Introduction to Econometrics (4th Edition)

by

James H. Stock and Mark W. Watson

Solutions to Odd-Numbered End-of-Chapter Exercises:
Chapter 18

(This version September 18, 2018)
18.1. (a) Suppose there are $n$ observations. Let $b_1$ be an arbitrary estimator of $\beta_1$. Given the estimator $b_1$, the sum of squared errors for the given regression model is

$$\sum_{i=1}^{n} (Y_i - b_1 X_i)^2.$$ \(\hat{b}^{RLS}_1\), the restricted least squares estimator of $\beta_1$, minimizes the sum of squared errors. That is, $\hat{b}^{RLS}_1$ satisfies the first order condition for the minimization which requires the derivative of the sum of squared errors with respect to $b_1$ equals zero:

$$\sum_{i=1}^{n} 2(Y_i - b_1 X_i)(-X_i) = 0.$$

Solving for $b_1$ from the first order condition leads to the restricted least squares estimator

$$\hat{b}^{RLS}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}.$$

(b) We show first that $\hat{b}^{RLS}_1$ is unbiased. We can represent the restricted least squares estimator $\hat{b}^{RLS}_1$ in terms of the regressors and errors:

$$\hat{b}^{RLS}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} = \frac{\sum_{i=1}^{n} X_i (\beta_1 X_i + u_i)}{\sum_{i=1}^{n} X_i^2} = \beta_1 + \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}.$$

Thus

$$E(\hat{b}^{RLS}_1) = \beta_1 + E \left( \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2} \right) = \beta_1 + E \left[ \frac{\sum_{i=1}^{n} X_i E(u_i|X_i, X_n)}{\sum_{i=1}^{n} X_i^2} \right] = \beta_1,$$

where the second equality follows by using the law of iterated expectations, and the third equality follows from

$$\frac{\sum_{i=1}^{n} X_i E(u_i|X_i, X_n)}{\sum_{i=1}^{n} X_i^2} = 0.$$
18.1 (continued)

because the observations are i.i.d. and \(E(u_i|X_i) = 0\). (Note, \(E(u_i|X_1, \ldots, X_n) = E(u_i|X_i)\) because the observations are i.i.d.

Under assumptions 1-3 of Key Concept 18.1, \(\hat{\beta}_1^{RLS}\) is asymptotically normally distributed. The large sample normal approximation to the limiting distribution of \(\hat{\beta}_1^{RLS}\) follows from considering

\[
\hat{\beta}_1^{RLS} - \beta_1 = \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2} = \frac{1}{n} \sum_{i=1}^{n} X_i u_i.
\]

Consider first the numerator which is the sample average of \(v_i = X_i u_i\). By assumption 1 of Key Concept 18.1, \(v_i\) has mean zero:

\(E(X_i u_i) = E[X_i E(u_i|X_i)] = 0\). By assumption 2, \(v_i\) is i.i.d. By assumption 3, var(\(v_i\)) is finite. Let \(\bar{v} = \frac{1}{n} \sum_{i=1}^{n} X_i u_i\), then \(\sigma_v^2 = \sigma_v^2 / n\). Using the central limit theorem, the sample average

\[
\frac{\bar{v}}{\sigma_v} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_i \xrightarrow{d} N(0, 1)
\]

or

\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i u_i \xrightarrow{d} N(0, \sigma_v^2).
\]

For the denominator, \(X_i^2\) is i.i.d. with finite second variance (because \(X\) has a finite fourth moment), so that by the law of large numbers

\[
\frac{1}{n} \sum_{i=1}^{n} X_i^2 \xrightarrow{d} E(X^2).
\]

Combining the results on the numerator and the denominator and applying Slutsky’s theorem lead to

(continued on the next page)
18.1 (continued)

\[
\sqrt{n}(\hat{\beta}_1^{RLS} - \beta_1) = \frac{1}{\sqrt{n}} \frac{1}{\frac{1}{n} \sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2} \Rightarrow N \left( 0, \frac{\text{var}(X_i u_i)}{E(X_i^2)} \right). 
\]

(c) \( \hat{\beta}_1^{RLS} \) is a linear estimator:

\[
\hat{\beta}_1^{RLS} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} = \sum_{i=1}^{n} a_i Y_i, \quad \text{where} \quad a_i = \frac{X_i}{\sum_{i=1}^{n} X_i^2}.
\]

The weight \( a_i \) \( (i = 1, \ldots, n) \) depends on \( X_1, \ldots, X_n \) but not on \( Y_1, \ldots, Y_n \).

