What Students Expect and What They See: Ideology, Identity and the Double Auction Classroom Experiment

by

D. Andrew Austin
(CERGE-EI, Prague)

and

Nathaniel T. Wilcox
(University of Houston)

.

January 2002

Abstract:
Many economists find that classroom experiments using the Double Auction (DA) trading institution are an effective pedagogical tool in introductory economics classes. Results of such experiments reliably illustrate the concepts and descriptive relevance of the theory of competitive equilibrium (or CE). However, we have noticed that the degree to which students are "surprised" by the CE theory's ability to predict DA outcomes seems to vary from class to class, and especially across classes at markedly different universities. We speculate that this is due to differences in students' ideological leanings and that these, in turn, are related to various socioeconomic or "identity" variables, such as class and race, that may vary systematically across universities. This paper reports some initial experimental results that explore this hypothesis. We find that only a few socioeconomic variables significantly predict students' ideology, and that at least one measure of ideology is a robust predictor of students' prior expectations and posterior evaluations of the predictive performance of CE theory. Several other variables, including sex, union status and work experience, also help predict students' expectations or evaluations; but none of these is as strong or robust as ideology itself

Abstrakt:
Mnoho ekonomů považuje experimenty založené na dvojité aukci (DA) za efektivní pedagogický nástroj a zařazuje je do úvodních přednášek ekonomie. Výsledky těchto experimentů totiž spolehlivě ilustrují koncepty používané teorií soutěžní rovnováhy a dokumentují deskriptivní platnost této teorie. Lze si však povšimnout, že míra do které jsou studenti „překvapeni" schopností teorie soutěžní rovnováhy predikovat výsledky DA se liší od třídy k třídě, zvláště pokud se jedná o třídy na značně rozdílných univerzitách. Domníváme se, že to může být způsobeno ideologickou pozicí studentů (liberální vs. konzervativní), která naopak může odrážet ekonomicko-sociální charakteristiky studentů (sociální vrstva, rasa, apod.), jenž se mohou univerzita od univerzity značně lišit. K ověření této hypotézy jsme provedli několik experimentů, jejichž předběžné výsledky shrnuje tento článek. Zjistili jsme, že jen několik socioekonomických charakteristik významně ovlivňuje ideologickou pozici studentů. Alespoň jeden ze způsobů, jimiž jsme tuto pozici měřili, systematicky ovlivňuje apriorní očekávání studentů o predikčních možnostech teorie soutěžní rovnováhy i jejich aposteriorní hodnocení těchto možností. Dalšími, ale méně významnými proměnnými, které ovlivňují očekávání a hodnocení studentů jsou pohlaví, pracovní zkušenosti a členství v odborových organizacích.

0. Introduction

Many economists, including us, have found that classroom experiments on the Double Auction (DA) trading institution are an effective pedagogical tool in introductory economics classes. Results of such experiments reliably illustrate the concepts and descriptive relevance of the theory of competitive equilibrium (CE). However, we have noticed in our own teaching that the degree to which CE theory's ability to predict DA outcomes is "surprising" to students seems to depend on the student audience. In particular, we have noticed that classes at different universities seem differentially impressed by the results of a DA experiment performed before their eyes. We speculate that this occurs because students at different universities have systematically different characteristics, such as ideology and "identity" (that is, socioeconomic characteristics such as age, gender, race, class and life history), which cause students to have different prior expectations of economic theory's descriptive relevance.[1] This paper reports the results of an experiment and two follow-up studies that explore these conjectures.

We find that some identity variables are correlated with students' ideological leanings, as almost anyone would expect. Specifically, race, union household status (in particular, whether a student lives in a household with at least one union member) and a history of household calamities such as unemployment spells, serious illness or natural disasters are significantly related to measured ideology. However, identity variables cannot account for the bulk of between-subjects variance in ideology. We also find that at least one measure of students' ideology is a strong predictor of students' prior expectations of the predictive performance of CE theory. Students' posterior evaluations of the performance of CE theory are also related to

---

[1] We borrow the term "identity" from the contemporary humanities, where age, gender, race and class—collectively referred to as "identity"—are viewed as the primary determinants of values and beliefs in the realm of the sociology of knowledge. This is a central view of what is commonly called postmodernism. An unusually clear definition of postmodernism, with standard references, can be found in the section "Postmodernism: Main Ideas" of Agger (2001). Later in this paper, we focus on just these particular sociodemographic characteristics in our data analysis, referring to them as "primary identity variables."

ideology, the prediction errors they observe in a classroom Double Auction Experiment, and a small number of identity variables. The findings are consistent with the notion that different student audiences experience different degrees of "surprise" because of their different ideological leanings; but the independent contribution of identity, above and beyond identity's relationship to ideology, is less robust. Moreover, identity variables leave a great deal of the variance in ideology unexplained, which suggests that average sociodemographic differences between classes would have to be quite pronounced in order for the typical pedagogical DA user to notice different degrees of "surprise" in different classes.[2]

Correct or not, economics has a reputation for being a conservative discipline, and many observers regard its theoretical world as more the product of ideology than empiricism. For our purposes, whether this view is correct does not matter. What matters is how widespread this view is, and whether the prevalence of this view correlates with ideology in the population at large. Then a tendency for "left-leaning" students to distrust what economists teach them relatively more than do "right-leaning" students would be unsurprising. This is the assumption that inspires the experiment we report on here, and it seems to be confirmed by the evidence.

I. Experimental Design, Procedures, Measurements and Variables

The subjects for this experiment were students enrolled in six introductory courses in microeconomics and macroeconomics at the University of Houston in the summer of 2000. Experimental sessions took place in these classes at their regularly scheduled times. At the University of Houston, the syllabi of both courses contain an introduction to supply and demand and CE theory. We arranged with several course instructors to visit classes shortly after CE theory had been introduced. In introductory remarks, we told students in each class that most of

---

[2] The first author initially noticed the difference in comparing his experience with classroom experiments at the University of Houston and Bowdoin College. The former is a large urban state university—according to *U.S. News*

the session would be devoted to an experiment that is commonly used across the country as a pedagogical tool, in classes just like theirs. They were also advised that certain aspects of the class session were for research purposes.

Students received a playing card, and were instructed to use their card as their "name" when filling out forms throughout the session.[3] Students then received the Initial Survey (see Appendix A1) which included standard socioeconomic or "identity" variables, such as age, gender, race or ethnicity, family income and so forth. We also included questions that we thought might help to explain differences in students' ideological leanings, such as whether their families own stock, include union members or have suffered from various adverse events or calamities in the past ten years. Table 1 contains information about variables constructed from survey responses, such as response categories and coding procedures, as well as their sample distributions.

The survey also elicited students' ideological leanings by two different methods. First, students provided an ideological self-report, by placing themselves on a seven-point scale from very liberal (scale value = 1) to very conservative (scale value = 7). The National Election Surveys use this scale, which many political scientists analyze in studies of voting behavior. We call this variable IDEOL_SR (for "Ideology—Self Report").

Second, students answer a battery of eleven statements about general economic policy issues that we selected from surveys used by the National Opinion Research Center. These are questions 20 through 30 in the survey (see Appendix A1). Students are asked to express their

---

and *World Report's* annual issue on colleges and universities, the third "most diverse" campus in the United States. The latter school, by contrast, is an elite private liberal arts college in Maine.

[3] The motivation for this procedure was explained to students as follows. We wanted them to remain anonymous throughout and after the experiment, so that they would feel free to answer all survey questions and to express their expectations and evaluations of the predictive performance of CE theory frankly (and without worry of academic consequences). But since three separate forms were used, we needed a consistent identifier for each subject—hence, each subject's card would be her "name" during the session.

agreement or disagreement with each of these statements on a five-point scale, and five[4] of these eleven responses are combined to create an aggregate "economic ideology" scale. We call this variable IDEOL_MQ (for "Ideology—Multiple Questions").

When every student had finished the Initial Survey, they were collected and it was announced that twelve of the students present would be asked to volunteer to be the "subjects" in the DA trading experiment that would be conducted shortly. DA trading experiments are so well-known that we will not describe the details of such experiments here (see Davis and Holt 1993); the market rules are a stylized version of those found in multilateral "open outcry" markets such as the futures markets of the Chicago Mercantile Exchange. In preparation for this, all students were asked to follow along as the experimenter described how these volunteers, in randomly assigned roles as buyers or sellers, would make money buying and selling units of an imaginary good in the coming market. The experimenter did this in the style of a lecture, using overhead projector slides to illustrate how volunteers could make money when they managed to make a sale (if they were sellers) or make a purchase (if they were buyers) at some contract price. The rules of the DA market , including the way in which pairs of volunteers could agree to a contract price, were then explained in a similar manner.

