PLSC 503: "Multivariate Analysis for Political Research"

Dichotomous Covariates
February 21, 2017

Introduction

Today we'll talk about dichotomous covariates, also known as "indicator" variables, "binary" variables, and/or "dummy" variables. This sounds like a simple topic, but in fact it's a good bit more complicated than you might think (hence, spending an entire class on it).

Why So Dumb?

Dummy variables exist for many reasons.

1. Some things are naturally dichotomous, like gender. Likewise, the presence or absence of something can be thought of as a dummy variable (e.g., majority status for coalitions, etc.).

2. It is also the case that *any* variable – at any level of measurement – can be dichotomized into one or more dummy variables.

   - A *polychotomous* (that is, *nominal-level*) variable with j categories can always be broken down into j separate dummy variables, one for each category. So, for a variable coded:

$$\texttt{partyid} = \begin{cases} 0 = \text{Labor} \\ 1 = \text{Liberal} \\ 2 = \text{Conservative} \end{cases}$$

     we can create:

$$\texttt{labor} = \begin{cases} 1 \text{ if Labor} \\ 0 \text{ otherwise} \end{cases}$$

$$\texttt{liberal} = \begin{cases} 1 \text{ if Liberal} \\ 0 \text{ otherwise} \end{cases}$$

$$\texttt{conserv} = \begin{cases} 1 \text{ if Conservative} \\ 0 \text{ otherwise} \end{cases}$$

- One can do a similar thing for categorical ordinal variables (that is, create j separate indicator variables, one for each category). So:

$$\texttt{agreement} = \begin{cases} -1 = \text{Disagree} \\ 0 = \text{Neutral} \\ 1 = \text{Agree} \end{cases}$$

becomes:

$$\texttt{disagree} = \begin{cases} 1 \text{ if ``Disagree''} \\ 0 \text{ otherwise} \end{cases}$$

$$\texttt{neutral} = \begin{cases} 1 \text{ if ``Neutral} \\ 0 \text{ otherwise} \end{cases}$$

$$\texttt{agree} = \begin{cases} 1 \text{ if ``Agree''} \\ 0 \text{ otherwise} \end{cases}$$

As we will see, at times there are advantages to doing this with ordinal covariates, in that it can extract additional information about the relationship between the various levels and the response variable Y.

- We can also always create one (or more) dummy variable(s) out of *continuous* data, simply by "splitting" the variable at one or more cutoff points. So, if we have

$$\texttt{thermometer} \in [0, 100]$$

we can create (e.g.):

$$\texttt{thermdummy} = \begin{cases} 0 \text{ if } \texttt{thermometer } \leq \tau \\ 1 \text{ if } \texttt{thermometer } > \tau \end{cases}$$

The important question then, of course, is how to define τ. Standard options include

  - Using $\tau = \text{median}(X)$,

  - $\tau = \bar{X}$, or
  - Choosing a value of τ on the basis of some substantively meaningful criterion (say, dichotomizing the "number of *amicus curiae* briefs" into "0" and "greater than zero").

  A caution about this practice, however: Dichotomizing continuous variables is nothing less than *throwing away data.* It should generally be undertaken only if there is some strong substantive reason to favor using a dichotomous version over a continuous one (e.g., if there are strong expectations of "threshold effects" on Y).

3. There are also a few other circumstances in which dummy variables occur "naturally;" these include (but are not limited to):

   - *Structural breaks.*
     - In time-series data, we often use dummy variables to denote specific time periods.
     - So, for example, if we had annual data on U.S. use of force (Y), and we believe that American attitudes toward that use changed in a discrete way after the country's experience in Vietnam, then one way to code that would be to have an indicator coded zero prior to 1974 (or whenever) and one after that.
   - *Proper nouns.*
     - It may or may not be a good (substantive) idea to do so, but it is possible to code (essentially) "unit-specific" dummy variables.
     - So, for example, a model of cabinet turnovers might include a dummy variable for Italy, since that country has (historically speaking) been notoriously unstable in its governments.

Coding Dichotomous Variables

Most political science applications use the sort of coding scheme described above, where the two values of a dichotomous variable are coded as 0 and 1:

$$\texttt{female} = \begin{cases} 0 \text{ if male} \\ 1 \text{ if female} \end{cases}$$

This is known as "dummy coding." There are, however, literally an infinite variety of ways to code dichotomous variables; in fact, any two distinct values will do.

