Using Invention Activities to Teach Econometrics

DRAFT: February 22, 2019

Douglas McKee and George Orlov

An invention activity is a teaching technique that involves giving students a difficult substantive problem that cannot be readily solved with any methods they have already learned. The work of Dan Schwartz and colleagues (Schwartz & Bransford, 1998; Schwartz & Martin, 2004), suggests that such activities prepare students to learn the "expert's solution" better than starting directly with a lecture on that solution. In this paper we present six new invention activities appropriate for a college econometrics course. We describe how we introduce each activity, guide students as they work, and wrap up the activity with a short lecture.

1. Introduction

An invention activity is a teaching technique that involves giving students a difficult substantive problem that cannot be readily solved with any methods they have learned up to that point. The work of Dan Schwartz and colleagues (Schwartz & Bransford, 1998; Schwartz & Martin, 2004), suggests that such activities prepare students to learn the "expert's solution" better than starting directly with a lecture on that solution. In particular, they find that students are better able to transfer their learning to new contexts and retain what they've learned for a longer period of time.

In Spring 2018, we developed and fielded eight new invention activities in an applied econometrics course, and based on our experience, we fielded refined versions of six in Fall 2018. In these activities, students were given carefully scaffolded problems related to topics including categorical independent variables, interactions of independent variables, difference-in-differences, and fixed effects. We believe we are the first to report the use of invention activities in an economics course.

In Section 2 we review the empirical and theoretical literature on the effectiveness of invention activities at the high school and college levels in a range of disciplines. Section 3 presents in detail each of the six invention activities that we currently use in our courses. We describe how we introduce each activity, guide students as they work through the problems, and wrap up the activity with a short lecture. In Section 4 we share our experience fielding the activities during two semesters and share student feedback on them. Section 5 concludes.

2. Literature Review

The key element that differentiates an invention activity from other kinds of small group classroom activities is that the instructor asks students to try to solve a problem *before* explicitly teaching them the methods required (Schwartz & Bransford, 1998). It is important that the goal of the activity be clear and free of jargon, and students are usually given several cases with different characteristics with which to evaluate their solution. While students work on the problem, instructors circulate around the room and ask groups to articulate their proposed solution. The beauty of an invention activity is that students are not required to successfully solve the problem to benefit from the experience. Instructors gently nudge them toward a good solution solely by pointing out interesting features and potential shortcomings of their work. The final stage of the activity is a brief explanation that provides a conceptual framework for the problem and the consensus expert's solution. The instructor may also present a few notable student solutions.

There are a variety of theories that explain why and how invention activities are effective, and this is an active area for research. The primary benefit, according to Schwartz & Martin (2004), is that invention activities prepare students for future learning. Specifically, they help students identify the important pieces of information involved and organize them in their mind. Without preparation, students often skip this step and simply memorize the solution without understanding why and in what contexts it applies. The contrasting cases students work with allow students to immediately evaluate and understand the expert's solution when it's presented. Invention activities also encourage students to think creatively in an environment where they are primarily asked to apply one of a finite set of methods to solve a problem.

There is a growing empirical literature that shows the impact invention activities have on student performance. Students that participate in these activities do not always score higher on conventional assessments that involve applying the methods in contexts they have seen before, but they have been shown to do substantially better at higher level tasks such as learning similar ideas and applying what they've learned to new situations.

The empirical research spans a wide range of disciplines and grade levels. Schwartz & Bransford (1998) show that college students who first analyze simplified experimental designs and data from classic psychology experiments and then hear a lecture more accurately predict results of hypothetical experiments than students who instead read about the cases and hear the lecture. Schwartz & Martin (2004) find that ninth grade statistics students that participate in invention activities perform better on tests that require learning from a resource during the test. Jarosz, Goldenberg, & Wiley (2017) find similar results in a college statistics course. Taylor, Smith, van Stolk, & Spiegelman (2010) find that invention activities induce college students in an introductory biology course to more quickly engage with future problems and produce a higher number of reasonable hypotheses. In addition, eighth grade physics students (Schwartz, Chase, Oppezzo, & Chin, 2011) as well as college physics students (Roll, Holmes, Day, & Bonn, 2012) have been shown to benefit from invention activities. Holmes et al (2014) have also demonstrated that providing appropriate scaffolding for invention activities improved students' conceptual understanding in an assessment administered two months after the activity concluded. We believe our work is the first application of invention activities in economics.

