# More Regression Applications

Lecture 24

Reading: None

1

# Outline

- Discuss three case studies – empirical papers published in academic journals – to reinforce key concepts from Chapters 14, 18 – 21
  - Tekin and Mocan (2010): "Ugly Criminals"
  - Deryugina and Shurchkov (2015): "Does Beauty Matter in Undergraduate Education?"
  - Andreoni and Vesterlund (2001): "Which is the fair sex? Gender differences in altruism"

2

# "Ugly Criminals"

**Abstract:** Being very attractive reduces a young adult's propensity for criminal activity and being unattractive increases it. Being very attractive is also positively associated with wages and with adult vocabulary test scores, which implies that beauty may have an impact on human capital formation. The results suggest that a labor market penalty provides a direct incentive for unattractive individuals toward criminal activity. The level of beauty in high school is associated with criminal propensity seven to eight years later, which seems to be due to the impact of beauty in high school on human capital formation, although this avenue seems to be effective for females only.

Tekin and Mocan (2010), *The Review of Economics and Statistics,* https://www.mitpressjournals.org/doi/10.1162/rest.2009.11757, copy on Readings page in Quercus (optional)

3

## “Ugly Criminals”: Data

- Uses data from National Longitudinal Study of Adolescent Health
  - Wave III interviews in 2001/02: respondents are 18 – 26 years old
  - Asked many questions
  - Interviewer answered (discretely): “How physically attractive is the respondent?”

Dist. of Attractiveness Ratings (%) among Young Adults (18 – 26)

| Category | Males | Females |
|---|---|---|
| **1.** Very unattractive | 1.37 | 2.44 |
| **2.** Unattractive | 5.22 | 4.81 |
| **3.** About average | 51.82 | 40.55 |
| **4.** Attractive | 33.66 | 38.00 |
| **5.** Very attractive | 7.92 | 14.19 |
| N | 7,159 | 8,020 |

4

Variable Definitions:

- *Wage* is hourly wage rate in dollars (mean ~ $11 and s.d. ~ $7)
- *Test score* is the percentile score for the Peabody Picture Vocabulary test (mean ~ 50 and s.d. ~ 29)
- “*Very Attractive* captures individuals who received the highest rating of 5; *Unattractive,* those with a rating of 1 or 2” p. 16
- “*Personal characteristics* are age, race/ethnicity, non-wage income, self-reported health status, whether he or she was born in the United States, birth weight, and religious affiliation.” “[*Family characteristics* are] the mother’s education, whether the family was on welfare, family income, whether the father was biological or a stepfather, the age of the mother at birth, whether the father was in jail, and birth weight.” p. 17

How many of the personal characteristic variables are dummies?

5

**Table 7. Effect of Beauty on Wages and Test Scores**

| | | Females | | Males | |
|---|---|---|---|---|---|
| | | *Log Wages* | *Test Score* | *Log Wages* | *Test Score* |
| *Very Attractive* | What is reference | 0.065*** (0.014) | 2.999*** (0.906) | 0.107*** (0.024) | 3.706*** (1.163) |
| *Unattractive* | (omitted) category? | -0.043** (0.020) | -2.330* (1.210) | -0.041* (0.025) | -1.800 (1.326) |
| *Control Variables: Personal and family attributes* | | Yes | Yes | Yes | Yes |
| *Interviewer fixed effects* | | Yes | Yes | Yes | Yes |
| *N* | | 3,730 | 5,954 | 3,521 | 5,209 |

*Note:* * Estimated coefficient is statistically different from 0 at the 10% level, ** significant at 5%, and *** significant at 1% or better

After controlling for personal/family attributes and interviewers’ tastes, very attractive females on average have wages that are about ___________ compared to average/attractive females.

6

“Does Beauty Matter in Undergraduate Education?”

