After introducing the necessary notation, the authors begin explaining their estimation technique in earnest.

> The first step of the estimation procedure is equivalent to a Maximum Likelihood estimator applied to the individual location decisions. The estimator is based simply on maximizing the probability that the model correctly matches each household observed in the sample with its chosen house type. In particular, for any combination of the heterogeneous parameters in λ and mean indirect utilities (δ_h), the model predicts the probability that each household *i* chooses house type *h*.

As Professors Bayer and McMillan report, they use a procedure that is, in essence, the maximum likelihood technique, which, by the way, is one of several popular methods of estimating parameters (another you may have heard of is ordinary least squares). The authors then describe the second step of the procedure in a similar manner.

The length and detail with which you describe your methods will be determined by the complexity of your analysis. Did you need to estimate a price function or some other kind of function? Did you have to control for certain fixed effects in order to avoid potential biases (such as biases arising from omitted variables)? Did you have to weight the data? What about any corrections you had to make for autocorrelation, which arises when regression residuals for certain observations are related to residuals for other observations? Or for multicollinearity, a condition in which your explanatory variables are not independent of each other? Or for heteroskedasticity (i.e., changes in the variance of errors)? You are expected to describe the corrections you made for any of those problems.

18e. Reporting—and Interpreting—Your Results

The results section of an empirical paper is usually the longest. In an empirical economics paper, you test a model with data; in the

results section, you report the outcome of that test. What are the answers to your research questions? What is the relationship between your dependent variable and the several independent variables you have chosen to examine? Does the model "fit" the observed data?

In most cases, when you report the results of your analysis, you are at the same time referring the reader to a table in which the results are presented. When you present information in a table, there are at least two expectations that you need to fulfill. The first is that you explicitly introduce the table. You are expected to point out to your readers that the table exists and indicate, briefly, its general content. Usually, those two things can be accomplished in a single sentence: "Table 1 shows the incomes earned by full-time workers in the United States," or "In table 1, I present the results of the three regressions that explore the relationship between income and education." Once you introduce the table and briefly describe its general contents, you can discuss the table more particularly. That brings us to our second expectation.

The second expectation is that you should, in your narrative, identify the main points made by the data in the table, the points that most closely correspond to your research question. *The table cannot, and should not be expected to, "speak for itself."* Rather, you should explicitly tell your readers the important realities that the data show: "Table 1 reveals several significant characteristics of our sample that could affect our results: one-third of women in the sample had less than a high-school education; nearly two-thirds were unmarried; and exactly one-half had at least one child under 3," or "As expected, the coefficient on education is, in every regression, significant and positive." You may also wish to point out any counterintuitive results or results that are especially large or small. Please note, however, that you are not expected to comment on or restate every piece of information that a table contains; but you are expected to point out to your readers the "meaning" or your interpretation of the data in it. What do you most want your reader to take away from the table?

In stating your results, you should draw your reader's attention to the applicable numerical figure in the table. "As seen in column 1, the coefficient on education is 0.583 and is statistically significant at the 5 percent level." The reader should be able to look at column 1 in the table and find that figure for education.

All of this is to say that you have to describe the contents of the table in the text. You cannot simply refer to a table (or worse, not refer to it at all!) and leave it at that.

There is a useful discussion of the process in *The Student's Guide to Writing Economics*, by Robert Neugeboren. Here is the situation that Professor Neugeboren sets up; I'm quoting him almost verbatim:

> Suppose you are writing about the effect of education on wages. Your main regression places an individual's wage on the left-hand side and regressors such as education, race, gender, seniority at the individual's job, labor market experience, and state of residence on the right-hand side. You believe that the regressor of interest—education—is correlated with the error term of the wage equation: that is, more "able" people earn more at their jobs and also obtain more education. Because of this correlation between the error term and education, the measured effect of education in the regression will reflect not only the true causal effect of education on wages but also some of the effect of ability on wages. To circumvent this "ability bias" you use a separate measure as a proxy for ability. Though such a proxy is not available, assume for the sake of exposition that a special data set contains an individual's evaluation by his or her second-grade teacher. When presenting your results, you want to focus only on the estimates of the education effect and the ability effect. (39–40)

Professor Neugeboren draws up a table with hypothetical results:

Table 1 OLS Estimates of the Effect of Education on Wages. Dependent Variable: Log of Yearly Earnings, 1985–1995

	1	2	3	4
Years of Education	.091 (.001)	.031 (.003)	.086 (.002)	.027 (.005)
Ability Dummy		.251 (.010)		.301 (.010)
State Dummies Included?	No	No	Yes	Yes
No. of Observations	35,001	35,001	19,505	18,505
No. of Persons	5,505	5,505	4,590	4,590
Adjusted R^2	.50	.55	.76	.79

Note: Standard errors are in parentheses. Data are from the Tennessee Second Grade Ability Survey and Wage Follow-up, and include individuals evaluated between 1962 and 1971. The "ability dummy" equals 1 if the individual's second-grade teacher classified the individual as "able," and 0 otherwise. Each regression also includes yearly dummies, ten one-digit industry and twenty Census-defined occupation dummies, labor market experience (defined as one's age minus 6), experience squared, seniority on the current job, seniority squared, Census region of current residence, marital status, race, gender, and a dummy variable denoting whether the individual lives in a city of more than 100,000 persons. Columns 3 and 4 have fewer observations because the state of residence is not available for some individuals.

How would a discussion of the results presented in this table likely go? Here is one possibility, as presented in Professor Neugeboren's book (again, I am quoting almost verbatim):

> Table 1 presents the OLS estimates of the effect of education on wages. It shows that including a measure of ability in the wage equation dramatically lowers the predicted effect of education on

earnings. Column 1 does not include an ability measure and indicates that a year of education raises wages by 9.1 percent. Column 2 adds the ability measure; the education effect now drops to 3.1 percent. Columns 3 and 4 show that this general pattern is repeated even when state-level dummy variables are included. The estimates in table 1 are therefore consistent with the hypothesis that the OLS estimates suffer from an upward ability bias.

A few points about the preceding example are instructive. First, the discussion begins by introducing the table and indicating its content ("Table 1 presents the OLS estimates . . ."). Second, the meaning or conclusion to be drawn from the table is explicitly stated ("It shows that including a measure . . ."). Indeed, the conclusion is even restated in a different way at the end of the discussion ("The estimates in table 1 are therefore . . ."). And third, the discussion does not mention every single piece of data in the table. Instead, it selects for discussion only those data that are important for the task at hand.

19. Writing Introductions

If there is one section of an economics paper that seems to give writers the most trouble, it is the introduction. Every economics paper contains an introduction, a section that brings your reader into your paper. A good introduction gives your reader a context, a frame, for ordering and understanding the information you present in the body of your paper.

Introductions should normally answer the following questions: What is the purpose of the paper? That is, what does the paper "do"? What important economic question does it try to answer, or what issue does it try to shed light on? What contribution does the paper make, and how does it relate to previous work on the topic?

In their excellent book *Academic Writing for Graduate Students*, John Swales and Christine Feak suggest that when it comes to economics papers, it is helpful to think of your