Thus

\[
\hat{\beta}_1^{RLS} = \beta_1 + \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}.
\]

\( \hat{\beta}_1^{RLS} \) is conditionally unbiased because

\[
E(\hat{\beta}_1^{RLS}|X_1, \ldots, X_n) = E \left( \beta_1 + \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}|X_1, \ldots, X_n \right) \\
= \beta_1 + E \left( \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}|X_1, \ldots, X_n \right) \\
= \beta_1.
\]

The final equality used the fact that

\[
E \left( \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} X_i^2}|X_1, \ldots, X_n \right) = \frac{\sum_{i=1}^{n} X_i E(u_i|X_1, \ldots, X_n)}{\sum_{i=1}^{n} X_i^2} = 0
\]

because the observations are i.i.d. and \( E(u_i|X_i) = 0 \).

(continued on the next page)
18.1 (continued)

(d) The conditional variance of \( \hat{\beta}_1^{RLS} \), given \( X_1, \ldots, X_n \), is

\[
\text{var}(\hat{\beta}_1^{RLS}|X_1, \ldots, X_n) = \text{var} \left( \beta_1 + \frac{\sum^n_{i=1} X_i u_i}{\sum^n_{i=1} X_i^2} | X_1, \ldots, X_n \right)
\]

\[
= \frac{\sum^n_{i=1} X_i^2 \text{var}(u_i|X_1, \ldots, X_n)}{\left(\sum^n_{i=1} X_i^2\right)^2}
\]

\[
= \frac{\sum^n_{i=1} X_i^2 \sigma^2_u}{\left(\sum^n_{i=1} X_i^2\right)^2}
\]

\[
= \frac{\sigma^2_u}{\sum^n_{i=1} X_i^2}.
\]

(e) The conditional variance of the OLS estimator \( \hat{\beta}_1 \) is

\[
\text{var}(\hat{\beta}_1|X_1, \ldots, X_n) = \frac{\sigma^2_u}{\sum^n_{i=1} (X_i - \bar{X})^2}.
\]

Since

\[
\sum^n_{i=1} (X_i - \bar{X})^2 = \sum^n_{i=1} X_i^2 - 2\bar{X}\sum^n_{i=1} X_i + n\bar{X}^2 = \sum^n_{i=1} X_i^2 - n\bar{X}^2 < \sum^n_{i=1} X_i^2,
\]

the OLS estimator has a larger conditional variance:

\[
\text{var}(\hat{\beta}_1|X_1, \ldots, X_n) > \text{var}(\hat{\beta}_1^{RLS}|X_1, \ldots, X_n).
\]

The restricted least squares estimator \( \hat{\beta}_1^{RLS} \) is more efficient.

(f) Under assumption 5 of Key Concept 18.1, conditional on \( X_1, \ldots, X_n \), \( \hat{\beta}_1^{RLS} \) is normally distributed since it is a weighted average of normally distributed variables \( u_i \):

\[
\hat{\beta}_1^{RLS} = \beta_1 + \frac{\sum^n_{i=1} X_i u_i}{\sum^n_{i=1} X_i^2}.
\]

(continued on the next page)
18.1 (continued)

Using the conditional mean and conditional variance of \( \hat{\beta}_1^{RLS} \) derived in parts (c) and (d) respectively, the sampling distribution of \( \hat{\beta}_1^{RLS} \), conditional on \( X_1, \ldots, X_n \), is

\[
\hat{\beta}_1^{RLS} \sim N \left( \beta_1, \frac{\sigma_u^2}{\sum_{i=1}^{n} X_i^2} \right).
\]

(g) The estimator

\[
\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i} - \beta_1 + \frac{\sum_{i=1}^{n} u_i}{\sum_{i=1}^{n} X_i}
\]

