The F-test (to check the goodness of linear models)

We have the model, $Y = X\beta + \epsilon$, $X_{n\times p}$ of rank $r \le p$ and with $\epsilon \sim N_n(0, \sigma^2 I_n)$. Suppose we want to test $H_0 : A\beta = c$, $A_{q\times p}$ of rank $q \le r$, and c is given. Then

$$\begin{aligned}
\mathrm{RSS} = \mathrm{SSE} &= (Y - X\hat\beta)'(Y - X\hat\beta) = Y'(I-P)Y \\
\mathrm{RSS}_{H_0} &= (Y - X\hat\beta_{H_0})'(Y - X\hat\beta_{H_0}), \text{ where} \\
\hat\beta_{H_0} &= \hat\beta + (X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1}\left\{c - A\hat\beta\right\}.
\end{aligned}$$

Theorem. Under the above mentioned assumptions, we have,
(i) $\mathrm{RSS} \sim \sigma^2\chi^2_{n-r}$;
(ii) $\mathrm{RSS}_{H_0} - \mathrm{RSS} = (A\hat\beta - c)'\left(A(X'X)^{-}A'\right)^{-1}(A\hat\beta - c)$;
(iii) $E\left(\mathrm{RSS}_{H_0} - \mathrm{RSS}\right) = q\sigma^2 + (A\beta - c)'\left(A(X'X)^{-}A'\right)^{-1}(A\beta - c)$;
(iv) under $H_0 : A\beta = c$,

$$F = \frac{(\ \mathrm{RSS}_{H_0}\text{ - RSS })/q}{\mathrm{RSS}\ /(n-r)} \sim F_{q,n-r};$$

(v) when $c = 0$,

$$F = \left(\frac{n-r}{q}\right)\frac{Y'(P - P_{H_0})Y}{Y'(I_n - P)Y},$$

where P_{H_0} is symmetric idempotent and $P_{H_0}P = PP_{H_0} = P_{H_0}$.

Proof. (i) Already known.
(ii) Note that

$$\begin{aligned}
\mathrm{RSS}_{H_0} &= (Y - X\hat\beta_{H_0})'(Y - X\hat\beta_{H_0}) \\
&= (Y - X\hat\beta + X\hat\beta - X\hat\beta_{H_0})'(Y - X\hat\beta + X\hat\beta - X\hat\beta_{H_0}) \\
&= (Y - X\hat\beta)'(Y - X\hat\beta) + (X\hat\beta - X\hat\beta_{H_0})'(X\hat\beta - X\hat\beta_{H_0}) \\
&\quad +2(X\hat\beta - X\hat\beta_{H_0})'(Y - X\hat\beta) \\
&= \mathrm{RSS} + (\hat\beta - \hat\beta_{H_0})'X'X(\hat\beta - \hat\beta_{H_0}),
\end{aligned}$$

since $(X\hat\beta - X\hat\beta_{H_0})'(Y - X\hat\beta) = (\hat\beta - \hat\beta_{H_0})'(X'Y - X'X\hat\beta) = 0$. Now from an earlier result, $(\hat\beta - \hat\beta_{H_0})'X'X(\hat\beta - \hat\beta_{H_0}) = (A\hat\beta - c)'\left(A(X'X)^{-}A'\right)^{-1}(A\hat\beta - c)$.
(iii) $A\hat\beta = MX\hat\beta = MPY \sim N_q(A\beta, \sigma^2(A(X'X)^{-}A'))$, so that $E(A\hat\beta - c) = A\beta - c$ and $Cov(A\hat\beta) = \sigma^2 A(X'X)^{-}A'$. Therefore,

$$\begin{aligned}
&E(\mathrm{RSS}_{H_0} - \mathrm{RSS}) \\
&= E\left\{(A\hat\beta - c)'\left(A(X'X)^{-}A'\right)^{-1}(A\hat\beta - c)\right\} \\
&= (A\beta - c)'\left(A(X'X)^{-}A'\right)^{-1}(A\beta - c) + \mathrm{tr}\left\{\sigma^2 A(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1}\right\} \\
&= q\sigma^2 + (A\beta - c)'\left(A(X'X)^{-}A'\right)^{-1}(A\beta - c),
\end{aligned}$$

which is large if $A\beta$ is far from c.
(iv) Note that

$$\text{RSS}_{H_0} - \text{RSS} = (A\hat{\beta} - c)' \left(A(X'X)^{-}A'\right)^{-1} (A\hat{\beta} - c) \sim \sigma^2\chi^2_q,$$

under H_0 since $A\hat{\beta} - c \sim N_q(A\beta - c, \sigma^2(A(X'X)^{-}A')) = N_q(0, \sigma^2 A(X'X)^{-}A')$. Also, RSS $\sim \sigma^2\chi^2_{n-r}$ from (i). Further, RSS is independent of $X\hat{\beta} = PY$. Since $A\beta$ is estimable, $A = MX$, so that $A\hat{\beta} = MX\hat{\beta} = MPY$, which is independent of RSS.
(v) If $c = 0$, we have,

$$\begin{aligned}
X\hat{\beta}_{H_0} &= X\left\{\hat{\beta} - (X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A\hat{\beta}\right\} \\
&= X\left\{(X'X)^{-}X'Y - (X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X'Y\right\} \\
&= \left\{X(X'X)^{-}X' - X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X'\right\} Y \\
&= (P - P_1)Y = P_{H_0}Y.
\end{aligned}$$

Clearly, P_{H_0} is symmetric. Further, P_1 is symmetric, $P_1^2 =$
$X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X'X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X' =$
$X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} \{A(X'X)^{-}X'X(X'X)^{-}A'\} \left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X'$
$= X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} \{A(X'X)^{-}A'\} \left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X'$
$= X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X' = P_1$, since the term in the middle of the expression,
$A(X'X)^{-}X'X(X'X)^{-}A' = MX(X'X)^{-}X'X(X'X)^{-}X'M' = MP^2M'$
$= MPM' = A(X'X)^{-}A'$. Also,
$P_1P = X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X'X(X'X)^{-}X'$
$= X(X'X)^{-}A'\left(A(X'X)^{-}A'\right)^{-1} A(X'X)^{-}X' = P_1$,
since $X'X(X'X)^{-}X' = X'P = X'P' = (PX)' = X'$. Note, $P_1 = (P_1)' = (P_1P)' = PP_1$. Therefore,
$P_{H_0}^2 = (P - P_1)^2 = P^2 - PP_1 - P_1P + P_1^2 = P - 2P_1 + P_1 = P - P_1 = P_{H_0}$
and $P_{H_0}P = (P - P_1)P = P - P_1 = P_{H_0} = PP_{H_0}$. Therefore,

$$\begin{aligned}
\text{RSS}_{H_0} &= ||Y - X\hat{\beta}_{H_0}||^2 = (Y - X\hat{\beta}_{H_0})'(Y - X\hat{\beta}_{H_0}) \\
&= (Y - P_{H_0}Y)'(Y - P_{H_0}Y) = Y'(I - P_{H_0})Y
\end{aligned}$$

and

$$\text{RSS}_{H_0} - \text{RSS} = Y'(I - P_{H_0})Y - Y'(I - P)Y = Y'(P - P_{H_0})Y.$$