ECO220Y1Y, Test #4, Prof. Murdock

March 31, 2023, 9:10 – 11:00 am

U of T E-MAIL:	**@MAIL.UTORONTO.CA**
SURNAME (FAMILY NAME):	
GIVEN NAME (FIRST NAME):	
UTORID: (e.g. LIHAO118)	

Instructions:

- You have 110 minutes. Keep these test papers and the *Supplement* closed and face up on your desk until the start of the test is announced. You must stay for a minimum of 60 minutes.
- You may use a **non-programmable calculator**.
- There are <u>5 questions</u> (all with multiple parts) with varying point values worth a total of <u>95 points</u>.
- This test includes these 8 pages plus the *Supplement*. The *Supplement* contains the aid sheets and statistical tables (Standard Normal, Student t, and F) and readings, figures, tables, and other materials for test questions. For each question referencing the *Supplement*, carefully review *all* materials. ***The Supplement will <u>NOT</u> be collected:*** write your answers on these test papers. When we announce the end of the test, hand these test papers to us (you keep the *Supplement*).
- Write your answers clearly, completely, and concisely in the designated space provided immediately after each question. An <u>answer guide</u> ends each question to let you know what is expected. For example, a <u>quantitative analysis</u>, a <u>fully labelled graph</u>, and/or <u>sentences</u>. Any answer guide asking for a <u>quantitative analysis</u> *always* automatically means that you must show your work and make your reasoning clear.
    - Anything requested by the question and/or the answer guide is required. Similarly, limit yourself to the answer guide. For example, if the answer guide does not request sentences, provide <u>only</u> what is requested (e.g. quantitative analysis).
    - Marking TAs are instructed to accept all reasonable rounding.
- ***Your entire answer must fit in the designated space provided immediately after each question.*** No extra space/pages are possible. You *cannot* use blank space for other questions, nor can you write answers on the *Supplement*. ***Write in <u>PENCIL</u> and use an <u>ERASER</u> as needed*** so that you can fit your final answer (including work and reasoning) in the appropriate space. Questions give more blank space than is needed for an answer (with typical handwriting) worth full marks. ***Follow the <u>answer guides</u> and avoid excessively long answers.***

(1) See ***Supplement for Question (1)***: *How Much Energy Do Building Energy Codes Save? Evidence from California Houses.*

(a) [10 pts] In the Stata regression output in the row for `constr_50_59`, compute the one hidden value. For that hidden value, write the associated hypotheses. Finally, *interpret* the conclusion of that hypothesis test *in this context*. Answer with hypotheses in formal notation, a quantitative analysis & 1 – 2 sentences.

(b) [5 pts] In the Stata regression output in the row for `constr_05_08`, compute the two hidden values. Answer with a quantitative analysis.

(c) [6 pts] What is the predicted 2009 annual electricity usage in MMBTUs for a house that is 1,750 square feet, has 3 residents, does not have central air conditioning, is built in 1994, and is in climate zone 1? (Note: To get the final answer in MMBTUs, recall that $e^{\ln(x)} = x$.) Answer with a quantitative analysis.

(d) [5 pts] A new regression is just like the one in the *Supplement* except that the new regression does not control for the climate zone that each house is located in. Answer by filling in the five blanks.

In the new regression the value of k is ____ [numeric value] versus a k of ____ [numeric value] in the regression in the Supplement. Further, in the new regression, the R^2 is ______________ [smaller / larger / the same], the SST is ______________ [smaller / larger / the same], and the SSE is ______________ [smaller / larger / the same] compared to the regression in the Supplement.

(2) See ***Supplement for Question (2)***: *2022 World Happiness Report: Mexico from 2005 through 2021.*

(a) [5 pts] Using the given Excel output, what is the P-value to assess if the regression is statistically significant overall? Of the 4 common significance levels (10%, 5%, 1%, and 0.1%), which does it meet? Answer with hypotheses in formal notation & 1 sentence.

(b) [9 pts] Draw a figure to *effectively communicate* happiness trends in Mexico. Answer with a fully labelled graph & show your work.

(c) [8 pts] From the aid sheets, use this formula $\sqrt{\frac{SSE}{n-k-1}}$. With relevant results from the Excel output, show how to plug into that formula and then compute its numeric value. *Interpret* that numeric value in this context. Assess if it is large or small in this context. <u>Answer with a quantitative analysis & 2 – 3 sentences.</u>

(3) See ***Supplement for Question (3)***: *Entry and Exit of Informal Firms and Development: Case Study of Vietnam.*

(a) [6 pts] For **Figure 3**, define $share_{it}$ as the share of households operating a farm or non-farm business in province i and year t. Define $rpipc_{it}$ as real provincial income per capita in 2006 in millions of VND in province i. Define $d06_{it}$ as 1 for 2006 and 0 otherwise and $d18_{it}$ as 1 for 2018 and 0 otherwise. Consider the estimated coefficients b_0, b_1, b_2, and b_3 for the model: $share_{it} = \beta_0 + \beta_1 rpipc_{it} + \beta_2 d06_{it} + \beta_3 d06_{it} \times rpipc_{it} + \varepsilon_{it}$. <u>Answer by filling in the eight blanks.</u>

For each point estimate, for the second blank choose from: intercept for 2006 / intercept for 2018 / slope for 2006 / slope for 2018 / difference in intercept for 2006 versus 2018 / difference in slope for 2006 versus 2018.

b_0 is ______________ [positive / negative] and is the ____.

b_1 is ______________ [positive / negative] and is the ____.

b_2 is ______________ [positive / negative] and is the ____.

b_3 is ______________ [positive / negative] and is the ____.