**ABSTRACT:** Physically attractive individuals achieve greater success in terms of earnings and status than those who are less attractive. However, whether this “beauty premium” arises primarily because of differences in ability or confidence, bias, or sorting remains unknown. We use a rich dataset from a women’s college to evaluate each of these three mechanisms at the college level. We find that students judged to be more attractive perform significantly worse on standardized tests but, conditional on test scores, are not evaluated more favorably at the point of admission, suggesting that more attractive people do not possess greater abilities at the beginning of college. Controlling for test scores, more attractive students receive only marginally better grades in some specifications, and the magnitudes of the differences are very small. Finally, there is substantial beauty-based sorting into areas of study and occupations.

What are the research questions? Observational data?

Deryugina and Shurchkov (2015): Copy on Readings page in Quercus (optional).

7

**Excerpt (p. 942):** Our dataset consists of 794 alumnae who graduated from an anonymous women’s college between the years 2002 and 2011. To measure attractiveness, we use pictures [from student ID cards] taken [by campus officials] when the alumnae were first-year students. The pictures were subsequently rated by current male and female students from a college in another state. Each picture was rated by at least 25 male and 25 female raters. [We combine these to form an attractiveness rating, which we standardize so that a rating of 1 means the person is 1 standard deviation above average.]

8

**Excerpt (p. 942):** The attractiveness rating is then matched to the alumna’s academic record, which includes her GPA, major, SAT scores [a test most students in the U.S. take in high school], race, non-merit-based financial aid awards, and scores from a quantitative reasoning (QR) test that all first-year students are required to take. Like the SAT, the QR test is scored blindly, without observing the test taker’s appearance. Finally, we observe each student’s admission rating, as assigned by three or more application reviewers. The college uses a “holistic” approach to assign admission ratings, considering each student’s academic record (including high school GPA, SAT and other standardized test scores), extracurricular activities, recommendation letters, two essays, and, in some cases, artwork or music. [There are no photos or interviews for admission so physical attractiveness cannot directly affect the rating.]

9

# Basic Summary Statistics

**Table 1: Summary Statistics**

| | *Above Median Attractiveness Rating (Attractive)* | | | | | *Below Median Attractiveness Rating (Unattractive)* | | | | | *Entire Sample* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | Max | Obs. | Mean | SD | Min | Max | Obs. | Mean | SD |
| Standardized attractiveness rating | 0.70 | 0.54 | -0.03 | 2.42 | 397 | -0.69 | 0.47 | -2.69 | -0.03 | 397 | 0 | 1 |
| Admissions rating | 6.34 | 1.36 | 0 | 10 | 397 | 6.62 | 1.29 | 1.67 | 10 | 395 | 6.48 | 1.34 |
| cGPA | 3.48 | 0.28 | 2.5 | 3.98 | 396 | 3.48 | 0.29 | 2.3 | 4 | 396 | 3.47 | 0.31 |
| Math SAT score | 678 | 62 | 510 | 800 | 387 | 689 | 57 | 490 | 800 | 378 | 684 | 60 |
| Verbal SAT score | 696 | 61 | 490 | 800 | 387 | 712 | 59 | 450 | 800 | 378 | 704 | 61 |
| QR test score | 13.08 | 2.65 | 2 | 18 | 397 | 13.42 | 2.55 | 4.5 | 18 | 397 | 13.25 | 2.60 |

10

**Table 2: Attractiveness and Test Scores**

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | Standardized Math SAT | | Standardized Verbal SAT | | Standardized QR test | |
| *Specification:* | *(1)* | *(2)* | *(3)* | *(4)* | *(5)* | *(6)* |
| *Explanatory variables:* | | | | | | |
| Standardized attractiveness rating | -0.10<br>(0.03) | | -0.14<br>(0.03) | | -0.20<br>(0.09) | |
| Attractiveness quintile = 2 | | 0.04<br>(0.10) | | -0.14<br>(0.11) | | 0.00<br>(0.27) |
| Attractiveness quintile = 3 | | -0.12<br>(0.10) | | -0.27<br>(0.11) | | -0.20<br>(0.29) |
| Attractiveness quintile = 4 | | -0.08<br>(0.10) | | -0.30<br>(0.11) | | -0.31<br>(0.27) |
| Top attractiveness quintile | | -0.29<br>(0.10) | | -0.40<br>(0.11) | | -0.55<br>(0.27) |
| Observations | 764 | 764 | 764 | 764 | 793 | 793 |
| $R^2$ | 0.22 | 0.22 | 0.11 | 0.11 | 0.12 | 0.12 |

*Notes:* Robust standard errors in parentheses. All specifications include year-of-enrollment and race fixed effects, as well as controls for the amount of financial aid received.