A particular coding that one sees used once in a while is called "effect coding." This coding uses values of -1 and 1 instead of 0 and 1:

$$\texttt{female} = \begin{cases} -1 \text{ if male} \\ 1 \text{ if female} \end{cases}$$

The advantages of this approach in a regression context is twofold:

1. Assuming an equal number of 0s and 1s, the mean of an effect-coded variable is zero. In an ANOVA / regression context, that means that the resulting intercept will be equal to the overall (grand) mean of Y.

2. Coding a variable this way means that the effects are symmetric; that is, the resulting coefficient measures the change around the mean for members of each group. So, for the variable above, a coefficient estimate of $\hat{\beta}_{\text{female}} = 0.5$ would mean that females were (on average) 0.5 units higher on Y, and males were 0.5 units lower.

Finally: *Be smart when you name and label your dummy variables*! Variable names like `sex`, `race`, and `partyid` are bad, because they are uninformative; those like `female`, `white`, and `gop` are good, because they have a natural interpretation.

Dummy Variables in Regression Models

We often use dichotomous covariates in regression models of various sorts. The simplest such case is that where we have a single dichotomous variable on the right-hand side of the model:

$$Y_i = \beta_0 + \beta_1 D_i + u_i \tag{1}$$

This amounts to saying that the expected value of Y varies depending on whether or not some trait D is present or absent. In this model,

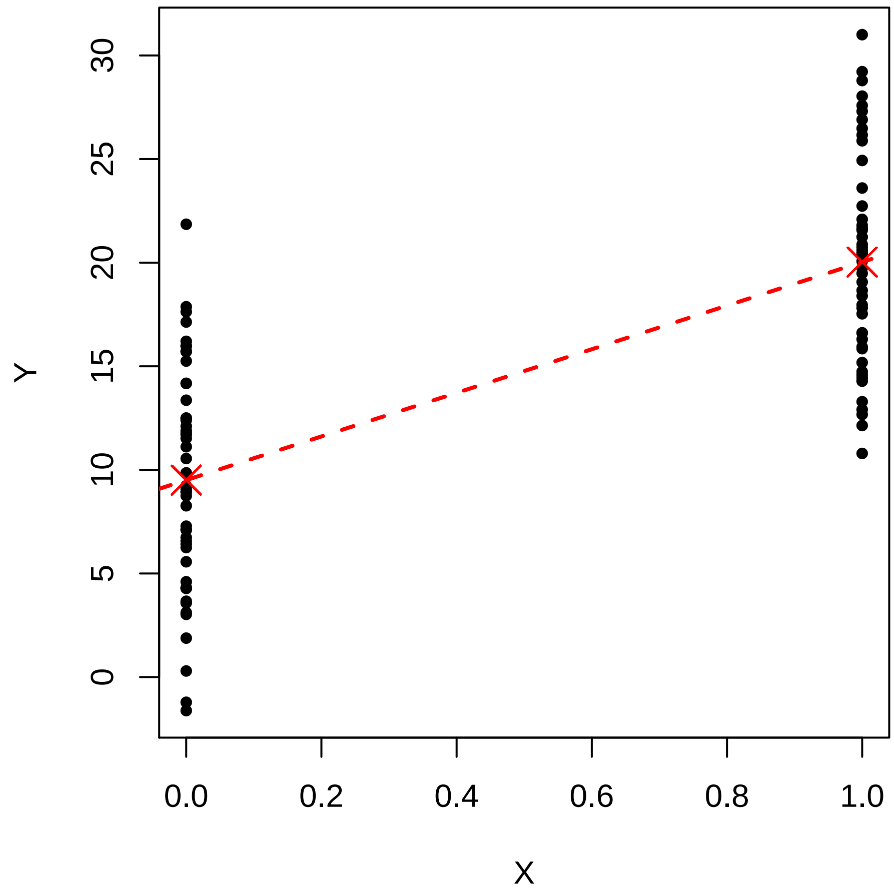
$$\text{E}(Y|D = 0) = \beta_0$$

and

$$\text{E}(Y|D = 1) = \beta_0 + \beta_1$$

Likewise, a test for $\beta_1 = 0$ is equivalent to test for different levels of $\bar{Y}$ across the two groups defined by D – that is, it is the same as a difference of means test. (That ain't exactly rocket science). Graphically, we can think of a scatterplot of Y against a dichotomous D as:

Figure 1: Scatterplot of Y Against Dichotomous D, with Regression Line



This logic extends readily to a model with ℓ separate dummy variables:

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + ... + \beta_\ell D_{\ell i} + u_i \tag{2}$$

Here,

- $\mathrm{E}(Y|D_k = 0) \,\forall\, k \in \ell = \beta_0$,
- Otherwise, $\mathrm{E}(Y) = \beta_0 + \sum_{k=1}^{\ell} \beta_k \,\forall\, k \; s.t. \; D_k = 1$.
- That is, the expectation is just the sum of the coefficient estimates for all those variables that have $D_\ell = 1$.