We then told the class that CE theory, which they had recently studied in their class, could be used to predict what would happen in the coming experiment. Specifically, the students were reminded that CE theory makes predictions about trading prices, the number of units traded, and the total surplus (i.e. total trading profits) earned by market participants. A sealed envelope was then produced, and the experimenter said that he had used CE theory to make those predictions for the experiment, and that these had been sealed in the envelope. This was handed to one student in the class for safekeeping until after the experiment was complete.

---

[4] We discuss our reasons for using only five of these eleven responses in the next section.

Twelve volunteers were then recruited, randomly sorted to roles as buyers or sellers, and given forms to record their trades, corresponding to what they had seen during the earlier instructions. Each buyer's form showed her own unit value, and each seller's form showed his own unit cost (each subject could buy, or sell, at most one unit per trading period). Space was provided for recording contract prices and for calculating any resulting profits. The record forms provided space for one practice round (during which no money would be at stake) and for five subsequent trading rounds for actual money.

The practice round of the experiment was then conducted, and immediately after this the Expectations Survey (see Appendix A2) was handed out to all students in the room—subjects and observers alike. This survey asked students to indicate how accurately they thought CE theory would predict

(a) the average trading price;

(b) the number of units traded; and

(c) the efficiency of the market in the final trading period.

The CE prediction of market efficiency was phrased in terms of maximization of surplus or of potential trading profits.

Students were asked to make their judgment on a four-point scale for each of these three predictions. These judgments are coded as 1 (very inaccurate), 2 (somewhat inaccurate), 3 (somewhat accurate) and 4 (very accurate). The sum of these coded expectations, which ranges from a low of 6 to a high of 12 in our sample, is one of our chief dependent variables. We call this variable ANTESUM (the sum of ex ante evaluations) and Table 2 shows its distribution.

When students had completed the practice period and Expectations Survey, five real-money trading rounds commenced. After the conclusion of the fifth round, the end of trading

5

was announced.[5] The student holding the predictions envelope was then asked to open the envelope and read the predictions. Then the experimenter briefly reviewed how CE theory was used to make these predictions, and compared those predictions to what had actually happened in the fifth and final round, noting where CE theory had been right and, where it was not, saying exactly how wrong it had been. Finally an Evaluations Survey was distributed to all students (see Appendix A3). This instrument elicited the same information as the previous Expectations Survey, using the same response categories and virtually the same wording (except in a retrospective rather than prospective format). The sum of students' coded ex post evaluations of the predictive performance of CE theory is POSTSUM (sum of ex post evaluations). In our sample POSTSUM ranges from 6 to 12. This dependent variable plays an important role in our analysis, and its distribution is shown in Table 2.

We use four other variables in this study. One of these is DELTASUM, which is simply the difference between POSTSUM and ANTESUM: It is the change in students' judgments brought about by the DA experiment. Second, we also know whether a subject was an observer or a subject or "player" in his classroom DA experiment: The dummy variable PLAYER takes the value of 1 when a student was one of these players. Finally, two dummy variables ERROR1 and ERROR2 represent what students actually saw in their classroom. If CE theory made at least one mistake, ERROR1 took the value of 1. That is, if any of the price, quantity or total surplus predictions of CE theory was incorrect in the session that a subject observed (this was true in four of the six classes), ERROR1 = 1 and is zero otherwise. Similarly, if CE theory made

---

[5] The five-round total was chosen for two reasons, one practical and one methodological. Practically speaking, doing the experiment with more than five trading rounds would have been difficult in the 110 minutes available in a single summer class period. From the viewpoint of received results on DA markets, five periods is not, generally speaking, quite enough time for inexperienced subjects to reliably converge on CE predictions. This makes it likely that students will actually see variations in the evidence for the CE theory across classrooms in the final trading period. While this increases variance, which is a nuisance, it also provides a reasonable "sanity test" of sorts. We can check whether students' revisions of their opinions are reasonable in other respects than how ideology is related to them. If not—if for instance students' revisions of their opinions are unrelated to any prediction errors they see— then suspicion about the experimental procedures in a more general way would be warranted. Later in the paper we show that students' ex post evaluations do relate to observed errors in a minimally reasonable way.

two mistakes, ERROR2 took the value of 1 (this occurred in two of the six classes). CE theory was never wrong about all three predictions during these classroom sessions. Table 2 shows the distribution of DELTASUM, PLAYER, ERROR1 and ERROR2 in the sample.

II. Results

We begin by examining our measure of economic ideology IDEOL_MQ from a psychometric perspective. We believe this examination establishes that this scale has desirable psychometric properties of unidimensionality, internal consistency, reliability and two forms of validity. We then examine a few simple cross-tabulations of several variables. This allows us to make two general points. First, it will be clear that students respond to the <u>presence</u> of evidence against CE theory in a reasonable manner, though they seem insensitive to the <u>magnitude</u> of that evidence. Second, ideology clearly affects students' expectations and evaluations of CE theory's predictive performance. Finally, ideology has no clear effect on the change in students' judgments of CE theory's empirical relevance. We then turn to a more detailed multivariate analysis and note some relationships between ideology, identity variables and student opinions. In general, this analysis supports the conclusion that ideology is a strong predictor of students' opinions about CE theory's relevance.

II-A. Psychometric Properties of the Measure of Economic Ideology

We begin by discussing how we arrived at our economic ideology measure IDEOL_MQ, which is based on five of the eleven items (questions 20 through 30) shown in Appendix A1. The five items actually summed to construct IDEOL_MQ (questions 20, 22, 23, 24 and 28) appear in bold italic in Appendix A1. Our selection of these five items out of the eleven available ones was based on various psychometric criteria (Robinson et al. 1991) which we discuss below.

Note that an "economically conservative" ideology would sometimes be expressed by agreement and sometimes by disagreement, with the statements in questions 20 through 30. Therefore, we start our analysis by recoding subjects' responses (on the scale 1 to 5) so the "most conservative" response is 5 and the "least conservative" response is 1.[6] This yields an inter-item correlation matrix for the eleven item responses with mostly positive correlations. The few negative correlations are statistically insignificant.

A scale that measures a single underlying personality trait of subjects is usually considered better than one that measures a combination of different traits. Formally, this means that a factor analysis of a set of items should find that all items "load" on a single "important" factor. The "importance" of underlying factors is usually judged by examining the eigenvalues of the correlation matrix of the items (these eigenvalues are the measures of each potential factor). One rule of thumb is that the important factors are those associated with eigenvalues greater than one. By this rule of thumb, the eleven ideological items we used in our survey appear to load on three important factors. Responses to questions 20, 22, 23, 24 and 28 load most heavily on the primary factor among these three important factors, while the answers to the other six questions load most heavily on one of the two less important factors. Moreover, a second factor analysis of just the responses to questions 20, 22, 23, 24 and 28 shows that these items load on just one important factor. This suggests these five items mostly measure a single underlying personality trait, while the remaining items may mostly measure other traits.

Another general criterion for evaluating a set of items destined for scale construction is internal consistency. We note two common measures of internal consistency: (a) The average inter-item correlation in that set, and (b) Cronbach's α, a measure of inter-item covariances such that $\alpha = 1$ when all items covary perfectly and $\alpha = 0$ when all item covariances are zero

---

[6] It is tempting to refer to low scores as "liberal" and high scores as "conservative," and shortly we will adopt this convention for convenience. However, we caution readers that the assumption that the personality traits "liberal"

(Cronbach 1951).[7] The average inter-item correlation among the eleven item responses from our survey is just 0.09, whereas it is 0.22 among the five item responses we have selected for our scale. Average inter-item correlations below 0.10 are considered poor, while those in excess of 0.20 are considered good (0.30 or greater is considered excellent; see Robinson et al. 1991). Cronbach's α among the eleven survey items is 0.48, while it is 0.59 among the five selected items. A value of Cronbach's α near 0.60 seems to be a minimal desirable standard for a set of items destined for scale construction (Robinson et al. 1991). By both of these internal consistency measures, then, we feel more confident building a scale from the five selected items than from all eleven items.

