Negative Examples in Lecture Improve Student Learning[*]

Austin Boyle and William L. Goffe[†]
Department of Economics
Pennsylvania State University

July 21, 2025

Abstract

This paper explores the impact of teaching three commonly misunderstood concepts in a macroeconomics principles class with "negative examples." Utilizing a crossover design and a dataset of 1,229 students, this paper finds that using negative examples improves student learning by approximately 21 percentage points over not using negative examples.

Version History

1. Presented at the Conference on Teaching and Research in Economic Education (CTREE) in Denver, May 2025
2. Literature review expanded and presented at JET SET in St. Louis, July 2025.

1 Introduction

Explaining definitions, concepts, and procedures is a core element of teaching. One might be teaching a large principles class with active learning, leading a small senior seminar, or talking with a student during office hours, but in every case the instructor will certainly be explaining. Economists might not give much thought to how they craft their explanations, but others have, and published considerable research. This paper explores one aspect of explanations – "negative examples" and tests their usefulness in a macroeconomics principles class.

This paper is organized as follows. Research on explanations is described in the following section, and next is how negative explanations were utilized in a macroeconomics principles class. Afterward, the experimental design is described, and then the results are explained, followed by a conclusion.

2 Research on Explanations by Instructors

Walstad and Allgood (1999) found that economics courses appear to not be imparting much knowledge to the students who take them. College seniors who had an economics course performed little better on an economics assessment they developed (mean of 62% of questions answered correctly) than those who did not take an economics course (mean of 48% of questions answered correctly). They also note that the mean score on the economics portion of the "Major Field Test in Business II (MFTB)," which was given to some 12,000 graduating business majors, who presumably had taken several economics courses, was 41%. There are many possible reasons for these results, and perhaps these might include how students are

[*]This study was approved by the Penn State Office for Research Protections, STUDY00013162. We would like to thank Laura Cruz for her expert help with navigating the IRB process and to Carlos Cortinhas for his thoughtful comments at the 2025 Conference on Teaching and Research in Economic Education (CTREE).

[†]aboyle@psu.edu, bill.goffe@psu.edu.

instructed. This paper explores one part of economics instruction–how economics concepts are explained to students by instructors.

Engelmann and Carnine (1982) is an influential and formal treatment of the use of examples to explain concepts to students.[1] It is the foundational text for "Direct Instruction" (or DI; the words are capitalized, and it is distinct from "direct instruction," which is described below (Alessandri et al., 2025, p.1)). They describe in great detail how one might "faultlessly" explain concepts to students with positive and negative examples.[2] Their definition of a negative example (sometimes "non-example") is straightforward: "A new example is positive if it has the same quality(ies) possessed by all the examples presented earlier. It is a negative example if it does not have the same quality(ies)." (Engelmann and Carnine, 1982, p.4) They later point out (p. 37) that "Fact 1. It is impossible to teach a concept through the presentation of one example." To illustrate this point, the text has a picture of a pencil, and the authors state it is an example of a "glerm." This term might mean a straight object, something made of wood, an item that is horizontal as in the picture, etc. They go on to state, "Fact 2. it is impossible to present a group of positive examples that communicates only one interpretation." but "... a set of positive examples is always capable of generating more than one possible explanation. Therefore the sequence of examples must contain negatives as well." The authors seem to have in mind something like the following: even if a teacher gave a very young student (or English language learner) innumerable examples of a dog, it is difficult to imagine that the student would realize that a cat is not a dog unless they were given the negative example of a cat.

Apparently the only published paper that explores the efficacy of negative examples and Direct Instruction in college teaching is Alessandri et al. (2025). Their study consists of 30 students in a lab (not a classroom) who were taught about the psychological concept of reinforcement. They conclude their paper with "A practical implication of our results suggests that only providing a definition and an example of a concept is not sufficient for learning. For efficient teaching of concept *[sic]*, it is necessary to provide both multiple examples and corresponding nonexamples."