(b) [7 pts] For **Figure 3** and **2018**, for the formula $\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}}$ consider $\boldsymbol{x_g} = \mathbf{20}$. Roughly, what is the value of $\hat{\boldsymbol{y}}_{\boldsymbol{x_g}}$? What is the *interpretation* of the value of $\hat{\boldsymbol{y}}_{\boldsymbol{x_g}}$? Answer with a numeric estimate & 1 – 2 sentences.

(c) [6 pts] In **Figure 5**, for the **Entry** line the OLS coefficient estimates are $b_0 = 27.31074379$ and $b_1 = -0.55966942$. *Interpret* the point estimate of the *slope.* Answer with 1 precise sentence.

(4) See ***Supplement for Question (4)***: *Why Have College Completion Rates Increased?*

(a) [8 pts] Using **Column (1)** in **Table 7**, how do average first-year GPAs compare for the year **2000** versus the year **1990**? Next, *interpret* your estimate and *a*ssess *economic significance*. Answer with a quantitative analysis & 2 – 3 sentences.

(b) [8 pts] In an alternate reality version of the U.S. from 1990 to 2000 suppose that these two things are true:

1. There is no grade inflation: every year students get the grades they deserve. For example, work submitted in 1990 that earns a grade of "C" (a 2.0 on the GPA scale) would still get exactly a "C" if submitted in 2000.
2. Students on average tend to get academically stronger each year from 1990 to 2000, which is clear from rising marks on standardized tests written in high school.

In this alternate reality, what key result should we expect to see in **Column (1)** of **Table 7**? *Explain.* In this alternate reality, what key result should we expect to see in **Column (7)** of **Table 7**? *Explain.* Answer with 2 – 3 sentences.

(5) See ***Supplement for Question (5)***: *Consumer Demand with Social Influences: Evidence from an E-Commerce Platform.*

(a) [4 pts] Consider the number of explanatory (x) variables for the regressions in Table 1. Answer by filling in the four blanks.

The number of x variables in the regression in Column (1) is ____ [numeric value], in Column (2) is ____ [numeric value], and in Column (3) is ____ [numeric value]. Compared to the regression in Column (3), in the regression in Column (4) the number of x variables is _____________ [smaller / larger / the same].

(b) [8 pts] For **Table 1**, how do the results for "Red (indicator)" differ in **Column (1)** and **Column (3)**? Make sure to both *interpret* them and to explain *why* they differ. Answer with 2 – 3 sentences.

This *Supplement* has the aid sheets and statistical tables (Standard Normal, Student t, F) and readings, figures, tables, and other materials for test questions. For each question referencing this *Supplement*, carefully review *all* materials.

Sample mean: $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ **Sample variance:** $s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{X})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2}{n-1} - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n(n-1)}$ **Sample s.d.:** $s = \sqrt{s^2}$

Sample coefficient of variation: $CV = \frac{s}{\bar{X}}$ **Sample covariance:** $s_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{X})(y_i-\bar{Y})}{n-1} = \frac{\sum_{i=1}^{n} x_i y_i}{n-1} - \frac{(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n(n-1)}$

Sample interquartile range: $IQR = Q3 - Q1$ **Sample coefficient of correlation:** $r = \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^{n} z_{x_i} z_{y_i}}{n-1}$

Addition rule: $P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$ **Conditional probability:** $P(A|B) = \frac{P(A\ and\ B)}{P(B)}$

Complement rules: $P(A^C) = P(A') = 1 - P(A)$ $P(A^C|B) = P(A'|B) = 1 - P(A|B)$

Multiplication rule: $P(A\ and\ B) = P(A|B)P(B) = P(B|A)P(A)$

Expected value: $E[X] = \mu = \sum_{all\ x} xp(x)$ **Variance:** $V[X] = E[(X-\mu)^2] = \sigma^2 = \sum_{all\ x}(x-\mu)^2 p(x)$

Covariance: $COV[X,Y] = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY} = \sum_{all\ x}\sum_{all\ y}(x-\mu_X)(y-\mu_Y)p(x,y)$

<u>Laws of expected value</u>:

$E[c] = c$

$E[X+c] = E[X] + c$

$E[cX] = cE[X]$

$E[a+bX+cY] = a + bE[X] + cE[Y]$

<u>Laws of variance</u>:

$V[c] = 0$

$V[X+c] = V[X]$

$V[cX] = c^2V[X]$

$V[a+bX+cY] = b^2V[X] + c^2V[Y] + 2bc * COV[X,Y]$

$V[a+bX+cY] = b^2V[X] + c^2V[Y] + 2bc * SD(X) * SD(Y) * \rho$

where $\rho = CORRELATION[X,Y]$

<u>Laws of covariance</u>:

$COV[X,c] = 0$

$COV[a+bX, c+dY] = bd * COV[X,Y]$

Combinatorial formula: $C_x^n = \frac{n!}{x!(n-x)!}$ **Binomial probability:** $p(x) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$ for $x = 0,1,2,\dots,n$

If X is Binomial $(X \sim B(n,p))$ **then** $E[X] = np$ **and** $V[X] = np(1-p)$

If X is Uniform $(X \sim U[a,b])$ **then** $f(x) = \frac{1}{b-a}$ **and** $E[X] = \frac{a+b}{2}$ **and** $V[X] = \frac{(b-a)^2}{12}$

<u>Sampling distribution of $\bar{X}$</u>:

$\mu_{\bar{X}} = E[\bar{X}] = \mu$

$\sigma_{\bar{X}}^2 = V[\bar{X}] = \frac{\sigma^2}{n}$

$\sigma_{\bar{X}} = SD[\bar{X}] = \frac{\sigma}{\sqrt{n}}$

<u>Sampling distribution of $\hat{P}$</u>:

$\mu_{\hat{P}} = E[\hat{P}] = p$

$\sigma_{\hat{P}}^2 = V[\hat{P}] = \frac{p(1-p)}{n}$

$\sigma_{\hat{P}} = SD[\hat{P}] = \sqrt{\frac{p(1-p)}{n}}$

<u>Sampling distribution of $(\hat{P}_2 - \hat{P}_1)$</u>:

$\mu_{\hat{P}_2-\hat{P}_1} = E[\hat{P}_2 - \hat{P}_1] = p_2 - p_1$

$\sigma_{\hat{P}_2-\hat{P}_1}^2 = V[\hat{P}_2 - \hat{P}_1] = \frac{p_2(1-p_2)}{n_2} + \frac{p_1(1-p_1)}{n_1}$

$\sigma_{\hat{P}_2-\hat{P}_1} = SD[\hat{P}_2 - \hat{P}_1] = \sqrt{\frac{p_2(1-p_2)}{n_2} + \frac{p_1(1-p_1)}{n_1}}$

<u>Sampling distribution of $(\bar{X}_1 - \bar{X}_2)$, independent samples</u>:

$\mu_{\bar{X}_1-\bar{X}_2} = E[\bar{X}_1 - \bar{X}_2] = \mu_1 - \mu_2$

$\sigma_{\bar{X}_1-\bar{X}_2}^2 = V[\bar{X}_1 - \bar{X}_2] = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

$\sigma_{\bar{X}_1-\bar{X}_2} = SD[\bar{X}_1 - \bar{X}_2] = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

<u>Sampling distribution of $(\bar{X}_d)$, paired $(d = X_1 - X_2)$</u>:

$\mu_{\bar{X}_d} = E[\bar{X}_d] = \mu_1 - \mu_2$

$\sigma_{\bar{X}_d}^2 = V[\bar{X}_d] = \frac{\sigma_d^2}{n} = \frac{\sigma_1^2 + \sigma_2^2 - 2*\rho*\sigma_1*\sigma_2}{n}$

$\sigma_{\bar{X}_d} = SD[\bar{X}_d] = \frac{\sigma_d}{\sqrt{n}} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2*\rho*\sigma_1*\sigma_2}{n}}$

Inference about a population proportion:

***z* test statistic:** $z = \frac{\hat{P}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ **CI estimator:** $\hat{P} \pm z_{\alpha/2}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$

Inference about comparing two population proportions:

***z* test statistic under Null hypothesis of no difference:** $z = \frac{\hat{P}_2-\hat{P}_1}{\sqrt{\frac{\bar{P}(1-\bar{P})}{n_1}+\frac{\bar{P}(1-\bar{P})}{n_2}}}$ **Pooled proportion:** $\bar{P} = \frac{X_1+X_2}{n_1+n_2}$

CI estimator: $(\hat{P}_2 - \hat{P}_1) \pm z_{\alpha/2}\sqrt{\frac{\hat{P}_2(1-\hat{P}_2)}{n_2} + \frac{\hat{P}_1(1-\hat{P}_1)}{n_1}}$

Inference about the population mean:

***t* test statistic:** $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$ **CI estimator:** $\bar{X} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$ **Degrees of freedom:** $\nu = n-1$

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

***t* test statistic:** $t = \frac{(\bar{X}_1-\bar{X}_2)-\Delta_0}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$ **CI estimator:** $(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Degrees of freedom: $\nu = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1}\left(\frac{s_1^2}{n_1}\right)^2+\frac{1}{n_2-1}\left(\frac{s_2^2}{n_2}\right)^2}$

Inference about a comparing two population means, independent samples, assuming equal variances:

***t* test statistic:** $t = \frac{(\bar{X}_1-\bar{X}_2)-\Delta_0}{\sqrt{\frac{s_p^2}{n_1}+\frac{s_p^2}{n_2}}}$ **CI estimator:** $(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$ **Degrees of freedom:** $\nu = n_1 + n_2 - 2$

Pooled variance: $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$

Inference about a comparing two population means, paired data: (n is number of pairs and $d = X_1 - X_2$)

***t* test statistic:** $t = \frac{\bar{d}-\Delta_0}{s_d/\sqrt{n}}$ **CI estimator:** $\bar{X}_d \pm t_{\alpha/2}\frac{s_d}{\sqrt{n}}$ **Degrees of freedom:** $\nu = n-1$