Another desirable characteristic of a scale is high test-retest reliability. This is simply the correlation between two identical versions of a scale administered at two different time periods to the same group of subjects. After our experiment was completed, we measured this characteristic in a new sample of forty-one similar subjects (University of Houston undergraduates in introductory economics classes). The eleven items in questions 20 through 30 were administered to these subjects twice, separated by a period of seven to ten days. Test-retest reliability of the scale constructed from all eleven items was 0.81, while for the scale constructed from the five item subset it was 0.87. So again, by this criterion, the scale based on the five-item subset seems preferable to that based on all eleven items.

Finally, a good scale should have "validity." Several types of validity are discussed in the literature on testing and measurement. One of these is known-groups validity, which is the property that a scale can significantly predict group membership for at least some "known groups." In our sample, logit models containing the lone predictor IDEOL_MQ significantly

---

and "conservative" are two ends of one continuum is controversial (Knight 1999).

[7] The formula for Cronbach's α is $\alpha = [k/(k-1)]\{1-[(\Sigma S^2_i)/S^2_{sum}]\}$, where k = the number of items, S^2_i = the sample variance of item i, and S^2_{sum} = the sample variance of the sum of the k items. Cronbach's α is mathematically equivalent to the following (perhaps more intuitive) average: Split a set of items into two groups; compute the sum

predict whether subjects are white, black, asian, or from union households (p-values of less than 0.0001, 0.06, 0.01 and 0.01), though it does not significantly predict their sex. IDEOL_MQ also significantly predicts subjects' self-reported ideology IDEOL_SR ($p<.0001$) in an ordered logit model. However, the correlation between IDEOL_MQ and IDEOL_SR is just 0.35. While highly significant ($p<.0001$) this correlation is low enough that the two measures of ideology may perform quite differently in the following analysis.

Convergent validity is the property that a scale actually performs its theoretically specified task. In our case, this is whether IDEOL_MQ predicts our subjects' prior beliefs about the descriptive relevance of CE theory. Much of the rest of this paper is devoted to this question and related issues. To close this subsection, we think that our scale IDEOL_MQ passes some commonly accepted psychometric criteria quite well, though we are not wholly satisfied with it (for instance, the value of Cronbach's α of 0.59 is only minimally acceptable). Moreover, we caution readers that scales such as this are not usually regarded as solid until they have proven their convergent validity in several studies.[8] Finally, we have not addressed several interesting psychometric issues here, though we will return to some of these in our conclusions. Some of our results depend on whether we base our analysis on the five-item scale or the scale based on all eleven questions 20 to 30. Where this is the case, we mention it in footnotes to follow.

II-B. The Dependent Measures and their Basic Relationships

First we ask whether, and to what extent, the students behaved in a reasonable manner as regards updating their beliefs in the light of evidence. In Table 3, we cross-tabulated the number

---

of the item responses in each group; calculate the correlation between these two sums; and average this correlation against all possible divisions of the original set into two groups.

[8] After beginning our work we discovered one attractive and well-tested alternative to our own scale. This is the "Capitalist Values Scale" of McClosky and Zaller (1984). The scale intrigues us and it has good psychometric properties—perhaps the best amongst those scales that attempt to measure economic values (Feldman 1999). However, this scale is based on 26 items, making it potentially unwieldy for a research task like ours which involves many other measurements and procedures in a limited time period.

of prediction errors students saw with their posterior evaluations of CE theory (the variable POSTSUM). Students who saw no errors were much happier with CE theory's performance than were students who saw some errors. For instance 63% of the students who saw no errors gave CE theory the highest "grade" possible (POSTSUM=12) while only 18% of the students who saw some errors did. However, errors beyond the first one seem to have little effect, as the distribution of "grades" given by students who saw just one error and students who saw two errors are not very different. Chi-square tests confirm these impressions (see the notes to Table 3). We conclude that students respond to the <u>presence</u> of error in a minimally reasonable way, which suggests that students' judgments are not frivolous or completely random. However, the second error, when observed by students, seems to have little marginal effect. (Three errors were never observed.) This implies that most students view the three predictions of CE theory as a whole, and judge them collectively as "correct" or "incorrect" according to whether the theory makes any prediction errors.

Table 4 is a similar cross-tabulation of the number of errors students observed; here, however, the column variable is the change in students' judgments of the empirical relevance of CE theory (the variable DELTASUM). In looking at how students respond to observed errors, this analysis may have an advantage over the previous one since it tries to remove the effect of students' prior judgments on their posterior judgments. However, the results in Table 4 largely echo those in Table 3. Almost half (45%) of those students who saw no errors changed the "grade" they gave CE theory by 3 or more scale values, while only 21% of those students who saw one or more errors changed their "grade" by this large of a value. Witnessing more than one error apparently had little effect on changes in "grades." And once again, chi-square tests confirm these impressions (see the notes to Table 4). Therefore, we draw the same conclusions as before: students respond to the presence of error in a reasonable fashion, but they seem insensitive to the exact number of errors they have observed.

Tables 5, 6 and 7 examine one of our central questions—whether students' judgments depend on the ideology they bring with them to class. In this initial analysis, we focus on the economic ideology scale we constructed from the battery of eleven questions in the initial survey—the variable IDEOL_MQ. In order to construct cross-tabulations, we divide the range of this scale into three categories: Scale values less than or equal to 13 (40 observations), values from 14 to 16 inclusive (49 observations) and values greater than or equal to 17 (39 observations).[9] For convenience we refer to these ideology categories as "liberals, moderates and conservatives," though perhaps some other labeling would be more satisfactory.[10]

If ideology effects show up anywhere, they most likely (and indeed most reasonably) will show up in students' prior expectations of the predictive performance of CE theory. Table 5 cross-tabulates students' ideology categories with their prior expectations of the predictive performance of CE theory (the variable ANTESUM). Inspection of Table 5 suggests that ideology correlates with students' prior expectations of the predictive performance of CE theory. First note that in the entire sample, the median value of ANTESUM is 9. The percentage of students whose value of ANTESUM falls above this median is 13% for the liberals, 35% for the moderates and 54% for the conservatives. The more conservative students are more likely to believe that CE theory will do well in the coming experiment, and chi-square tests show that this is a significant effect (see the notes to Table 4).

Table 6 shows a similar cross-tabulation, except that ideology categories are compared according to the distribution of students' ex post evaluations of the predictive performance of CE theory (the variable POSTSUM). Since expectations seem related to ideology, and since students do seem to update their expectations in response to the prediction errors they have

---

[9] We first attempted to choose these ranges to split the total observations into three roughly equal subgroups; but because this was not possible, the upper and lower groups ("conservatives" and "liberals") were chosen to be of nearly equal size. It is interesting, but only serendipitous, that this results in the central group ("moderates") being centered on the value 15, which is also the mean value of the scale for a student who, on average, neither agrees nor disagrees with the eleven statements about economic policy and responsibility.

observed in a minimally reasonable way, one might expect to see some effect of ideology on ex post evaluations. Table 6 confirms this effect. Again consider the percentage of students whose ex post evaluations lie above the median prior expectation of 9 (from Table 5). This is 58% for the liberals, 84% for the moderates and 95% for the conservatives. The conservatives seem much more generous in their ex post evaluations of CE theory, and chi-square tests bear this out (see the notes to Table 6). We conclude that the relationship between ideology and ex post evaluations is strong, just as it is for ex ante expectations.[11]

To conclude our analysis of simple cross-tabulations, Table 7 shows the relationship between ideology categories and the change in students' judgments (the variable DELTASUM). Students' ideological stances appear to have little relationship with how they changed their evaluation of the DA. For instance, consider the percentage of students who reported a positive change in their judgments (students for whom DELTASUM > 0): This is 68% of the liberals, 69% of the moderates and 71% of the conservatives. Thus, while Table 4 shows that changes in judgment are sensitive to observed prediction errors, Table 7 suggests that changes in judgment are insensitive to ideology. Chi-square tests bear this out (see the notes to Table 7).

We have not said anything so far about our second ideology measure—the ideological self-report variable IDEOL_SR. As will become clear in the next subsection, this is because the measure we have been using, IDEOL_MQ, performs much better in our multivariate analyses. This is no surprise; it is well known that scales meant to represent underlying attributes of subjects are more reliable when they are constructed from a battery of questions rather than from a single question. And as mentioned earlier, while the two ideology measures are significantly

---

[10] See footnote 6.

[11] The relationship between ideology and ex post evaluations virtually disappears, however, if all eleven items in questions 20 to 30 are used to construct IDEOL_MQ: In particular, the chi-square p-value is just 0.11 when just only liberals and conservatives are compared. See subsection II-A above for a discussion of the psychometric advantages of the five item scale and the disadvantages of the eleven item scale.

correlated, that correlation is a relatively low 0.35—low enough that it should be no surprise that one performs much better as a predictor of judgments concerning CE theory.