While the above take a formal or empirical approach, the idea of negative examples has a long and deep history in teaching:

> One of the mad things about examples/non-examples is that it's almost a unified theory of learning in the sense that so many theorists from a range of different traditions have advocated for it in one form or another. Socrates, Aristotle, Vygotsky, Bruner, Skinner, Ausubel, Sweller all basically say the same thing on this which is that learning is driven by clear distinctions—knowing what something is requires knowing what it isn't. (Hendrick, 2025)

Further, while Direct Instruction has been influential, the above-mentioned direct instruction, or "explicit instruction" has been even more influential, and it too emphasizes crystal clear instruction (and thus implicitly, negative examples). It has more of a basis in cognitive science than does Direct Instruction and it goes beyond using examples. Kirschner et al. (2006) and Rosenshine (2012) are two papers that describe directinstruction.[3] The former argues that classroom instruction should provide "information that fully explains the concepts and procedures that students are required to learn as well as learning strategy support that is compatible with human cognitive architecture."[4] They are writing in opposition to "discovery-based learning," where students ferret out principles on their own.[5] They come to this

[1]These authors take a very formal approach to explanations in teaching, and as a result the book is actually rather dense. Needham (2019) is an accessible summary and has a strong focus on negative examples—in its 17 pages, negative examples are mentioned 24 times.

[2]Their examples are often topics seen in early elementary school, but their logic certainly holds for all students.

[3]Note there are more than 11,000 citations to the former in Google Scholar as of July 2025.

[4]Thus, in this meaning, "direct instruction" is distinct from a typical lecture, though sometimes those who lecture refer to lecture as direct instruction; they seem not to be familiar with the work of Kirschner et al. (2006) and Rosenshine (2012). Proponents of direct instruction would argue that it is much more than the typical lecture.

[5]It is worth noting that their focus is largely on K-12 based teaching, but as it is based on how humans learn, it presumably is applicable to college instruction.

conclusion, in part, by noting that humans have a very limited working memory (Miller, 1956) and that classroom instruction can easily saturate it, which renders learning much more difficult. Finally, it appears that direct instruction and negative examples were introduced to economists by Hamid (2022, p. 71).[6]

"Contrasting cases" are similar to negative examples in that they focus on the types of examples that instructors present to their students. Schwartz et al. (2016, pp. 26-38) note that they are "close examples that help people discern what make each instance distinctive." They also note that "Experts develop their precision by comparing many examples over the years. Contrasting cases shorten the time to learn by using carefully juxtaposed examples." While contrasting cases are broader than negative examples, the authors do point out that "learning what a thing is also depends on learning what it is not."

Mathematics educators sometimes use "Variation Theory," (Kullberg et al., 2017). At least at a surface level, it seems little different than contrasting cases. They point out:

> For instance, to understand the concept of a linear function y = mx + b one needs to know how it differs from non-linear functions. Otherwise it is merely a synonym for 'function'. Similarly, a triangle must be compared to a circle or any other shape to have a meaning of its own.

They go on to point out that after making contrasts between concepts, the instructor must generalize to deepen students' understanding.

Generalizing from negative examples, contrasting cases, and variation theory, cognitive scientists have argued for nearly a century, starting with Bartlett (1932), that humans generally remember facts, concepts, and procedures not in isolation, but in terms of other things they know. The collection of related items in our memories are referred to as "schemas" (Anderson, 1977). It is much easier to remember a new topic that connects to one's existing schemas (e.g. a policy change by the Fed as heard by a monetary economist), and much harder when there is no connection (e.g. anyone trying to recall a password consisting of arbitrary letters and numbers). Thus, student memory is enhanced when they see a wide variety of examples that illustrate a concept, be they framed as negative or positive examples, contrasting cases, or variations. All should enhance students' schemas. Willingham (2021, p. 95) phrases it this way (with an additional point): "We understand new things in the context of things we already know, and most of what we know is concrete."

3 Implementing Negative Examples in a Macro Principles Class

This study used students from the second author's macroeconomics principles classes in the spring and fall of 2024 (three sections in each semester). As described in Boyle and Goffe (2018), clickers are used extensively by this author to poll students. Students in that paper used dedicated "hardware" clickers, while in this study iClicker Cloud was used. It can be installed on phones, tablets, and computers. One useful feature for researchers is that iClicker Cloud records each students' individual answer for a given polling question. This feature is used extensively in this paper.