For (3), is $k$ = 1?

11

**Table 3: Selection Into Subject Areas**

| | *Dependent variable is percentage of all courses that the student took that are in:* | | |
|---|---|---|---|
| | the sciences | the humanities | economics |
| *Specification:* | *(7)* | *(8)* | *(9)* |
| *Explanatory variables:* | | | |
| Standardized attractiveness rating | -1.92<br>(0.62) | -0.05<br>(0.63) | 1.59<br>(0.45) |
| Standardized Math SAT score | 4.30<br>(0.75) | -3.72<br>(0.75) | 3.23<br>(0.54) |
| Standardized Verbal SAT score | -1.86<br>(0.70) | 2.26<br>(0.71) | -1.84<br>(0.51) |
| Admission rating | 1.04<br>(0.57) | -0.25<br>(0.57) | 0.08<br>(0.41) |
| Observations | 762 | 762 | 762 |
| $R^2$ | 0.12 | 0.15 | 0.14 |

*Notes:* Robust standard errors in parentheses. All specifications include year-of-enrollment and race fixed effects, as well as controls for the amount of financial aid received.

After discussing the results in Table 3, on page 952 the authors conclude: "Thus, there is substantial beauty-based course selection."
What do they mean by "substantial"?

12

**Table 4: Attractiveness and Admissions Ratings & Attractiveness and cGPA**

| | *Dependent variable:* | | |
|---|---|---|---|
| | Admission Rating | Admission Rating | cGPA |
| *Specification:* | *(10)* | *(11)* | *(12)* |
| *Explanatory variables:* | | | |
| Standardized attractiveness rating | -0.131 | -0.038 | 0.016 |
| | (0.049) | (0.042) | (0.010) |
| Standardized Math SAT score | | 0.432 | 0.017 |
| | | (0.050) | (0.012) |
| Standardized Verbal SAT score | | 0.408 | 0.005 |
| | | (0.045) | (0.011) |
| Admission rating | | | 0.059 |
| | | | (0.009) |
| Observations | 791 | 762 | 760 |
| $R^2$ | 0.13 | 0.35 | 0.18 |

*Notes:* Robust standard errors in parentheses. All specifications include year-of-enrollment and race fixed effects, as well as controls for the amount of financial aid received.

Conclusions comparing (10) and (11)?

Are admissions ratings a *significant* factor in explaining variation in cGPA across students?

13

## "Which is the fair sex? Gender differences in altruism"

**ABSTRACT [1st sentence]:** We study gender differences in altruism by examining a modified dictator game with varying incomes and prices.

- A modified dictator game?
- Participants (students) are all in a lecture hall
- You each make eight decisions to allocate tokens between yourself and another *anonymous* person in the room: your partner
- How many points each token is worth to you and your partner varies across eight decisions
- Each point is always worth 10 cents to all players
- How many tokens will you keep? Pass to your partner?

14

**DECISION SHEET**

For each of the eight decisions, the number of tokens you choose to *Hold* plus the number you choose to *Pass* must equal the total tokens you are asked to divide.

While the points-per-token vary, remember that each *point* is always worth $0.10.

Write your token allocations in the blank spaces. Fill in all 16 blanks.

