Quadratics -- annoying but necessary

Formula for the quadratic: \( y = ax^2 + bx + c \)

Symmetrical but any position.

\( \frac{dy}{dx} = b_1 + 2b_2x \) Formula for the slope

If you have to poke around to understand.

If it is very useful to consider the formula for the slope.
Formula for the quadratic

"slope of the quadratic"

\[ \text{Slope} \]

1. Every parabola has a maximum or minimum point.
   Where is it? It is the \( x \) where slope is zero.

   If \( \text{Slope} = b_1 + 2b_2x \)

   then \( 0 = b_1 + 2b_2x \implies x = \frac{-b_1}{2b_2} \)

   but slope tomen is more generally useful - e.g.
   what is slope at \( x = 0 \)?

   \[ \text{Slope} = b_1 \]  so \( b_1 \) tells you slope at \( x = 0 \)

sign of \( x^2 \) tells you "U" or "\( V \)"

\[ b_2 < 0 \implies \bigwedge \]

\[ b_2 > 0 \implies \bigvee \]
Table 2. Estimated Racial Achievement Gap over the First Four Years of School, Math

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>-0.665 (0.025)</td>
<td>-0.724 (0.027)</td>
<td>-0.748 (0.029)</td>
<td>-0.882 (0.031)</td>
<td>-0.999 (0.026)</td>
<td>-0.209 (0.028)</td>
<td>-0.279 (0.031)</td>
<td>-0.382 (0.033)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.738 (0.024)</td>
<td>-0.681 (0.025)</td>
<td>-0.568 (0.026)</td>
<td>-0.539 (0.026)</td>
<td>-0.197 (0.024)</td>
<td>-0.189 (0.026)</td>
<td>-0.122 (0.027)</td>
<td>-0.078 (0.028)</td>
</tr>
<tr>
<td>Asian</td>
<td>0.11 (0.058)</td>
<td>0.088 (0.056)</td>
<td>-0.025 (0.052)</td>
<td>0.066 (0.054)</td>
<td>0.258 (0.050)</td>
<td>0.226 (0.050)</td>
<td>0.992 (0.047)</td>
<td>0.163 (0.049)</td>
</tr>
<tr>
<td>Other race</td>
<td>-0.495 (0.047)</td>
<td>-0.481 (0.048)</td>
<td>-0.497 (0.050)</td>
<td>-0.541 (0.050)</td>
<td>-0.158 (0.040)</td>
<td>-0.175 (0.043)</td>
<td>-0.21 (0.046)</td>
<td>-0.244 (0.046)</td>
</tr>
<tr>
<td>Age (in months)</td>
<td>0.038 (0.002)</td>
<td>0.053 (0.002)</td>
<td>0.037 (0.002)</td>
<td>0.019 (0.002)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
</tr>
<tr>
<td>Birth weight (in ounces)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
<td>0.003 (0.000)</td>
</tr>
<tr>
<td>Female</td>
<td>0.005 (0.017)</td>
<td>-0.005 (0.017)</td>
<td>-0.044 (0.018)</td>
<td>-0.175 (0.018)</td>
<td>0.006 (0.001)</td>
<td>0.006 (0.001)</td>
<td>0.005 (0.000)</td>
<td>0.006 (0.001)</td>
</tr>
<tr>
<td>Number of children's books</td>
<td>0.006 (0.001)</td>
<td>0.006 (0.001)</td>
<td>0.005 (0.001)</td>
<td>0.006 (0.001)</td>
<td>-0.021 (0.002)</td>
<td>-0.02 (0.003)</td>
<td>-0.019 (0.003)</td>
<td>-0.020 (0.003)</td>
</tr>
<tr>
<td>Number of children's books (squared) (+1000)</td>
<td>-0.165 (0.026)</td>
<td>0.107 (0.025)</td>
<td>0.086 (0.022)</td>
<td>0.083 (0.024)</td>
<td>0.396 (0.016)</td>
<td>0.282 (0.015)</td>
<td>0.256 (0.015)</td>
<td>0.288 (0.015)</td>
</tr>
<tr>
<td>Mother over 30 at first birth</td>
<td>0.165 (0.026)</td>
<td>0.107 (0.025)</td>
<td>0.086 (0.022)</td>
<td>0.083 (0.024)</td>
<td>0.396 (0.016)</td>
<td>0.282 (0.015)</td>
<td>0.256 (0.015)</td>
<td>0.288 (0.015)</td>
</tr>
<tr>
<td>Socioeconomic status measure</td>
<td>-0.212 (0.021)</td>
<td>-0.191 (0.022)</td>
<td>-0.19 (0.023)</td>
<td>-0.208 (0.024)</td>
<td>-0.114 (0.021)</td>
<td>-0.118 (0.022)</td>
<td>-0.131 (0.025)</td>
<td>-0.132 (0.025)</td>
</tr>
<tr>
<td>Mother receives Wic benefits</td>
<td>0.307 (0.013)</td>
<td>0.304 (0.013)</td>
<td>0.286 (0.012)</td>
<td>0.275 (0.012)</td>
<td>-4.357 (0.154)</td>
<td>-3.952 (0.160)</td>
<td>-2.795 (0.166)</td>
<td>-1.576 (0.168)</td>
</tr>
<tr>
<td>Mother a teenager at first birth</td>
<td>0.32 (0.011)</td>
<td>0.29 (0.012)</td>
<td>0.24 (0.013)</td>
<td>0.26 (0.014)</td>
<td>11201</td>
<td>11201</td>
<td>11201</td>
<td>11201</td>
</tr>
<tr>
<td>Constant</td>
<td>0.11</td>
<td>0.11</td>
<td>0.1</td>
<td>0.12</td>
<td>0.32</td>
<td>0.29</td>
<td>0.24</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is the math test score at various points in the student's career. Test scores are IRT scores, normalized to have a mean of 0 and a standard deviation of 1 in the full, unweighted sample. Non-Hispanic Whites are the omitted race category, so all of the race coefficients are relative to that group. The unit of observation is a student. Standard errors in parentheses. Estimation is done using weighted least squares, using sample weights provided in the data set.
4th class  Part 1