II-C. Multivariate Analysis: Ideology, Identity and Judgment

We now turn to a multivariate analysis of our dependent measures. We would like to know three things. First, can socioeconomic or "identity" variables predict ideology, and to what extent? Second, to what extent are expectations and ex post evaluations of CE theory's predictive success related to ideology and identity variables? Finally, do identity variables tell us anything about expectations and ex post evaluations above and beyond ideology alone?

To answer these questions, we will examine our dependent measures using two different sets of regressors. In each case, Model 1 uses a subset of the variables found in Table 1, which we call the primary identity variables. These are linear and quadratic terms in AGE, the dummy variables WHITE, BLACK and ASIAN (for race), the dummy variable FEMALE (for gender) and the income class and union household dummy variables INCOME1 to INCOME4 and UNION (for class).[12] "Model 2" is more expansive, containing all of the variables listed in Table 1. The purpose of Model 2 is to see whether any effects of ideology on beliefs weaken or disappear when more explanatory variables are added to Model 1. Where the dependent measure is a belief, or changes in beliefs, the ideology measures IDEOL_MQ and IDEOL_SR are regressors in both models; and where the dependent measure involves ex post evaluations, the observed error dummies ERROR1 and ERROR2 are also included as regressors. The results are summarized in Tables 8, 9, 10 and 11, which also display p-values for several tests of joint significance for various subsets of regressors.

---

[12] Much contemporary thinking in humanities departments views age, race, gender and class as the main determinants of values, ideology and beliefs. Indeed, our frequent use of the word "identity" to describe sociodemographic characteristics derives from the academic humanities, and Model 1 assesses the importance of these "primary identities" to ideology and beliefs.

Our ideology scale takes many values in our sample, and few observations are even close to its lower and upper bounds of 5 and 25 (no observations are exactly at these bounds, either). As a result, we think it is reasonable to treat IDEOL_MQ as a continuous and unlimited dependent variable and hence to model it using ordinary least squares. Our belief measures ANTESUM and POSTSUM, on the other hand, take on just the seven discrete values of the integers from 6 to 12; moreover, there are significant numbers of observations that are equal to the upper bound of 12 for both of these measures. Therefore, we argue that these variables cannot reasonably be treated as continuous or unbounded, so that OLS methods of analysis are inappropriate. Therefore, we use an ordered logit model to examine these dependent variables. The ordered logit takes care of the heaping of observations at upper or lower bounds in much the same general statistical manner that a Tobit model does for an interval scale measure. It is also appropriate for discrete ordinal measures such as our belief measures.[13]

II-C-1. Ideology

We begin by asking: Can identity variables predict ideology, and to what extent? The first three columns of Table 8 report the results of an OLS regression of IDEOL_MQ on the primary identity variables defined above. Racial differences in ideology are significant as a group, and are substantial: The estimated difference between the white and asian students on this ideology scale is 2.56 points, and the standard deviation of the scale is 3.18. Also note the significance of the estimated effect of having a union member in one's family, which fits with common expectations: These union household students are less conservative than other students. Besides race and union status, no other identity variables are significant as a group. For instance, while two of the four income range dummies are individually significant, the four dummies are

---

[13] We attempted one treatment using a odds-ratio scale for ex ante and ex post responses. However, students seemed unfamiliar with the odds concept, although some research from the 1960s recommended this approach.

not jointly significant and the estimates themselves suggest no obvious pattern of relating income and ideology.

The last three columns of Table 2 show our results for the more expansive Model 2. The union status effect on ideology is not robust to the inclusion of the thirteen extra variables in Model 2, though race is. Finally, the variable CALAMITY appears as a significant predictor of ideology. CALAMITY is a count of the number of disruptive events a students' household has endured during the previous ten years. These include serious health problems, crimes, unemployment spells, natural disasters, fires, deaths in the family and divorces. On balance, the more of these a student's household has endured, the less conservative the student will be.

We conclude that our measure of ideology is significantly related to at most a few identity variables, as well as to whether a student's family has suffered various misfortunes. We also note that at most about half of the explainable variance in ideology is explained by sociodemographic variables. The explainable variance of a scale, in terms of maximum achievable R^2, is equal to that scale's true reliability. Cronbach's α is a lower bound on true reliability, while test-retest reliability is an upper bound on true reliability. Even in Model 2, where <u>all</u> the variables in Table 1 (except IDEOL_SR) are included, the R^2 is just 0.33. Given our estimates of Cronbach's α of 0.59 and test-retest reliability of 0.87, this means that at most 56%, and perhaps as little as 38%, of the explainable variance in IDEOL_MQ can be predicted from our sociodemographic measures. From a purely statistical viewpoint, then, it will not be surprising if IDEOL_MQ predicts beliefs even when all other variables are included in our models of beliefs—to which we now turn.

II-C-2. Expectations

Does ideology and identity predict students' prior expectations of CE theory's predictive performance (ANTESUM) and if so, to what extent? The first three columns of Table 9 report

our Model 1 ordered logit analysis of ANTESUM, where the primary identity variables and our ideology measures IDEOL_MQ and IDEOL_SR are the potential predictors. The significant predictors of ANTESUM are IDEOL_MQ and FEMALE. As Table 5 suggests, the more conservative a student is, the more they expect of CE theory. And women are more skeptical than men are, even controlling for ideology. No other variables are significant, however, either individually or in groups. In particular, note that IDEOL_SR adds no explanatory power to the model, suggesting that self-reported ideology is not the right measure for our purposes.

The last three columns of Table 9 show that the effects of IDEOL_MQ and FEMALE are robust to the inclusion of extra regressors. But some of these extra regressors are significant in their own right: These are WORKEXP, VOLEXP, UNDER and APATHY, all of which are positively related to ANTESUM. WORKEXP and VOLEXP are dummy variables representing whether a student has held a full-time paid job and any volunteer job: These students expect CE theory to do better than other students do. UNDER is a dummy variable representing first and second-year undergraduates (underclassmen); perhaps the juniors and seniors are more cynical than these newer students. Finally, APATHY is a dummy variable that codes for students who decline to categorize themselves as liberal, moderate or conservative: They choose the response "Don't Know or Haven't Thought Much About This" in the ideological self-placement question. Political scientists have found that these respondents are not simply "moderate" or "middle of the road" in their behavior, so it is perhaps unsurprising that they differ in some way here as well.

The significance and robustness of FEMALE is worth considering closely, especially since FEMALE was not a significant predictor of IDEOL_MQ (see Table 8). This is the sole element of identity which predicts prior beliefs, even when ideology is controlled for; and it is not itself a predictor of our ideology measure. This suggests that womens' and mens' prior beliefs differ in some way that cannot be reduced to, or summarized by, differences in their economic ideologies.

17

II-C -3. Ex Post Evaluations

Are students' ex post evaluations of CE theory's performance (the variable POSTSUM) related to the number of prediction errors they observe, their ideology and identity, and if so, to what extent? The first three columns of Table 10 report our Model 1 ordered logit analysis of POSTSUM, where the error dummies ERROR1 and ERROR2, the primary identity variables and our ideology measures IDEOL_MQ and IDEOL_SR are the right-hand side variables. The three significant variables of ex post evaluations here are IDEOL_MQ, ERROR1 and UNION. As in the models of expectations, IDEOL_SR does not improve goodness of fit for the estimation of ex post evaluations when IDEOL_MQ is in the model. As expected from the results in Table 3, ERROR1 predicts harsher evaluations of CE theory, while ERROR2 does not. That is, the presence of prediction errors, and not the magnitude, governs students' evaluations of CE theory. UNION is the only significant predictor of POSTSUM among the primary identity variables.

The last three columns in Table 10 show that all of these results are robust to the inclusion of extra regressors. Additionally, two of the extra regressors—CALAMITY and PLAYER—are themselves significant predictors of ex post evaluations. CALAMITY predicts more generous ex post evaluations of CE theory, while PLAYER predicts harsher ones. We are puzzled by the CALAMITY effect. The PLAYER effect, on the other hand, is intriguing. Social psychologists routinely draw a distinction between the ways in which actors and observers see the social world. Here, the players are students who are the buyers and sellers in the experiment while other students are observers. Players may have richer experiences of the Double Auction than observers. As a result players may be less willing to pin the cause of its outcomes on any single set of influences—including those influences specified by CE theory. This may lead players to be less generous in their ex post evaluations of the theory than are the students who simply observe the action.