The conditional variance is

\[
\text{var}(\tilde{\beta}_1|X_1,\ldots,X_n) = \text{var}\left( \beta_1 + \frac{\sum_{i=1}^{n} u_i}{\sum_{i=1}^{n} X_i} |X_1,\ldots,X_n \right)
\]

\[
= \sum_{i=1}^{n} \text{var}(u_i|X_1,\ldots,X_n)
\]

\[
= \frac{\sum_{i=1}^{n} \text{var}(u_i|X_1,\ldots,X_n)}{(\sum_{i=1}^{n} X_i)^2}
\]

\[
= \frac{n\sigma_u^2}{(\sum_{i=1}^{n} X_i)^2}.
\]

The difference in the conditional variance of \( \tilde{\beta}_1 \) and \( \hat{\beta}_1^{RLS} \) is

\[
\text{var}(\tilde{\beta}_1|X_1,\ldots,X_n) - \text{var}(\hat{\beta}_1^{RLS}|X_1,\ldots,X_n) = \frac{n\sigma_u^2}{(\sum_{i=1}^{n} X_i)^2} - \frac{\sigma_u^2}{\sum_{i=1}^{n} X_i^2}.
\]

In order to prove \( \text{var}(\tilde{\beta}_1|X_1,\ldots,X_n) \geq \text{var}(\hat{\beta}_1^{RLS}|X_1,\ldots,X_n) \), we need to show

\[
\frac{n}{(\sum_{i=1}^{n} X_i)^2} \geq \frac{1}{\sum_{i=1}^{n} X_i^2}
\]

or equivalently

(continued on the next page)
18.1 (continued)

\[ n \sum_{i=1}^{n} X_i^2 \geq \left( \sum_{i=1}^{n} X_i \right)^2. \]

This inequality comes directly by applying the Cauchy-Schwarz inequality

\[ \left[ \sum_{i=1}^{n} (a_i \cdot b_i) \right]^2 \leq \sum_{i=1}^{n} a_i^2 \cdot \sum_{i=1}^{n} b_i^2 \]

which implies

\[ \left( \sum_{i=1}^{n} X_i \right)^2 = \left( \sum_{i=1}^{n} 1 \cdot X_i \right)^2 \leq \sum_{i=1}^{n} 1^2 \cdot \sum_{i=1}^{n} X_i^2 = n \sum_{i=1}^{n} X_i^2. \]

That is \( n \sum_{i=1}^{n} X_i^2 \geq (\sum_{i=1}^{n} X_i)^2 \), or \( \text{var}(\tilde{\beta}_1 | X_1, \ldots, X_n) \geq \text{var}(\hat{\beta}_1^{RLS} | X_1, \ldots, X_n) \).

Note: because \( \tilde{\beta}_1 \) is linear and conditionally unbiased, the result \( \text{var}(\tilde{\beta}_1 | X_1, \ldots, X_n) \geq \text{var}(\hat{\beta}_1^{RLS} | X_1, \ldots, X_n) \) follows directly from the Gauss-Markov theorem.
18.3. (a) Using Equation (18.19), we have

\[ \sqrt{n}(\hat{\beta}_1 - \beta_1) = \sqrt{n} \frac{1}{\sigma_u} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( X_i - \bar{X} \right) u_i \]

\[ = \sqrt{n} \frac{1}{\sigma_u} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[ (X_i - \mu_X) - (\bar{X} - \mu_X) \right] u_i \]

\[ = \sqrt{n} \frac{1}{\sigma_u} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( X_i - \bar{X} \right)^2 \]

\[ = \sqrt{n} \frac{1}{\sigma_u} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left( X_i - \bar{X} \right)^2 \]

\[ = \frac{1}{\sigma_u} \frac{\sqrt{n}}{\sqrt{n}} \sum_{i=1}^{n} v_i \]

by defining \( v_i = (X_i - \mu_X) u_i \).