**DECISIONS:**

**1.** Divide 40 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 3 points each.

**2.** Divide 60 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 2 points each.

**3.** Divide 75 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 2 points each.

**4.** Divide 60 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 1 point each.

**5.** Divide 100 tokens: *Hold* _____ @ 1 point each, and *Pass* _____ @ 1 point each.

**6.** Divide 60 tokens: *Hold* _____ @ 2 points each, and *Pass* _____ @ 1 point each.

**7.** Divide 75 tokens: *Hold* _____ @ 2 points each, and *Pass* _____ @ 1 point each.

**8.** Divide 40 tokens: *Hold* _____ @ 3 points each, and *Pass* _____ @ 1 point each.

15

## ECO220Y (2014, 2015, 2016)

**Table 2: Mean Payoff to Other Party (Canadian $s)**

| Budget | Token endowment | Income $m$ | $p_o/p_s$ | All subjects (n=868) | Males (n=334) | Females (n=534) | $t$-stat |
|---|---|---|---|---|---|---|---|
| 1 | 40 | 4.00 | 1/3 | 4.74 | 5.39 | 4.34 | 3.71 |
| 2 | 60 | 6.00 | 1/2 | 4.83 | 5.41 | 4.47 | 3.66 |
| 3 | 75 | 7.50 | 1/2 | 5.97 | 6.56 | 5.60 | 3.03 |
| 4 | 60 | 6.00 | 1 | 2.17 | 1.98 | 2.30 | -3.17 |
| 5 | 100 | 10.00 | 1 | 3.36 | 2.98 | 3.60 | -3.80 |
| 6 | 60 | 12.00 | 2 | 1.91 | 1.64 | 2.08 | -3.80 |
| 7 | 75 | 15.00 | 2 | 2.32 | 2.00 | 2.52 | -3.63 |
| 8 | 40 | 12.00 | 3 | 1.18 | 1.01 | 1.28 | -3.14 |
| Average | | | | 3.31 | 3.37 | 3.27 | 1.17 |

16

## "Which is the fair sex? Gender differences in altruism"

**ABSTRACT:** We study gender differences in altruism by examining a modified dictator game with varying incomes and prices. Our results indicate that the question "which is the fair sex?" has a complicated answer—when altruism is expensive, women are kinder, but when it is cheap, men are more altruistic. That is, we find that the male and female "demand curves for altruism" cross, and that men are more responsive to price changes. Furthermore, men are more likely to be either perfectly selfish or perfectly selfless, whereas women tend to be "equalitarians" who prefer to share evenly.

Observational or experimental data? y variable? x variables?

Andreoni and Vesterlund (2001), *The Quarterly Journal of Economics,* https://doi.org/10.1162/003355301556419, copy on Readings page in Quercus (optional)

17

## Link: Chapters 14 & 21

- "Comparing Two Means" (Chap. 14), 3 cases:
  - Two independent samples, unequal variances
    - E.g. Money passed by males (n=334) v. females (n=534) in Budget 4 (and separately for other 7 budgets)
  - Two independent samples, equal variances
    - E.g. Assuming variance for males equals that of females
  - Paired data
    - E.g. Money passed by people (n=868) in Budget 5 vs. 4
- Special cases of regression analysis

18

# Andreoni and Vesterlund (2001)

**Table 1: Mean Payoff to Other Party (U.S. $s)**

| Budget | Token endowment | Income $m$ | $p_o/p_s$ | All subjects (n=142) | Males (n=95) | Females (n=47) | $t$-stat |
|---|---|---|---|---|---|---|---|
| 1 | 40 | 4.00 | 1/3 | 3.79 | 4.18 | 3.01 | 1.96 |
| 2 | 60 | 6.00 | 1/2 | 4.03 | 4.30 | 3.49 | 1.48 |
| 3 | 75 | 7.50 | 1/2 | 4.68 | 5.00 | 4.03 | 1.53 |
| 4 | 60 | 6.00 | 1 | 1.54 | 1.36 | 1.91 | -2.26 |
| 5 | 100 | 10.00 | 1 | 2.52 | 2.33 | 2.92 | -1.42 |
| 6 | 60 | 12.00 | 2 | 1.42 | 1.21 | 1.82 | -2.07 |
| 7 | 75 | 15.00 | 2 | 1.71 | 1.42 | 2.29 | -2.35 |
| 8 | 40 | 12.00 | 3 | 0.89 | 0.67 | 1.32 | -2.97 |
| Average | | | | 2.57 | 2.56 | 2.60 | -0.24 |

