C15.7 (i) As usual, if $unem_t$ is correlated with e_t, OLS will be biased and inconsistent for estimating β_1.

(ii) If $E(e_t|inf_{t-1},unem_{t-1}, \ldots) = 0$ then $unem_{t-1}$ is uncorrelated with e_t, which means $unem_{t-1}$ satisfies the first requirement for an IV in

$$\Delta inf_t = \beta_0 + \beta_1 unem_t + e_t.$$

(iii) The second requirement for $unem_{t-1}$ to be a valid IV for $unem_t$ is that $unem_{t-1}$ must be sufficiently correlated. The regression $unem_t$ on $unem_{t-1}$ yields

$$\widehat{unem}_t = 1.57 + .732\ unem_{t-1}$$
$$(0.58)\quad (.097)$$
$$n = 48,\ R^2 = .554.$$

Therefore, there is a strong, positive correlation between $unem_t$ and $unem_{t-1}$.

(iv) The expectations-augmented Phillips curve estimated by IV is

$$\widehat{\Delta inf}_t = .694 - .138\ unem_t$$
$$(1.883)\quad (.319)$$
$$n = 48,\ R^2 = .048.$$

The IV estimate of β_1 is much lower in magnitude than the OLS estimate (–.543), and $\hat{\beta}_1$ is not statistically different from zero. The OLS estimate had a t statistic of about –2.36 [see equation (11.19)].

C15.8 (i) The OLS results are

$$\widehat{pira} = -.198 + .054\ p401k + .0087\ inc - .000023\ inc^2 - .0016\ age + .00012\ age^2$$
$$(.069)\ (.010)\quad (.0005)\quad (.000004)\quad (.0033)\quad (.00004)$$

$n = 9{,}275,\ R^2 = .180$

The coefficient on *p401k* implies that participation in a 401(k) plan is associate with a .054 higher probability of having an individual retirement account, holding income and age fixed.

(ii) While the regression in part (i) controls for income and age, it does not account for the fact that different people have different taste for savings, even within given income and age categories. People that tend to be savers will tend to have both a 401(k) plan as well as an IRA. (This means that the error term, u, is positively correlated with *p401k*.) What we would like to know is, for a given person, if that person participates in a 401(k) does it make it less likely or

more likely that the person also has an IRA. This ceteris paribus question is difficult to answer by OLS without many more controls for the taste for saving.

(iii) First, we need *e*401*k* to be partially correlated with *p*401*k*; not surprisingly, this is not an issue, as being eligible for a 401(k) plan is, by definition, necessary for participation. (The regression in part (iv) verifies that they are strongly positively correlated.) The more difficult issue is whether *e*401*k* can be taken as exogenous in the structural model. In other words, is being *eligible* for a 401(k) correlated with unobserved taste for saving? If we think workers that like to save for retirement will match up with employers that provide vehicles for retirement saving, then *u* and *e*401*k* would be positively correlated. Certainly we think that *e*401*k* is less correlated with *u* than is *p*401*k*. But remember, this alone is not enough to ensure that the IV estimator has less asymptotic bias than the OLS estimator; see page 519.

(iv) The reduced form equation, estimated by OLS but with heteroskedasticity-robust standard errors, is

$$\underset{(.046)}{\widehat{p401k} = .059} + \underset{(.008)}{.689\, e401k} + \underset{(.0003)}{.0011\, inc} - \underset{(.0000027)}{.0000018\, inc^2} - \underset{(.0022)}{.0047\, age} + \underset{(.000026)}{.000052\, age^2}$$

$n = 9{,}275,\ R^2 = .596$

The *t* statistic on *e*401*k* is over 85, and its coefficient estimate implies that, holding income and age fixed, eligibility in a 401(k) plan increases the probability of participation in a 401(k) by .69. Clearly, *e*401*k* passes one of the two requirements as an IV for *p*401*k*.

(v) When *e*401*k* is used as an IV for *p*401*k* we get the following, with heteroskedasticity-robust standard errors:

$$\underset{(.065)}{\widehat{pira} = -.207} + \underset{(.013)}{.021\, p401k} + \underset{(.0005)}{.0090\, inc} - \underset{(.000004)}{.000024\, inc^2} - \underset{(.0032)}{.0011\, age} + \underset{(.00004)}{.00011\, age^2}$$

$n = 9{,}275,\ R^2 = .180$

The IV estimate of β_{p401k} is less than half as large as the OLS estimate, and the IV estimate has a *t* statistic roughly equal to 1.62. The reduction in $\hat{\beta}_{p401k}$ is what we expect given the unobserved taste for saving argument made in part (ii). But we still do not estimate a tradeoff between participating in a 401(k) plan and participating in an IRA. This conclusion has prompted some in the literature to claim that 401(k) saving is additional saving; it does not simply crowd out saving in other plans.

(vi) After obtaining the reduced form residuals from part (iv), say $\hat{v}_i$, we add these to the structural equation and run OLS. The coefficient on $\hat{v}_i$ is .075 with a heteroskedasticity-robust *t* = 3.92. Therefore, there is strong evidence that *p401k* is endogenous in the structural equation (assuming, of course, that the IV, *e*401*k*, is exogenous).