3. Invention Activities

In this section we present six invention activities that we have used successfully in two iterations of a course in applied econometrics. For each activity, we provide its learning goals, explain how we introduce the activity to the class, and present the questions that students will try to answer during the activity itself. We also share advice for guiding students through the activity and wrapping up the activity with a short lecture.

3.1. Bivariate Regression

3.1.1. Activity Learning Goals

- Understand and apply the Ordinary Least Squares (OLS) estimation method.
- Understand and apply the Least Absolute Deviation (LAD) estimation method.
- Recognize situations where these two methods work well and do not work well.

3.1.2. Introducing the Activity

We start the activity by writing down a simple bivariate regression model ($y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$) and giving a few examples of what it can be used to describe. This might be wages as a function of years of schooling or demand for ice cream as determined by outside temperature. We then draw x and y axes and show that if we ignore the error term, we get a line that represents on average what we expect y to be given x. Because the model does contain an error term, the observed data are actually random deviations from this line. We draw some dots near the line to represent the observed data. We then erase the line since in the real world we, usually, do not know the true values of the β's. Finally, we raise the question of how we might estimate the β's (i.e., the line) using the observed data (i.e., the dots).

3.1.3. The Activity

Students receive a printed worksheet containing the six different scatter plots shown in Figure 1 and the following questions:

Q1: *How do the scatter plots differ from each other?*

The first plot is the simplest one, and students should be encouraged to compare the other figures to it. Plots 2 and 3 are identical but with the addition of a few outliers. Plot 4 is exactly like the first except with a negative slope. Plot 5 has the same general slope as the first, but contains more noise, and the last plot is the same as the fifth but with a negative slope. We have found that students are quite good at identifying these differences.

Q2: *Write down a procedure (i.e., sets of steps) for fitting a line ($\hat{y}_i = b_0 + b_1 x_i$) through the data (i.e., a set of n points x_i, y_i).*

Students will often initially write down procedures that are not well-defined. For example, we've seen many groups include a step calling for outliers to be removed. Instructors circulating around the classroom should ask for clarification in these cases.

Q3: *Write down another procedure for fitting a line through the data.*

The students who remembered the method of Ordinary Least Squares from another class are forced to be creative here.

Q4: *How do you think the results of each procedure compare in each of the above data sets?*

This is the most important question in the whole activity as students learn to identify the contexts where their method works well and where it does not. Often a method that works well when there is a strong positive correlation (e.g., "Connect the bottom left point to the upper right point") works poorly when there are outliers or a strong negative correlation.

Q5: *Which of your procedures better represents the average linear relationship between x and y?*

This is difficult and motivates the idea that there isn't a single method that is the "best" in all contexts. It can also lead to a good discussion of how one might quantify the uncertainty in our estimates using standard errors or confidence intervals.

3.1.4. Wrapping up the Activity

We select 2-4 examples of student work, take pictures of them, and share them with the class. We point out where procedures are well-defined and ill-defined, and we show cases (scatter plots) where procedures give good and poor results. Now that the students have identified several important features of bivariate data and have practiced evaluating their own algorithms, they are ready to be taught the methods of Ordinary Least Squares (OLS) and Least Absolute Deviations (LAD). The last question (about which procedure is best) can be used to motivate a presentation of the Gauss-Markov Theorem that says OLS is the Best Linear Unbiased Estimator (BLUE).

3.2. Categorical Independent Variables

3.2.1. Learning Goals

- Incorporate categorical independent variables into linear regression models as sets of dummy variables.
- Interpret coefficients on dummy variables as expected changes in the conditional mean of the dependent variable relative to a reference category.
- Recognize and avoid the "dummy variable trap" of including dummy variables for every possible value of a categorical independent variable.

3.2.2. Introducing the Activity

Imagine that you run a local coffee shop and are trying to understand the determinants of your customers' demand for coffee. Over the past year you have randomly varied the price you charge for coffee each week (p_i) and recorded how many cups you sell each week (q_i). You have also created a variable ($season_i$) that is coded as 1 for spring, 2 for summer, 3 for fall, and 4 for winter.