19

**Table 3: Mean Payoff to Other Party: A&V (2001) versus ECO220Y (2014, 2015 and 2016)**

| | *Males* | | | | *Females* | | | |
|---|---|---|---|---|---|---|---|---|
| Bud. | A&V Mean (s.d.) | ECO220 Mean (s.d.) | Diff. (s.e.) | P-value (2-tailed) | A&V Mean (s.d.) | ECO220 Mean (s.d.) | Diff. (s.e.) | P-value (2-tailed) |
| 1 | 4.18 (4.22) | 5.39 (4.44) | -1.21 (0.50) | 0.0157 | 3.01 (2.83) | 4.34 (3.43) | -1.33 (0.44) | 0.0036 |
| 2 | 4.30 (3.77) | 5.41 (4.10) | -1.12 (0.45) | 0.0134 | 3.49 (2.63) | 4.47 (3.01) | -0.98 (0.40) | 0.0190 |
| 3 | 5.00 (4.67) | 6.56 (5.02) | -1.56 (0.55) | 0.0052 | 4.03 (2.77) | 5.60 (3.70) | -1.56 (0.43) | 0.0006 |
| 4 | 1.36 (1.48) | 1.98 (1.58) | -0.62 (0.17) | 0.0005 | 1.91 (1.31) | 2.30 (1.20) | -0.38 (0.20) | 0.0596 |
| 5 | 2.33 (2.51) | 2.98 (2.50) | -0.65 (0.29) | 0.0272 | 2.92 (2.27) | 3.60 (2.07) | -0.68 (0.34) | 0.0537 |
| 6 | 1.21 (1.57) | 1.64 (1.75) | -0.43 (0.19) | 0.0245 | 1.82 (1.68) | 2.08 (1.54) | -0.26 (0.25) | 0.3126 |
| 7 | 1.42 (1.96) | 2.00 (2.19) | -0.57 (0.23) | 0.0154 | 2.29 (2.12) | 2.52 (1.88) | -0.23 (0.32) | 0.4687 |
| 8 | 0.67 (1.11) | 1.01 (1.28) | -0.34 (0.13) | 0.0121 | 1.32 (1.27) | 1.28 (1.20) | 0.03 (0.19) | 0.8588 |
| Obs. | 95 | 334 | - | - | 47 | 534 | - | - |

# ECO220Y (2014, 2015, 2016)

$$H_0: \mu_{M4} - \mu_{F4} = 0; H_1: \mu_{M4} - \mu_{F4} \neq 0$$

**Unequal variances (general):**

$$t = \frac{(\bar{X}_{M4} - \bar{X}_{F4}) - \Delta_0}{\sqrt{\frac{s_{M4}^2}{n_M} + \frac{s_{F4}^2}{n_F}}}$$

$$t = \frac{(1.98 - 2.30) - 0}{\sqrt{\frac{1.58^2}{334} + \frac{1.20^2}{534}}}$$

$$t = \frac{-0.32}{0.101} = -3.17$$

**Assuming equal variances:**

$$t = \frac{(\bar{X}_{M4} - \bar{X}_{F4}) - \Delta_0}{\sqrt{\frac{s_{p4}^2}{n_M} + \frac{s_{p4}^2}{n_F}}}$$

$$s_{p4}^2 = \frac{(n_M - 1)s_{M4}^2 + (n_F - 1)s_{F4}^2}{n_M + n_F - 2}$$

$$= \frac{(333)1.58^2 + (533)1.20^2}{334 + 534 - 2} = 1.846$$

$$t = \frac{-0.32}{0.095} = -3.37$$

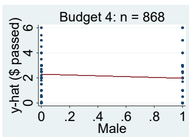
21

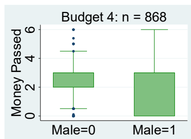
# Homoscedasticity => Equal Variances

```
. regress money_passed male if budget4==1;

      Source |       SS       df       MS              Number of obs =     868
-------------+------------------------------           F(  1,   866) =   11.34
       Model |  20.9206814     1  20.9206814           Prob > F      =  0.0008
    Residual |  1597.25572   866  1.84440615           R-squared     =  0.0129
-------------+------------------------------           Adj R-squared =  0.0118
       Total |  1618.17641   867  1.86640877           Root MSE      =  1.3581

------------------------------------------------------------------------------
money_passed |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        male |  -.3190832   .0947424    -3.37   0.001    -.5050347   -.1331316
       _cons |   2.295131   .0587703    39.05   0.000     2.179782     2.41048
------------------------------------------------------------------------------
```