To explore teaching with negative examples, three concepts that principles students find difficult were taught with and without negative examples: capital, technology, and money. These were selected based on the second author's experience as an instructor, and their difficulty was confirmed in this study (described below).

To test the usefulness of negative examples, some sections were taught with a lecture with only positive examples of a concept, while the others were taught with both positive and negative examples. As detailed below, a crossover design was used with the three sections for the three concepts. The design is easiest to implement with three teaching approaches, so there were two approaches to positive and negative examples: straight lecture and active learning.

[6] This is where the authors first came across negative examples.

The following images of PowerPoint slides used in class[7] for capital illustrate the three teaching approaches. First, all sections were introduced to the concept, illustrated for capital with Figure 1. Then, each of the three separate sections were taught with one of the following: positive examples (P), Figure 2; positive and negative examples (P/N), Figure 3; or positive and negative examples with a clicker question (P/N/C), Figure 4. Ali (1981) offers empirically derived suggestions on how to present positive and negative examples. First, it is best to pair positive and negative examples, but in this case, there were no obvious pairs, so the positive ones were placed first (as suggested). Second, various suggestions are given on sequencing the examples, such as their difficulty, concreteness, or how personal they might be, but these concepts did not map well into the examples used here. Third, they do suggest more positive than negative examples. However, the negative examples used here were chosen to represent common misconononceptions. Finally, all sections saw Figure 5 at the next class meeting, typically five days later, to assess their understanding with a clicker poll. The other two concepts, technology and money, were taught similarly.

Figure 1: Introduction to capital (all sections).

Figure 2: Positive examples for capital (P).

[7] While the entire slide is shown here, in class PowerPoint's animation feature was used to sequence the parts of the slides. First was the title of the slide, then the definition, then the first example and its picture, and then the second example and its picture. This approach, which helps students focus on the point at hand, was used for all class slides and is consistent with Mayer (2002).

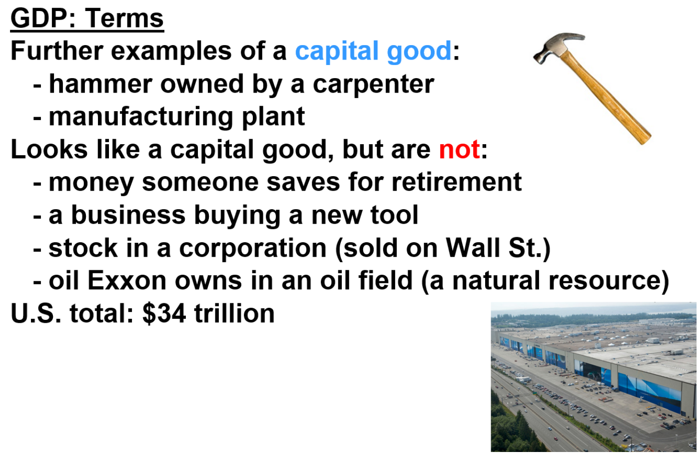

Figure 3: Positive and negative examples for capital (P/N).

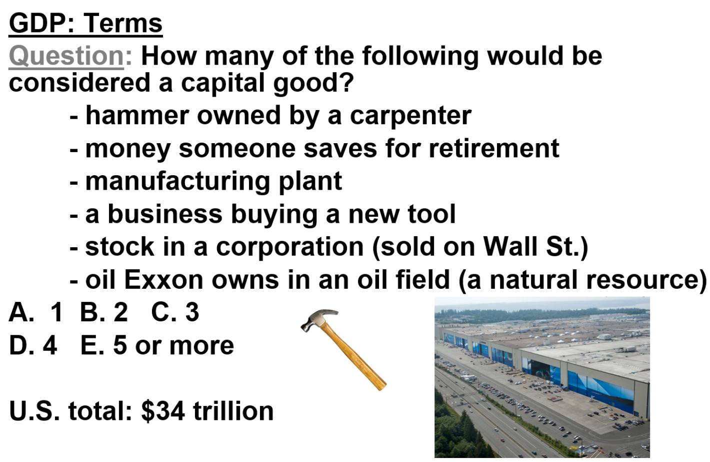

Figure 4: Positive and negative examples for capital with a clicker question (N/P/C).

