(1) First, let’s expand the traditional formula. Note that
\[
\sum_{i} (x_i - \overline{x})(y_i - \overline{y}) = \sum_{i} x_i y_i - n\overline{xy} = \sum_{i} x_i y_i - \frac{1}{n} (\sum_{i} x_i)(\sum_{i} y_i).
\]
Likewise, for the denominator:
\[
\sum_{i} (x_i - \overline{x})^2 = \sum_{i} x_i^2 - n\overline{x^2} = \sum_{i} x_i^2 - \frac{1}{n} (\sum_{i} x_i)^2.
\]
Multiplying each of these by \(n\), it follows that the traditional formula for the OLS slope estimator is equivalent to:
\[
\hat{\beta}^2 = \frac{n \sum_{i} x_i y_i - (\sum_{i} x_i)(\sum_{i} y_i)}{n \sum_{i} x_i^2 - (\sum_{i} x_i)^2}.
\]
Now, direct application of our \((X'X)^{-1}X'y\) formula gives:
\[
\hat{\beta} = \left[ \frac{n \sum_{i} x_i y_i - (\sum_{i} x_i)(\sum_{i} y_i)}{n \sum_{i} x_i^2 - (\sum_{i} x_i)^2} \right]^{-1} \left[ \frac{\sum_{i} y_i}{\sum_{i} x_i y_i} \right].
\]
Inverting the 2 \times 2 matrix and picking of the second element of \(\hat{\beta}\) shows that
\[
\hat{\beta}_2 = \frac{n \sum_{i} x_i y_i - (\sum_{i} x_i)(\sum_{i} y_i)}{n \sum_{i} x_i^2 - (\sum_{i} x_i)^2},
\]
which is the same as the above.

(2a) To accomplish this, first define \(A\) as a \(k \times k\) matrix that permutes the columns of \(X\). The matrix \(A\) will take the form that each row of \(A\) will contain only one 1, and all other entries will be zero. Similarly, each column of \(A\) will contain only one 1, and all the other elements will be zero.

Just to provide a specific example, suppose that \(X\) is an \(n \times 3\) matrix. And, let \(A\) be defined as follows:
\[
A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.
\]
If we let
\[
Z = XA,
\]
it follows that $Z$ is simply an $n \times 3$ matrix that puts the third column of $X$ as the first column of $Z$, the first column of $X$ as the second column of $Z$, and the second column of $X$ as the third column of $Z$. So, $A$ is constructed to simply permute the column order of $X$ in this particular way. This generalizes beyond the three-variable case and $A$ can be constructed to accomplish column rearranging of any desired sort. Finally, note that, however constructed,

$$AA' = A'A = I_k$$

so that $A^{-1} = A'$. Now, let us consider the OLS estimator of a regression of $y$ on $Z$ and call this estimator $\hat{\theta}$:

$$\hat{\theta} = (Z'Z)^{-1}Z'y = ([XA][XA])^{-1}[XA]'y = (A'(X'X)A)^{-1}A'X'y = A^{-1}(X'X)^{-1}A'X'y = A^{-1}(X'X)^{-1}X'y = A'\hat{\beta}$$

This also implies

$$A\hat{\theta} = \hat{\beta},$$

where $\hat{\beta}$ is the OLS estimator of a regression of $X$ on $y$. Thus, the OLS estimator produced when rearranging the columns of $X$ is simply a rearrangement of the initial OLS estimates.

(2b) No, this does not violate the rank condition. $r$ and $x$ will be correlated typically, but this construction does not imply that any one of the columns of the $X$ matrix can be expressed as a linear combination of the other columns of $X$.

(2c) This follows similarly to (2a) with the construction

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where it follows that

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$
or

\[
\begin{bmatrix}
\hat{\pi}_0 \\
\hat{\pi}_1 \\
\hat{\pi}_2 \\
\hat{\pi}_3
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & -1 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_0 \\
\hat{\beta}_1 \\
\hat{\beta}_2 \\
\hat{\beta}_3
\end{bmatrix} = 
\begin{bmatrix}
\hat{\beta}_0 \\
\hat{\beta}_1 \\
\hat{\beta}_2 - \hat{\beta}_1 \\
\hat{\beta}_3
\end{bmatrix}.
\]

This makes sense since, upon subbing \( w_i = r_i - x_i \), we get

\[
y_i = \beta_0 + \beta_1 (r_i - x_i) + \beta_2 x_i + \beta_3 z_i + u_i
\]

This substitution gives the same relationship between the coefficients \( \beta \) and \( \pi \) as suggested among the estimates \( \hat{\pi} \) and \( \hat{\beta} \) above.