QUESTION 1

(a) For the $E_i^2$ variable,

$$t = \frac{-0.0061}{0.0009} \approx -6.78 < -1.96$$

so it is significant at the 5% significance level.

(b) To proceed, note that, taking the derivative of $\hat{W}$ with respect to $E$ and setting it equal to zero, we get

$$\frac{\partial \hat{W}}{\partial E} = 0.298 + (-0.0061)2E = 0.298 - 0.0122E = 0.$$  

Solving for $E$ in the above equation yields

$$E = \frac{0.298}{0.0122} \approx 24.43$$

so that the marginal effect of work experience on wage starts to turn negative after one has accumulated about 24.43 years of work experience.

(c) Yes, there is diminishing effect of work experience on wage since

$$\frac{\partial^2 \hat{W}}{\partial E^2} = -0.0122.$$  

QUESTION 2

(a) Note that under the stated conditions, $Cov(x_i, u_i) = 0$; hence, $\beta$ can be estimated consistently by running OLS on the first equation. It follows that there is no identification issue with respect to the parameter $\beta$. 
(b) Substituting the right-hand side of the first equation into the second equation for $z_i$, we obtain

\[ y_i = \pi (\beta x_i + u_i) + v_i \]
\[ = \pi \beta x_i + (v_i + u_i \pi) \]
\[ = \varphi x_i + \eta_i \]

where

\[ \varphi = \pi \beta, \]
\[ \eta_i = v_i + u_i \pi, \]

This is a reduced form representation for $y_i$ since

\[ \text{Cov}(x_i, \eta_i) = \text{Cov}(x_i, v_i) + \pi \text{Cov}(x_i, u_i) = 0 \]

given that $E[v_i|x_i] = 0$ and $E[u_i|x_i] = 0$ with probability one imply that $\text{Cov}(x_i, v_i) = 0$ and $\text{Cov}(x_i, u_i) = 0$, respectively. It follows that we can estimate $\varphi$ consistently by running OLS on this equation. Moreover, because $\beta \neq 0$, we can write

\[ \pi = \frac{\varphi}{\beta} \]

so that $\pi$ is a ratio of parameters that we can estimate consistently. Hence, $\pi$ is identified. In fact, in this case, the equation

\[ y_i = \pi z_i + v_i \]

is the structural equation whereas

\[ z_i = \beta x_i + u_i \]

is the so-called “first-stage” equation. $x_i$ is a valid instrument for estimating the coefficient $\pi$ of the structural equation because it satisfies both the exogeneity condition (i.e., $E[v_i|x_i] = 0$ with probability one) and the relevance condition since $\pi \neq 0$.

(c) In light of the discussion in parts (a) and (b) above, both $\beta$ and $\pi$ are identified. We can estimate $\beta$ using the OLS estimator

\[ \hat{\beta} = \frac{\sum_{i=1}^{n} x_i z_i}{\sum_{i=1}^{n} x_i^2} \]

and $\pi$ using the 2SLS estimator

\[ \hat{\pi} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i z_i}. \]
QUESTION 3

(a) Note that the underlying regression model is of the form

$$\log(salary_i) = \beta_0 + \beta_1 \log(sales_i) + \beta_2 roe_i + \beta_3 D^F_i + \beta_4 D^{CP}_i + \beta_5 D^U_i + \varepsilon_i$$

It follows that the (approximate) percentage difference in estimated salary between a CEO in the utilities industry vis-à-vis one in the transportation industry, holding sales and roe fixed, is given by

$$E[\log(salary_i) | D^U_i = 1, D^F_i = 0, D^{CP}_i = 0, \log(sales_i), roe_i] - E[\log(salary_i) | D^U_i = 0, D^F_i = 0, D^{CP}_i = 0, \log(sales_i), roe_i] = \beta_0 + \beta_1 \log(sales_i) + \beta_2 roe_i + \beta_5 - (\beta_0 + \beta_1 \log(sales_i) + \beta_2 roe_i) = \beta_5$$

