Economic, or Practical, versus Statistical Significance

Because we have emphasized *statistical significance* throughout this section, now is a good time to remember that we should pay attention to the magnitude of the *coefficient* estimates in addition to the size of the t statistics. The statistical significance of a variable x_j is determined entirely by the size of $t_{\hat{\beta}_j}$, whereas the **economic significance** or **practical significance** of a variable is related to the size (and sign) of $\hat{\beta}_j$.

Recall that the t statistic for testing H_0: $\beta_j = 0$ is defined by dividing the estimate by its standard error: $t_{\hat{\beta}_j} = \hat{\beta}_j / \text{se}(\hat{\beta}_j)$. Thus, $t_{\hat{\beta}_j}$ can indicate statistical significance either because $\hat{\beta}_j$ is "large" or because $\text{se}(\hat{\beta}_j)$ is "small." It is important in practice to distinguish between these reasons for statistically significant t statistics. Too much focus on statistical significance can lead to the false conclusion that a variable is "important" for explaining y even though its estimated effect is modest.

EXAMPLE 4.6 PARTICIPATION RATES IN 401(k) PLANS

In Example 3.3, we used the data on 401(k) plans to estimate a model describing participation rates in terms of the firm's match rate and the age of the plan. We now include a measure of firm size, the total number of firm employees (*totemp*). The estimated equation is

$$\widehat{prate} = \underset{(0.78)}{80.29} + \underset{(0.52)}{5.44}\, mrate + \underset{(.045)}{.269}\, age - \underset{(.00004)}{.00013}\, totemp$$
$$n = 1{,}534,\ R^2 = .100.$$

The smallest t statistic in absolute value is that on the variable *totemp*: $t = -.00013/.00004 = -3.25$, and this is statistically significant at very small significance levels. (The two-tailed p-value for this t statistic is about .001.) Thus, all of the variables are statistically significant at rather small significance levels.

How big, in a practical sense, is the coefficient on *totemp*? Holding *mrate* and *age* fixed, if a firm grows by 10,000 employees, the participation rate falls by $10{,}000(.00013) = 1.3$ percentage points. This is a huge increase in number of employees with only a modest effect on the participation rate. Thus, although firm size does affect the participation rate, the effect is not practically very large.

The previous example shows that it is especially important to interpret the magnitude of the coefficient, in addition to looking at t statistics, when working with large samples. With large sample sizes, parameters can be estimated very precisely: Standard errors are often quite small relative to the coefficient estimates, which usually results in statistical significance.

Some researchers insist on using smaller significance levels as the sample size increases, partly as a way to offset the fact that standard errors are getting smaller. For example, if we feel comfortable with a 5% level when n is a few hundred, we might use the 1% level when n is a few thousand. Using a smaller significance level means that economic and statistical significance are more likely to coincide, but there are no guarantees: In the previous example, even if we use a significance level as small as .1% (one-tenth of 1%), we would still conclude that *totemp* is statistically significant.

Most researchers are also willing to entertain larger significance levels in applications with small sample sizes, reflecting the fact that it is harder to find significance with smaller sample sizes (the critical values are larger in magnitude, and the estimators are less precise). Unfortunately, whether or not this is the case can depend on the researcher's underlying agenda.

EXAMPLE 4.7 EFFECT OF JOB TRAINING ON FIRM SCRAP RATES

The scrap rate for a manufacturing firm is the number of defective items—products that must be discarded—out of every 100 produced. Thus, for a given number of items produced, a decrease in the scrap rate reflects higher worker productivity.

We can use the scrap rate to measure the effect of worker training on productivity. Using the data in JTRAIN.RAW, but only for the year 1987 and for nonunionized firms, we obtain the following estimated equation:

$$\widehat{\log(scrap)} = \underset{(5.69)}{12.46} - \underset{(.023)}{.029}\, hrsemp - \underset{(.453)}{.962} \log(sales) + \underset{(.407)}{.761} \log(employ)$$

$$n = 29, R^2 = .262.$$

The variable *hrsemp* is annual hours of training per employee, *sales* is annual firm sales (in dollars), and *employ* is the number of firm employees. For 1987, the average scrap rate in the sample is about 4.6 and the average of *hrsemp* is about 8.9.

