Omitted Variables Problem

True Model:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_r X_{rt} + \beta_{r+1} X_{r+1t} \ + \cdots + \beta_K X_{Kt} + \varepsilon_t$$

Misspecified Model:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_r X_{rt} + u_t$$

where $u_t = \beta_{r+1} X_{r+1t} + \cdots + \beta_K X_{Kt} + \varepsilon_t$

Consider the regressions of the <u>excluded</u> variables on all of the <u>included</u> variables (in deviation form):

$$\begin{aligned}
x_{r+1t} &= \gamma_{r+1,1\cdot 2,3,\ldots,r} x_{1t} + \gamma_{r+1,2\cdot 1,3,\ldots,r} x_{2t} + \cdots + \gamma_{r+1,r\cdot 1,2,\ldots,r-1} x_{rt} + v_{r+1t} \\
&. \\
&. \\
&. \\
x_{Kt} &= \gamma_{K,1\cdot 2,3,\ldots,r} x_{1t} + \gamma_{K,2\cdot 1,3,\ldots,r} x_{2t} + \cdots + \gamma_{K,r\cdot 1,2,\ldots,r-1} x_{rt} + v_{Kt}
\end{aligned}$$

where for example, $\gamma_{r+1,1\cdot 2,3,\ldots,r}$ is the regression coefficient on x_{1t} from the regression of x_{r+1t} on x_{1t} , $x_{2t}, \ldots, x_{rt}$.

It can be shown that

$$\begin{aligned}
E(\tilde{\beta}_1) &= \beta_1 + \beta_{r+1}\gamma_{r+1,1\cdot 2,3,\ldots,r} + \beta_{r+2}\gamma_{r+2,1\cdot 2,3,\ldots,r} + \cdots + \beta_K \gamma_{K,1\cdot 2,3,\ldots,r} \\
&. \\
&. \\
&. \\
E(\tilde{\beta}_r) &= \beta_r + \beta_{r+1}\gamma_{r+1,r\cdot 1,2,\ldots,r-1} + \beta_{r+2}\gamma_{r+2,r\cdot 1,2,\ldots,r-1} + \cdots + \beta_K \gamma_{K,r1,\cdot 2,3,\ldots,r-1}
\end{aligned}$$

Bias Expressions:

$$E(\tilde{\beta}_1) - \beta_1 = \underbrace{\beta_{r+1}\gamma_{r+1,1\cdot 2,3,\ldots,r} + \beta_{r+2}\gamma_{r+2,1\cdot 2,3,\ldots,r} + \cdots + \beta_K\gamma_{K,1\cdot 2,3,\ldots,r}}_{\text{Bias}}$$

$$.$$

$$.$$

$$.$$

$$E(\tilde{\beta}_r) - \beta_r = \underbrace{\beta_{r+1}\gamma_{r+1,r\cdot 1,2,\ldots,r-1} + \beta_{r+2}\gamma_{r+2,r\cdot 1,2,\ldots,r-1} + \cdots + \beta_K\gamma_{K,r1,\cdot 2,3,\ldots,r-1}}_{\text{Bias}}$$