Hence, our point estimate of this difference is

$$\hat{\beta}_5 = -0.283.$$  

Moreover, testing whether this difference is statistically significant is really testing the null hypothesis

$$H_0 : \beta_5 = 0.$$  

Given our estimation results, the value of the t-statistic for testing this $$H_0$$ is given by

$$t = \frac{-0.283}{0.099} \approx -2.86$$  

Note further that, in large sample, the t-statistic has a standard normal distribution under $$H_0$$, so our decision rule is to reject $$H_0$$ if $$|t| > 1.96$$. Since in this case

$$|t| \approx 2.86 > 1.96,$$

we reject $$H_0$$ and conclude that the difference is statistically significant.

(b) Note that the (approximate) percentage difference in estimated salary between a CEO in the consumer products industry vis-à-vis one in the financial industry, holding sales and roe fixed, is given by

$$E[\log(salary_i) | D^U_i = 0, D^F_i = 0, D^{CP}_i = 1, \log(sales_i), roe_i] - E[\log(salary_i) | D^U_i = 0, D^F_i = 1, D^{CP}_i = 0, \log(sales_i), roe_i] = \beta_0 + \beta_1 \log(sales_i) + \beta_2 roe_i + \beta_4 - (\beta_0 + \beta_1 \log(sales_i) + \beta_2 roe_i + \beta_3) = \beta_4 - \beta_3$$

Hence, our point estimate of this difference is given by

$$\hat{\beta}_4 - \hat{\beta}_3 = 0.181 - 0.158 = 0.023.$$
To test the statistical significance of this difference is the same as testing the null hypothesis

\[ H_0 : \beta_4 - \beta_3 = 0. \]

This can be done by transforming the above regression as follows

\[
\log (salary_i) = \beta_0 + \beta_1 \log (sales_i) + \beta_2 \text{roe}_i + \beta_3 D_{Fi}^F + \beta_4 D_{Ci}^{CP} + \beta_5 D_{Ui}^U + \varepsilon_i
\]

\[
= \beta_0 + \beta_1 \log (sales_i) + \beta_2 \text{roe}_i + \beta_3 (D_{Fi}^F + D_{Ci}^{CP}) + (\beta_4 - \beta_3) D_{Ci}^{CP} + \beta_5 D_{Ui}^U + \varepsilon_i
\]

\[
= \beta_0 + \beta_1 \log (sales_i) + \beta_2 \text{roe}_i + \beta_3 W_i + \gamma D_{Ci}^{CP} + \beta_5 D_{Ui}^U + \varepsilon_i \tag{2}
\]

where

\[
W_i = D_{Fi}^F + D_{Ci}^{CP},
\]

\[
\gamma = \beta_4 - \beta_3.
\]

It follows that one can test the significance of this difference by estimating (2) above and then test the null hypothesis

\[ H_0 : \gamma = 0 \]

using a large sample t-test.

**QUESTION 4**

(a) We should not be surprised by this result since the J-test can only be used to test the exogeneity of the extra instruments. In this case, we only have one instrument and one endogenous regressor so we do not have any extra instruments and we are at best just identified. As a result, the J-statistic is always equal to zero in this case, and so the procedure is not informative about the possible exogeneity of this one instrument.

(b) In this case, we are possibly overidentified since we have four instruments and only one endogenous regressor. Assuming that at least one of the instruments is valid, we can then use the J-test to test the exogeneity of the extra instruments. Here, under the null hypothesis is that all instruments are exogenous, the large sample distribution of the J-statistic is a chi-square distribution with \# of instruments − \# of endogenous regressors = 4 − 1 = 3 degrees of freedom. At the 5% significance level, the critical value under this distribution is 7.81. Since

\[
J = 12.01 > 7.81,
\]

we reject the null hypothesis.
QUESTION 5

To proceed, note first that averaging equation (3) across time, we obtain

\[
\frac{1}{2} \sum_{t=1}^{2} Y_{it} = \frac{1}{2} \sum_{t=1}^{2} [\alpha_i + \beta X_{it} + u_{it}]
= \alpha_i + \beta \frac{1}{2} \sum_{t=1}^{2} X_{it} + \frac{1}{2} \sum_{t=1}^{2} u_{it}
\]