II-C-4. Changes in Judgment

Finally we ask: Are changes in students' judgments of CE theory's predictive performance (DELTASUM) predictable from the errors they observe, their ideology and their identity, and if so to what extent? In a sense, using DELTASUM controls for individual idiosyncracies by by differencing ANTESUM and POSTSUM. Thus one might regard inclusion of subject-specific characteristics a model of DELTASUM as pointless. However, this view only makes sense under the restrictive assumption that subject-specific effects appear as constants added to both ANTESUM and POSTSUM. Moreover, different subject characteristics are significant predictors of these dependent measures. For instance, FEMALE significantly predicts ANTESUM but not POSTSUM, while UNION significantly predicts POSTSUM but not ANTESUM; and there are six other similar predictors (WORKEXP, VOLEXP, CALAMITY, PLAYER, UNDER and APATHY). If the asymmetric influence of these eight variables on ANTESUM and POSTSUM is real, we should expect all eight of them to show up as significant predictors of DELTASUM. So one way to view the modeling of DELTASUM is as a robustness check for the apparent asymmetric influence of these effects.

The first three columns of Table 11 show the results for the Model 1 ordered logit analysis of DELTASUM. As was suggested in the analysis of Table 7, the ideology variables IDEOL_MQ and IDEOL_SR are neither individually nor jointly significant in this model. This encourages the view that subjects are at least minimally reasonable: While ideology may provide reasonable heuristic guidance for prior beliefs, we think it would be harder to make such an argument for changes in beliefs. As in the previous analysis of POSTSUM, ERROR1 is significant and of the expected sign, while ERROR2 is insignificant. And finally, both FEMALE and UNION are significant predictors of changes in beliefs, as their asymmetric influences on ANTESUM and POSTSUM suggest they should be. We can safely say that

women change their beliefs more than men, largely because they start out more skeptical than the men but end up with the same posterior beliefs.[14] Similarly, we can conclude that students from union households change their beliefs less than others, not because they start out less skeptical than others, but because they end up more skeptical than others.

The last three columns of Table 11 show that these conclusions are robust to the addition of extra regressors. They also show that the asymmetric work experience effects WORKEXP and VOLEXP (significant predictors of ANTESUM but not POSTSUM) are both significant and of the signs expected on the basis of the results in Tables 9 and 10. We surmise that work experience, whether paid or on a volunteer basis, predicts smaller changes in beliefs, not because these students end up with harsher posterior beliefs but because they begin with more generous prior beliefs. Finally, we note that no other predictors are significant—including the remaining four predictors that had asymmetric effects on ANTESUM and POSTSUM. Because of this, we view the significance of the predictors CALAMITY and PLAYER in the model of POSTSUM, as well as the significance of the predictors UNDER and APATHY in the model of ANTESUM, as not predictive of changes in beliefs—even though they seem to predict either prior or posterior beliefs. As a result, we view these four variables as less robust predictors of beliefs.

II-C-5. Follow-Up Evidence on Two Potential Problems With Our Ideology Measure

We now address two potential problems with our ideology measure. First, it is possible that ideology as we measure it is correlated with some other unobservable trait that is the true determinant of students' beliefs. One such possible trait is trust in authority. The extensive research literature on authoritarianism and authoritarian personality (Christie 1991) does not have a scale that captures exactly what we have in mind. We have in mind a personality trait

---

[14] Ortmann and Tichy (1999) observed that in the first round of a prisoners' dilemma game women played the cooperative strategy more than men, but this difference disappeared in later rounds.

which is the degree to which an individual is willing to trust a person who, according to traditional hierarchical notions, they should trust—parents, experts, religious leaders and teachers, for instance. Clearly, this could have something to do with students' willingness to trust what any professor teaches them (much less an economist). If conservatives are more likely to trust traditional sources of authority, it is possible that our ideology measure is simply an instrument for this trust in authority, and this could be the reason that our ideology measure predicts beliefs about CE theory in an economist's classroom.

Another possible concern is bias due to Socially Desirable Responding or SDR that afflicts much survey work and many personality measurements (Paulhus 1991). Suppose for instance that students vary in the degree to which they wish to please us, the economists conducting the experiment. Also suppose that all students believe that economists like us will be most pleased by (a) relatively conservative responses to items, and (b) relatively generous beliefs about CE theory. In this case, the relationship between measured ideology and measured beliefs would be exactly the same as in our study; but this would be a pure artifact of students' differing motivations to please us. Thus different propensities for SDR—rather than an indication of any relationship between their actual ideology and actual belief – could explain this relationship. In essence, a student with a high propensity for SDR lies to us in order to please us.

To explore these two possibilities, we administered our five ideology scale items (questions 20, 22, 23, 24 and 28 in Appendix A1) to a new group of students enrolled in introductory economics classes at the University of Houston in the summer of 2001. These students also responded to sixteen additional items, shown in Appendix A4. The first six items were written by us in an attempt to elicit students' trust for traditional sources of authority. For example, one of these items is the statement "Most 'authorities' are just people with some credentials like a degree or a license, and that doesn't mean they know anything special." As with our ideology items, students responded to these statements on a five point scale from

"strongly agree" to "strongly disagree," and we coded these responses from 1 to 5, where 1 = least trust in authority and 5 = most trust in authority, depending on the wording of each statement. We found that responses to one of these trust items—one having to do with trust for religious leaders and philosophers on ethical matters (see Appendix A4)—was completely unrelated to responses to the other items, so we exclude this item from our analysis. We sum the remaining five coded responses to create an exploratory Trust in Authority or TA scale. We also found that responses to two particular items—the one on general trust for authorities quoted above, and the one on particular trust for psychologists and social scientists—are the only two items that load strongly on a single factor. Therefore, we sum responses to these two items to create an alternative exploratory Trust in Authority scale which we label TA*.

The second set of ten items comes from the image management or IM component of the Balanced Inventory of Desirable Responding (Paulhus 1984). The IM component contains statements about socially undesirable behavior that virtually everyone has engaged in from time to time. The IM scale is essentially a count of extreme denials of these types of behavior by each student, and is thus regarded as a measure of each student's propensity to manage their image in a socially desirable manner. Each item in the IM component is a statement about the respondent herself, such as "I have never dropped litter on the street" or "I have some pretty awful habits." Students respond to these statements on a five-point scale from "not true" to "very true." Coding of responses to these statements proceeds in two steps. First, responses are coded on a five-point scale, where 5 is the most socially desirable response and 1 is the least socially desirable response, depending on the wording of each statement. Subjects then get 1 point for responses coded 5 and ½ point for responses coded 4. These points are summed to create the IM scale measuring each student's propensity for socially desirable image management.

The correlation between our ideology scale IDEOL_MQ and the IM scale of desirable responding in the new sample is −0.01 (p=.96). Therefore, we feel confident that socially

desirable responding is not affecting our measurement of ideology. This does <u>not</u> prove that socially desirable responding is absent from our data. Socially desirable responding could still explain subjects' reported beliefs to some extent; and if it does, including an IM scale would increase the power of our experimental design. But the observed relationship between ideology and beliefs is clearly not an artifact of socially desirable responding.

The evidence on trust in authority is less clear. The correlation between IDEOL_MQ and TA is a positive but insignificant 0.09 (p=.39). However, the correlation between IDEOL_MQ and TA[*], the version of the trust scale based only on the two most "psychometrically well-behaved" trust items, is a marginally significant 0.18 (p=.07). This weak correlation does not threaten our results, though another scale of trust in authority based on a larger, psychometrically sound set of items might be an effective predictor of students' beliefs, and could undermine the predictive value of ideology itself. Future research should examine these possibilities.

III.   Discussion and Conclusions

Does ideology matter? We think that, at least when it comes to what students expect and what they end up believing, the answer is clearly yes. Using one's ideology to form priors is far from silly: ideology can be a useful heuristic for making judgments in the absence of good evidence. The Bayesian tradition gives little guidance regarding the source of priors. In semi-Bayesian terms, if students begin with different priors and view similar evidence, one would expect their posterior judgments to differ as well. In that case, whatever individual differences are related to priors will also be related to posteriors; but these individual differences will not be related to individual changes in beliefs when appropriately measured. Our evidence concerning ideology is consistent with this semi-Bayesian story, while our evidence concerning three other variables—gender, union status and work experiences—is not. However, it is probably a mistake to push this Bayesian analogy too far with these data. We do not, after all, measure

23

beliefs as odds ratios and test Bayesian updating of beliefs here. Our measures of beliefs are crude, categorical and ordinal measures, rather than exact probability or odds judgments. One avenue for a future experiment would be to attempt a more precise measurement of students' prior and posterior beliefs.