3.2.3. The Activity

Q1: *How would you interpret the coefficient on season in the following model?*

$$q_i = \beta_0 + \beta_1 p_i + \beta_2 season_i + \varepsilon_i$$

At this point in the course, most students can interpret a coefficient on a count variable: β_2 represents the expected difference in quantity sold between one season and the following season.

Q2: *What assumption are you making about the effects of the different seasons in this model?*

The expected difference between spring and summer is the same as the difference between summer and fall and the difference between fall and winter. These are clearly not reasonable assumptions.

Q3: *Can you think of a better way to control for season in your model?*

Students usually come up with a variety of ideas on their own, but if a group is stuck, you can suggest that they try defining a new variable (or set of variables) based on season and include those variables instead.

3.2.4. Wrapping up the Activity

Some students will create a single dummy variable for a season. Their model tells them nothing about expected differences in sales between the other seasons, and in essence, this solution throws away important information. Some students will put all four dummy variables in the model. Here we point out that it is difficult to interpret the intercept because it isn't possible for all four dummy variables to equal zero. It's also difficult to interpret the coefficients on the other dummy variables. This may or may not be an appropriate time to point out that this model suffers from perfect multicollinearity. Finally, we present the expert's solution: Choose a reference category and include all the other season dummy variables. Now we can clearly interpret all the model coefficients. We finish by showing that the choice of reference category has no effect on predicted differences between categories.

3.3. Heterogeneous Effects

3.3.1. Learning Goals

- Use interactions in multiple regression models to allow effects of variables to depend on the values of other variables.
- Correctly interpret coefficients on interactions of continuous and dummy explanatory variables.

3.3.2. Introducing the Activity

Suppose a university is considering increasing the number of tutors it hires, but it wants a good estimate of the effect of tutoring on student outcomes first. The university chooses a representative sample comprised of 100 students and randomly assigns a tutor to half of them. tut_i is a dummy variable equal to 1 if a tutor was assigned to student *i* and 0 otherwise. The university also collects data on test scores (y_i), student gender ($male_i$), and grade point average (GPA), recorded in the preceding term.

3.3.3. The Activity

Q1: *The administrators start their analysis by estimating the following model:*

$$y_i = \beta_0 + \beta_1 tut_i + \beta_2 male_i + \beta_3 GPA_i + \varepsilon_i$$

How should we interpret β_1, the coefficient on the tutor dummy variable? Is β_1 an unbiased estimate of the Average Treatment Effect (ATE)? Why or why not?

This question reviews material they have seen before, and most students should recognize that the coefficient on the tutor dummy does indeed represent the causal effect of a student having a tutor on test scores because tutors were randomly assigned. When talking to students, it may be worthwhile to verify that they understand that the estimate of β_1 is the ATE only under the assumption of perfect compliance (all students who had tutors assigned use the services of these tutors). You may also want to point out that controlling for gender and GPA is unnecessary for getting an unbiased estimate in this case, but it should result in a more precise estimate of the tutoring effect.

Q2: *The university wants to know if the effect of a tutor is different for male students relative to female students. The original regression model assumes effects for each of these groups (i.e., males and females) are the same. Suppose you estimate the following model separately for males and females:*

$$y_i = \beta_0 + \beta_1 tut_i + \beta_2 GPA_i + \varepsilon_i$$

All we are doing here is introducing the idea that the effect of something (like tutoring) might differ for different groups. You should point out that estimating the original model using the whole sample estimates the average effect for the whole population.

Q2a: *How do you interpret your two sets of estimates of* β_1 *and* β_2*?*

We expect students to recognize that the estimates of β_1 represents the effects of tutoring specifically for males and females. The coefficients on GPA should not be interpreted causally—Instead, β_2 represents the expected difference in test scores between two students (male for one estimate, female for the other) who have GPA's that differ by one unit.

Q2b: *Write down a regression model that would be estimated on the whole sample that allows the effect of tutoring to differ for males and females but assumes the effect of GPA is the same for males and females. Interpret the coefficients of your new model.*

This is where the students try to invent something they've never seen before. Some groups succeed by adding an interaction between male and tutor to their model. The groups that do not succeed still benefit from the exercise as they learn why it might be useful to include an interaction in a model.

