beta_xz*z_val
generat pi1_hat = 1/(1+exp(-xb_hat))

generat marginal = gamma_hat*(1 - pi1_hat)*(pi1_hat)
summarize marginal, detail
```

Graphing Logit Marginal Effects in Stata I

```
logit y x z xz

collapse (mean) x (min) z_min=z (max) z_max=z
expand 1000
```

Graphing Logit Marginal Effects in Stata I

```
logit y x z xz

collapse (mean) x (min) z_min=z (max) z_max=z
expand 1000

generat z = z_min + (z_max - z_min)*(_n-1)/_N
generat xz = x*z
expand 1000
```

Graphing Logit Marginal Effects in Stata II

```
drawnorm beta_x beta_z beta_xz alpha,
  means(e(b)) cov(e(V))

generat xb_hat = alpha + beta_x*x_val + beta_z*z_val
   + beta_xz*x_val*z_val
generat gamma_hat = beta_x + beta_xz*z_val
generat pi1_hat = 1/(1+exp(-xb_hat))

generat marginal = gamma_hat*(1 - pi1_hat)*(pi1_hat)
```

Graphing Logit Marginal Effects in Stata III

```
collapse (mean) marginal (p5) marg_lb=marginal (p95)
  marg_ub=marginal, by(z)
```

Graphing Logit Marginal Effects in Stata III

```
collapse (mean) marginal (p5) marg_lb=marginal (p95)
  marg_ub=marginal, by(z)

twoway line marginal marg_lb marg_ub z, sort
   lpattern(solid dash dash) yline(0)
```

Graphing Logit Marginal Effects in Stata III

```
collapse (mean) marginal (p5) marg_lb=marginal (p95)
  marg_ub=marginal, by(z)

twoway line marginal marg_lb marg_ub z, sort
   lpattern(solid dash dash) yline(0)

twoway lowess marginal z, sort lcolor(black)
   || lowess marg_lb z, lpattern(dash) lcolor(gs6)
   || lowess marg_ub z, lpattern(dash) lcolor(gs6)
   yline(0)
   ytitle("Marginal Effect of X on P(Y=1|X,Z)")
   xtitle("Value of Z")
```

Conclusion

- Lots of bias can emerge if constitutive terms not included.

Conclusion

- Lots of bias can emerge if constitutive terms not included.
- Even if you have a theory, probably best to include them.

Conclusion

- Lots of bias can emerge if constitutive terms not included.
- Even if you have a theory, probably best to include them.
- Many ways to assess significance.

Conclusion

- Lots of bias can emerge if constitutive terms not included.
- Even if you have a theory, probably best to include them.
- Many ways to assess significance.
- Same principle allows calculation for any estimator or form of interactions.