Note that, in instances where the D_ℓ are mutually exclusive and exhaustive:

- The expected values are the same as the within-group means. Moreover,

- One has the alternative of either

  1. omitting one of the categorical dummy variables and including a constant term β_0 (in which case $\bar{Y}$ for the omitted category equals the intercept, and all other estimates are relative to that "omitted" group), or
  2. leaving all ℓ dummy variables in the model, and omitting a constant term – in which case the coefficient estimates $\hat{\beta}_\ell$ are themselves equal to the within-group means $\bar{Y}_\ell$.

We'll do an illustration of this latter distinction a bit later.

Ordinal Covariates and their Effects

The model in (2) also suggests the potential usefulness of "dummying out" ordinal-level variables. It's often the case that we're not sure whether or not we can treat an ordinal-level variable as interval-level – that is, whether the effects on Y of changes in such a variable are constant across categories. Dummy variables are (one) useful way of seeing whether or not this is the case.

For example, suppose I want to measure the impact of party identification on survey respondents' "closeness" to former President Bush, where the latter is measured as a standard "feeling thermometer" (0-100) scale, and

$$\texttt{gopscale} = \begin{cases} -2 = \text{Strong Democrat} \\ -1 = \text{Weak Democrat} \\ 0 = \text{Independent} \\ 1 = \text{Weak Republican} \\ 2 = \text{Strong Republican} \end{cases}$$

One way to do this is to treat `gopscale` as an interval-level variable, and include it in the model; that might yield results like:

$$\texttt{closeness}_i = 46.0 + 17.5(\texttt{gopscale}_i) + u_i$$

But this assumes that the impact on `closeness` of a one-unit change in `gopscale` is the same across all five categories – that is, that going from "strong democrat" to "weak democrat" moves one close to Bush to the same extent that going from "independent" to "weak Republican" does. If – alternatively – you think that might not be the case, an alternative would be to "dummy out" the `gopscale` variable into five categories, and estimate a model like:

$$\texttt{closeness}_i = \beta_0 + \beta_1(\texttt{strongdem}_i) + \beta_2(\texttt{weakdem}_i) + \beta_3(\texttt{weakgop}_i) + \beta_4(\texttt{stronggop}_i) + u_i$$

where the variables are naturally coded and the "omitted" category is independents. We might then obtain estimates that looked like this:

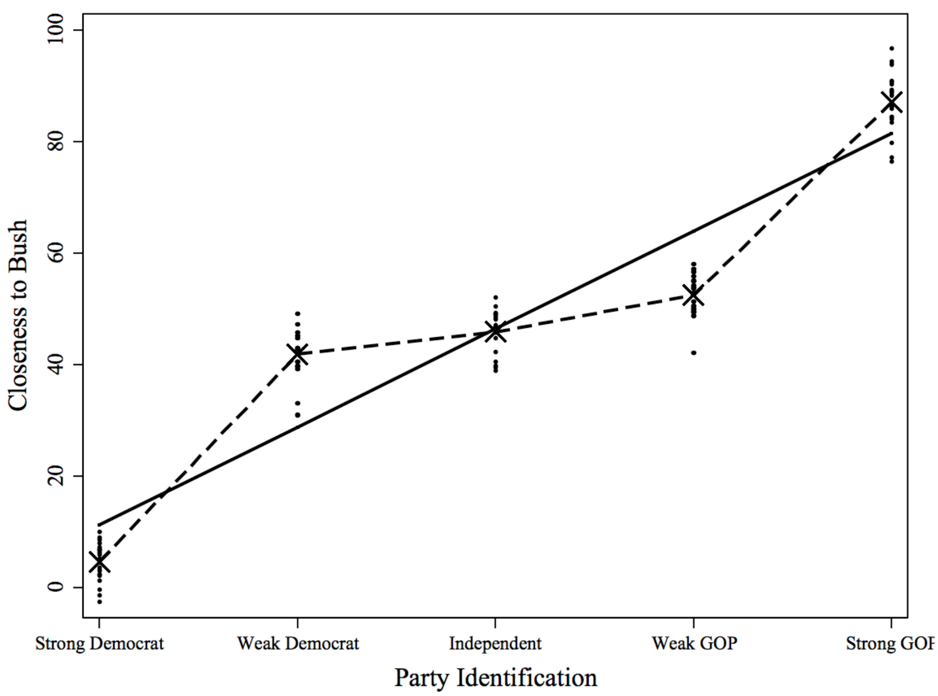
$$\texttt{closeness}_i = 45.5 - 40(\texttt{strongdem}_i) - 6(\texttt{weakdem}_i) + 7(\texttt{weakgop}_i) + 42(\texttt{stronggop}_i) + u_i$$

