1. 

(a) Estimator:
\[ \hat{\beta} = \left( \sum_{i=1}^{n} x_i x_i' \right)^{-1} \left( \sum_{i=1}^{n} x_i y_i \right) \]
\[ \hat{e}_i = y_i - x_i' \hat{\beta} \]
\[ \hat{\mu}_3 = \frac{1}{n} \sum_{i=1}^{n} \hat{e}_i^3 \]

(b) Percentile Bootstrap

i. Draw an observation \((y_i^*, x_i^*)\) randomly from the observed sample \(\{y_i, x_i\}\)

ii. Repeat this \(n\) times, to obtain a sample \(\{y_i^*, x_i^*\}, i = 1, ..., n\)

iii. Compute the estimator \(\hat{\mu}_3\) on this bootstrap sample. This is
\[ \hat{\beta}^* = \left( \sum_{i=1}^{n} x_i^* x_i'^* \right)^{-1} \left( \sum_{i=1}^{n} x_i^* y_i^* \right) \]
\[ \hat{e}_i^* = y_i^* - x_i'^* \hat{\beta}^* \]
\[ \hat{\mu}_3^* = \frac{1}{n} \sum_{i=1}^{n} \hat{e}_i^3 \]

iv. Repeat this \(B\) times, to obtain \(\{\hat{\mu}_3^b\}, b = 1, ..., B\).

v. Let \(\hat{q}_{.05}\) and \(\hat{q}_{.95}\) be the 5% and 95% empirical quantiles of \(\hat{\mu}_3\). These can be computed as the \([.05(B+1)]^{\text{th}}\) and \([.95(B+1)]^{\text{th}}\) order statistics of \(\hat{\mu}_3\)

vi. The Efron percentile confidence interval for \(\mu_3\) is \([\hat{q}_{.05}, \hat{q}_{.95}]\)

2. 

(a) The estimated covariance matrix for \(\hat{\beta}_1, \hat{\beta}_2\) is
\[ \hat{V}_\beta = \begin{bmatrix} s(\hat{\beta}_1)^2 & \hat{\rho}s(\hat{\beta}_1)s(\hat{\beta}_2) \\ \hat{\rho}s(\hat{\beta}_1)s(\hat{\beta}_2) & s(\hat{\beta}_2)^2 \end{bmatrix} \]
(recall that the standard errors are the square roots of the diagonal elements of the estimated covariance matrix).

The variance estimate for \(\hat{\theta} = \hat{\beta}_1 - \hat{\beta}_2\) is
\[ \hat{V}_\theta = \begin{pmatrix} 1 & -1 \end{pmatrix} \hat{V}_\beta \begin{pmatrix} 1 & -1 \end{pmatrix} = s(\hat{\beta}_1)^2 - 2\hat{\rho}s(\hat{\beta}_1)^2s(\hat{\beta}_2)^2 + s(\hat{\beta}_2)^2 \]
the standard error is
\[ s(\hat{\theta}) = \sqrt{s(\hat{\beta}_1)^2 - 2\hat{\rho}s(\hat{\beta}_1)^2s(\hat{\beta}_2)^2 + s(\hat{\beta}_2)^2} \]
An asymptotic 95% confidence interval for \(\theta\) is
\[ \hat{C} = \hat{\theta} \pm 2s(\hat{\theta}) = \hat{\beta}_1 - \hat{\beta}_2 \pm 2\sqrt{s(\hat{\beta}_1)^2 - 2\hat{\rho}s(\hat{\beta}_1)^2s(\hat{\beta}_2)^2 + s(\hat{\beta}_2)^2} \]

(b) No. All we know is \(-1 < \hat{\rho} < 1\)
(c) At the observed values, the confidence interval for $\theta$ is

$$\hat{C} = 0.2 \pm 2 \cdot (.07) \sqrt{2(1 - \hat{p})} \simeq 0.2 \pm 0.2\sqrt{(1 - \hat{p})}$$

We don’t know $\hat{p}$, but the question is whether or not we can reject $\theta = 0$, or equivalently if $\hat{C}$ includes 0. Looking at the above interval, this occurs if and only if $\hat{p} > 0$. Therefore the reported evidence, by itself, does not support the conclusion that $\beta_2$ exceeds $\beta_1$. While it may be true that $\hat{p} > 0$ (and then we can reject $\beta_1 = \beta_2$), this is unknown from the reported information, so the author’s claim is (at best) unsupported.