Our primary purpose was to ask whether beliefs about how markets work are related to ideology and identity, and whether identity governs ideology. From the viewpoint of mainstream economics, the latter question is probably the most natural one. We think it is plausible to most economists that normative ideology is influenced, though not determined, by self interest and group interest. The adage "Where you stand depends upon where you sit" encapsulates an old idea associated with Karl Marx, which most economists probably subscribe to—whether or not they are Marxists. If socioeconomic characteristics explain interests, then it is natural to ask whether such characteristics can predict normative ideology.

Many economists part company with Marx, however, over the issue of whether ideology determines beliefs about descriptive or positive matters. To Marx, ideology is a belief system, including beliefs about facts. This epistemological tradition, which proliferates in many segments of the humanities today, has it that ideology can be "false" in many ways: An ideology usually will contain some beliefs that are false and not in the best interests of those who hold those false beliefs. A proponent of high economic rationality as a descriptive theory cannot hold this view: Put simply, it is not in the best interests of a rational agent to hold false beliefs. Indeed many economists sell their introductory economics classes to their students with this sort of promise: "Regardless of what it is you value, a clear understanding of how the economic world really works is in your best interest."[15]

We do not disagree with this statement or this sort of salesmanship, but the potential relationship between ideology and belief in the classrooms is more complex for two reasons.

First, the predictions of CE theory used here require a minimal degree of individual rationality and convexity in production. While experimenters understand that the standard DA design satisfies these conditions, subjects who are new to economics may base their expectations of the DA upon their perceptions or experiences with more complicated markets that may function less smoothly. For example, markets fail to achieve stable equilibrium outcomes or efficient allocations because of production nonconvexities. In addition, markets that work poorly or erratically may receive disproportionate attention, while well-functioning markets attract little attention. Media coverage of business focuses on stock markets, where speculative bubbles may separate asset prices from fundamental values for long periods of time. Of course such bubbles are unlikely where resale is not allowed and finite time horizons for trading are clear as in the classroom DA experiment. Thus subjects may base their expectations of how the DA performs on skewed information, at least from the view of economists who see the world as satisfying standard convexity assumptions most of the time.[16]

The second issue is the reasonable one of whom to trust, rather than an irrational one of holding false beliefs that are harmful to oneself. In the classroom context, the teaching economist is ideally perceived as the representative of a descriptive enterprise whose own theoretical world is the product of honest empiricism. The extent to which this ideal is true or, more importantly, _believed_ to be true by students is the issue. If economists themselves are widely perceived to be ideologically tainted (and from our experience, this perception is common), it will be completely unsurprising—and perhaps rational, given that perception—for those with different ideological positions to trust economists' truth claims differently. Indeed, classicists such as David Hume believed that whatever warrant we have for trusting another's

---

[15] John Matsusaka reminded us of this common piece of pedagogical salesmanship in a conversation.
[16] This could be explored by using a DA experiment with non-convex production sets, such as used in Van Boening and Wilcox (1996). Those DA outcomes exhibit considerable instability and volatility.

testimony derives from information about the testifier that is external to the testimony itself—the testifier's veracity and sincerity as judged by other information.

Economic ideology—at least as we have measured it—cannot explain all of the variation in students' beliefs about the descriptive truth of economic theory. As we have noticed, sex, union status and work experience are all related to beliefs in ways that cannot be reduced to variations in measured ideology. Yet this is a very limited set of significant effects, and two of them—being from a union family and work experience—are matters of experience that are not confined to any one age group, gender, race or class. Thus, this quartet of usual postmodern suspects adds little to an understanding of belief above and beyond what ideology alone can explain.[17] With the sole exception of gender, these postmodern predictors are insignificant predictors of beliefs when ideology itself is part of the equation.

---

[17] The second note to Tables 8, 9, 10 and 11 show that while the eleven primary identity variables are jointly significant predictors of ideology itself, they are not jointly significant predictors of either prior beliefs or changes in belief—though, as a group, they are significant predictors of posterior beliefs. As we noted in subsection II-C-4, when variables predict either prior or posterior beliefs but not both; and when those same variables do not predict changes in belief; we regard the effects of those variables on beliefs as something less than robust.

References

Agger, Ben. 2001. Critical theory, poststructuralism, postmodernism: Their sociological relevance. Univ. of Texas at Arlington Manuscript: www.uta.edu/huma/illuminations/agger2.htm.

Christie, Richard. 1991. Authoritarianism and related constructs. In J. P. Robinson et al., eds., Measures of Personality and Social Psychological Attitudes, pp. 501-572. London and San Diego: Academic Press.

Cronbach, L. 1951. Coefficient alpha and the internal structure of tests. Psychometrika 31:93-96.

Davis, Doug, and Charles Holt. 1993. Experimental Economics. Princeton, NJ: Princeton University Press.

Feldman, Stanley. 1999. Economic values and inequality. In J. P. Robinson et al., eds., Measures of Political Attitudes, pp. 159-202. London and San Diego: Academic Press.

Knight, Kathleen. 1999. Liberalism and conservatism. In J. P. Robinson et al., eds., Measures of Political Attitudes, pp. 59-158. London and San Diego: Academic Press.

McClosky, H. and J. Zaller. 1984. The American Ethos. Cambridge, MA: Harvard University Press.

Ortmann, Andreas and Lisa K. Tichy. 1999. Gender differences in the laboratory: evidence from prisoners' dilemma games. Journal of Economic Behavior and Organization (39):327-339.

Paulhus, Delroy L. 1984. Two-component models of socially desirable responding. Journal of Personality and Social Psychology 46:598-609.

_____. 1991. Measurement and control of response bias. In J. P. Robinson et al., eds., Measures of Personality and Social Psychological Attitudes, pp. 17-60. London and San Diego: Academic Press.

Robinson, John P., Phillip R. Shaver and Lawrence S. Wrightman. 1991. Criteria for scale selection and evaluation. In J. P. Robinson et al., eds., Measures of Personality and Social Psychological Attitudes, pp. 1-16. London and San Diego: Academic Press.

Van Boening, M. and N. Wilcox. 1996. Avoidable cost: Ride a double auction roller coaster. American Economic Review 86: 461-477.

Table 1

Coding of Variables from the Survey (For Variables Used in this Study)

Variable Name	Response Categories (if any)		Numerical Coding	Sample Percent
(1) AGE:			Years of age	age≤21…52%
				age≥30…10%
(2) FEMALE:	Female		1	51%
	Male		0	49%
(3) Race Dummy Variables:				
WHITE	White		1	39%
BLACK	Black		1	18%
ASIAN	Asian		1	22%
(Other—mostly Hispanic—when all three dummies are zero)				21%
(4) UNDER	Freshman or Sophomore		1	34%
	Junior, Senior or Post-Bacc.		0	56%
(5) Major Dummy Variables:				
SOCSCI	Social Science		1	23%
NATSCI	Natural Science		1	12%
PREPROF	Pre-professional		1	49%
(Other major or undecided when all three dummies are zero)				16%
(6) ECON	Had Previous Econ Course		1	76%
	This is First Econ Course		0	24%
(7) WORKEXP	Have Had Full Time Job		1	73%
	Never Had Full Time Job		0	27%
(8) VOLEXP	Experience As Volunteer		1	37%
	No Volunteer Experience		0	63%
(9) UNION	From Union Family		1	12%
	Not From Union Family		0	88%
(10) Family Income Dummies:				
INCOME1	Between 22K and 50K		1	26%
INCOME2	Between 50K and 75K		1	22%
INCOME3	Between 75K and 100K		1	13%
INCOME4	Greater than 100K		1	26%
(Less than 22K when all four dummies are zero)				13%

Table 1 (Continued)

Coding of Variables from the Survey (Used in this Study)

Variable Name	Response Categories (if any)	Numerical Coding	Sample Percent
(11) ASSET	Family Owns Stocks	1	56%
	No Stocks Owned	0	44%

(12)   Parental Education Dummies (degree of most educated parent):

COLL	College Degree	1	31%
GRAD	Graduate Degree	1	22%
(At most some college courses if both dummies are zero)			47%

(13)   CALAMITY: Count of the following number of items that respondent said his/her family had been affected by during the previous decade: natural disaster; more than one month of unemployment; major crime; major health problem; major accident such as auto accident or fire; death in the family; or divorce.

