Question 4

(b).

From the result of (a). we know that for optimization problem

$$\min_{\beta} E[q(Y-\beta)^{+} + (1-q)(Y-\beta)^{-}] \qquad (1)$$

The optimal solution (β^{*}) is the q-th quantile y_q of the random variable Y as any point that satisfies the equation $F(y_q) = q$, where $F(y)$ is the CDF of random variable Y.
According to the definition of CDF, $F(y) = Prob(Y \leq y)$.

For question (b), we are looking for the optimal solution of optimization problem

$$\min_{\{\beta(x):\, R^{d} \to R\}} E[q(Y-\beta(X))^{+} + (1-q)(Y-\beta)^{-}] \qquad (2)$$

As $\beta(x)$ is a function that map x from R^{d} to R, and $X \in R^{d}$, expression (1) and (2) are equivalent. While the q-th quantile y_q of the random variable Y is the optimal solution for (1), given condition $X = x$, the conditional q-quantile $\beta^{*}(x)$ of random variable Y is the optimal solution for expression (2).
Therefore, q-quantile $\beta^{*}(x)$ assess how much y will change for distribution y as x change by 1 unit at quantile point q for a given set of other covariates.

(c).

The goal is to find w that $\min_{w \in R^{d+1}} [\![y - Mw]\!]_1 + (2q-1)1^{T}(y - Mw)$

For $[\![y - Mw]\!]_1$, if $y > Mw, [\![y - Mw]\!]_1 = Y - Mw$
if $y < Mw, [\![y - Mw]\!]_1 = Mw - Y$

Therefore,

$$\min_{w \in R^{d+1}} [\![y - Mw]\!]_1 + (2q-1)1^{T}(y - Mw)$$

$$= min \sum_{i \in y_i > M_i w_i}^{N} y_i - M_i w_i + (2q-1)(y_i - M_i w_i) + \sum_{i \in y_i < M_i w_i}^{N} M_i w_i - y_i + (2q-1)(y_i - M_i w_i)$$

$$= min \sum_{i \in y_i > M_i w_i}^{N} y_i - M_i w_i + (2q)(y_i - M_i w_i) - (1)(y_i - M_i w_i) + \sum_{i \in y_i < M_i w_i}^{N} -(y_i - M_i w_i) + (2q)(y_i - M_i w_i) - (1)(y_i - M_i w_i)$$

$$= min \sum_{i \in y_i > M_i w_i}^{N} (2q)(y_i - M_i w_i) + \sum_{i \in y_i < M_i w_i}^{N} (2q-2)(y_i - M_i w_i)$$

According to what is defined in the question,

$\beta(x)$ is restricted to be of the form $\beta(x) = \left[X_i^{T} 1\right] * w_i$, and M_i is $[X_i^{T} 1]$

We could rewrite the expression

$$min \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} (2q)(y_i - M_i w_i) \; + \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} (2q-2)(y_i - M_i w_i)$$

$$= \; min \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} (2q)\big(y_i - \beta(x_i)\big) \; + \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} (2q-2)\big(y_i - \beta(x_i)\big)$$

$$= \; min \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} q\big(y_i - \beta(x_i)\big) \; + \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} (q-1)\big(y_i - \beta(x_i)\big)$$

$$= \; min \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} q\big(y_i - \beta(x_i)\big) \; + \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} (1-q)(-(y_i - \beta(x_i)))$$

Define $(x)^+ \; := \max\{x, 0\}, when\ x < 0, (x)^+ = 0, and\ when\ x > 0, (x)^+ = x$

So

$$\sum_{i \,\epsilon\, y_i > M_i w_i}^{N} q\big(y_i - \beta(x_i)\big)$$

$$= \; q \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} 0 + \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} \big(y_i - \beta(x_i)\big)$$

$$= \; q \sum_{i}^{N} (y_i - \beta(x_i))^+$$

Define $(x)^- \; := \max\{-x, 0\}, when\ x < 0, (x)^+ = -x, and\ when\ x > 0, (x)^+ = 0$

So

$$\sum_{i \,\epsilon\, y_i < M_i w_i}^{N} (1-q)(-(y_i - \beta(x_i)))$$

$$= (1-q) \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} -\big(y_i - \beta(x_i)\big) + \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} 0$$

$$= (1-q) \sum_{i}^{N} (y_i - \beta(x_i))^-$$

Therefore

$$min \sum_{i \,\epsilon\, y_i > M_i w_i}^{N} q\big(y_i - \beta(x_i)\big) \; + \sum_{i \,\epsilon\, y_i < M_i w_i}^{N} (1-q)(-(y_i - \beta(x_i)))$$

$$= \min q \sum_{i}^{N} (y_i - \beta(x_i))^+ + (1-q) \sum_{i}^{N} (y_i - \beta(x_i))^-$$

$$= \min \sum_{i}^{N} [q(y_i - \beta(x_i))^+ + \; (1-q)(y_i - \beta(x_i))^- \,]$$

$$= \min E[q(Y - \beta(X))^+ + \; (1-q)(Y - \beta(X))^-]$$

Therefore, we prove that $\min E[q(Y - \beta(X))^+ + \; (1-q)(Y - \beta(X))^-]$

$$\approx \min_{w \,\epsilon\, R^{d+1}} [\![y - Mw]\!]_1 + (2q-1)1^T(y - Mw)$$