(b) The random variables \( u_1, \ldots, u_n \) are i.i.d. with mean \( \mu_u = 0 \) and variance \( 0 < \sigma_u^2 < \infty \). By the central limit theorem,

\[ \frac{\sqrt{n}(\bar{u} - \mu_u)}{\sigma_u} = \frac{\sqrt{n}}{\sigma_u} \sum_{i=1}^{n} u_i \xrightarrow{d} N(0, 1). \]

The law of large numbers implies \( \bar{X} \xrightarrow{P} \mu_X \), or \( \bar{X} - \mu_X \xrightarrow{P} 0 \). By the consistency of sample variance, \( \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \) converges in probability to population variance, \( \text{var}(X_i) \), which is finite and non-zero. The result then follows from Slutsky’s theorem.

(c) The random variable \( v_i = (X_i - \mu_X) u_i \) has finite variance:

\[ \text{var}(v_i) = \text{var}[(X_i - \mu_X) u_i] \]

\[ \leq E[(X_i - \mu_X)^2 u_i^2] \]

\[ \leq \sqrt{E[(X_i - \mu_X)^4]} E[(u_i)^4] < \infty. \]

The inequality follows by applying the Cauchy-Schwartz inequality, and the second inequality follows because of the finite fourth moments for \( (X_i, u_i) \). The finite variance along with the fact that \( v_i \) has mean zero (by assumption 1 of Key Concept 18.1) and \( v_i \) is i.i.d. (by assumption 2) implies that the sample average \( \bar{v} \) satisfies the requirements of the central limit theorem. Thus,

(continued on the next page)
18.3 (continued)

\[ \frac{V}{\sigma_v} = \sqrt{\frac{n}{n}} \sum_{i=1}^{n} v_i \]

satisfies the central limit theorem.

(d) Applying the central limit theorem, we have

\[ \frac{\sqrt{\frac{1}{n}} \sum_{i=1}^{n} v_i}{\sigma_v} \xrightarrow{d} N(0, 1). \]

Because the sample variance is a consistent estimator of the population variance, we have

\[ \frac{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2}{\text{var}(X_i)} \xrightarrow{d} 1. \]

Using Slutsky’s theorem,

\[ \frac{\frac{1}{n} \sum_{i=1}^{n} v_i}{\sigma_v} \xrightarrow{d} N(0, 1), \]

or equivalently

\[ \frac{\sqrt{\frac{1}{n}} \sum_{i=1}^{n} v_i}{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2} \xrightarrow{d} N\left(0, \frac{\text{var}(v_i)}{[\text{var}(X_i)]^2}\right). \]

Thus

\[ \sqrt{n}(\hat{\beta}_1 - \beta_1) = \frac{\sqrt{\frac{1}{n}} \sum_{i=1}^{n} v_i}{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2} - \frac{(\bar{X} - \mu_x)\sqrt{\frac{1}{n}} \sum_{i=1}^{n} u_i}{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2} \xrightarrow{d} N\left(0, \frac{\text{var}(v_i)}{[\text{var}(X_i)]^2}\right). \]
since the second term for $\sqrt{n}(\hat{\beta}_1 - \beta_1)$ converges in probability to zero as shown in part (b).
18.5. Let $a = W^2$. Then $\text{var}(a) = \text{E}(a^2) - \text{E}(a)^2 = \text{E}(W^4) - \text{E}(W^2)^2 \geq 0$.

Thus $\text{E}(W^2) \leq [\text{E}(W^4)]^{1/2}$, and the result follows.
18.7. (a) The joint probability distribution function of $u_i, u_j, X_i, X_j$ is $f(u_i, u_j, X_i, X_j)$. The conditional probability distribution function of $u_i$ and $X_i$ given $u_j$ and $X_j$ is $f(u_i, X_i|u_j, X_j)$. Since $u_i, X_i, i = 1, \ldots, n$ are i.i.d., $f(u_i, X_i|u_j, X_j) = f(u_i, X_i)$. By definition of the conditional probability distribution function, we have

$$f(u_i, u_j, X_i, X_j) = f(u_i, X_i|u_j, X_j)f(u_j, X_j)$$
$$= f(u_i, X_i)f(u_j, X_j).$$

(b) The conditional probability distribution function of $u_i$ and $u_j$ given $X_i$ and $X_j$ equals

$$f(u_i, u_j|X_i, X_j) = \frac{f(u_i, u_j, X_i, X_j)}{f(X_i, X_j)} = \frac{f(u_i, X_i)f(u_j, X_j)}{f(X_i)f(X_j)} = f(u_i|X_i)f(u_j|X_j).$$