The main variable of interest is *hrsemp*. One more hour of training per employee lowers log(*scrap*) by .029, which means the scrap rate is about 2.9% lower. Thus, if *hrsemp* increases by 5—each employee is trained 5 more hours per year—the scrap rate is estimated to fall by 5(2.9) = 14.5%. This seems like a reasonably large effect, but whether the additional training is worthwhile to the firm depends on the cost of training and the benefits from a lower scrap rate. We do not have the numbers needed to do a cost benefit analysis, but the estimated effect seems nontrivial.

What about the *statistical significance* of the training variable? The t statistic on *hrsemp* is $-.029/.023 = -1.26$, and now you probably recognize this as not being large enough in magnitude to conclude that *hrsemp* is statistically significant at the 5% level. In fact, with $29 - 4 = 25$ degrees of freedom for the one-sided alternative, H_1: $\beta_{hrsemp} < 0$, the 5% critical value is about -1.71. Thus, using a strict 5% level test, we must conclude that *hrsemp* is not statistically significant, even using a one-sided alternative.

Because the sample size is pretty small, we might be more liberal with the significance level. The 10% critical value is -1.32, and so *hrsemp* is almost significant against the one-sided alternative at the 10% level. The p-value is easily computed as $\mathrm{P}(T_{25} < -1.26) = .110$. This may be a low enough p-value to conclude that the estimated effect of training is not just due to sampling error, but opinions would legitimately differ on whether a one-sided p-value of .11 is sufficiently small.

Remember that large standard errors can also be a result of multicollinearity (high correlation among some of the independent variables), even if the sample size seems fairly large. As we discussed in Section 3.4, there is not much we can do about this problem other than to collect more data or change the scope of the analysis by dropping or combining certain independent variables. As in the case of a small sample size, it can be hard to precisely estimate partial effects when some of the explanatory variables are highly correlated. (Section 4.5 contains an example.)

We end this section with some guidelines for discussing the economic and statistical significance of a variable in a multiple regression model:

1. Check for statistical significance. If the variable is statistically significant, discuss the magnitude of the coefficient to get an idea of its practical or economic importance. This latter step can require some care, depending on how the independent and dependent variables appear in the equation. (In particular, what are the units of measurement? Do the variables appear in logarithmic form?)
2. If a variable is not statistically significant at the usual levels (10%, 5%, or 1%), you might still ask if the variable has the expected effect on y and whether that effect is practically large. If it is large, you should compute a p-value for the t statistic. For small sample sizes, you can sometimes make a case for p-values as large as .20 (but there are no hard rules). With large p-values, that is, small t statistics, we are treading on thin ice because the practically large estimates may be due to sampling error: A different random sample could result in a very different estimate.
3. It is common to find variables with small t statistics that have the "wrong" sign. For practical purposes, these can be ignored: we conclude that the variables are statistically insignificant. A significant variable that has the unexpected sign and a practically large effect is much more troubling and difficult to resolve. One must usually think more about the model and the nature of the data to solve such problems. Often, a counterintuitive, significant estimate results from the omission of a key variable or from one of the important problems we will discuss in Chapters 9 and 15.

REMARK: "Statistically significant" does not necessarily imply "economically significant." For example, suppose the CEO of a supermarket chain plans a certain course of action *if* $\beta_2 \neq 0$. Furthermore, suppose a large sample is collected from which we obtain the estimate $b_2 = 0.0001$ with $\text{se}(b_2) = 0.00001$, yielding the t-statistic $t = 10.0$. We would reject the null hypothesis that $\beta_2 = 0$ and accept the alternative that $\beta_2 \neq 0$. Here $b_2 = 0.0001$ is statistically different from zero. However, 0.0001 may not be "economically" different from zero, and the CEO may decide not to proceed with the plans. The message here is that one must think carefully about the importance of a statistical analysis before reporting or using the results.