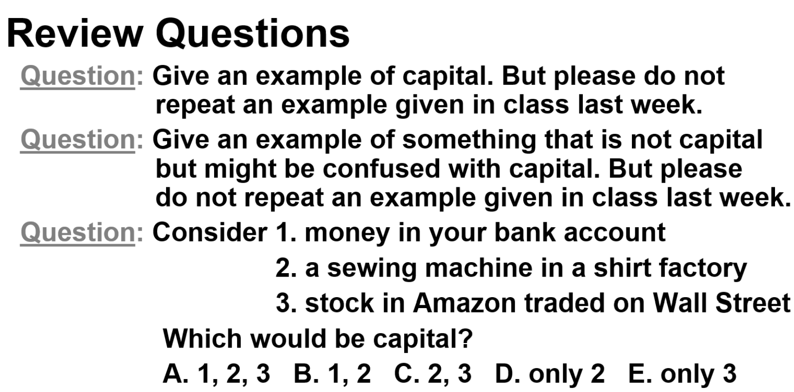

Figure 5: Assessment of knowledge about capital (5 days later).

4 Experimental Design

Many teaching studies are done comparing results in one section to those in another; these include Mikek (2023), Settlage and Wollscheid (2019), and Eisenkopf and Sulser (2016). However, this approach assumes that there are no differences between sections. Our results shown in the Appendix 7.7 indicate that the sections in this study indeed had differing abilities. This study employs a crossover design in order to control for the differences between sections. This design also controls for differences in difficulty between the three concepts. Each semester, three concepts were used to test three methods of instruction to see their effect on correctly answering a question on that concept. Instruction occurred on a Thursday (of a Tuesday-Thursday class) and the question used for assessment of each approach took place the following Tuesday[8], presumably with minimal studying by students in between. Thus, studying by students is unlikely to confound the results of the different teaching approaches, which addresses a point raised by Allgood (2001). He suggests that students with grade targets might study less when classroom instruction imparts more knowledge.

The crossover implementation is shown in Table 1 with the numbers of the section of the course preceded by an S for spring or F for fall semester.

Table 1: Crossover Design

Concept\Treatment	P	P/N	P/N/C
Capital	S6, F2	S7, F5	S4, F6
Money	S7, F5	S4, F6	S6, F2
Technology	S4, F6	S6, F2	S7, F5

Table 1 shows the rotation of treatments such that each section received all three treatments and each of the three concepts also received all three treatments.

To assess the impact of the teaching method, student responses were collected on the Tuesday assessment only for students who were in class on the preceding Thursday when the concept was introduced. Clicker data were used to determine presence in class. For the P and P/N treatments, students needed to answer at least one clicker question for the day the concept was introduced, at day's class to be counted as present. For the P/N/C treatment, students needed to answer the clicker question specifically about the concept being studied in order to count as being present. For the responses to Tuesday's assessment, only responses from students who were present on the preceding Thursday are used in the analysis.

5 Results

The main result of this research is that adding negative examples (with or without an accompanying clicker question) increased the rate of correct responses on the corresponding assessment asked five days later. The magnitude of this effect is about 21 percentage points improvement comparing classes that received the negative examples to those that did not after controlling for concept and section.[9] This finding is robust to multiple methods of analysis including examination of summary statistics and the following regression models:

Model 1: panel random effects model controlling for section
Model 2: panel student-level fixed effects

[8]However, capital was tested a full week after instruction for the spring of 2024 as classes were canceled on Tuesday due to snow.

[9]Sometimes the section is controlled for indirectly, such as when student-level fixed effects are used. No students switched sections during the semester.