What we find is that the more extreme partisan feelings are having a bigger impact; rather than an "even" 17.5 point shift in `closeness` for each unit increase in `partyid`, the expected values of closeness go from 5.5 $\rightarrow$ 39.5 $\rightarrow$ 45.5 $\rightarrow$ 52.5 $\rightarrow$ 87.5 across each of the five categories. In the meantime, the model that treats `gopscale` as interval-level winds up

- overestimating closeness for strong Democrats and weak Republicans, while
- underestimating it for strong Republicans and weak Democrats.

In other words, the simple linear/interval model treats weak party identifiers too much like strong ones, and vice-versa; this is illustrated in Figure 2.

Figure 2: Scatterplot of `closeness` Against `gopscale`, with Regression Lines



Solid line is fitted values from treating `partyid` as continuous; dashed line is fitted values for separate dichotomous variables for each category of `partyid`.

Combined Dichotomous and Continuous Variables

Now consider a simple model where we have both a continuous variable and a dummy variable among the covariates:

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + u_i \tag{3}$$

In this model, the expected value of Y is:

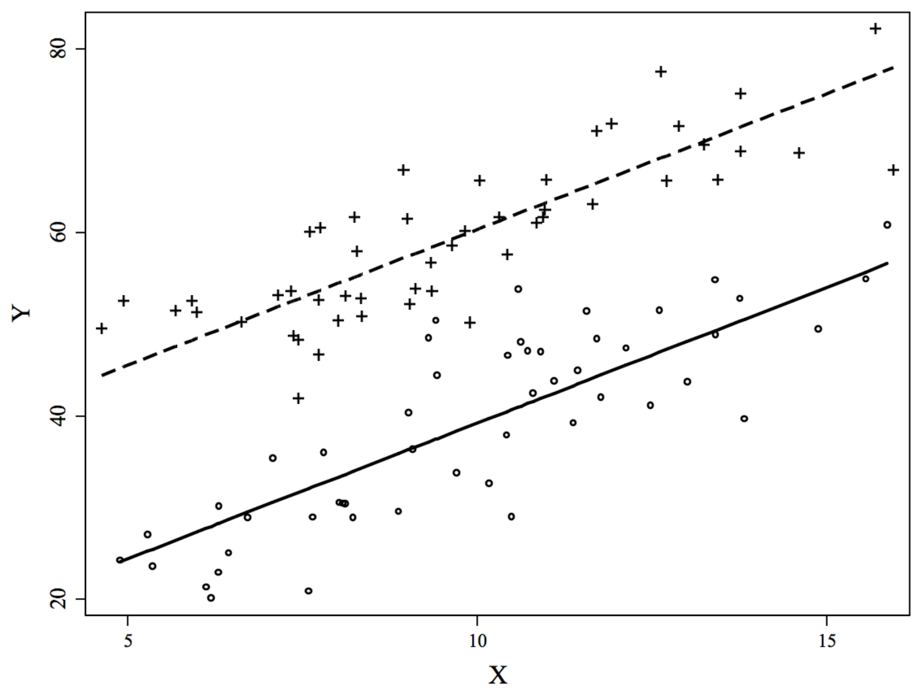
$$\mathrm{E}(Y|X, D = 0) = \beta_0 + \beta_2 X_i$$

and

$$\mathrm{E}(Y|X, D = 1) = (\beta_0 + \beta_1) + \beta_2 X_i$$

In other words, this regression assumes that observations with different values of D have different intercepts, but the same slopes for X. Graphically:

Figure 3: Scatterplot and Regression Lines of Y on X for $D = 0$ and $D = 1$



Circles are values of X and Y for $D = 0$; crosses are those for $D = 1$. Solid line is the fitted regression of Y on X for $D = 0$; dashed line is for $D = 1$.

