Proof. We first prove $\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}(y_i-\mu(\mathbf{x}))(\mu(\mathbf{x})-\widehat{\mu}_{\mathbf{S}_j}(\mathbf{x}_i))=0$
Let $P=\cup_{i=1}^{k}A_i$ where $A_i=\{u: u\in P, x_u=x_i\}$ and the unique value of $\mathbf{x}$ is $\mathbf{x}_1,\ldots,\mathbf{x}_k$.
$|A_i|=n_i$ for all $1\le i\le k$

$$\begin{aligned}
&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}(y_i-\mu(\mathbf{x}_i))(\mu(\mathbf{x}_i)-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))\\
=&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}[\sum_{i\in\mathcal{A}_1}(y_i-\mu(\mathbf{x}_i))(\mu(\mathbf{x}_i)-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))+\cdots+\sum_{i\in\mathcal{A}_k}(y_i-\mu(\mathbf{x}_i))(\mu(\mathbf{x}_i)-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))]
\end{aligned}$$

We define $Ave_{i\in A_k}y_i=\bar{y}_{A_k}$
For $i\in A_m$ where $1\le m\le k$, we have,

$$\begin{aligned}
&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}[\sum_{i\in A_m}y_i\bar{y}_{A_m}-y_i\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)-\bar{y}_{A_m}^2+\bar{y}_{A_m}\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)]\\
=&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}n_m\bar{y}_{A_m}^2-\frac{1}{N_{\mathcal{S}}N}\sum_{i\in A_m}y_i\sum_{j=1}^{N_{\mathcal{S}}}\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)-\frac{n_m}{N}\bar{y}_{A_m}^2+\frac{1}{N_{\mathcal{S}}N}\sum_{j=1}^{N_{\mathcal{S}}}\bar{y}_{A_m}\sum_{i\in A_m}\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)\\
=&\frac{n_m}{N}\bar{y}_{A_m}^2-\frac{n_m}{N}\bar{y}_{A_m}^2+\frac{1}{N}[\sum_{i\in A_m}\bar{y}_{A_m}\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)-\sum_{i\in A_m}\bar{y}_{A_m}y_i\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)]\\
=&\frac{1}{N}[\sum_{i\in A_m}\bar{y}_{A_m}\bar{\hat{\mu}}(\mathbf{x}_i)-\sum_{i\in A_m}y_i\bar{\hat{\mu}}(\mathbf{x}_i)]\\
&\text{for all i}\in A_m, \bar{\hat{\mu}}(\mathbf{x}_i)\text{ is a constant say c}\\
=&\frac{\bar{y}_{A_m}\times n_m\times c}{N}-\frac{c\times n_m\times\bar{y}_{A_m}}{N}\\
=&0
\end{aligned}$$

$\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}[\sum_{i\in\mathcal{A}_1}(y_i-\mu(\mathbf{x}))(\mu(\mathbf{x})-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))+\cdots+\sum_{i\in\mathcal{A}_k}(y_i-\mu(\mathbf{x}))(\mu(\mathbf{x})-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))]=0$

Hence,

$$
\begin{aligned}
&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}(y_i-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))^2\\
=&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}[(y_i-\mu(\mathbf{x}_i))-(\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)-\mu(\mathbf{x}_i))]^2\\
=&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}(y_i-\mu(\mathbf{x}_i))^2-2(y_i-\mu(\mathbf{x}))(\mu(\mathbf{x})-\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i))+(\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)-\mu(\mathbf{x}_i))^2\\
=&\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}(y_i-\mu(\mathbf{x}_i))^2+\frac{1}{N_{\mathcal{S}}}\sum_{j=1}^{N_{\mathcal{S}}}\frac{1}{N}\sum_{i\in\mathcal{P}}(\widehat{\mu}_{\mathcal{S}_j}(\mathbf{x}_i)-\mu(\mathbf{x}_i))^2
\end{aligned}
$$

$\square$