(14) IDEOL_SR	Very Liberal	1	2%
	Liberal	2	10%
	Somewhat Liberal	3	16%
	Moderate or Don't Know[*]	4	39%
	Somewhat Conservative	5	15%
	Conservative	6	14%
	Very Conservative	7	5%

(15) APATHY	[*]Chose "Don't Know" for the above ideology item	1	11%
	Otherwise	0	89%

(16)   IDEOL_MQ: The sum of a subject's coded responses to the items contained in questions 20, 22, 23, 24 and 28 from the survey—see Appendix section A1). The mean of the scale in the sample is 14.81. The distribution of this scale in the sample, into rough quintiles, is as follows:

Less than or equal to 12:	24.2%
Between 13 and 14, inclusive:	16.4%
Equal to 15 (mode and median):	19.5%
Between 16 and 17, inclusive:	19.5%
Greater than or equal to 18:	19.5%

Table 2

Other Variables Used in this Study
(From Elicitation of Expectations, Ex Post Evaluations and Other Recorded Data)

Variable Name	Definition	Observed Values	Sample Percent

(17) ANTESUM	The sum of students' accuracy expectations for CE theory's predictions of the price, quantity, and total surplus observed in the last period of the DA experiment they are about to see. For each prediction, the numerical coding is as follows:
	    4 = very accurate;
	    3 = somewhat accurate;
	    2 = somewhat inaccurate; and
	    1 = very inaccurate.
	The numbers above are summed for each of the three predictions.
		12
		11
		10
		9
		8
		7
		6
			6%
			11%
			17%
			42%
			17%
			3%
			4%

(18) POSTSUM	The sum of students' accuracy evaluations for CE theory's predictions of the price, quantity, and total surplus observed in the last period of the DA experiment they have just seen. See the definition of ANTESUM above for coding details.	12 11 10 9 8 7 6	39% 20% 19% 17% 2% 2% 1%
(19) DELTASUM	The difference between POSTSUM and ANTESUM: The amount by which students "update their prior expectations."	5 4 3 2 1 0 -1 -2	2% 6% 24% 16% 21% 21% 6% 3%
(20) PLAYER	All students in each class fill out the survey and provide their expectations and ex post evaluations, but just twelve students in each class serve as the "subjects" in the DA experiment itself. This dummy variable equals 1 for these students.	1 0	53% 47%

Table 2 (Continued)

Other Variables Used in this Study
(From Elicitation of Expectations, Ex Post Evaluations and Other Recorded Data)

	Variable Name	Definition	Observed	Sample
(21)	ERROR1	A dummy variable taking the value of 1 if the predictions of CE theory were incorrect for AT LEAST ONE of price, quantity and/or total surplus (in the classroom experiment that the subject saw).	1 0	53% 47%
(22)	ERROR2	A dummy variable taking the value of 1 if the predictions of CE theory were incorrect for exactly TWO of price, quantity and/or total surplus (in the classroom experiment that the subject saw).	1 0	27% 73%

Table 3

Cross-Tabulation of Students' Ex Post Evaluations and the Errors They Saw

Number of Prediction Errors Seen	Ex Post Evaluation of CE theory's accuracy (values of POSTSUM)							Number of Observations
	6	7	8	9	10	11	12	
0	.000	.000	.000	.150	.117	.100	.633	60
Exactly 1	.029	.059	.059	.176	.176	.294	.206	34
Exactly 2	.000	.000	.000	.206	.353	.294	.147	34

Notes: The numbers in the table are row proportions. In this unaggregated cross-tabulation, expected cell frequencies under the null hypothesis of row homogeneity are too small (in some cells) for a chi-square test to be valid. However, when the columns are aggregated into four columns (values ≤ 9 and the values 10, 11 and 12), row homogeneity is soundly rejected ($\chi^2_6 = 32.5, p < .0001$). This result is also true when the top row is compared to either the middle or the bottom row ($\chi^2_3 = 16.7, p = .0008$ and $\chi^2_3 = 22.4, p < .0001$, respectively), but not when the middle and bottom rows are compared ($\chi^2_3 = 3.22, p = .36$).

Table 4

Cross-Tabulation of Students' Changes in Judgment and the Errors They Saw

Number of Prediction Errors Seen	Ex post evaluations minus expectations (values of DELTASUM)								Number of Observations
	-2	-1	0	1	2	3	4	5	
0	.017	.017	.200	.200	.117	.317	.083	.050	60
Exactly 1	.088	.059	.294	.147	.176	.176	.059	.000	34
Exactly 2	.000	.147	.147	.294	.235	.147	.029	.000	34

Notes: The numbers in the table are row proportions. In this unaggregated cross-tabulation, expected cell frequencies under the null hypothesis of row homogeneity are too small (in some cells) for a chi-square test to be valid. However, when the columns are aggregated into four columns (values ≤ 0, the values 1 and 2, and values ≥ 3), row homogeneity is rejected ($\chi^2_6 = 12.9, p < .0443$). This result is also true, though somewhat weaker, when the top row is compared to either the middle or the bottom row ($\chi^2_3 = 6.62, p = .0849$ and $\chi^2_3 = 7.67, p = .0532$, respectively), but not when the middle and bottom rows are compared ($\chi^2_3 = 3.24, p = .36$).

Table 5

Cross-Tabulation of Students' Expectations and Ideology

Ideology Category (according to IDEOL_MQ)	Expectation of CE theory's accuracy (values of ANTESUM)							Number of Observations
	6	7	8	9	10	11	12	
Liberal	.100	.025	.225	.525	.025	.075	.025	40
Moderate	.020	.020	.143	.469	.204	.102	.041	49
Conservative	.000	.051	.154	.256	.256	.154	.128	39

Notes: The numbers in the table are row proportions. In this unaggregated cross-tabulation, expected cell frequencies under the null hypothesis of row homogeneity are too small (in some cells) for a chi-square test to be valid. However, when the columns are aggregated into four columns to solve this problem (values $\leq$ 8, the values 9 and 10, and values $\geq$ 11), row homogeneity is strongly rejected ($\chi^2_6 = 18.05$, $p = .0061$). This result remains true if the middle row is omitted so that only liberals and conservatives are compared ($\chi^2_3 = 16.16$, $p = .0011$).

Table 6

Cross-Tabulation of Students' Ex Post Evaluations and Ideology

Ideology Category (according to IDEOL_MQ)	Ex post evaluations of CE theory's accuracy (values of POSTSUM)							Number of Observations
	6	7	8	9	10	11	12	
Liberal	.025	.050	.025	.325	.125	.175	.275	40
Moderate	.000	.000	.020	.143	.245	.245	.347	49
Conservative	.000	.000	.000	.051	.205	.179	.564	39

Notes: The numbers in the table are row proportions. In this unaggregated cross-tabulation, expected cell frequencies under the null hypothesis of row homogeneity are too small (in some cells) for a chi-square test to be valid. However, when the columns are aggregated into four columns to solve this problem (values $\leq$ 9 and the values 10, 11 and 12), row homogeneity is strongly rejected ($\chi^2_6 = 20.88$, $p = .0019$). This result remains true if the middle row is omitted, so that only liberals and conservatives are compared ($\chi^2_3 = 16.19$, $p = .0010$).

Table 7

Cross-Tabulation of Ideology and the Difference Between Students'
Ex Post Evaluations and Ex Ante Expectations

Ideology Category (according to IDEOL_MQ)	Ex post evaluations minus expectations (values of DELTASUM)								Number of Observations
	-2	-1	0	1	2	3	4	5	
Liberal	.100	.025	.200	.175	.150	.225	.125	.000	40
Moderate	.000	.082	.224	.245	.122	.265	.041	.020	49
Conservative	.000	.077	.205	.205	.231	.205	.026	.051	39

Notes: The numbers in the table are row proportions. In this unaggregated cross-tabulation, expected cell frequencies under the null hypothesis of row homogeneity are too small (in some cells) for a chi-square test to be valid. However, when the columns are aggregated into four columns to solve this problem (values ≤ 0, the values 1 and 2, and values ≥ 3), row homogeneity is accepted ($\chi^2_6 = 2.56$, $p = .86$). This result remains true if the middle row is omitted, so that only liberals and conservatives are compared ($\chi^2_3 = 1.18$, $p = .76$).

