EXAMPLE 13.8 Effect of Enterprise Zones on Unemployment Claims

Papke (1994) studied the effect of the Indiana enterprise zone (EZ) program on unemployment claims. She analyzed 22 cities in Indiana over the period from 1980 to 1988. Six enterprise zones were designated in 1984, and four more were assigned in 1985. Twelve of the cities in the sample did not receive an enterprise zone over this period; they served as the control group.

A simple policy evaluation model is

$$\log(uclms_{it}) = \theta_t + \beta_1 ez_{it} + a_i + u_{it},$$

where $uclms_{it}$ is the number of unemployment claims filed during year t in city i. The parameter θ_t just denotes a different intercept for each time period. Generally, unemployment claims were falling statewide over this period, and this should be reflected in the different year intercepts. The binary variable ez_{it} is equal to one if city i at time t was an enterprise zone; we are interested in β_1. The unobserved effect a_i represents fixed factors that affect the economic climate in city i. Because enterprise zone designation was not determined randomly—enterprise zones are usually economically depressed areas—it is likely that ez_{it} and a_i are positively correlated (high a_i means higher unemployment claims, which lead to a higher chance of being given an EZ). Thus, we should difference the equation to eliminate a_i:

$$\Delta\log(uclms_{it}) = \alpha_0 + \alpha_1 d82_t + \ldots + \alpha_7 d88_t + \beta_1 \Delta ez_{it} + \Delta u_{it}. \quad [13.32]$$

The dependent variable in this equation, the change in $\log(uclms_{it})$, is the approximate annual growth rate in unemployment claims from year $t-1$ to t. We can estimate this equation for the years 1981 to 1988 using the data in EZUNEM; the total sample size is $22 \cdot 8 = 176$. The estimate of β_1 is $\hat{\beta}_1 = -.182$ (standard error = .078). Therefore, it appears that the presence of an EZ causes about a 16.6% $[\exp(-.182) - 1 \approx -.166]$ fall in unemployment claims. This is an economically large and statistically significant effect.

There is no evidence of heteroskedasticity in the equation: the Breusch-Pagan F test yields $F = .85$, p-value = .557. However, when we add the lagged OLS residuals to the differenced equation (and lose the year 1981), we get $\hat{\rho} = -.197$ ($t = -2.44$), so there is evidence of minimal negative serial correlation in the first-differenced errors. Unlike with positive serial correlation, the usual OLS standard errors may not greatly understate the correct standard errors when the errors are negatively correlated (see Section 12-1). Thus, the significance of the enterprise zone dummy variable will probably not be affected.

EXAMPLE 13.9 County Crime Rates in North Carolina

Cornwell and Trumbull (1994) used data on 90 counties in North Carolina, for the years 1981 through 1987, to estimate an unobserved effects model of crime; the data are contained in CRIME4. Here, we estimate a simpler version of their model, and we difference the equation over time to eliminate a_i, the unobserved effect. (Cornwell and Trumbull use a different transformation, which we will cover in Chapter 14.) Various factors including geographical location, attitudes toward crime, historical records, and reporting conventions might be contained in a_i. The crime rate is number of crimes per person, *prbarr* is the estimated probability of arrest, *prbconv* is the estimated probability of conviction (given an arrest), *prbpris* is the probability of serving time in prison (given a conviction), *avgsen* is the average sentence length served, and *polpc* is the number of police officers per capita. As is standard in criminometric studies, we use the logs of all variables to estimate elasticities. We also include a full set of year dummies to control for state trends in crime rates. We can use the years 1982 through 1987 to estimate the differenced equation. The quantities in parentheses are the usual OLS standard errors; the quantities in brackets are standard errors robust to both serial correlation and heteroskedasticity:

$$
\begin{aligned}
\widehat{\Delta\log(crmrte)} = \; & .008 - .100\, d83 - .048\, d84 - .005\, d85 \\
& (.017) \; (.024) \quad (.024) \quad (.023) \\
& [.014] \; [.022] \quad [.020] \quad [.025] \\
& + .028\, d86 + .041\, d87 - .327\, \Delta\log(prbarr) \\
& \quad (.024) \quad (.024) \quad (.030) \\
& \quad [.021] \quad [.024] \quad [.056] \\
& - .238\, \Delta\log(prbconv) - .165\, \Delta\log(prbpris) \\
& \quad (.018) \qquad (.026) \\
& \quad [.040] \qquad [.046] \\
& - .022\, \Delta\log(avgsen) + .398\, \Delta\log(polpc) \\
& \quad (.022) \qquad (.027) \\
& \quad [.026] \qquad [.103] \\
& n = 540, R^2 = .433, \bar{R}^2 = .422.
\end{aligned} \quad [13.33]
$$

The three probability variables—of arrest, conviction, and serving prison time—all have the expected sign, and all are statistically significant. For example, a 1% increase in the probability of arrest is predicted to lower the crime rate by about .33%. The average sentence variable shows a modest deterrent effect, but it is not statistically significant.

The coefficient on the police per capita variable is somewhat surprising and is a feature of most studies that seek to explain crime rates. Interpreted causally, it says that a 1% increase in police per capita *increases* crime rates by about .4%. (The usual t statistic is very large, almost 15.) It is hard to believe that having more police officers causes more crime. What is going on here? There are at least two possibilities. First, the crime rate variable is calculated from *reported* crimes. It might be that, when there are additional police, more crimes are reported. Second, the police variable might be endogenous in the equation for other reasons: counties may enlarge the police force when they expect crime rates to increase. In this case, (13.33) cannot be interpreted in a causal fashion. In Chapters 15 and 16, we will cover models and estimation methods that can account for this additional form of endogeneity.

The special case of the White test for heteroskedasticity in Section 8-3 gives $F = 75.48$ and p-value = .0000, so there is strong evidence of heteroskedasticity. (Technically, this test is not valid if there is also serial correlation, but it is strongly suggestive.) Testing for AR(1) serial correlation yields $\hat{\rho} = -.233$, $t = -4.77$, so negative serial correlation exists. The standard errors in brackets adjust for serial correlation and heteroskedasticity. [We will not give the details of this; the calculations are similar to those described in Section 12-5 and are carried out by many econometric packages. See Wooldridge (2010, Chapter 10) for more discussion.] No variables lose statistical significance, but the t statistics on the significant deterrent variables get notably smaller. For example, the t statistic on the probability of conviction variable goes from -13.22 using the usual OLS standard error to -6.10 using the fully robust standard error. Equivalently, the confidence intervals constructed using the robust standard errors will, appropriately, be much wider than those based on the usual OLS standard errors.

Naturally, we can apply the Chow test to panel data models estimated by first differencing. As in the case of pooled cross sections, we rarely want to test whether the intercepts are constant over time; for many reasons, we expect the intercepts to be different. Much more interesting is to test whether slope coefficients have changed over time, and we can easily carry out such tests by interacting the explanatory variables of interest with time-period dummy variables. Interestingly, while we cannot estimate the slopes on variables that do not change over time, we can test whether the partial effects of