# But, with Robust Standard Errors…

```
.  regress money_passed male if budget4==1, robust;

Linear regression                                      Number of obs =     868
                                                       F(  1,   866) =   10.02
                                                       Prob > F      =  0.0016
                                                       R-squared     =  0.0129
                                                       Root MSE      =  1.3581

------------------------------------------------------------------------------
             |               Robust
money_passed |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        male |  -.3190832   .1007783    -3.17   0.002    -.5168815   -.1212848
       _cons |   2.295131   .0519189    44.21   0.000      2.19323    2.397033
------------------------------------------------------------------------------
```

These standard errors are *robust* to violations of Assumption #3 (the homoscedasticity assumption).



# ECO220Y (2014, 2015, 2016)

**Table 2: Mean Payoff to Other Party**

| Budget | Token endowment | Income $m$ | $p_o/p_s$ | All subjects (n=868) | Males (n=334) | Females (n=534) | $t$-stat |
|---|---|---|---|---|---|---|---|
| 1 | 40 | 4.00 | 1/3 | 4.74 | 5.39 | 4.34 | 3.71 |
| 2 | 60 | 6.00 | 1/2 | 4.83 | 5.41 | 4.47 | 3.66 |
| 3 | 75 | 7.50 | 1/2 | 5.97 | 6.56 | 5.60 | 3.03 |
| 4 | 60 | 6.00 | 1 | 2.17 | 1.98 | 2.30 | -3.17 |
| 5 | 100 | 10.00 | 1 | 3.36 | 2.98 | 3.60 | -3.80 |
| 6 | 60 | 12.00 | 2 | 1.91 | 1.64 | 2.08 | -3.80 |
| 7 | 75 | 15.00 | 2 | 2.32 | 2.00 | 2.52 | -3.63 |
| 8 | 40 | 12.00 | 3 | 1.18 | 1.01 | 1.28 | -3.14 |
| Average | | | | 3.31 | 3.37 | 3.27 | 1.17 |

## Reshape Data: Unit of Observation is a Decision by a Student

```
. regress money_passed male, robust;

Linear regression                               Number of obs =     6944
                                                F(  1,  6942) =     1.38
                                                Prob > F      =   0.2401
                                                R-squared     =   0.0002
                                                Root MSE      =   3.1823
```

What does 6944 mean?

Why use robust standard errors in this case?

| money_passed | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| male | .0983501 | .0837065 | 1.17 | 0.240 | -.0657404 | .2624405 |
| _cons | 3.273057 | .0427419 | 76.58 | 0.000 | 3.18927 | 3.356844 |

**Inference about a comparing two population means, independent samples, unequal variances:**

This formula on your Aid Sheets for Chapter 14 has a *robust standard error* in the denominator.

$$t = \frac{(\bar{X}_M - \bar{X}_F) - \Delta_0}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}}$$

25

```
. summarize money_passed if male==1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
money_passed |      2672    3.371407    3.720474          0         15

. summarize money_passed if male==0

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
money_passed |      4272    3.273057    2.793558          0         15
```

$H_0: (\mu_M - \mu_F) = 0$

$H_1: (\mu_M - \mu_F) \neq 0$

If you use the "assuming equal variances" formula on your Aid Sheets for Chapter 14 instead, you get the regular (not robust) se.

$$t = \frac{(\bar{X}_M - \bar{X}_F) - \Delta_0}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}}} = \frac{(3.37141 - 3.27306)}{\sqrt{\frac{3.72047^2}{2672} + \frac{2.79356^2}{4272}}} = \frac{0.09835}{0.08371} = 1.17$$

*Robust* standard error

26

## Full Set of Budget Dummies

```
. regress money_passed budget1-budget3 budget5-budget8, robust;

Linear regression                               Number of obs =     6944
                                                F(  7,  6936) =   341.06
                                                Prob > F      =   0.0000
                                                R-squared     =   0.2493
                                                Root MSE      =   2.7587
```

2.17235 + 2.570968 = 4.743318, where have we seen 4.74 before?