SIMPLE REGRESSION:

Model: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ **OLS line:** $\hat{y}_i = b_0 + b_1 x_i$ $b_1 = \frac{s_{xy}}{s_x^2} = r\frac{s_y}{s_x}$ $b_0 = \bar{Y} - b_1\bar{X}$

Coefficient of determination: $R^2 = (r)^2$ **Residuals:** $e_i = y_i - \hat{y}_i$

Standard deviation of residuals: $s_e = \sqrt{\frac{SSE}{n-2}} = \sqrt{\frac{\sum_{i=1}^{n}(e_i-0)^2}{n-2}}$ **Standard error of slope:** $s.e.(b_1) = s_{b_1} = \frac{s_e}{\sqrt{(n-1)s_x^2}}$

Inference about the population slope:

t **test statistic:** $t = \frac{b_1 - \beta_{10}}{s.e.(b_1)}$ **CI estimator:** $b_1 \pm t_{\alpha/2} s.e.(b_1)$ **Degrees of freedom:** $\nu = n - 2$

Standard error of slope: $s.e.(b_1) = s_{b_1} = \frac{s_e}{\sqrt{(n-1)s_x^2}}$

Prediction interval for y at given value of x (x_g):

$\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{1 + \frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}}$ **or** $\hat{y}_{x_g} \pm t_{\alpha/2} \sqrt{\left(s.e.(b_1)\right)^2 (x_g - \bar{X})^2 + \frac{s_e^2}{n} + s_e^2}$

Degrees of freedom: $\nu = n - 2$

Confidence interval for predicted mean at given value of x (x_g):

$\hat{y}_{x_g} \pm t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{(x_g - \bar{X})^2}{(n-1)s_x^2}}$ **or** $\hat{y}_{x_g} \pm t_{\alpha/2} \sqrt{\left(s.e.(b_1)\right)^2 (x_g - \bar{X})^2 + \frac{s_e^2}{n}}$ **Degrees of freedom:** $\nu = n - 2$

SIMPLE & MULTIPLE REGRESSION:

Model: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i$

$SST = \sum_{i=1}^{n} (y_i - \bar{Y})^2 = SSR + SSE \quad SSR = \sum_{i=1}^{n} (\hat{y}_i - \bar{Y})^2 \quad SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

$s_y^2 = \frac{SST}{n-1} \quad MSE = \frac{SSE}{n-k-1} \quad Root\ MSE = \sqrt{\frac{SSE}{n-k-1}} \quad MSR = \frac{SSR}{k}$

$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad Adj.R^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)} = \left(R^2 - \frac{k}{n-1}\right)\left(\frac{n-1}{n-k-1}\right)$

Residuals: $e_i = y_i - \hat{y}_i$ **Standard deviation of residuals:** $s_e = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{\frac{\sum_{i=1}^{n}(e_i - 0)^2}{n-k-1}}$

Inference about the overall statistical significance of the regression model:

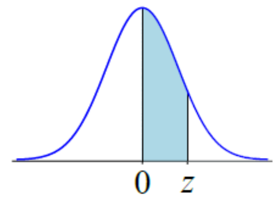
$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{(SST-SSE)/k}{SSE/(n-k-1)} = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE}$

Numerator degrees of freedom: $\nu_1 = k$ **Denominator degrees of freedom:** $\nu_2 = n - k - 1$

Inference about the population slope for explanatory variable j:

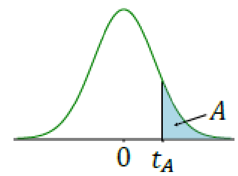
t **test statistic:** $t = \frac{b_j - \beta_{j0}}{s_{b_j}}$ **CI estimator:** $b_j \pm t_{\alpha/2} s_{b_j}$ **Degrees of freedom:** $\nu = n - k - 1$

Standard error of slope: $s.e.(b_j) = s_{b_j}$ (for multiple regression, must be obtained from technology)



The Standard Normal Distribution:

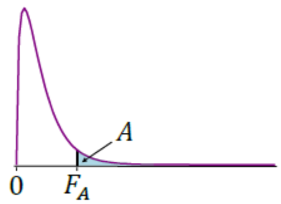
	Second decimal place in z									
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998
3.6	0.4998	0.4998	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999	0.4999



Critical Values of Student t Distribution:

ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.001}$	$t_{0.0005}$	ν	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	$t_{0.001}$	$t_{0.0005}$
1	3.078	6.314	12.71	31.82	63.66	318.3	636.6	38	1.304	1.686	2.024	2.429	2.712	3.319	3.566
2	1.886	2.920	4.303	6.965	9.925	22.33	31.60	39	1.304	1.685	2.023	2.426	2.708	3.313	3.558
3	1.638	2.353	3.182	4.541	5.841	10.21	12.92	40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610	41	1.303	1.683	2.020	2.421	2.701	3.301	3.544
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869	42	1.302	1.682	2.018	2.418	2.698	3.296	3.538
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959	43	1.302	1.681	2.017	2.416	2.695	3.291	3.532
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408	44	1.301	1.680	2.015	2.414	2.692	3.286	3.526
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041	45	1.301	1.679	2.014	2.412	2.690	3.281	3.520
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781	46	1.300	1.679	2.013	2.410	2.687	3.277	3.515
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587	47	1.300	1.678	2.012	2.408	2.685	3.273	3.510
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	48	1.299	1.677	2.011	2.407	2.682	3.269	3.505
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318	49	1.299	1.677	2.010	2.405	2.680	3.265	3.500
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221	50	1.299	1.676	2.009	2.403	2.678	3.261	3.496
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140	51	1.298	1.675	2.008	2.402	2.676	3.258	3.492
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073	52	1.298	1.675	2.007	2.400	2.674	3.255	3.488
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015	53	1.298	1.674	2.006	2.399	2.672	3.251	3.484
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965	54	1.297	1.674	2.005	2.397	2.670	3.248	3.480
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922	55	1.297	1.673	2.004	2.396	2.668	3.245	3.476
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883	60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850	65	1.295	1.669	1.997	2.385	2.654	3.220	3.447
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819	70	1.294	1.667	1.994	2.381	2.648	3.211	3.435
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792	75	1.293	1.665	1.992	2.377	2.643	3.202	3.425
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768	80	1.292	1.664	1.990	2.374	2.639	3.195	3.416
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745	90	1.291	1.662	1.987	2.368	2.632	3.183	3.402
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725	100	1.290	1.660	1.984	2.364	2.626	3.174	3.390
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707	120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690	140	1.288	1.656	1.977	2.353	2.611	3.149	3.361
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674	160	1.287	1.654	1.975	2.350	2.607	3.142	3.352
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659	180	1.286	1.653	1.973	2.347	2.603	3.136	3.345
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646	200	1.286	1.653	1.972	2.345	2.601	3.131	3.340
31	1.309	1.696	2.040	2.453	2.744	3.375	3.633	250	1.285	1.651	1.969	2.341	2.596	3.123	3.330
32	1.309	1.694	2.037	2.449	2.738	3.365	3.622	300	1.284	1.650	1.968	2.339	2.592	3.118	3.323
33	1.308	1.692	2.035	2.445	2.733	3.356	3.611	400	1.284	1.649	1.966	2.336	2.588	3.111	3.315
34	1.307	1.691	2.032	2.441	2.728	3.348	3.601	500	1.283	1.648	1.965	2.334	2.586	3.107	3.310
35	1.306	1.690	2.030	2.438	2.724	3.340	3.591	750	1.283	1.647	1.963	2.331	2.582	3.101	3.304
36	1.306	1.688	2.028	2.434	2.719	3.333	3.582	1000	1.282	1.646	1.962	2.330	2.581	3.098	3.300
37	1.305	1.687	2.026	2.431	2.715	3.326	3.574	∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Degrees of freedom: ν



The F Distribution:

ν_1	1	2	3	4	5	6	7	8	9	10	11	12	15	20	30	∞
ν_2	**Critical Values of F Distribution for $A = 0.10$:**															
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.28	3.27	3.24	3.21	3.17	3.10
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.30	2.28	2.24	2.20	2.16	2.06
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.04	2.02	1.97	1.92	1.87	1.76
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.91	1.89	1.84	1.79	1.74	1.61
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77	1.72	1.67	1.61	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.74	1.71	1.66	1.61	1.54	1.38
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66	1.60	1.54	1.48	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.63	1.60	1.55	1.48	1.41	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.57	1.55	1.49	1.42	1.34	1.00
ν_2	**Critical Values of F Distribution for $A = 0.05$:**															
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.62	4.56	4.50	4.36
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.85	2.77	2.70	2.54
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.40	2.33	2.25	2.07
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.20	2.12	2.04	1.84
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.01	1.93	1.84	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.92	1.84	1.74	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.84	1.75	1.65	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.87	1.83	1.75	1.66	1.55	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75	1.67	1.57	1.46	1.00
ν_2	**Critical Values of F Distribution for $A = 0.01$:**															
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.96	9.89	9.72	9.55	9.38	9.02
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.77	4.71	4.56	4.41	4.25	3.91
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.52	3.37	3.21	2.87
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.29	3.23	3.09	2.94	2.78	2.42
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84	2.70	2.55	2.39	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66	2.52	2.37	2.20	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.35	2.20	2.03	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34	2.19	2.03	1.86	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18	2.04	1.88	1.70	1.00
ν_2	**Critical Values of F Distribution for $A = 0.001$:**															
5	47.2	37.1	33.2	31.1	29.8	28.8	28.2	27.6	27.2	26.9	26.6	26.4	25.9	25.4	24.9	23.8
10	21.0	14.9	12.6	11.3	10.5	9.93	9.52	9.20	8.96	8.75	8.59	8.45	8.13	7.80	7.47	6.76
15	16.6	11.3	9.34	8.25	7.57	7.09	6.74	6.47	6.26	6.08	5.94	5.81	5.54	5.25	4.95	4.31
20	14.8	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24	5.08	4.94	4.82	4.56	4.29	4.00	3.38
30	13.3	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39	4.24	4.11	4.00	3.75	3.49	3.22	2.59
40	12.6	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02	3.87	3.75	3.64	3.40	3.14	2.87	2.23
60	12.0	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69	3.54	3.42	3.32	3.08	2.83	2.55	1.89
120	11.4	7.32	5.78	4.95	4.42	4.04	3.77	3.55	3.38	3.24	3.12	3.02	2.78	2.53	2.26	1.54
∞	10.83	6.91	5.42	4.62	4.10	3.74	3.47	3.27	3.10	2.96	2.84	2.74	2.51	2.27	1.99	1.00

