1. We know that both $\hat{\beta}$ and $\tilde{\beta}$ are unbiased for $\beta$, $\hat{\beta} - \beta = (X'X)^{-1} (X'e)$ and $\tilde{\beta} - \beta = (X'D^{-1}X)^{-1} (X'D^{-1}e)$

(a) Note that

$$\hat{\beta} - \tilde{\beta} = (\hat{\beta} - \beta) - (\tilde{\beta} - \beta) = (X'X)^{-1} (X'e) - (X'D^{-1}X)^{-1} (X'D^{-1}e)$$

and

$$E\left( (\hat{\beta} - \beta) | X \right) = 0.$$ 

Thus

$$Cov\left( (\hat{\beta} - \beta, \tilde{\beta} | X \right) = Cov\left( (\hat{\beta} - \beta | X \right) = Cov\left( (\hat{\beta} | X \right) - Cov\left( (\tilde{\beta} | X \right)$$

and

$$E\left( (\hat{\beta} - \tilde{\beta} | X \right) = 0.$$ 

(b) Since

$$0 = Cov\left( (\hat{\beta} - \tilde{\beta}, \tilde{\beta} | X \right) = Cov\left( (\hat{\beta}, \tilde{\beta} | X \right) - Cov\left( (\tilde{\beta} | X \right)$$

it follows from our result in part (a) that $Cov\left( (\hat{\beta}, \tilde{\beta} | X \right) = Var\left( \tilde{\beta} | X \right)$. 

\[ \]
(c) Using part (b),

\[ \text{Var} \left( \hat{\beta} - \tilde{\beta} \mid X \right) = \text{Var} \left( \hat{\beta} \mid X \right) + \text{Var} \left( \tilde{\beta} \mid X \right) + \text{Cov} \left( \hat{\beta}, \tilde{\beta} \mid X \right) \\
= \text{Var} \left( \hat{\beta} \mid X \right) + \text{Var} \left( \tilde{\beta} \mid X \right) - \text{Var} \left( \hat{\beta} \mid X \right) - \text{Var} \left( \tilde{\beta} \mid X \right) \\
= \text{Var} \left( \hat{\beta} \mid X \right) - \text{Var} \left( \tilde{\beta} \mid X \right) \]

(d) Using part (b) and the known variances of \( \hat{\beta} \) and \( \tilde{\beta} \),

\[ V_n = \text{Var} \left( \hat{\beta} - \tilde{\beta} \mid X \right) \\
= \text{Var} \left( \hat{\beta} \mid X \right) - \text{Var} \left( \tilde{\beta} \mid X \right) \\
= \left( X'X \right)^{-1} \left( X'DX \right) \left( X'X \right)^{-1} - \left( X'D^{-1}X \right)^{-1} . \]

Note 1: Perhaps an easier method of proof would be to first prove part (b), and use this to prove part (a). The first argument works as follows:

\[ \text{Cov} \left( \hat{\beta}, \tilde{\beta} \mid X \right) = E \left[ \left( \hat{\beta} - \beta \right) \left( \tilde{\beta} - \beta \right) ' \mid X \right] \\
= E \left[ \left( X'X \right)^{-1} \left( X'e \right) \left( e'D^{-1}X \right) \left( X'D^{-1}X \right)^{-1} \mid X \right] \\
= \left( X'X \right)^{-1} X'E \left[ \left( \left| x \right| \leq c \right) \mid X \right] D^{-1}X \left( X'D^{-1}X \right)^{-1} \\
= \left( X'X \right)^{-1} X'DD^{-1}X \left( X'D^{-1}X \right)^{-1} \\
= \left( X'X \right)^{-1} X'X \left( X'D^{-1}X \right)^{-1} \\
= \left( X'D^{-1}X \right)^{-1} \\
= \text{Var}(\tilde{\beta} \mid X) . \]

Note 2: This result holds quite generally when one estimator is efficient. In general, for any unbiased estimators \( \hat{\beta} \) and \( \tilde{\beta} \), where \( \tilde{\beta} \) is efficient, then \( \text{Var} \left( \hat{\beta} - \tilde{\beta} \mid X \right) = \text{Var} \left( \hat{\beta} \mid X \right) - \text{Var} \left( \tilde{\beta} \mid X \right) \).