Q2c: *State a hypothesis in terms of your regression coefficients that you would use to test whether the effect of tutoring differs for males and females.*

Answering this question requires students to think hard about the interpretation of the coefficient on the interaction.

3.3.4. Wrapping up the Activity

This activity leads naturally to a brief lecture on why you might include an interaction in a model and how you interpret it. This usually includes a discussion of dummy-dummy, dummy-continuous, and continuous-continuous interactions.

3.4. Difference-in-Differences

3.4.1. Learning Goals

- Estimate causal effects by applying difference-in-differences (DD) estimation to aggregate level data.
- Understand and evaluate the parallel trends assumption of DD in different empirical contexts.

3.4.2. Introducing the Activity

Do free laptop computers improve student outcomes? Suppose São Paulo, the capitol of Brazil instituted a free laptop program in all of its elementary schools in 2009. Suppose also that Rio de Janeiro, another large city a few hundred miles up the coast, did NOT implement the program. While this scenario is hypothetical, the government of Uruguay implemented a One-Laptop-Per-Child program across their country in 2009, and many schools in the US have also distributed free computers to their students. These programs are expensive, and it is important to have good estimates of their benefits.

3.4.3. The Activity

Q1: *You have average elementary school test scores in São Paulo and Rio de Janeiro for the end of the 2009 school year. Why is the difference between them a poor measure of the effect of the program?*

When students are having trouble getting started, we ask more pointed questions: What does this difference capture above and beyond the effect of the program? Are there other differences between São Paolo and Rio de Janeiro that could explain some of the observed differences in test scores?

Q2: *You get the average test score for São Paulo students in 2008. Why is the difference between this and the average São Paulo score in 2009 a poor estimate of the effect of the program?*

We hope that students will recognize that there may be other changes that occurred between these two years that could explain the difference in test scores.

Q3: *Suppose you have the average test scores for both São Paulo and Rio in 2008 and 2009. Can you use these together to improve upon the estimate suggested in Q1? How about Q2? Hint: Think about the Q1 and Q2 differences in terms of Treatment on the Treated and Selection Bias.*

In our course, we introduce the vocabulary of treatment effects earlier in the semester, and we encourage students to think about the difference between the outcomes of two groups in a non-experimental context as the sum of the Treatment on the Treated and Selection Bias. If these terms are not familiar to your students, you can instead suggest that the simple differences presented in Q1 and Q2 are sums of the causal effect of the treatment and another part that represents pre-existing differences. The key is to encourage students to look for a new difference that can be used as an estimate of the second part and then subtracted from the combined effects to isolate the effect of the program.

3.4.4. Wrapping up the Activity

In our experience, many students are able to discover the method of difference-in-differences through the activity. This allows us to give a very concise lecture summarizing the method and explaining how the parallel trends assumption relies on the difference across time in the control group being a good approximation of what would have happened in the treatment group in the absence of the treatment. Equivalently, we explain that the difference between the control group and treatment group in the pre-treatment period must approximate the difference that would exist between the two groups in the post-treatment period if the treatment had never been applied.

3.5. Regression Discontinuity

3.5.1. Learning Goals

- Understand and explain how the Regression Discontinuity (RD) method works.
- Judge situations where RD can and cannot be applied:
    - Treatment must depend on whether the assignment variable is above or below a threshold.
    - The relationship of the assignment variable to the outcome must be continuous in the absence of treatment.
- Estimate causal effects using linear and non-linear parametric RD models.

3.5.2. Introducing the Activity

The Adams Scholarship was launched in Massachusetts in 2005. It gave small awards to students who exceeded a particular district-specific test score if they attended a public 4-year college in Massachusetts. In the scatter plots shown in Figure 2 (reproduced from Goodman, 2008), GAP represents the number points above (+) or below (-) the required score. The y-axis is the enrollment rate for students that have a particular GAP. The plots show how the overall, public, and private enrollment rates varied with test score before and after the program was implemented.

3.5.3. The Activity

Q1: *What explains the upward trend in the upper left figure?*

There will be at least a few groups that need help interpreting the graphs, but once they are clear, most students quickly recognize that higher test scores make admission to college more likely.