Model 3: pooled cross-sectional OLS
Model 4: panel probit
Model 5: pooled cross-sectional probit
Model 6: pooled cross-sectional logit

The dependent variable in every model is whether or not students correctly answered the clicker assessment on Tuesday. There were 1,229 students who fit the criteria of being present and then answering the assessment the following class period. The minimum number of assessments answered by a student is 1, the maximum is 3, and the average is 2.5, giving a total of 3,043 observations used in this analysis.

Some questions were more difficult than others. For example, only 49% of students correctly answered the technology question in spring 2024, but 81% correctly answered the capital question. Some sections had lower accuracy than others. For example, only 59% of clicker questions across these 3 concepts were correctly answered by section S4 while section F2 answered 74% correctly. More differences between sections and concepts can be seen in appendix 7.7.

The coefficients on the P/N and P/N/C variables in models 1-3 show how much more likely a student is to correctly answer a assessment in one of these treatments compared to receiving only positive examples. Results of these linear models are shown in Table 2.[10] These results show that there is roughly a 23-25% increase in correctly answering a question (95% confidence intervals between 17% - 28%) when negative examples are added. The addition of a clicker question did not seem to help at the margin. The difference between P/N and P/N/C was not statistically significant in any model.

Table 2: Regression coefficients for negative examples with and without a clicker question

	Model 1	Model 2	Model 3
P/N	0.2334	.2326	.2339
	(12.52)	(11.86)	(12.16)
P/N/C	.2502	.2350	.2523
	(13.26)	(11.86)	(12.96)
R^2 overall	0.1199	0.1084	0.1173

t-scores in parentheses

6 Conclusion

This paper provides evidence that negative examples aid student understanding of difficult concepts in a macroeconomics principles class. It thus confirms the findings of cognitive scientists who found this result in other domains. More broadly, it suggests that examples employed by economics instructors should be "rich" and cover a wide variety of cases. Note that adding negative examples to a lecture is a low-cost intervention, in the spirit of Lang (2021).

This study may provide other contributions. These include a straightforward design that controls for section and concept effects, as well as focusing on measuring teaching innovations that are not confounded by students studying the concept, as they would if learning is assessed with a test or exam.

One puzzle from this study is that active learning (the P/N/C teaching method) did not outperform lecture (P/N). However, keep in mind that widely cited papers on active learning, (Freeman et al., 2014) (Kozanitis and Nenciovici, 2023) generally took place over much longer time spans (like semesters) and

[10] Full regression results for each model can be seen in the Appendix.

active learning was very broadly defined as something that was not lecture. Perhaps the design used here can be used to explore active learning in future studies.

Finally, as economics education research continues to develop, perhaps more effort should be spent on incorporating the work of cognitive scientists, as was done here.

References

Alessandri, J., J. Nuchadee, C. Nikolov, B. Leroy, C. Cançado, and V. Rivière (2025). Effects of multiple examples and nonexamples on learning the concept of reinforcement by undergraduate students. *European Journal of Behavior Analysis*, 1–24.

Ali, A. M. (1981). The use of positive and negative examples during instruction: Some important issues related to the design and development of instructional materials. *Journal of instructional development 5*(1), 2–7.

Allgood, S. (2001). Grade targets and teaching innovations. *Economics of Education Review 20*(5), 485–493.

Anderson, R. C. (1977). The notion of schemata and the educational enterprise: General discussion of the conference. In *Schooling and the acquisition of knowledge*, pp. 415–431. Routledge.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge University Press.

Boyle, A. and W. L. Goffe (2018). Beyond the flipped class: The impact of research-based teaching methods in a macroeconomics principles class. In *AEA Papers and Proceedings*, Volume 108, pp. 297–301. American Economic Association.

Eisenkopf, G. and P. A. Sulser (2016). Randomized controlled trial of teaching methods: Do classroom experiments improve economic education in high schools? *The Journal of Economic Education 47*(3), 211–225.

Engelmann, S. and D. Carnine (1982). *Theory of instruction: Principles and applications.* Irvington Publishers New York.

Freeman, S., S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the national academy of sciences 111*(23), 8410–8415.