3.

(a) Using matrix notation

$$\hat{\beta} = (X'ZW'X)^{-1} (X'ZW'Y)$$

(b) By the law of iterated expectations

$$E(z_i e_i) = E(E(z_i e_i | z_i)) = E(z_i E(e_i | z_i)) = E(z_i E(\delta n^{-1/2} + u_i | z_i)) = E(z_i \delta n^{-1/2}) = \mu \delta n^{-1/2} \neq 0$$

The third step using $e_i = \delta n^{-1/2} + u_i$ and the fourth using $E(u_i | z_i) = 0$. This contradicts (1).

(c) Since $Y = X \beta + e$

$$\sqrt{n} \left( \hat{\beta} - \beta \right) = \sqrt{n} \left( X'ZW'X \right)^{-1} \left( X'ZW' \epsilon \right)$$

Then using $e = \delta n^{-1/2} + u$

$$\sqrt{n} \left( \hat{\beta} - \beta \right) = \sqrt{n} \left( X'ZW'X \right)^{-1} \left( X'ZW \left( Z' \delta n^{-1/2} + Z' u \right) \right) = \left( \frac{1}{n} X'ZW \frac{1}{n} Z'X \right)^{-1} \left( \frac{1}{n} X'ZW \left( \frac{1}{n} Z' \delta + \frac{1}{\sqrt{n}} Z' u \right) \right)$$

(d) By the WLLN,

$$\frac{1}{n} \sum_{i=1}^{n} z_i x_i' \rightarrow E (z_i x_i') = Q$$

$$\frac{1}{n} \sum_{i=1}^{n} z_i \rightarrow E (z_i) = \mu_z$$

and by the CLT

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i u_i \rightarrow N(0, \Omega)$$

where

$$\Omega = E \left( z_i z_i' u_i^2 \right)$$

Thus

$$\frac{1}{n} Z' \delta + \frac{1}{\sqrt{n}} Z' u \rightarrow N(\mu_z \delta, \Omega)$$
and
\[
\sqrt{n} (\hat{\beta} - \beta) = \left( \frac{1}{n} X' Z W \frac{1}{n} Z' X \right)^{-1} \left( \frac{1}{n} X' Z W \left( \frac{1}{n} Z' \delta + \frac{1}{\sqrt{n}} Z' u \right) \right)
\rightarrow_d (Q' W Q)^{-1} Q' W N(\mu_\beta, \Omega)
\]
\[
N((Q' W Q)^{-1} Q' W \mu_\beta, (Q' W Q)^{-1} Q' W \Omega W Q (Q' W Q)^{-1})
\]
The asymptotic distribution is non-central normal with a classic covariance matrix. Thus the GMM estimator is asymptotically biased.

4.