Q2: *Why would regressing 2005 enrollment on a dummy for receipt of the scholarship in 2005 give a poor estimate of the program's effect?*

We nudge groups that are stuck on this question by asking "If there were no effect of the program at all, what would you expect the sign on this dummy variable to be?"

Q3: *What's true about A, C, and E but isn't true for D and F?*

We want students to notice the discontinuous jump up at the eligibility threshold in the post-treatment period, that does not exist before the program goes into effect. Some groups need to be encouraged to compare C to D and then E to F.

Q4: *Based on figures B, D, and F, what are the effects of the program?*

Most groups that answer Q3 correctly also recognize that the magnitude of the jump across the threshold is an estimate of the program effect. We ask groups that answer this quickly to think about whether this is an estimate of the Average Treatment Effect (ATE) or whether it is only applicable to students near the threshold. This primes them for a later discussion of the Local Average Treatment Effect (LATE).

Q5: *Write down a regression model that allows a linear effect of GAP and a potential discontinuous jump at the eligibility threshold (GAP=0). Which coefficient represents the effect of the program?*

While the first set of questions involve building intuition for RD, the second set has students explore models that allow for formal estimation. The simplest is the one we are looking for here:

$$y_i = \beta_0 + \beta_1 GAP_i + \beta_2 D_i^{GAP>0}$$

y_i $r$epresents the probability of attending college, $D_i^{GAP>0}$ is 1 when the test score is below the threshold, and $D_i^{GAP>0}$ is 0 when it is below. When groups are struggling, we ask them what terms would capture a linear effect of GAP and a discontinuous jump at the threshold. We want students to recognize that in any of the models they write down for Q5, Q6, and Q7, the effect of the program is the coefficient on the threshold dummy variable.

Q6: *Write down a regression model that allows for a quadratic effect of GAP and a potential discontinuous jump at the eligibility threshold (GAP=0). Which coefficient represents the effect of the program?*

We are looking for students to add a quadratic term to the specification developed above:

$$y_i = \beta_0 + \beta_1 GAP_i + \beta_2 GAP_i^2 + \beta_3 D_i^{GAP>0}$$

Q7: *Write down a regression model that allows for a linear effect of GAP, a potential discontinuous jump at the eligibility threshold (where GAP=0), **<u>and</u>** allows the slope to be different on each side of the threshold. Which coefficient represents the effect of the program?*

To answer this question, students must combine what they've learned so far about RD with what they've learned about interaction terms. Specifically, they must recognize that an interaction can be used to let the effect of GAP differ for students above and below the threshold. Some students simply add the interaction:

$$y_i = \beta_0 + \beta_1 GAP_i + \beta_2 D_i^{GAP>0} + \beta_3 GAP_i D_i^{GAP>0}$$

We ask students what the slope of the regression line is on each side of the threshold (β_1 and $\beta_1 + \beta_3$ in this case) and make sure they recognize that the effect of the treatment is still the coefficient on the threshold dummy variable. Some students write down an equivalent model that is somewhat easier to interpret:

$$y_i = \beta_0 + \beta_1 GAP_i D_i^{GAP\leq 0} + \beta_2 GAP_i D_i^{GAP>0} + \beta_3 D_i^{GAP>0}$$

Here the slope to the left of the threshold is β_1 and the slope to the right is β_2, while the effect of the program is still the coefficient on the threshold dummy variable.

3.5.4. Wrapping up the Activity

We usually implement this activity in two stages. We start by giving them a worksheet containing the figure and the first four questions and focus on building intuition. At the end of the first stage, we make it very clear that the discontinuous jump is our RD estimate of the effect of the program. We also discuss the substance of this particular study: The Adams Scholarship induced a fair amount of switching of students from private to public colleges, but it did not result in an increase in the total number of high school graduates attending a 4-year college.

After the second stage of the activity (Q5-Q7) we carefully write down correct models for each of the questions and interpret their coefficients. We finish the activity with a discussion of what would be different if the threshold was not zero. Suppose x_i is the test score and x_0 is the eligibility threshold. We need a new model in order to allow the slope to differ on each side of the threshold:

$$y_i = \beta_0 + \beta_1(x_i - x_0)D_i^{GAP\leq 0} + +\beta_2(x_i - x_0)D_i^{GAP>0} + \beta_3 D_i^{GAP>0}$$

Explaining why it is necessary to subtract x_0 from $x_i$is far easier once students have a solid understanding of the case where the threshold is zero.