Hamid, Y. (2022). *How to Teach Economics.* John Catt Educational.

Hendrick, C. (2025). The power of examples and non-examples. `https://carlhendrick.substack.com/p/the-power-of-examples-and-non-examples`. The Learning Dispatch, Accessed: 2025-06-26.

Kirschner, P. A., J. Sweller, and R. E. Clark (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist 41*(2), 75–86.

Kozanitis, A. and L. Nenciovici (2023). Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: A meta-analysis. *Higher Education 86*(6), 1377–1394.

Kullberg, A., U. Runesson Kempe, and F. Marton (2017). What is made possible to learn when using the variation theory of learning in teaching mathematics? *ZDM—Mathematics Education 49*, 559–569.

Lang, J. M. (2021). *Small teaching: Everyday lessons from the science of learning.* John Wiley & Sons.

Mayer, R. E. (2002). Multimedia learning. In *Psychology of learning and motivation*, Volume 41, pp. 85–139. Elsevier.

Mikek, P. (2023). A flipped classroom experiment in growth theory. *Eastern Economic Journal 49*(3), 433.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review 63*(2), 81.

Needham, T. (2019). Teaching through examples. In A. Boxer (Ed.), *The researchED Guide to Explicit and Direct Instruction: An Evidence-Informed Guide for Teachers*, pp. 47–65. Hachette UK.

Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. *American educator 36*(1), 12.

Schwartz, D. L., J. M. Tsang, and K. P. Blair (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them.* WW Norton & Company.

Settlage, D. and J. Wollscheid (2019). The invisible hand in action–an interactive classroom experience. *Journal of Economics Teaching 4*(1), 27–39.

Walstad, W. and S. Allgood (1999). What do college seniors know about economics? *American Economic Review 89*(2), 350–354.

Willingham, D. T. (2021). *Why don't students like school?: A cognitive scientist answers questions about how the mind works and what it means for the classroom.* John Wiley & Sons.

7 Appendix

7.1 Model 1: Random Effects

```
Random-effects GLS regression                   Number of obs     =      3,043
Group variable: student_num                     Number of groups  =      1,229

R-squared:                                      Obs per group:
     Within  = 0.1642                                         min =          1
     Between = 0.0652                                         avg =        2.5
     Overall = 0.1199                                         max =          3

                                                Wald chi2(9)      =     432.00
corr(u_i, X) = 0 (assumed)                      Prob > chi2       =     0.0000

------------------------------------------------------------------------------
correct_cli~r | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
--------------+----------------------------------------------------------------
          P/N |   .2334296   .0186395    12.52   0.000     .1968969    .2699624
        P/N/C |   .2502224   .0188641    13.26   0.000     .2132496    .2871953
      capital |   .2441717   .0188597    12.95   0.000     .2072073     .281136
        money |   .1821875   .0191571     9.51   0.000     .1446404    .2197347
      sec4_sp |  -.0449992   .0365446    -1.23   0.218    -.1166253     .026627
      sec7_sp |   .0670837   .0370617     1.81   0.070     -.005556    .1397233
      sec2_fa |   .0972151   .0247678     3.93   0.000     .0486712     .145759
      sec5_fa |   .0706123   .0252852     2.79   0.005     .0210541    .1201705
         FA24 |  -.0040837   .0328671    -0.12   0.901     -.068502    .0603346
        _cons |    .339377   .0318152    10.67   0.000     .2770204    .4017335
--------------+----------------------------------------------------------------
      sigma_u |  .11774007
      sigma_e |  .41934838
          rho |  .07307101   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

7.2 Model 2: Fixed Effects

```
Fixed-effects (within) regression               Number of obs     =      3,043
Group variable: student_num                     Number of groups  =      1,229

R-squared:                                      Obs per group:
     Within  = 0.1646                                         min =          1
     Between = 0.0477                                         avg =        2.5
     Overall = 0.1084                                         max =          3

                                                F(4, 1810)        =      89.15
corr(u_i, Xb) = -0.0025                         Prob > F          =     0.0000