Table 8

Regressions of IDEOL_MQ on Various Socioeconomic or "Identity" Variables

Regressor	Model 1: Primary Identity Variables Only. $R^2 = 0.259$			Model 2: All Variables $R^2 = 0.329$		
	Estimate	Stand. Error	Joint F-Test p-value	Estimate	Stand. Error	Joint F-Test p-value
Intercept	11.64**	5.40	—	11.75**	6.22	—
AGE	0.15	0.39		0.19		
AGE^2	−0.0032	0.0071	0.8368	−0.0040	0.0077	0.7976
FEMALE	0.64	0.54	—	0.55	0.60	—
WHITE	1.67**	0.74		2.21**	0.86	
BLACK	−0.56	0.82	0.0023	−0.47	0.95	0.0009
ASIAN	−0.89	0.81		−0.77	0.87	
INCOME1	0.52	0.87		0.24	0.92	
INCOME2	1.91**	0.91		1.45	1.05	
INCOME3	0.23	1.01	0.1250	−0.60	1.22	0.2592
INCOME4	1.55*	0.91		0.96	1.09	
UNION	−1.67**	0.84	—	−1.28	0.91	—
VOLEXP				0.51	0.61	
WORKEXP				0.38	0.65	
COLL				−0.26	0.75	
GRAD				−1.32	0.84	
ECON				0.36	0.69	
CALAMITY				−0.60**	0.23	
PLAYER		—		−0.0087	0.55	0.6281
ASSET				0.38	0.67	
SOCSCI				−0.29	0.99	
NATSCI				−1.01	1.14	
PREPROF				−0.02	0.84	
UNDER				−0.10	0.64	
APATHY				−0.69	0.92	

Notes: (1) Superscripts *, ** and *** indicate significance at 10%, 5% and 1%, respectively.
(2) The primary identity variables in Model 1 (all of the variables) are jointly significant $(F_{11,116} = 3.86, p = .0002)$.

Table 9

**Ordered Logit Model of ANTESUM on Ideology Measures and
Various Socioeconomic or "Identity" Variables**

Regressor	Model 1: Primary Identity Variables Only. Generalized R^2 = 0.189 Kendall's tau-a = 0.269			Model 2: All Variables Generalized R^2 = 0.324 Kendall's tau-a = 0.354		
	Estimate	Standard Error	Joint Wald χ^2 test p-value	Estimate	Standard Error	Joint Wald χ^2 test p-value
IDEOL_SR	−0.093	0.137		−0.070	0.145	
IDEOL_MQ	0.21***	0.06	0.0047	0.23***	0.067	0.0027
AGE	−0.33	0.25		−0.37	0.27	
AGE^2	0.0064	0.0045	0.3066	0.0077	0.0049	0.1241
FEMALE	−0.85**	0.36	—	−1.12***	0.40	—
WHITE	−0.0044	0.4870		−0.32	0.57	
BLACK	0.54	0.53	0.6537	0.90	0.61	0.1579
ASIAN	−0.060	0.518		−0.29	0.56	
INCOME1	−0.86	0.57		−0.39	0.60	
INCOME2	−0.75	0.60		0.13	0.68	
INCOME3	0.18	0.65	0.3163	1.11	0.78	0.3321
INCOME4	−0.31	0.59		0.22	0.70	
UNION	−0.22	0.55	—	−0.36	0.59	—
VOLEXP				0.89**	0.40	
WORKEXP				0.93**	0.43	
COLL				0.079	0.478	
GRAD				0.62	0.54	
ECON				0.51	0.45	
CALAMITY				0.12	0.15	
PLAYER		—		−0.14	0.35	0.0886
ASSET				−0.23	0.43	
SOCSCI				−1.03	0.64	
NATSCI				−0.88	0.73	
PREPROF				−0.86	0.54	
UNDER				0.77*	0.42	
APATHY				1.20**	0.60	

Notes: (1) Superscripts *, ** and *** indicate significance at 10%, 5% and 1%, respectively.
(2) The primary identity variables in Model 1 (all variables except the two ideology measures) are jointly insignificant (Wald χ^2_{11} = 12.58, p = .3218).

Table 10

**Ordered Logit Model of POSTSUM on Ideology Measures, Observed Errors and
Various Socioeconomic or "Identity" Variables**

Regressor	Model 1: Primary Identity Variables Only. Generalized R^2 = 0.343 Kendall's tau-a = 0.382			Model 2: All Variables Generalized R^2 = 0.415 Kendall's tau-a = 0.428		
	Estimate	Standard Error	Joint Wald χ^2 test p-value	Estimate	Standard Error	Joint Wald χ^2 test p-value
IDEOL_SR	−0.13	0.15		−0.16	0.16	
IDEOL_MQ	0.16**	0.07	0.0555	0.23***	0.07	0.0072
ERROR1	−1.71***	0.47		−1.70***	0.55	
ERROR2	0.13	0.47	0.0001	0.027	0.533	0.0011
AGE	0.13	0.26		0.33	0.29	
AGE^2	−0.0024	0.0046	0.8698	−0.0054	0.0051	0.4889
FEMALE	0.19	0.37	—	−0.014	0.407	—
WHITE	−0.10	0.50		−0.62	0.63	
BLACK	0.44	0.58	0.5212	0.38	0.69	0.2449
ASIAN	−0.42	0.54		−0.73	0.60	
INCOME1	−0.61	0.59		−0.19	0.63	
INCOME2	−0.07	0.64		0.78	0.74	
INCOME3	0.65	0.70	0.2530	1.42*	0.86	0.1761
INCOME4	−0.54	0.62		0.15	0.74	
UNION	−1.80***	0.59	—	−2.05***	0.64	—
VOLEXP				−0.06	0.41	
WORKEXP				−0.30	0.44	
COLL				−0.20	0.53	
GRAD				0.03	0.58	
ECON				0.70	0.50	
CALAMITY			—	0.34**	0.17	0.3735
PLAYER				−0.71*	0.40	
ASSET				0.30	0.48	
SOCSCI				−0.30	0.72	
NATSCI				−0.33	0.79	
PREPROF				−0.47	0.60	
UNDER				0.63	0.45	
APATHY				1.06	0.65	

Notes: (1) Superscripts *, ** and *** indicate significance at 10%, 5% and 1%, respectively.
(2) The primary identity variables in Model 1 (all variables except the two ideology measures and error dummies) are jointly significant (Wald χ^2_{11} = 19.71, p = .0495).

Table 11

Ordered Logit Model of DELTASUM on Ideology Measures, Observed Errors and Various Socioeconomic or "Identity" Variables

Regressor	Model 1: Primary Identity Variables Only. Generalized R^2 = 0.143 Kendall's tau-a = 0.240			Model 2: All Variables Generalized R^2 = 0.241 Kendall's tau-a = 0.314		
	Estimate	Standard Error	Joint Wald χ^2 test p-value	Estimate	Standard Error	Joint Wald χ^2 test p-value
IDEOL_SR	0.036	0.134		−0.0091	0.1413	
IDEOL_MQ	−0.059	0.060	0.6132	−0.016	0.064	0.9624
ERROR1	−1.09**	0.44		−0.82*	0.50	
ERROR2	0.20	0.46	0.0193	−0.086	0.511	0.0997
AGE	0.29	0.25		0.42	0.27	
AGE^2	−0.0056	0.0045	0.4304	−0.0083*	0.0049	0.1720
FEMALE	0.67*	0.34	—	0.63*	0.38	—
WHITE	−0.033	0.473		−0.029	0.564	
BLACK	0.033	0.533	0.9886	−0.14	0.62	0.9934
ASIAN	−0.14	0.51		−0.13	0.56	
INCOME1	0.03	0.55		0.0036	0.5777	
INCOME2	0.44	0.58		0.42	0.66	
INCOME3	0.35	0.65	0.7420	0.14	0.78	0.9422
INCOME4	−0.17	0.57		0.082	0.677	
UNION	−0.76	0.54	—	−1.15**	0.58	—
VOLEXP				−0.89**	0.39	
WORKEXP				−0.93**	0.41	
COLL				−0.25	0.46	
GRAD				−0.41	0.53	
ECON				0.20	0.45	
CALAMITY				0.21	0.15	
PLAYER		—		−0.24	0.37	0.3268
ASSET				0.12	0.42	
SOCSCI				1.00	0.63	
NATSCI				0.86	0.73	
PREPROF				0.58	0.53	
UNDER				−0.11	0.41	
APATHY				−0.064	0.577	

Notes: (1) Superscripts *, ** and *** indicate significance at 10%, 5% and 1%, respectively.
(2) The primary identity variables in Model 1 (all variables except the two ideology measures and error dummies) are jointly insignificant (Wald χ^2_{11} = 9.45, p = .5804).