That is, the impact of X on Y is the same irrespective of the value of D (that is, irrespective of the "type" of the observation). The same intuition extends fairly straightforwardly to cases where there are multiple dummy variables and/or multiple continuous (X) covariates: changes in the values of the various D_ℓ variables "shift" the regression line/plane up or down (by a value of β_ℓ), but don't affect its slope.[1]

A Few Practical Examples

I thought it would be useful to discuss some "tricks" for handling dichotomous variables, particularly in R; there are a list of comparable Stata commands at the end of these notes. To do that, I draw on some data on the Supreme Court, in particular, on the number of *amicus curiae* briefs filed in each of the 7,161 cases decided by the Court between 1953 and 1985. The data "look like" this:

```
> summary(SCOTUS)
       id             term           Namici          lctdiss          multlaw
 Min.   :   1   Min.   :53.00   Min.   : 0.000   Min.   :0.0000   Min.   :0.0000
 1st Qu.:1791   1st Qu.:64.00   1st Qu.: 0.000   1st Qu.:0.0000   1st Qu.:0.0000
 Median :3581   Median :72.00   Median : 0.000   Median :0.0000   Median :0.0000
 Mean   :3581   Mean   :71.12   Mean   : 0.842   Mean   :0.1509   Mean   :0.1490
 3rd Qu.:5371   3rd Qu.:79.00   3rd Qu.: 1.000   3rd Qu.:0.0000   3rd Qu.:0.0000
 Max.   :7161   Max.   :85.00   Max.   :39.000   Max.   :1.0000   Max.   :1.0000
                NA's   : 4.00                    NA's   :4.0000   NA's   :5.0000
    civlibs           econs           constit           lctlib
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :  0.0000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:  0.0000
 Median :1.0000   Median :0.0000   Median :0.0000   Median :  0.0000
 Mean   :0.5009   Mean   :0.1709   Mean   :0.2536   Mean   :  0.3742
 3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:1.0000   3rd Qu.:  1.0000
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :  1.0000
                                                    NA's   :120.0000
```

Note that all of the variables except `id`, `term` and `Namici` are dichotomous.

Generating Dummies

One of the "tricks" to dummy covariates is that, if the dummies are mutually exclusive, they can be combined into "larger" categories simply by adding them. So, if we wanted to create a variable equal to one for all civil liberties and economics cases and zero otherwise, we could do so by:

[1]The idea that the regression line ("slope") doesn't change across subgroups of our data is something that we've been assuming with all along, though we haven't been that explicit about it; we'll discuss this at much greater length a bit later in the course.

```
> SCOTUS$civil.econ<-SCOTUS$civlibs + SCOTUS$econs
```

Dummy variables are also known as "factors" or "factor variables" in statistics. R has a nifty little component called (appropriately) `factor`, that manages dichotomous (and, for that matter, unordered polytymous) variables. In particular, `factor()` will take a variable (say, Z) with k categories and generate k separate, dichotomous indicator variables, one for each value of Z:

```
> SCOTUS$termdummies<-factor(SCOTUS$term)
> is.factor(SCOTUS$termdummies)
[1] TRUE
> summary(SCOTUS$termdummies)
  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67
 126 109 128 162 196 165 157 160 148 189 223 156 187 201 285

  68  69  70  71  72  73  74  75  76  77  78  79  80  81
 207 185 227 262 269 267 223 253 254 244 244 221 255 269

  82  83  84  85 NA's
 277 298 301 309    4
```

The latter bit lists the frequencies of "ones" for each of the factor variables here (which are identical to the "`_I...`" variables that Stata generates with `-xi-`; see below).

Dummy Variable Regressions

Interpretation of a regression with a dummy covariate is, as we said, straightforward:

```
> fit1<-with(SCOTUS, lm(Namici~civlibs))
> summary(fit1)

Call:
lm(formula = Namici ~ civlibs)

Residuals:
   Min     1Q Median     3Q    Max
-0.918 -0.918 -0.766  0.082 38.234

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.91774    0.03661  25.069  < 2e-16 ***
civlibs     -0.15136    0.05173  -2.926  0.00344 **
```

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

Residual standard error: 2.189 on 7159 degrees of freedom
Multiple R-squared:  0.001195,Adjusted R-squared:  0.001055
F-statistic: 8.563 on 1 and 7159 DF,  p-value: 0.003442
```

This means that:

- The expected value (mean) of `Namici` for non-civil liberties cases is 0.92,
- The expected value in civil liberties cases is $0.92 - 0.15 \approx 0.77$; that means that, on average, civil rights and liberties cases had about 0.15 fewer *amicus curiae* briefs filed than did other sorts of cases.
- The significance test for $\beta_1 = 0$ is the same as that for a t-test for the difference in means across the groups defined by `civlibs`:

```
> with(SCOTUS, t.test(Namici~civlibs))

Welch Two Sample t-test

data:  Namici by civlibs
t = 2.9258, df = 7114.116, p-value = 0.003446
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.04995001 0.25277126
sample estimates:
mean in group 0 mean in group 1
      0.9177392       0.7663786
```

Note that if we recode `civlibs` according to *effect coding*, a couple things change:

```
> SCOTUS$civlibeffect<-SCOTUS$civlibs
> SCOTUS$civlibeffect[SCOTUS$civlibs==0]<-(-1)
> fit2<-with(SCOTUS, lm(Namici~SCOTUS$civlibeffect))
> summary(fit2)

Call:
lm(formula = Namici ~ SCOTUS$civlibeffect)

Residuals:
   Min     1Q Median     3Q    Max
-0.918 -0.918 -0.766  0.082 38.234
```

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.84206    0.02586  32.559  < 2e-16 ***
SCOTUS$civlibeffect -0.07568    0.02586  -2.926  0.00344 **
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.189 on 7159 degrees of freedom
Multiple R-squared:  0.001195,Adjusted R-squared:  0.001055
F-statistic: 8.563 on 1 and 7159 DF,  p-value: 0.003442
```

Note that:

- Now the intercept is very close to the overall mean; in fact, it would be exactly the mean if there were the same number of observations with `civlibeffect` $= -1$ and `civlibeffect` $= 1$.
- The coefficient estimate is exactly half the size it was before; that is because
- We would now interpret the effect of `civlibeffect` to be that civil liberties cases decrease the average number of briefs by 0.075, while non-civil liberties case increases that average by the same amount. However,
- The level of statistical significance of the effect is exactly the same as before.

Practically speaking, we rarely see effect coding used in political science (or most other social sciences – psychology being a notable exception).

In the context of multivariate regression, the same general kinds of interpretations apply for models with more than one dichotomous covariate. So, for example:

```
> fit3<-with(SCOTUS, lm(Namici~lctdiss+multlaw+civlibs+
+                          econs+constit+lctlib))
> summary(fit3)

Call:
lm(formula = Namici ~ lctdiss + multlaw + civlibs + econs + constit +
    lctlib)

Residuals:
   Min     1Q Median     3Q    Max
-2.582 -0.976 -0.472 -0.260 37.086
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.47245    0.05273   8.960  < 2e-16 ***
lctdiss      0.36760    0.07173   5.125 3.06e-07 ***
multlaw      0.61306    0.07445   8.235  < 2e-16 ***
civlibs     -0.21255    0.06022  -3.530 0.000419 ***
econs        0.08772    0.07652   1.146 0.251691
constit      0.53793    0.06372   8.442  < 2e-16 ***
lctlib       0.50309    0.05396   9.323  < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1

Residual standard error: 2.15 on 7033 degrees of freedom
  (121 observations deleted due to missingness)
Multiple R-squared:  0.05013,Adjusted R-squared:  0.04932
F-statistic: 61.86 on 6 and 7033 DF,  p-value: < 2.2e-16
```

Once again, the interpretation of the effects is straightforward:

- The intercept now tells us the expected number of *amicus* briefs when all the covariates are equal to zero.
- The expected value of Y is then just the sum of the coefficient estimates for the covariates which are "present" for that observation.
- So, the expected number of amici for a non-constitutional economics case decided liberally (but without dissent) in the lower court and that involved multiple legal provisions would be $0.472 + 0.613 + 0.088 + 0.503 = \mathbf{1.676}$.
- Inference, etc. proceeds just as it normally would.