2. To solve this question, you need to recognize the following. First, you can write the estimator as

\[ \tilde{\beta} - \beta = \left( \sum_{i=1}^{n} x_i x'_i 1 \left( |x_i| \leq c \right) \right)^{-1} \left( \sum_{i=1}^{n} x_i e_i 1 \left( |x_i| \leq c \right) \right) . \]

Second, \( x_i x'_i 1 \left( |x_i| \leq c \right) \) is an iid random variable with mean

\[ E \left( x_i x'_i 1 \left( |x_i| \leq c \right) \right) \equiv Q_c . \]
Third, \( x_i e_i 1(|x_i| \leq c) \) is an iid random variable with mean zero:

\[
E(x_i e_i 1(|x_i| \leq c)) = E(x_i 1(|x_i| \leq c)) E(e_i | x_i)) = 0
\]

(using the law of iterated expectations), and has variance

\[
E[(x_i e_i 1(|x_i| \leq c)) (x_i e_i 1(|x_i| \leq c))] - E(x_i x_i' e_i^2 1(|x_i| \leq c)) = \Omega_c.
\]

Note that \( Q_c \neq Q = E(x_i x_i') \) and \( \Omega_c \neq \Omega = E(x_i x_i' e_i^2) \).

(a) By the WLLN,

\[
\frac{1}{n} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \leq c) \rightarrow_p E(x_i x_i' 1(|x_i| \leq c)) = Q_c
\]

and

\[
\frac{1}{n} \sum_{i=1}^{n} x_i e_i 1(|x_i| \leq c) \rightarrow_p E(x_i e_1 1(|x_i| \leq c)) = 0.
\]

Hence, if \( Q_c > 0 \), then by the CMT,

\[
\tilde{\beta} - \beta = \left( \frac{1}{n} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \leq c) \right)^{-1} \left( \frac{1}{n} \sum_{i=1}^{n} x_i e_i 1(|x_i| \leq c) \right) \rightarrow_p Q_c^{-1} 0 = 0.
\]

(b) By the CLT,

\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i e_i 1(|x_i| \leq c) \rightarrow_d N(0, \Omega_c).
\]

Hence

\[
\sqrt{n} (\tilde{\beta} - \beta) = \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i x_i' 1(|x_i| \leq c) \right)^{-1} \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i e_i 1(|x_i| \leq c) \right) \rightarrow_d Q_c^{-1} N(0, \Omega_c) = N(0, Q_c^{-1} \Omega_c Q_c^{-1}).
\]

(c) Bonus Question: The condition \( E(x_i e_i) = 0 \) is not sufficient for

\[
E(x_i e_1 1(|x_i| \leq c)) = 0.
\]

Thus \( \tilde{\beta} \) will not necessarily be consistent for \( \beta \). In fact,

\[
\tilde{\beta} \rightarrow_p \beta + Q_c^{-1} E(x_i e_1 1(|x_i| \leq c)).
\]
(a) These results follow from the fact that the sample mean is an unbiased estimator of the population mean, regardless of the distribution the data. Observe that

$$E(\bar{y}) = Ey_i = \mu.$$  

Hence

$$\tau_n = E(\bar{y} - \mu) = \mu - \mu = 0.$$  

Let $y_i^*$ be a random variable with distribution $F_n$, and $\bar{y}^*$ the sample mean of a random sample $\{y_1^*, ..., y_n^*\}$. By linearity,

$$E\bar{y}^* = E\bar{y}_i^* = \sum_{j=1}^{n} y_j P(y_i^* = y_j) = \frac{1}{n} \sum_{j=1}^{n} y_j = \bar{y}.$$  

The bootstrap estimate of bias treats $\hat{\mu} = \bar{y}$ as the true value. Hence

$$\tau_n^* = E\bar{y}^* - \hat{\mu} = \bar{y} - \bar{y} = 0.$$  

(b) The bias is $E\hat{\mu}^2 - \mu^2$. This is the variance of $\hat{\mu} = \bar{y}$, which we know is $n^{-1}\sigma^2$. In detail,

$$\tau_n = E\hat{\mu}^2 - \mu^2 = E(\hat{\mu} - \mu)^2 = E\left(\frac{1}{n} \sum_{i=1}^{n} (y_i - \mu)\right)^2 = \frac{1}{n^2} \sum_{i=1}^{n} E(y_i - \mu)^2 = \frac{\sigma^2}{n}.$$  

(c) Bonus Question: The variance of the EDF is $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2$. Thus the bootstrap estimate $\tau_n^*$ of $\tau_n$ is

$$\tau_n^* = \frac{\hat{\sigma}^2}{n}.$$  

In this case, you do not need to do a simulation to calculate $\tau_n^*$.  

4