Leviit u. Fryer

Fall 1K math:

\[ T = a + 0.006 \text{ books} - 0.021 \text{ books}^2 \]

\[ (0.001) \quad (0.002) \]

"*1000" means that 

\[ 0.021 \text{ is really } 0.000021 \]

\[ 0.002 \text{ is } 0.000002 \]

Never enter 

\[ 0.001 \text{ into a table} \]

\[ (0.000) \Rightarrow \text{completely} \]

\[ \text{unimportant} \]

Figure out how to judge useful numbers

First class

\[ Cd = a + 0.14 \text{ Fair Income} \]

\[ (0.02) \]

Income in 10,000s = 0.14

\[ (0.02) \]

Income in dollars = 0.00014

\[ (0.00002) \]

\[ 1.4 \times 10^{-4} \]

\[ (0.02) \times 10^{-4} \]
Return to

\[ T = a + 0.006B - 0.021B^2 \]

(1) Negative effect on squared term indicated

(2) Where is max?

\[ \frac{-b_1}{2b_2} = \frac{-0.006}{0.00042} = 142 \quad \text{(near 60)} \]

(3) What is slope at zero?

First, effect in B \( \Rightarrow \) 0.006

\[ = \] first book is assigned with 0.006 std increase in test score

First 10 books, n = 0.06 std

(4) What is value of function at B = 0?

\[ T = a + 0.006(142) - 0.021(142)^2 \]

\[ a + 0.429 \text{ std} \]

\[ a + 0.429 \text{ std} = \]

\[ \Rightarrow \text{ANNOYING!} \]
latest growth models -- often quadratic

level 1 (child) \[ T = a + b_1 \text{Age} + b_2 \text{Age}^2 \]

\( \Rightarrow \) fit a quadratic to each child

\[ b_1 = a + b, \text{Demog} \]
\[ b_2 = c + d, \text{Demog} \]
\[ b_2 = e + f, \text{Demog} \]

even with \( b \)

in the aggregate, each seen

character for individual

\( \text{TV} \)

\( \text{Age} \)
4th class Part II

& max in Dummy variables -- interact.

Support two groups B(1,0), (W) when B

is F1 less advantageous for blacks?

Write: Ed = a + b, Inc

Blue: Ed = c + d, Inc Is c1 < b1?

Investigate with dummy variable interact.

Data set: my Blades at UCLA

Include both groups and means;

\[ Ed = \alpha + \beta_1 f_1 \text{Inc} + f_2 B + f_3 B \cdot \text{Inc} \]

Why is the line same?

For white: \[ Ed = E + f_1 \text{Inc} + 0 \]

\[ E = a \quad f_1 = b \]

For blacks: \[ Ed = (e + f_2) + (f_1 + f_3) \text{Inc} \]

\[ \frac{f_1 + f_3}{f_2} \]

\[ f_2 > 0 \]

\[ f_3 < 0 \]
4th class Part II

key is that $f_3$ produces exactly the sig
test that you want.

Breakdown groups:

Suppose that you pooled groups and ran

$E_d = C + f_1 Inc + f_3 B \times Inc$

(i.e. left out the main effect
of B)

What would the graph look like?