The first and third equalities used the definition of the conditional probability distribution function. The second equality used the conclusion the from part (a) and the independence between $X_i$ and $X_j$. Substituting

$$f(u_i, u_j|X_i, X_j) = f(u_i|X_i)f(u_j|X_j)$$

into the definition of the conditional expectation, we have

$$E(u_i, u_j|X_i, X_j) = \int \int u_iu_j f(u_i, u_j|X_i, X_j) \, du_i \, du_j$$
$$= \int \int u_iu_j f(u_i|X_i)f(u_j|X_j) \, du_i \, du_j$$
$$= \int u_i f(u_i|X_i) \, du_i \int u_j f(u_j|X_j) \, du_j$$
$$= E(u_i|X_i)E(u_j|X_j).$$

(c) Let $Q = (X_1, X_2, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$, so that $f(u_i|X_1, \ldots, X_n) = f(u_i|X_i, Q)$. Write
18.7 (continued)

\[
f(u_i | X_i, Q) = \frac{f(u_i, X_i, Q)}{f(X_i, Q)} \\
= \frac{f(u_i, X_i) f(Q)}{f(X_i) f(Q)} \\
= \frac{f(u_i, X_i)}{f(X_i)} \\
= f(u_i | X_i)
\]

where the first equality uses the definition of the conditional density, the second uses the fact that \((u_i, X_i)\) and \(Q\) are independent, and the final equality uses the definition of the conditional density. The result then follows directly.

(d) An argument like that used in (c) implies

\[
f(u_i u_j | X_i, X_j) = f(u_i u_j | X_i, X_j)
\]

and the result then follows from part (b).
18.9. We need to prove

\[ \frac{1}{n} \sum_{i=1}^{n} \left[ (X_i - \bar{X})^2 \hat{u}_i^2 - (X_i - \mu_X)^2 u_i^2 \right] \overset{P}{\to} 0. \]

Using the identity \( \bar{X} = \mu_X + (\bar{X} - \mu_X) \),

\[ \frac{1}{n} \sum_{i=1}^{n} \left[ (X_i - \bar{X})^2 \hat{u}_i^2 - (X_i - \mu_X)^2 u_i^2 \right] = (\bar{X} - \mu_X)^2 \frac{1}{n} \sum_{i=1}^{n} \hat{u}_i^2 - 2(\bar{X} - \mu_X) \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_X) \hat{u}_i^2 + \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu_X)^2 (\hat{u}_i^2 - u_i^2). \]

The definition of \( \hat{u}_i \) implies

\[ \hat{u}_i^2 = u_i^2 + (\hat{\beta}_0 - \beta_0)^2 + (\hat{\beta}_1 - \beta_1)^2 X_i^2 - 2u_i(\hat{\beta}_0 - \beta_0) \]

\[ - 2u_i(\hat{\beta}_1 - \beta_1)X_i + 2(\hat{\beta}_0 - \beta_0)(\hat{\beta}_1 - \beta_1)X_i. \]

Substituting this into the expression for \( \frac{1}{n} \sum_{i=1}^{n} [(X_i - \bar{X})^2 \hat{u}_i^2 - (X_i - \mu_X)^2 u_i^2] \) yields a series of terms each of which can be written as \( a_n b_n \) where \( a_n \overset{P}{\to} 0 \) and \( b_n = \frac{1}{n} \sum_{i=1}^{n} X_i^r u_i^s \) where \( r \) and \( s \) are integers. For example,

\[ a_n = (\bar{X} - \mu_X), \quad a_n = (\hat{\beta}_1 - \beta_1) \] and so forth. The result then follows from Slutsky’s theorem if \( \frac{1}{n} \sum_{i=1}^{n} X_i^r u_i^s \overset{d}{\to} d \) where \( d \) is a finite constant. Let \( w_i = X_i^r u_i^s \) and note that \( w_i \) is i.i.d. The law of large numbers can then be used for the desired result if \( E(w_i^2) < \infty \). There are two cases that need to be addressed. In the first, both \( r \) and \( s \) are non-zero. In this case write

\[ E(w_i^2) = E(X_i^{2r} u_i^{2s}) < \sqrt{[E(X_i^{4r})][E(u_i^{4s})]} \]

and this term is finite if \( r \) and \( s \) are less than 2. Inspection of the terms shows that this is true. In the second case, either \( r = 0 \) or \( s = 0 \). In this case the result follows directly if the non-zero exponent (\( r \) or \( s \)) is less than 4. Inspection of the terms shows that this is true.
18.11. \( \mu_{Y|X} = \mu_Y + (\sigma_{XY} / \sigma_X^2)(x - \mu_X) \).