3.6. Fixed Effects

3.6.1. Learning Goals

- Use fixed effects models in situations with time-invariant unobserved heterogeneity.
- Estimate fixed effects models using first differences.
- Estimate fixed effects models using within transformations.

3.6.2. Introducing the Activity

Do you believe getting married makes people less likely to commit crimes? Why? In this exercise we develop a new method that can be used to test this hypothesis. Suppose you have data containing the number of crimes committed in the previous year and current marital status for 500 individuals. Additionally, suppose you have two observations per individual spaced four years apart. Data where you have multiple observations per individual spread across time is called panel data or longitudinal data.

3.6.3. The Activity

Q1: *Consider the following model:*

$$crime_{it} = \beta_0 + \beta_1 married_{it} + \varepsilon_{it}$$

Suggest at least two omitted variables that could induce bias in your estimate of β_1.

Students are very good at coming up with possible confounders here. We have had students suggest that violent tendencies, risk aversion, and ability to earn a market wage are all correlated with marital status and could be predictors of criminal behavior.

Q2: *Suppose all of the omitted variable bias comes from variables whose values do not change across time. Let u_i in the following model represent the contribution of these variables. We will call this the "fixed effect."*

$$crime_{it} = \beta_0 + \beta_1 married_{it} + u_i + \varepsilon_{it}$$

We cannot estimate this model directly with OLS because we do not observe u_i, and the unobserved part of the equation ($u_i + \varepsilon_{it}$) may be correlated with marital status. That said, this equation must hold in both time period 1 and 2:

$$crime_{i1} = \beta_0 + \beta_1 married_{i1} + u_i + \varepsilon_{i1}$$

$$crime_{i2} = \beta_0 + \beta_1 married_{i2} + u_i + \varepsilon_{i2}$$

How might you combine these equations to get an equation that ***can*** *be estimated with OLS? Verify that each of the assumptions required by OLS holds and interpret β_1 in the context of your new model equation.*

Most students figure out that if they subtract one equation from the other, they get a new equation that does not contain the fixed effect. The key is for students to recognize that the error term in the new model ($\varepsilon_{i2} - \varepsilon_{i1}$) is mean zero and uncorrelated with the new explanatory variable ($married_{i2} - married_{i1}$).

Q3: *Now suppose you had three time periods of data. Propose another method that uses all of your data to estimate* β_1.

It is fairly unusual for students to come up with a within-difference model, and they more often difference the first two equations and the second and third equations.

3.6.4. Wrapping up the Activity

When we show them the first difference method, it usually looks very similar (if not identical) to what they've invented. The key is to point out that estimating this model requires regressing changes in criminal activity on changes in marital status. The model is identified by both marriages and marital dissolutions. That is, the model assumes that the effect of a marriage is exactly the negative of the effect of a divorce or widowhood. This is not always a reasonable assumption.

We also ask the class what it means that the differenced model does not have an intercept. We explain that this implies that the change across time (in this case during the 4 year period) will be on average zero if there is no change in marital status. In some situations this is realistic, and we discuss whether this is the case here. The answer hinges on whether we think an individual's propensity to commit crime changes as they age. To address this possibility, we introduce a time fixed effect into the model.

4. Implementation Experience

We developed and fielded our first version of the activities described above in Spring 2018. We knew students would be uncomfortable with this approach and explicitly explained what invention activities were and how they have been shown to improve learning in other courses. Reception by students was mixed as we did not always provide enough guidance or scaffolding in the activities. During a mid-semester focus group discussion one student reported that her group would often just sit there saying “I don’t know, do you know? No, I don’t know.” She went on to say that “it feels like not a good use of time. Some of the questions just seem too hard.” Other students suggested that breaking up questions into smaller questions, providing more guidance, or giving a hint after a certain amount of time would help them get “unstuck” on the activities. Students in the focus group also felt that the time we allocated to the activities was sometimes too long, with one student reporting that “if you don’t know at a certain point, more time isn’t going to help.”