Suppose 3 groups: white, Black($B$) and Hispanic($H$)

Pool and estimate

Child $E_d = g + h_1 FamInc + h_2 B + h_3 H$

$+ h_4 B \times FamInc + h_5 H \times FamInc$

For White: $E_d \approx g + h_2 FamInc$

Black: $E_d \approx (g + h_2) + (h_1 + h_4) FamInc$

Hispanics $E_d \approx (g + h_3) + (h_1 + h_5) FamInc$
4th class Part II

h4 and h5 test the interaction

Also might test $F_{HH}$

$B = H = 0$

Use dummy variables for missing data?

(b) (H)

Suppose Race/Ethnicity = White, Black, Hispanic, missing

(m)

Ed Lazear

Fryer – Levent

Test: $a + b_1 B + b_2 H + b_3 M$

$b_3$: do missing cases differ from units?

Useful for minor variables
Mediation and Moderation

Moderation -- interactions. Does the effect of \( X \) on \( Y \) depend on \( Z \)?

\[ X \rightarrow Y \]

\( Z \) is often an exogenous variable and it often a demographic measure.

Does the effect of family income \((X)\) on children's schooling \((Y)\) depend on demographic factors?

Missing moderation

Suppose treated \( X \) \( T \)
and outcome \( Y \)

and suppose we run \( Y = a + b_1T + \text{other controls} \)

and find that \( b_1 \) is insignificant -- nice try!

\( b_1 \) is the "main effect" of the experiment.
4th class Part III

Oftentimes null effects launch people off on a search for subgroups for which it does matter.

Two problems:

1. If you don’t measure exactly then your chance of finding a fluke interaction is high.

2. More -- Suppose two subgroups A: B
   If overall effect is null but effect for subgroup A is positive and significant then effect for subgroup B is negative.

   B is almost always except under the very

   An anomaly demands as much theory as negative effects for B as positive
   effects for A.

   e.g. ECE and disadvantaged children.
**4th class  Part III**

**Mediation**

Mediation was about regression models that differ in their effect of 
\[ x \rightarrow y \]
mediates the effect of \( x \) on \( y \).

Mediation of \( x \) through \( z \) where \( z \) is an exogenous variable that 
accounts for the effect of \( x \) on \( y \).

\[ z_1 \rightarrow y \]

\[ x \rightarrow z_2 \rightarrow y \]

\[ z_2 \rightarrow y \]

\[ z \rightarrow y \]

A mediated model is sometimes called a "causal model" in statistics, in a focus on

Causal model

Y in the presence and absence of \( x \) for

individuals accounts for

concepts sense -- what accounts for the causal effect of \( x \) on \( y \)

\[ x \rightarrow y \] "total effect"

\[ z \rightarrow y \] direct effect + indirect effect operating through \( z \).

Ideally, you want to know all 6 of the \( Z_i \)'s that

account for the effect and not have any

indirect effect.
 MEDULLAR WURSTELS are typically poured with SEM

SEM handles myriad cell---but if SEM
leaves SEM worlds can just request

... results...
in our model, the total effect of $X$ on $Y$ is the sum of the

- **Direct effect** ($d_2$) that operates independently of
  the mediator

- **Indirect effect** ($c_1d_1$) that operates through the
  mediator

Try it... it's exactly the same.

If more than one mediator -- $Z_1, Z_2, Z_3$ -- then

\[
\begin{array}{c}
\text{X} \\ \hline \hline \text{X} \\
\end{array}
\]

Still the case, the total effect of $X$ on $Y$ is the sum of the direct effect (3) and the two indirect effects ($1 \times 2$) and ($4 \times 5$)

Important things to note:

1. **Indirect effect under multiplication of path coefficients**
4th class Part III

\[ \begin{array}{c}
\times \\
0.25
\end{array} \]

Looks like a really important input offset = but if \(0.4 \times 0.4 = 0.16\)

Skill lurking model

\[ \begin{array}{c}
5_1 \rightarrow 5_2 \\
5_2 \rightarrow 5_3 \\
5_3 \rightarrow 5_4 \\
5_4 \rightarrow 5_5
\end{array} \]

But wait is the effect of \(5_1 \rightarrow 5_4\)

\[ b_1 \times b_2 \times b_3 \times b_4 \]

If \(b_i\) are .6, then \(b_1^4 = 0.13\) not so big

16, 4, W \(=\) .03

2. Just because, this is a causal model in a strict sense doesn't mean you shouldn't pay attention to omitted variable bias

\[ \begin{array}{c}
x \rightarrow 2 \\
\rightarrow Y
\end{array} \]

then omit effect attributes to 2 not right belong to other mediate.
more important: do not count odd if x odd. This should be just as many.

And even if x is randomly assigned, put in 2 to y is not considered.