Numerator degrees of freedom: ν_1; Denominator degrees of freedom: ν_2

Supplement for Question (1): Recall Levinson (2016) "How Much Energy Do Building Energy Codes Save? Evidence from California Houses." The Stata regression below uses the 2009 RASS data. The reference category for construction year is pre-1940. The omitted category for climate zone is climate zone 1. Some values are intentionally hidden.

```
      Source |       SS           df       MS      Number of obs   =     6,844
-------------+----------------------------------   F(23, 6820)     =    149.48
       Model |  602.975351        23  26.2163196   Prob > F        =    0.0000
    Residual |  1196.11028     6,820  .17538274    R-squared       =    0.3352
-------------+----------------------------------   Adj R-squared   =    0.3329
       Total |  1799.08564     6,843  .262908905   Root MSE        =    .41879

------------------------------------------------------------------------------
 ln_elec_mmbtu |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
    ln_sq_feet |   .4137596   .0143972    28.74   0.000     .3855367    .4419826
   ln_num_res |   .2681087   .0098682    27.17   0.000      .248764    .2874534
   central_ac |   .2460722   .0117783    20.89   0.000     .2229832    .2691613
  constr_40_49 |   .0238944   .0283619     0.84   0.400    -.0317037    .0794925
  constr_50_59 |    .040588    .022633     1.79   [redacted]    -.0037797    .0849558
  constr_60_69 |   .0683292   .0235543     2.90   0.004     .0221555    .1145029
  constr_70_74 |   .1064468   .0273128     3.90   0.000     .0529052    .1599884
  constr_75_77 |   .0934311   .0304441     3.07   0.002     .0337512     .153111
  constr_78_82 |   .1017169   .0271649     3.74   0.000     .0484653    .1549686
  constr_83_92 |   .0713506   .0243986     2.92   0.003     .0235218    .1191794
  constr_93_97 |    .054859   .0300657     1.82   0.068    -.0040791    .1137971
  constr_98_00 |   -.008971   .0334599    -0.27   0.789    -.0745629    .0566209
  constr_01_04 |  -.0368388   .0300877    -1.22   0.221    -.0958201    .0221425
  constr_05_08 |  -.1454542   .0330391    -4.40   0.000    [redacted]    [redacted]
climate_zone_2 |   .0625181   .0414352     1.51   0.131    -.0187078     .143744
climate_zone_3 |   .1656315   .0371343     4.46   0.000     .0928368    .2384262
climate_zone_4 |  -.0550539   .0339146    -1.62   0.105    -.1215371    .0114293
climate_zone_5 |  -.1345502   .0358016    -3.76   0.000    -.2047324   -.0643679
climate_zone_7 |     .07691   .0385818     1.99   0.046     .0012776    .1525424
climate_zone_8 |  -.0921371   .0337584    -2.73   0.006    -.1583141   -.0259601
climate_zone_9 |  -.0505617   .0346472    -1.46   0.145    -.1184811    .0173577
climate_zone_10 |   .0328076   .0338311     0.97   0.332    -.0335119    .0991272
climate_zone_13 |  -.1167857   .0329612    -3.54   0.000    -.1813999   -.0521715
         _cons |  -.3212621   .1087049    -2.96   0.003    -.5343576   -.1081666
------------------------------------------------------------------------------
```

Supplement for Question (2): The Excel regression below uses the *2022 World Happiness Report* data for Mexico from 2005 through 2021: there are 16 observations because 2006 is missing. The variable trend is 1 for 2005, 3 for 2007, 4 for 2008, and so on. The variable trend_sq is trend squared. The dependent variable cantril is the national mean reply each year of over 1,000 Mexicans to the happiness question where 0 is the worst possible life and 10 is the best possible life.

R Squared	0.627003529
Observations	16

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	1.592746063	0.796373031	10.92643832	0.001644677
Residual	13	0.947504494	0.072884961		
Total	15	2.540250557			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	6.345500925	0.25460366	24.92305458	2.32736E-12	5.795463157	6.895538693
trend	0.16692779	0.061495131	2.714487899	0.017698082	0.034075636	0.299779944
trend_sq	-0.011383772	0.003228906	-3.525582205	0.003726731	-0.018359398	-0.004408146

Supplement for Question (3): Recall McCaig and Pavcnik (2021) "Entry and Exit of Informal Firms and Development." They use extensive data from Vietnam. Vietnam has 58 provinces, and the currency of Vietnam is the dong (VND).

Excerpt: In **Figure 3** below, we find that the prevalence of operating any business, farm or non-farm, declines with development across provinces.