The course evaluations students completed at the end of the course gave us similar feedback with some students reporting that they really valued “the in class activities/worksheets that engaged us and kept us paying attention.” At the same time, one student found the activities uncomfortable, saying “I honestly didn't completely enjoy the group discussions during lecture even though my group was great.” Another student said “I found that a lot of times no one in my group really knew what to do or what the next step was and then the group activities weren't super productive.”

During the Summer of 2018, we took this feedback to heart and made serious refinements to most of the activities while abandoning a few of them. The major change was to listen to our students and the research of Holmes et al. (2014), and provide more explicit scaffolding in those activities where students had trouble getting started. For example, in the original version of the heterogeneous effects activity, we simply asked students to propose a model that allowed the effect of an explanatory variable to differ for different subpopulations.

The new version, described above, has students first interpret a model without an interaction, and then interpret its coefficients after estimating it separately for each subgroup.

The new activities were substantially more popular when we fielded them in Fall 2018. Another mid-semester focus group revealed none of the negative feedback we saw on the spring, with one student reporting that the invention activities were “more engaging” and that “you need more thinking than just a typical iclicker question.” In their course evaluations, most students were positive saying “having us sit in groups and giving us time to discuss answers to difficult questions helped me to better understand the material” and “the active learning activities are great!!!” At the same time, there is clearly some room for improvement in our implementation with one student saying “In class group work was annoying and took up way more time than it was worth. I would have rather gone over more examples as a whole class than spend 20 minutes waiting for the TA to come around.”

5. Conclusion

Active learning methods are primarily used in classrooms to evaluate students’ understanding of material and give them practice applying new methods and concepts. Invention activities augment this approach by preparing students to learn from the lecture more deeply than they would ordinarily. By attempting problems first and often grappling with a range of challenges, students develop knowledge structures that can be called upon when learning related new material.

In developing our own invention activities, we’ve learned that is critical to provide enough support, such that students do not get stuck, but, at the same time, not so much that they are simply following a set of steps to get to an answer. Our original implementation of the activities did not always hit the sweet spot, but the refined versions we have shared above were well-received by students in Fall 2018. In future work, we plan to compare student performance in a variety of areas between the courses where we used invention activities and the Fall 2017 iteration of the course that did not include invention activities at all.

While the specific invention activities presented here are likely only useful in a college econometrics course, we hope they will provide inspiration to other economists to create invention activities for other courses at both high school and college level. We believe many concepts in economics such as elasticity, supply/demand shocks, or even behavior of monopolists could be taught productively using these methods.

Bibliography

Goodman, J. (2008). Who merits financial aid?: Massachusetts' Adams Scholarship. *Journal of Public Economics*, *92*(10), 2121–2131. https://doi.org/10.1016/j.jpubeco.2008.03.009

Holmes, N. G., Day, J., Park, A. H. K., Bonn, D. A., & Roll, I. (2014). Making the failure more productive: scaffolding the invention process to improve inquiry behaviors and outcomes in invention activities. *Instructional Science*, *42*(4), 523–538. https://doi.org/10.1007/s11251-013-9300-7

Jarosz, A. F., Goldenberg, O., & Wiley, J. (2017). Learning by Invention: Small Group Discussion Activities That Support Learning in Statistics. *Discourse Processes: A Multidisciplinary Journal*, *54*(4), 285–302. https://doi.org/10.1080/0163853X.2015.1129593

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluating metacognitive scaffolding in Guided Invention Activities. *Instructional Science*, *40*(4), 691–710. https://doi.org/10.1007/s11251-012-9208-7

Schwartz, D. L., & Bransford, J. D. (1998). A Time for Telling. *Cognition and Instruction*, *16*(4), 475–522.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759–775. https://doi.org/10.1037/a0025140

Schwartz, D. L., & Martin, T. (2004). Inventing to Prepare for Future Learning: The Hidden Efficiency of Encouraging Original Student Production in Statistics Instruction. *Cognition and Instruction*, *22*(2), 129–184. https://doi.org/10.1207/s1532690xci2202_1