------------------------------------------------------------------------------
correct_cli~r | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
--------------+----------------------------------------------------------------
          P/N |   .2326282   .0196099    11.86   0.000     .1941677    .2710887
        P/N/C |   .2350457   .0198247    11.86   0.000      .196164    .2739274
      capital |   .2510568   .0200058    12.55   0.000     .2118199    .2902937
        money |   .1899291   .0201289     9.44   0.000     .1504507    .2294074
        _cons |   .3792841   .0184981    20.50   0.000     .3430042     .415564
--------------+----------------------------------------------------------------
      sigma_u |  .31099995
      sigma_e |  .41934838
          rho |  .35484292   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

7.3 Model 3: Pooled OLS

```
      Source |       SS           df       MS      Number of obs   =     3,043
-------------+----------------------------------   F(9, 3033)      =     45.93
       Model |  77.7332706         9  8.63703006   Prob > F        =    0.0000
    Residual |  570.361373     3,033  .188051887   R-squared       =    0.1199
-------------+----------------------------------   Adj R-squared   =    0.1173
       Total |  648.094643     3,042  .213048864   Root MSE        =    .43365

------------------------------------------------------------------------------
correct_cli~r | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
--------------+----------------------------------------------------------------
          P/N |   .2338616   .0192326    12.16   0.000     .1961513    .2715719
        P/N/C |   .2522588   .0194614    12.96   0.000     .2140998    .2904177
      capital |    .242817   .0194385    12.49   0.000     .2047031    .2809309
        money |   .1809637    .019757     9.16   0.000     .1422253    .2197021
     sec4_sp |  -.0460285   .0344791    -1.33   0.182    -.1136333    .0215764
     sec7_sp |   .0692521   .0349681     1.98   0.048     .0006885    .1378158
     sec2_fa |   .0997797   .0233545     4.27   0.000     .0539875    .1455719
     sec5_fa |    .072614   .0238967     3.04   0.002     .0257587    .1194693
         FA24 |  -.0062596   .0310187    -0.20   0.840    -.0670795    .0545603
        _cons |   .3412335   .0307215    11.11   0.000     .2809965    .4014705
------------------------------------------------------------------------------
```

7.4 Model 4: Panel Probit

```
Random-effects probit regression                    Number of obs    =  3,043
Group variable: student_num                         Number of groups =  1,229

Random effects u_i ~ Gaussian                       Obs per group:
                                                                 min =      1
                                                                 avg =    2.5
                                                                 max =      3

Integration method: mvaghermite                     Integration pts. =     12

                                                    Wald chi2(9)     = 301.26
Log likelihood = -1696.2509                         Prob > chi2      = 0.0000

------------------------------------------------------------------------------
correct_clicker | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
----------------+-------------------------------------------------------------
            P/N |   .7061811   .0632465    11.17   0.000     .5822202    .8301421
          P/N/C |    .755012   .0641734    11.77   0.000     .6292344    .8807896
        capital |   .7511158   .0653679    11.49   0.000      .622997    .8792346
          money |   .5391304   .0639745     8.43   0.000     .4137427    .6645181
        sec4_sp |  -.1061525   .1168757    -0.91   0.364    -.3352247    .1229196
        sec7_sp |   .2424137   .1189068     2.04   0.041     .0093606    .4754668
        sec2_fa |   .2567578   .0827135     3.10   0.002     .0946423    .4188734
        sec5_fa |   .1876059   .0841508     2.23   0.026     .0226734    .3525384
           FA24 |   .0339944   .1053589     0.32   0.747    -.1725053     .240494
          _cons |   -.521917   .1015957    -5.14   0.000     -.721041   -.3227931
----------------+-------------------------------------------------------------
       /lnsig2u |  -2.151529   .4118531                     -2.958746   -1.344311
----------------+-------------------------------------------------------------
        sigma_u |    .341037   .0702286                      .2277805    .5106067
            rho |   .1041885   .0384396                      .0493248     .206802
------------------------------------------------------------------------------
```

The effects of P/N and P/NC was found for capital, money, and technology using the average value of each section (weighting by their relative abundance in the sample). 95% confidence intervals had a low of 0.1475 and a high of 0.3155. z-scores are not reported below, but the smallest for any of these results was 11.02.