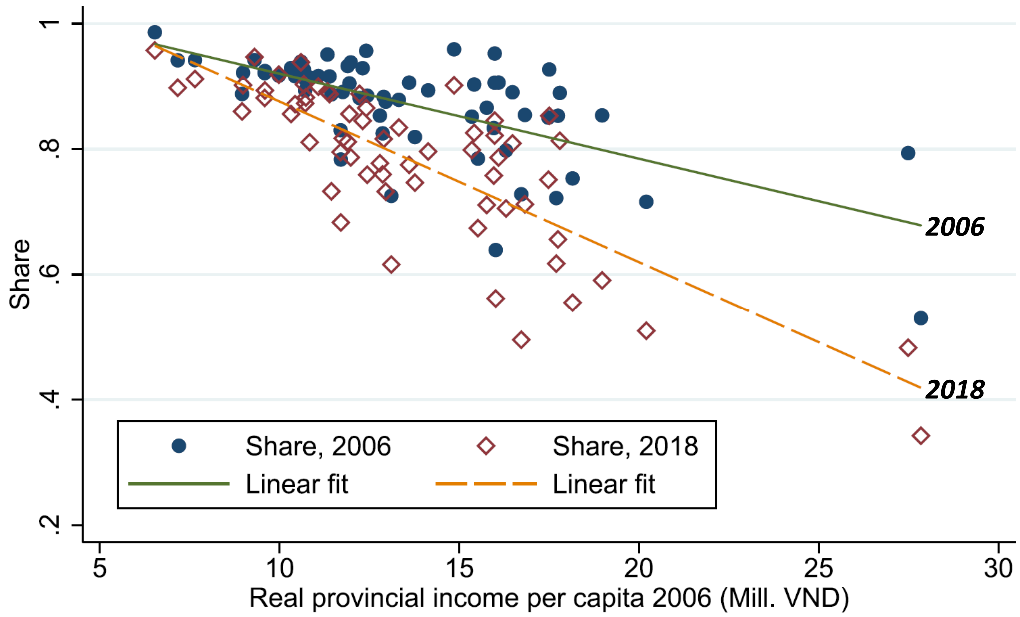


Figure 3. Share of households operating a farm or non-farm business

For the next part, note that annual entry rates are the fraction of businesses that are new (i.e. were not operating the previous year). The annual exit rate is the fraction that close down (i.e. were operating this year but not the next year).

Excerpt: In **Figure 5** below, we plot the annual exit and entry rates of informal businesses in a province relative to income per capita in 2006 across provinces. On average, more developed provinces experienced lower rates of entry and exit.

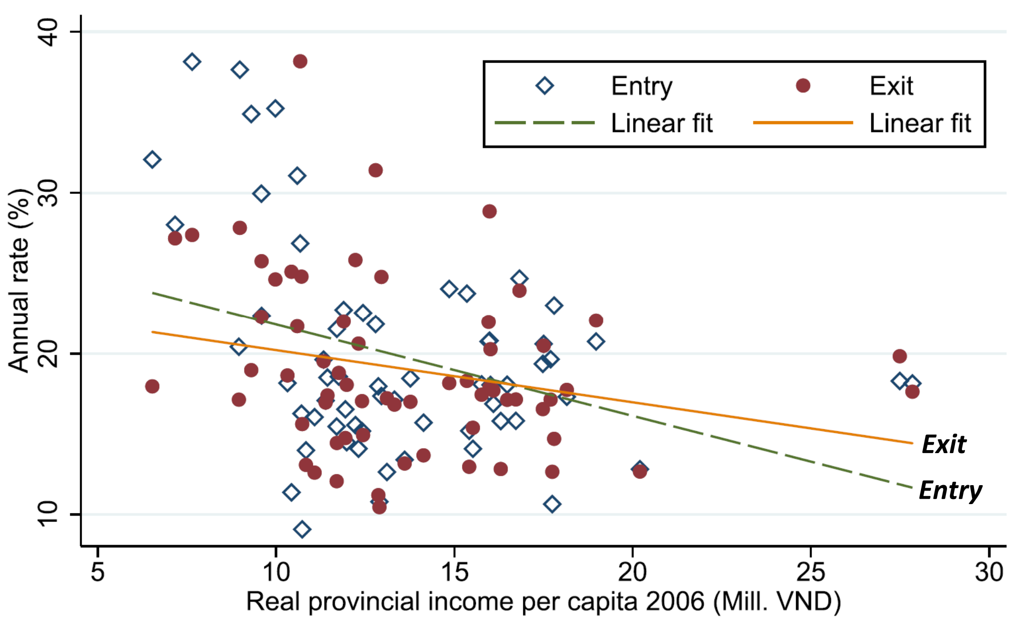


Figure 5. Annual informal entry and exit rates by province, 2006-08 panel

Supplement for Question (4): Recall Denning et al. (2022) "Why Have College Completion Rates Increased?" Recall that in the United States zip codes are like postal codes in Canada and are unique identifiers of small neighborhood areas. Also, the SAT is a standardized test that high school students write. Universities carefully consider SAT scores to make admission decisions because they are viewed as a predictor of academic success in university.

Excerpt, Abstract: We document that college completion rates have increased since the 1990s, after declining in the 1970s and 1980s. We find that most of the increase in graduation rates can be explained by grade inflation and that other factors, such as changing student characteristics and institutional resources, play little or no role. This is because GPA strongly predicts graduation, and GPAs have been rising since the 1990s. This finding holds in national survey data and in records from nine large public universities. We also find that at a public liberal arts college grades increased, holding performance on identical exams fixed.