(a) Using the hint and equation (18.38)

\[
f_{Y|X=x}(y) = \frac{1}{\sqrt{\sigma_Y^2(1 - \rho_{XY}^2)}} \frac{1}{\sqrt{2\pi}} \times \exp \left\{ -\frac{1}{2(1 - \rho_{XY}^2)} \left( \frac{x - \mu_X}{\sigma_X} \right)^2 - 2 \rho_{XY} \left( \frac{x - \mu_X}{\sigma_X} \right) \left( \frac{y - \mu_Y}{\sigma_Y} \right) + \left( \frac{y - \mu_Y}{\sigma_Y} \right)^2 \right\} + \frac{1}{2} \left( \frac{x - \mu_X}{\sigma_X} \right)^2.
\]

Simplifying yields the desired expression.

(b) The result follows by noting that \( f_{Y|X=x}(y) \) is a normal density (see equation (18.36)) with \( \mu = \mu_{Y|X} \) and \( \sigma^2 = \sigma_{Y|X}^2 \).

(c) Let \( b = \sigma_{XY} / \sigma_X^2 \) and \( a = \mu_Y - b \mu_X \).
18.13 (a) The answer is provided by equation (13.10) and the discussion following the equation. The result was also shown in Exercise 13.10, and the approach used in the exercise is discussed in part (b).

(b) Write the regression model as \( Y_i = \beta_0 + \beta_1 X_i + \nu_i \), where \( \beta_0 = E(\beta_0) \), \( \beta_1 = E(\beta_1) \), and \( \nu_i = u_i + (\beta_0 - \beta_0) + (\beta_1 - \beta_1)X_i \). Notice that

\[
E(\nu_i | X_i) = E(u_i | X_i) + E(\beta_0 - \beta_0 | X_i) + X_iE(\beta_1 - \beta_1 | X_i) = 0
\]

because \( \beta_0 \) and \( \beta_1 \) are independent of \( X_i \). Because \( E(\nu_i | X_i) = 0 \), the OLS regression of \( Y_i \) on \( X_i \) will provide consistent estimates of \( \beta_0 = E(\beta_0) \) and \( \beta_1 = E(\beta_1) \). Recall that the weighted least squares estimator is the OLS estimator of \( Y_i/\sigma_i \) onto \( 1/\sigma_i \) and \( X_i/\sigma_i \), where \( \sigma_i = \sqrt{\theta_0 + \theta_i X_i^2} \). Write this regression as

\[
Y_i / \sigma_i = \beta_0 (1/\sigma_i) + \beta_1 (X_i / \sigma_i) + \nu_i / \sigma_i.
\]

This regression has two regressors, \( 1/\sigma_i \) and \( X_i/\sigma_i \). Because these regressors depend only on \( X_i \), \( E(\nu_i | X_i) = 0 \) implies that \( E(\nu_i | 1/\sigma_i, X_i/\sigma_i) = 0 \). Thus, weighted least squares provides a consistent estimator of \( \beta_0 = E(\beta_0) \) and \( \beta_1 = E(\beta_1) \).
18.15

(a) Write \( W = \sum_{i=1}^{d} Z_i^2 \) where \( Z_i \sim N(0,1) \). From the law of large numbers

\[
\frac{W}{n} \xrightarrow{d} \frac{E(Z_i^2)}{n} = 1.
\]

(b) The numerator is \( N(0,1) \) and the denominator converges in probability to 1. The result follows from Slutsky’s theorem (equation (18.9)).

(c) \( V/m \) is distributed \( \chi^2_m / m \) and the denominator converges in probability to 1. The result follows from Slutsky’s theorem (equation (18.9)).