(a) No statistical model was specified. Therefore least-squares regression is estimating projections. The coefficient in (3) is the projection
\[
\beta = (E x_i x_i')^{-1} E x_i y_i
\]
which defines the error
\[
e_i = y_i - x_i' \beta
\]
The coefficient being estimated in (4) is the projection of \(e_i\) on \(z_i\)
\[
\gamma = (E z_i z_i')^{-1} E z_i e_i
\]
\[
= (E z_i z_i')^{-1} E z_i (y_i - x_i' \beta)
\]
\[
= (E z_i z_i')^{-1} \left( E z_i y_i - E z_i x_i' (E x_i x_i')^{-1} E x_i y_i \right)
\]
(b) We know that \(\hat{\beta} \rightarrow_p \beta\). Then
\[
\tilde{\gamma} = (Z' Z)^{-1} (Z' \hat{e})
\]
\[
= \left( \frac{1}{n} Z' Z \right)^{-1} \left( \frac{1}{n} Z' e \right) - \left( \frac{1}{n} Z' Z \right)^{-1} \left( \frac{1}{n} Z' X \right) (\hat{\beta} - \beta)
\]
\[
\rightarrow_p (E z_i z_i')^{-1} E z_i e_i = \gamma
\]
(c) This is
\[
W_n = n \tilde{\gamma}' \tilde{V}^{-1} \tilde{\gamma}
\]
where
\[
\tilde{V}_\gamma = \left( \frac{1}{n} Z' Z \right)^{-1} \Omega \left( \frac{1}{n} Z' Z \right)^{-1}
\]
\[
\tilde{\Omega} = \frac{1}{n} \sum_{i=1}^n z_i z_i' u_i^2
\]
Since this is a test on all coefficients, we can simplify this to
\[
W_n = \left( \frac{1}{\sqrt{n}} \tilde{e}' Z \right) \tilde{\Omega}^{-1} \left( \frac{1}{\sqrt{n}} Z' \tilde{e} \right)
\]
(d) When \(\gamma = 0\), \(E z_i e_i = 0\), so
\[
\frac{1}{\sqrt{n}} Z' e \rightarrow_d N(0, \Omega)
\]
where
\[
\Omega = E z_i z_i' e_i^2
\]
Then
\[
\frac{1}{\sqrt{n}} Z' \hat{e} = \frac{1}{\sqrt{n}} Z' e - \frac{1}{n} Z' X \sqrt{n} \left( \hat{\beta} - \beta \right)
\]
\[\to_d \ N(0, \Omega) + E(\hat{z}_i x_i') O_p(1)\]
which simplifies to \(N(0, \Omega)\) when \(E(\hat{z}_i x_i') = 0\). We find that
\[
W_n = \left( \frac{1}{\sqrt{n}} \hat{e}' Z \right) \tilde{\Omega}^{-1} \left( \frac{1}{\sqrt{n}} Z' \hat{e} \right)
\]
\[\to_d \ N(0, \Omega)^t \Omega^{-1} N(0, \Omega) = \chi^2 \]
A complete proof would show that \(\tilde{\Omega} \to_p \Omega\).
Even though the statistic \(W_n\) has ignored the two-step estimation, the asymptotic distribution is still standard. The reason is that the two steps are asymptotically independent when \(z_i\) and \(x_i\) are uncorrelated.

(e) If \(E(\hat{z}_i x_i') \neq 0\) then the distribution is no longer \(\chi^2\). The reason is that the sampling distribution of \(\sqrt{n} \left( \hat{\beta} - \beta \right)\) no longer drops out. For simplicity, suppose that the error is homoskedastic. Then
\[
\frac{1}{\sqrt{n}} Z' \hat{e} = \frac{1}{\sqrt{n}} Z' e - \frac{1}{n} Z' X \left( \frac{1}{n} X' X \right)^{-1} \left( \frac{1}{\sqrt{n}} X' e \right)
\]
\[\to_d \ (I - Q_{zz} Q_{xx}^{-1}) N \left( 0, \begin{bmatrix} Q_{zz} & Q_{zx} \\ Q_{xz} & Q_{xx} \end{bmatrix} \sigma^2 \right) \]
\[= N \left( 0, (Q_{zz} - Q_{zx} Q_{xz}^{-1} Q_{xx}) \sigma^2 \right) \]
and \(\Omega = Q_{zz} \sigma^2\). Hence
\[
W_n \to_d N(0, Q_{zz} - Q_{zx} Q_{xz}^{-1} Q_{xx}) Q_{xx}^{-1} N(0, Q_{zz} - Q_{zx} Q_{xz}^{-1} Q_{xx})
\]
which is not \(\chi^2\). Thus in the general case (where \(z_i\) and \(x_i\) are correlated), the statistic \(W_n\) has incorrectly ignored the two-step estimation problem.