The so-called “entity-demeaned” procedure seeks to estimate \( \beta \) by running OLS on the regression

\[
Y_{it} - \frac{1}{2} \sum_{t=1}^{2} Y_{it} = \alpha_i + \beta X_{it} + u_{it} - \left( \alpha_i + \beta \frac{1}{2} \sum_{t=1}^{2} X_{it} + \frac{1}{2} \sum_{t=1}^{2} u_{it} \right)
= \beta \left( X_{it} - \frac{1}{2} \sum_{t=1}^{2} X_{it} \right) + \left( u_{it} - \frac{1}{2} \sum_{t=1}^{2} u_{it} \right)
\]
or

\[
\tilde{Y}_{it} = \beta \tilde{X}_{it} + \tilde{u}_{it} \quad \text{for } i = 1, \ldots, n; \ t = 1, 2
\]

where

\[
\tilde{Y}_{it} = Y_{it} - \frac{1}{2} \sum_{t=1}^{2} Y_{it},
\]
\[
\tilde{X}_{it} = X_{it} - \frac{1}{2} \sum_{t=1}^{2} X_{it},
\]
\[
\tilde{u}_{it} = u_{it} - \frac{1}{2} \sum_{t=1}^{2} u_{it}.
\]

Applying OLS to (3) results in the estimator

\[
\hat{\beta} = \frac{\sum_{i=1}^{n} \sum_{t=1}^{2} \tilde{X}_{it} \tilde{Y}_{it}}{\sum_{i=1}^{n} \sum_{t=1}^{2} \tilde{X}_{it}^2}
\]

Note, however, that in this case

\[
\tilde{Y}_{it} = Y_{it} - \frac{1}{2} (Y_{i1} + Y_{i2}) = \begin{cases} -\frac{1}{2} \Delta Y_i & \text{for } t = 1 \\ \frac{1}{2} \Delta Y_i & \text{for } t = 2 \end{cases}
\]

Similarly, we have

\[
\tilde{X}_{it} = X_{it} - \frac{1}{2} (X_{i1} + X_{i2}) = \begin{cases} -\frac{1}{2} \Delta X_i & \text{for } t = 1 \\ \frac{1}{2} \Delta X_i & \text{for } t = 2 \end{cases}
\]
so that
\[ \sum_{t=1}^{2} \tilde{X}_{it} \tilde{Y}_{it} = \frac{1}{4} \Delta X_i \Delta Y_i + \frac{1}{4} \Delta X_i \Delta Y_i = \frac{1}{2} \Delta X_i \Delta Y_i \]
and
\[ \sum_{t=1}^{2} \tilde{X}_{it}^2 = \frac{1}{4} (\Delta X_i)^2 + \frac{1}{4} (\Delta X_i)^2 = \frac{1}{2} (\Delta X_i)^2 \]
from which it follows that
\[
\hat{\beta} = \frac{\sum_{i=1}^{n} \sum_{t=1}^{2} \tilde{X}_{it} \tilde{Y}_{it}}{\sum_{i=1}^{n} \sum_{t=1}^{2} \tilde{X}_{it}^2} = \frac{\sum_{i=1}^{n} \frac{1}{2} \Delta X_i \Delta Y_i}{\sum_{i=1}^{n} \frac{1}{2} (\Delta X_i)^2} = \frac{\sum_{i=1}^{n} \Delta X_i \Delta Y_i}{\sum_{i=1}^{n} (\Delta X_i)^2}
\]
but this is actually the OLS formula we would have obtained from doing the regression
\[ \Delta Y_i = \beta \Delta X_i + \Delta u_i \]
so we have shown the desired result.