Finally, note that we can use `factor` in a regression context as well. Suppose we wanted to test the hypothesis that the numbers of *amicus curiae* briefs filed in the Court were changing systematically over time (say, increasing). One way to do so would be to include the variable for `term` in the model:

```
> fit4<-with(SCOTUS, lm(Namici~lctdiss+multlaw+civlibs+
+                          econs+constit+lctlib+term))
> summary(fit4)

Call:
lm(formula = Namici ~ lctdiss + multlaw + civlibs + econs + constit +
    lctlib + term)
```

```
Residuals:
   Min     1Q Median     3Q    Max 
-2.968 -0.906 -0.428  0.143 36.958 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -2.726962   0.202367 -13.475  < 2e-16 ***
lctdiss      0.359494   0.070415   5.105 3.39e-07 ***
multlaw      0.649932   0.073109   8.890  < 2e-16 ***
civlibs     -0.289314   0.059295  -4.879 1.09e-06 ***
econs        0.199464   0.075419   2.645  0.00819 ** 
constit      0.515435   0.062559   8.239  < 2e-16 ***
lctlib       0.339891   0.053901   6.306 3.04e-10 ***
term         0.046142   0.002821  16.354  < 2e-16 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1 

Residual standard error: 2.11 on 7032 degrees of freedom
  (121 observations deleted due to missingness)
Multiple R-squared:  0.08493,Adjusted R-squared:  0.08402 
F-statistic: 93.24 on 7 and 7032 DF,  p-value: < 2.2e-16
```

That suggests that, during the 1953-1985 period, the average number of *amicus* briefs filed per case increased by about 0.05 per term. But suppose we were concerned that the change wasn't in fact, linear and/or monotonic (perhaps we believed that there was little or no change in the 1950s and 1960s, but a large increase in the 1970s and 1980s).

We could assess that hypothesis by creating a separate dummy variable for each term. That, in effect, allows the expected number of briefs to vary up or down by an undetermined amount each term. One can use `as.factor()` to include a covariate like Z in a regression (or other) model as a set of $k-1$ indicator variables (in a manner analogous to using `-xi-` in Stata):

```
> fit5<-with(SCOTUS, lm(Namici~lctdiss+multlaw+civlibs+
+                      econs+constit+lctlib+as.factor(term)))
> summary(fit5)

Call:
lm(formula = Namici ~ lctdiss + multlaw + civlibs + econs + constit + 
    lctlib + as.factor(term))

Residuals:
   Min     1Q Median     3Q    Max 
-3.064 -0.920 -0.384  0.106 36.831 
```

```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)        -0.16153    0.19530  -0.827 0.408200    
lctdiss             0.34558    0.07067   4.890 1.03e-06 ***
multlaw             0.64348    0.07334   8.774  < 2e-16 ***
civlibs            -0.27137    0.05967  -4.548 5.51e-06 ***
econs               0.20039    0.07581   2.643 0.008232 ** 
constit             0.54280    0.06297   8.620  < 2e-16 ***
lctlib              0.33863    0.05458   6.205 5.80e-10 ***
as.factor(term)54   0.26276    0.27934   0.941 0.346918    
as.factor(term)55   0.20958    0.26804   0.782 0.434309    
as.factor(term)56   0.12536    0.25126   0.499 0.617859    
as.factor(term)57   0.06432    0.24227   0.265 0.790654    
as.factor(term)58   0.08353    0.25274   0.331 0.741025    
as.factor(term)59   0.11212    0.25322   0.443 0.657942    
as.factor(term)60   0.32652    0.25172   1.297 0.194616    
as.factor(term)61   0.21610    0.25660   0.842 0.399710    
as.factor(term)62   0.44318    0.24452   1.812 0.069953 .  
as.factor(term)63   0.39759    0.23657   1.681 0.092879 .  
as.factor(term)64   0.52677    0.25443   2.070 0.038453 *  
as.factor(term)65   0.23301    0.24455   0.953 0.340723    
as.factor(term)66   0.29139    0.24046   1.212 0.225630    
as.factor(term)67   0.50656    0.22711   2.230 0.025748 *  
as.factor(term)68   0.45982    0.24022   1.914 0.055638 .  
as.factor(term)69   0.46774    0.24618   1.900 0.057480 .  
as.factor(term)70   0.45442    0.23585   1.927 0.054056 .  
as.factor(term)71   0.62313    0.23019   2.707 0.006806 ** 
as.factor(term)72   0.59503    0.22929   2.595 0.009476 ** 
as.factor(term)73   0.78179    0.22918   3.411 0.000650 ***
as.factor(term)74   0.53254    0.23636   2.253 0.024287 *  
as.factor(term)75   0.80353    0.23118   3.476 0.000513 ***
as.factor(term)76   0.49269    0.23138   2.129 0.033262 *  
as.factor(term)77   1.07725    0.23265   4.630 3.72e-06 ***
as.factor(term)78   1.04335    0.23243   4.489 7.27e-06 ***
as.factor(term)79   0.85363    0.23696   3.602 0.000318 ***
as.factor(term)80   1.21205    0.23183   5.228 1.76e-07 ***
as.factor(term)81   1.49347    0.22925   6.515 7.80e-11 ***
as.factor(term)82   1.46004    0.22858   6.388 1.79e-10 ***
as.factor(term)83   1.29417    0.22549   5.739 9.90e-09 ***
as.factor(term)84   1.23434    0.22517   5.482 4.36e-08 ***
as.factor(term)85   1.59037    0.22491   7.071 1.68e-12 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1   1
```