Table 3: Marginal effects of negative examples by concept

		Model 4	Model 5	Model 6
capital	P/N	0.1859	0.1861	0.1794
	P/NC	0.1987	0.2000	0.1909
money	P/N	0.2113	0.2118	0.2104
	P/NC	0.2260	0.2272	0.2239
technology	P/N	0.2500	0.2505	0.2574
	P/NC	0.2674	0.2688	0.2739

7.5 Model 5: Pooled Probit

```
Probit regression                                       Number of obs =  3,043
                                                        LR chi2(9)    = 355.85
                                                        Prob > chi2   = 0.0000
Log likelihood = -1700.0906                             Pseudo R2     = 0.0947

--------------------------------------------------------------------------------
correct_clicker | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
----------------+---------------------------------------------------------------
            P/N |   .6704907   .0599566    11.18   0.000      .552978    .7880034
          P/N/C |   .7193769   .0608689    11.82   0.000     .6000761    .8386777
        capital |    .708028   .0613586    11.54   0.000     .5877672    .8282887
          money |    .506687   .0609535     8.31   0.000     .3872204    .6261536
        sec4_sp |  -.1019242   .1055738    -0.97   0.334    -.3088451    .1049967
        sec7_sp |    .237353   .1074568     2.21   0.027     .0267414    .4479645
        sec2_fa |   .2472162   .0749276     3.30   0.001     .1003608    .3940717
        sec5_fa |   .1807513    .076374     2.37   0.018     .0310611    .3304415
           FA24 |   .0289758   .0951975     0.30   0.761    -.1576079    .2155595
          _cons |  -.4915186   .0929086    -5.29   0.000    -.6736161   -.3094212
--------------------------------------------------------------------------------
```

7.6 Model 6: Pooled Logit

```
Logistic regression                                     Number of obs =  3,043
                                                        LR chi2(9)    = 360.67
                                                        Prob > chi2   = 0.0000
Log likelihood = -1697.6802                             Pseudo R2     = 0.0960

--------------------------------------------------------------------------------
correct_clicker | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
----------------+---------------------------------------------------------------
            P/N |   1.124537   .1015003    11.08   0.000     .9255997    1.323473
          P/N/C |   1.196586   .1028708    11.63   0.000     .9949631    1.398209
        capital |   1.197522   .1047133    11.44   0.000     .9922875    1.402756
          money |   .8395759   .1016448     8.26   0.000     .6403558    1.038796
        sec4_sp |  -.1719291    .176867    -0.97   0.331    -.5185819    .1747238
        sec7_sp |   .3889991   .1796979     2.16   0.030     .0367977    .7412004
        sec2_fa |   .4094173   .1270673     3.22   0.001       .16037    .6584645
        sec5_fa |   .3298735   .1300336     2.54   0.011     .0750124    .5847346
           FA24 |   .0432771    .158609     0.27   0.785    -.2675908    .3541449
          _cons |  -.8291735   .1536462    -5.40   0.000    -1.130315   -.5280324
--------------------------------------------------------------------------------
```

7.7 Summary statistics tables

These tables show the average correct response rates for each concept in each section. The positive examples only treatment results are organized on the diagonal.

Table 4: Spring 2024 Results

%correct, SP24	Capital	Money	Technology	Avg Accuracy (section)
sec. 6	0.7582	0.6538	0.5357	0.6492
sec. 7	0.8788	0.5776	0.6907	0.7157
sec. 4	0.8074	0.7328	0.2522	0.5975
Avg. Accuracy (concept)	0.8148	0.6547	0.4929	

Table 5: Fall 2024 Results

%correct, FA24	Capital	Money	Technology	Avg Accuracy (section)
sec. 2	0.7077	0.7933	0.7130	0.7380
sec. 5	0.7686	0.6273	0.7696	0.7218
sec. 6	0.8498	0.8405	0.2000	0.6301
Avg. Accuracy (concept)	0.7754	0.7537	0.5609	