What do the *t* tests refer to?

| money_passed | Coef. | Robust Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| budget1 | 2.570968 | .139695 | 18.40 | 0.000 | 2.297123 | 2.844813 |
| budget2 | 2.658986 | .1274211 | 20.87 | 0.000 | 2.409202 | 2.908771 |
| budget3 | 3.795392 | .1525096 | 24.89 | 0.000 | 3.496426 | 4.094357 |
| budget5 | 1.187442 | .0897047 | 13.24 | 0.000 | 1.011594 | 1.363291 |
| budget6 | -.2599078 | .0723093 | -3.59 | 0.000 | -.4016562 | -.1181594 |
| budget7 | .1479263 | .0827918 | 1.79 | 0.074 | -.014371 | .3102235 |
| budget8 | -.9923963 | .0624764 | -15.88 | 0.000 | -1.114869 | -.8699235 |
| _cons | 2.17235 | .0463707 | 46.85 | 0.000 | 2.081449 | 2.263251 |

27

```
. xi: regress money_passed budget1-budget3 budget5-budget8 i.stud, robust;
                          (_Istud_1 for stud==A010ECO220Y, Feb. 14, 2014 omitted)

Linear regression                                      Number of obs =    6944
                                                       F(874,  6069) =   12.53
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.5516
                                                       Root MSE      =  2.2792

------------------------------------------------------------------------------
             |               Robust
money_passed |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     budget1 |   2.570968   .1130991    22.73   0.000     2.349253    2.792682
     budget2 |   2.658986   .0956745    27.79   0.000      2.47143    2.846542
     budget3 |   3.795392    .118522    32.02   0.000     3.563046    4.027737
     budget5 |   1.187442    .076437    15.53   0.000     1.037599    1.337286
     budget6 |  -.2599078   .0807719    -3.22   0.001    -.4182494   -.1015663
     budget7 |   .1479263   .0857755     1.72   0.085    -.0202242    .3160768
     budget8 |  -.9923963   .0785683   -12.63   0.000    -1.146418   -.8383746
    _Istud_2 |      3.375   1.836626     1.84   0.066    -.2254395     6.97544
    _Istud_3 |      3.875   .9187467     4.22   0.000      2.07393     5.67607
         ...
  _Istud_867 |      3.125   .9163724     3.41   0.001     1.328585    4.921415
  _Istud_868 |      3.875   .8388546     4.62   0.000     2.230547    5.519453
       _cons |  -.3885513   .7588858    -0.51   0.609    -1.876237    1.099134
------------------------------------------------------------------------------
```

Add a full set of student fixed effects: i.e. a dummy for each student (except one)

Note: (868 – 1) + (8 – 1) = 874 (i.e. $k = 874$)

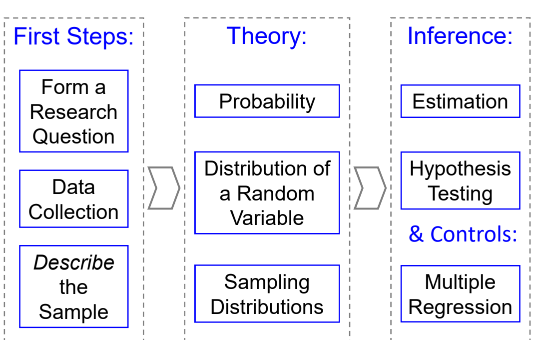
28

## Usual Purposes of Fixed Effects

- A full set of fixed effects is common with multi-dimensional observational data
  - Multiple subscripts: panel data or other types (e.g. each person makes 8 choices)
  - Idea: fixed effects can control for some lurking variables (e.g. differences across people)
    - In our experimental data, 8 budgets have *zero* correlation with individual characteristics so coefficients are unaffected by including fixed effects (but some s.e.'s do go down)
    - Observational data: budgets would differ by individual

29

## ECO220Y: Overview




"A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data" John P.A. Ioannidis, March 17, 2020 https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/

30