	Ex 2.4
	Wits E[Ric(Bir)] = E[Ric(Bir)]
	Note: Btr refers to the least so estimates from travers
	douter (DC, 41) (XN, 4N). Similarly, Bite FeFers to the
	Note: Btr refers to the least sq. estimates from training data (DC, y1) (XN, yN). Similarly, Bto refers to the least Sq. estimates from testing data (X, y,) (SCM, yM).
	Re($\hat{\beta}_{tr}$)= $\frac{1}{M}\sum_{i=1}^{M}(\hat{y}_{i}-\hat{\beta}_{tr}^{T}\hat{x}_{i})^{2} \geq \frac{1}{M}\sum_{i=1}^{M}(\hat{y}_{i}-\hat{\beta}_{te}^{T}\hat{x}_{i})^{2}$ (as $\hat{\beta}_{te}$ is defined as the $\hat{\beta}$ that minimises this eqn.)
	So, $E(R_{te}(\hat{\beta}_{tr})) \geq E(\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_{i}-\hat{\beta}_{te}\hat{y}_{i})^{2})$
	$= \mathbb{E}\left[\left(\tilde{y}_{i} - \hat{\beta}_{t}^{T} \tilde{X}_{i}\right)^{2}\right] $ (Since any arbitrary (\tilde{x}_{i} , \tilde{y}_{i}) Pentr (is drawn from the same Population
	1 is drawn from the same Population
	$= E \left[\frac{1}{N} \sum_{i=1}^{N} (\tilde{y}_i - \hat{\beta}_{i}^{T} \tilde{x}_i)^2 \right] $ (same logic)
	Since N test Points are drawn from the same Population
	as N train Points, this is equivalent to:
	= E[/N [(yi - β+r xi)2] = E[R+r(β+r)]
	$\Rightarrow E[R_{te}(\hat{\beta}_{tr})] \geq E[R_{te}(\hat{\beta}_{tr})]$
	- LINIE (Ptr)] > LINIE (Ptr)]
1	