Taylor, J. L., Smith, K. M., van Stolk, A. P., & Spiegelman, G. B. (2010). Using Invention to Change How Students Tackle Problems. *CBE—Life Sciences Education*, *9*(4), 504–512. https://doi.org/10.1187/cbe.10-02-0012

Figure 1: Scatter plots for the bivariate regression activity

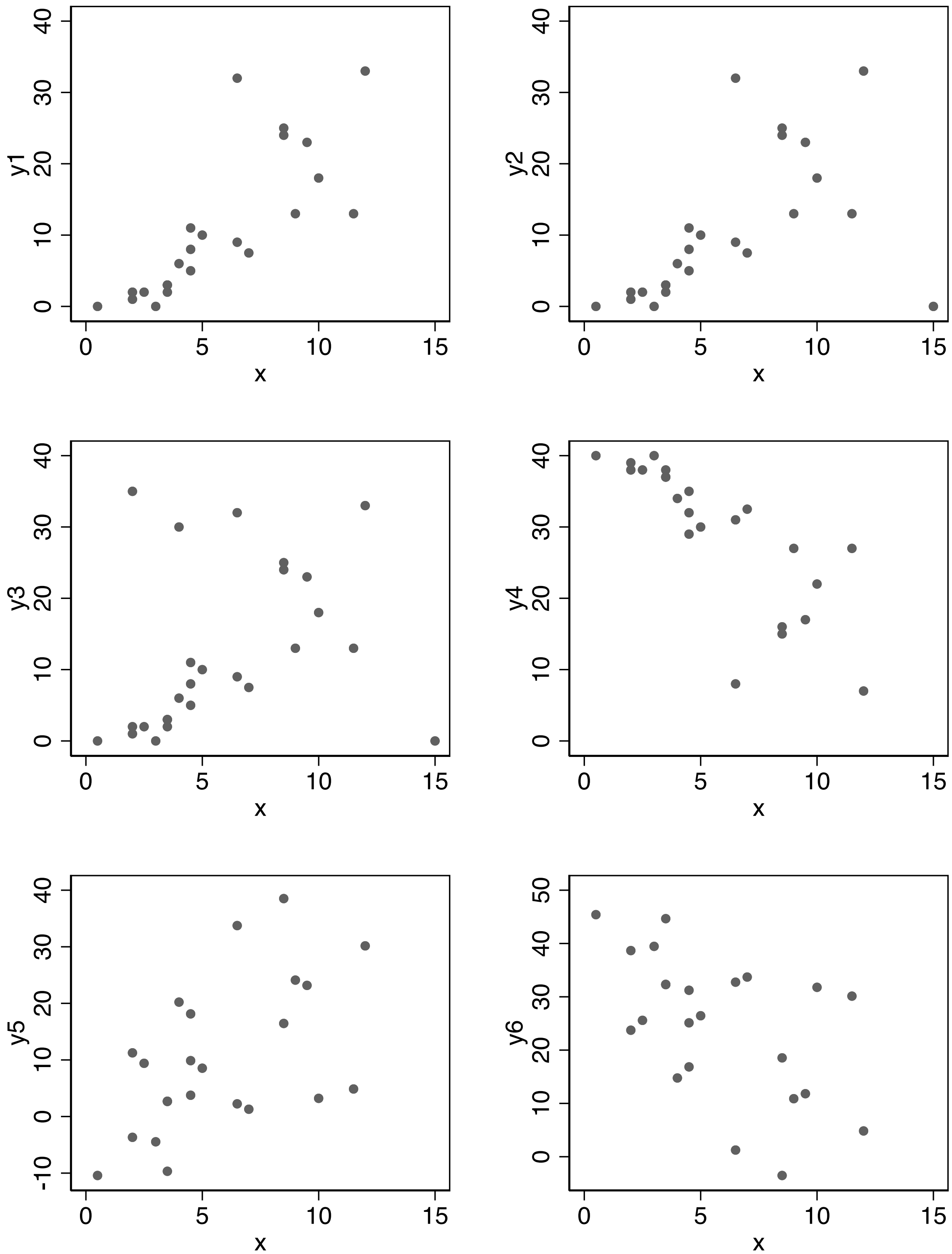

Figure 2: Plots for regression discontinuity activity reproduced from Goodman, 2008