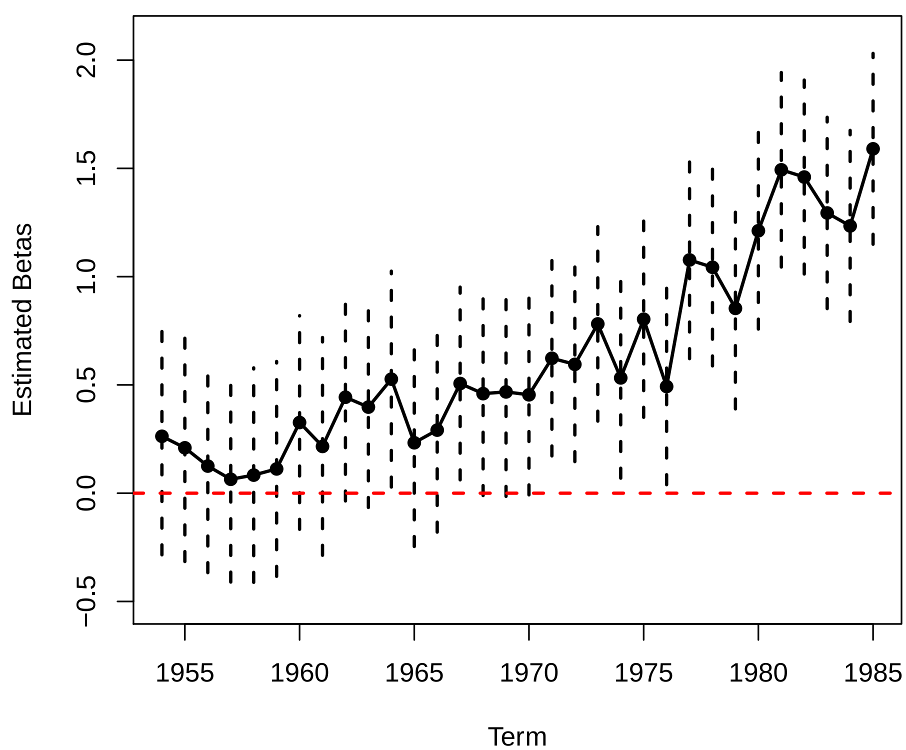
```
Residual standard error: 2.108 on 7001 degrees of freedom
  (121 observations deleted due to missingness)
Multiple R-squared:  0.0914,Adjusted R-squared:  0.08647
F-statistic: 18.53 on 38 and 7001 DF,  p-value: < 2.2e-16
```

This is revealing; note that:

- The "omitted" / reference category is the 1953 term (the first in the data); that means that
- All other `term` dummy variables indicate the (marginal) change in the mean number of *amicus* briefs filed relative to 1953. For example:
  - The 1963 term had about 0.4 more briefs per case, on average, than did 1953;
  - The 1985 term had 1.59 more briefs per case, on average, than in 1953.
  - Note that these average differences are after accounting for any changes in the other covariates that might have occurred over time during the same period.
- The pattern, while more or less monotonic, does vary somewhat over time.
- Finally, one could plot the coefficient estimates seen above (and their confidence intervals), for a clearer picture of what's changing...

```
> termbetas<-fit5$coefficients[8:39]
> SE5<-sqrt(diag(vcov(fit5)))[8:39]
> termUBs <- termbetas + 1.96*(SE5)
> termLBs <- termbetas - 1.96*(SE5)
> term<-seq(1954,1985)
>
> plot(term,termbetas, xlab="Term",ylab="Estimated Betas",
+       pch=19,ylim=c(-0.5,2.1))
> lines(term,termbetas,lwd=2)
> segments(term,termLBs,term,termUBs,lwd=2,lty=2)
> abline(h=0,lwd=2,lty=2,col="red")
```

Figure 4: Plot of Estimated βs and 95% C.I.s, By Term



That's probably enough for now. Next time, we'll begin to discuss interpretation of regression results in more detail.