Excerpt, p. 20: In Table 7 we examine the time trend in first-year GPA. Column 1 shows a statistically significant increase of 0.019 per year in first-year GPA between 1990 and 2000. Controlling for demographic characteristics, school attended, and home zip code leaves the coefficient unchanged. Including very flexible controls for SAT scores reduces the coefficient on year of entry only slightly, to 0.014. We also include fixed effects for major by institution to account for the potential of changing major composition. Last, we include fixed effects for all first-semester courses, and the coefficient is unchanged. We include these fixed effects to account for shifts in student course-taking that may explain changes in GPA but find that courses taken cannot explain the change in GPA. This evidence shows that rising grades cannot be meaningfully explained by demographics, preparation, courses, major, or school type. Put another way, equally prepared students in later cohorts from the same zip code, of the same gender and race, with the same initial courses, the same major, and at the same institution have higher first-year GPAs than earlier cohorts.

Table 7: Changes in GPA over Time, Large Public Universities

	Dependent variable: First-year GPA (4-point scale)						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Year of entry	0.019 (0.000)	0.019 (0.000)	0.018 (0.000)	0.018 (0.000)	0.014 (0.000)	0.015 (0.000)	0.014 (0.001)
Student characteristics		X	X	X	X	X	X
Home zip code fixed effects			X	X	X	X	X
University fixed effects				X	X	X	X
SAT math and verbal fixed effects					X	X	X
Major fixed effects						X	X
First-semester course fixed effects							X
Observations	411,951	411,951	411,951	411,951	411,951	411,951	411,951

Notes: This table reports the time trend in first-year GPA from 1990 to 2000 at Clemson, Colorado, Colorado State, Florida, Florida State, Georgia Tech, North Carolina State, Purdue, and Virginia Tech. "Student characteristics" include age and indicators for race and ethnicity, gender, transfer student, and US citizenship. "Home zip code fixed effects" include dummies for each zip code and a catch-all category for students without a reported US zip code. Indicators for each SAT math and SAT verbal score are interacted with indicators for the university. Institution-specific major fixed effects (defined at the end of the student's first year of college) and course fixed effects for every course in the student's first semester at the university are included in the final two specifications. Each observation is a unique student.

Supplement for Question (5): Recall the 2022 paper "Consumer Demand with Social Influences: Evidence from an E-Commerce Platform." See the excerpts and Table 1 below. Also, read a brief explanation of how quantity is measured.

Excerpt, Abstract: For some kinds of goods, rarity itself is valued. "Fashionable" goods are demanded in part because they are unique.

Excerpt, p. 2: We exploit a sale of outdoor slippers, or "slides," shown to the right. The product was offered in two colors – red and black – and in nine different adult sizes, creating nine distinct markets. The slides were sold directly to customers in an auction with a $50 reserve price. Total quantities for each color and size combination were announced ex ante [before the auctions began], and the auctions for each color-size product were run independently. Customers could submit at most one bid for each color-size combination. The red slide was considerably rarer than the black slide. The rareness of the red slide was not due to any anticipated difference in demand or difference in production cost for red versus black, which both cost $30 to manufacture. The product had never been sold before making it unlikely that inventory levels by shoe size and color were optimized.

The quantity of slides offered for each color and size combination is known. For example, there were 100 pairs of black slides in size 10 offered. For another example, there were 50 pairs of red slides in size 10 offered.

Note that Equation (1) below uses the common shorthand for a suite of dummies for more than two categories (like the size of the slide).

Excerpt, p. 6: To summarize the relationship between rarity and bidding behavior, we run regressions of the form:

$$\ln(\,b_i) = \alpha + \delta RED_i + \beta \ln(Q_i) + \gamma_{s_i} + \varepsilon_i, \tag{1}$$

where i denotes a bid. $\ln\,(b_i)$ is the natural log of the dollar amount a person bid for a pair of slides in the auction. RED_i is a dummy variable for the red slides. We also control for shoe size by including fixed effects for three size categories: small (5-7), medium (8-10), large (11-13). Results are presented in Table 1. Specifications (1) to (4) show the results based on Ordinary Least Squares (OLS) regressions. Column (1) only contains a constant and the dummy variable for red slides. On average, red slides receive bids that are 3.9% higher than black slides. Column (2) replaces the red dummy with the log of the inventory available for the shoe size-color combination corresponding to the bid. The estimated elasticity of bids to inventory levels is 4.8: doubling inventory levels reduces bids by 4.8% on average. This elasticity remains similar when we control for color (Column (3)) and fixed effects for the small, medium, and large shoe sizes (Column (4)).

Table 1: Effect of Rarity on Bids

	Dependent variable: Bid (log)			
	(1)	**(2)**	**(3)**	**(4)**
Constant	4.577 (0.008)	4.785 (0.047)	4.744 (0.068)	
Red (indicator)	0.039 (0.011)		0.013 (0.016)	0.011 (0.018)
Quantity (log)		-0.048 (0.012)	-0.039 (0.016)	-0.042 (0.022)
Shoe size FE	No	No	No	Yes
Observations	6,467	6,467	6,467	6,467

Notes: Standard errors are in parenthesis. Shoe size FE (fixed effects) are dummies for three size categories: small (5-7), medium (8-10), large (11-13).