Question 1: (from Cameron and Trivedi)

A leading simple example of IV is one where the instrument z is a **binary instrument**. Denote the subsample averages of y and x by $\bar{y}_1$ and $\bar{x}_1$, respectively, when $z = 1$ and by $\bar{y}_0$ and $\bar{x}_0$, respectively, when $z = 0$. Then $\Delta y/\Delta z = (\bar{y}_1 - \bar{y}_0)$ and $\Delta x/\Delta z = (\bar{x}_1 - \bar{x}_0)$, and (4.46) yields

$$\widehat{\beta}_{\text{Wald}} = \frac{(\bar{y}_1 - \bar{y}_0)}{(\bar{x}_1 - \bar{x}_0)}. \tag{4.48}$$

This estimator is called the **Wald estimator**, after Wald (1940), or the **grouping estimator**.

The Wald estimator can also be obtained from the formula (4.45). For the no-intercept model variables are measured in deviations from means, so $\mathbf{z}'\mathbf{y} = \sum_i (z_i - \bar{z})(y_i - \bar{y})$. For binary z this yields $\mathbf{z}'\mathbf{y} = N_1(\bar{y}_1 - \bar{y}) = N_1 N_0(\bar{y}_1 - \bar{y}_0)/N$, where N_0 and N_1 are the number of observations for which $z = 0$ and $z = 1$. This result uses $\bar{y}_1 - \bar{y} = (N_0\bar{y}_1 + N_1\bar{y}_1)/N - (N_0\bar{y}_0 + N_1\bar{y}_1)/N = N_0(\bar{y}_1 - \bar{y}_0)/N$. Similarly, $\mathbf{z}'\mathbf{x} = N_1 N_0(\bar{x}_1 - \bar{x}_0)/N$. Combining these results, we have that (4.45) yields (4.48).

For the earnings–schooling example it is being assumed that we can define two groups where group membership does not directly determine earnings, though it does affect level of schooling and hence indirectly affects earnings. Then the IV estimate is the difference in average earnings across the two groups divided by the difference in average schooling across the two groups.

Equation (4.45) is below:

For regression with scalar regressor x and scalar instrument z, the **instrumental variables (IV) estimator** is defined as

$$\widehat{\beta}_{\text{IV}} = (\mathbf{z}'\mathbf{x})^{-1}\mathbf{z}'\mathbf{y}, \tag{4.45}$$

Question 2:

c) , d), and e) Likelihood and log-likelihood functions, the first derivative of the log-likelihood function:

$\mathcal{L} = \prod_{i=1}^{n}[F(x\beta)]^{y_i} \prod_{i=1}^{n}[1 - F(x\beta)]^{1-y_i}, \;\; y_i = 0 \; or \; 1.$

$ln\, \mathcal{L} = \sum_{i=1}^{n} y_i \, ln[F(x\beta)] + \sum_{i=1}^{n}(1 - y_i)\, ln[1 - F(x\beta)], \;\;\; y_i = 0 \; or \; 1.$

$\frac{\partial ln\, \mathcal{L}(\beta|x)}{\partial \beta} = \sum_{i=1}^{n} \frac{y_i\, f(x\beta)\, x_i}{F(x\beta)} - \sum_{i=1}^{n} \frac{(1-y_i)\, f(x\beta)\, x_i}{1-F(x\beta)} = 0$, where $f(.)$ is the density function and $F(.)$ is the cumulative distribution function.

Part I.

a)

$$E(y|x) = P(y = 1|x) = F(x\beta) = \Lambda(x) = \frac{e^{x\beta}}{1 + e^{x\beta}}$$

b)

$$\frac{\partial E(y|x)}{\partial x_j} = \frac{\beta_j\, e^{x\beta}}{(1+e^{x\beta})^2}$$

Part II.

a) $E(y|x) = P(y=1|x) = F(x\beta) = \Phi(x\beta)$.

b)

$$\frac{\partial E(y|x)}{\partial x_j} = \varphi(x\beta)\, \beta_j$$

Question 3:

	True		
Classified	**1**	**0**	
+	84	18	102
-	126	72	198
	210	90	300

Percent correctly classified: (84+72)/300 = 52%