Question 5a) Discuss what would happen to the primary results if $q_p < q_\phi$. Prove your main results.

Recall that for a given violation $v \in [0,1]$, $q_p v$ is the probability with which nature reveals a violation after an employee reports the violation privately to the manager and the manager chooses not to fix. Likewise, $q_\phi v$ is the probability nature reveals a violation after an employee chooses to remain silent. So the assumption $q_p < q_\phi$ implies nature is less likely to reveal a violation if the employee reports privately. As a consequence, we now expect that both types of managers will always want an employee to report privately rather than stay silent, irrespective of whether they plan to fix a violation. That proof follows.

By Lemma 1a, we have the first cutpoint v-hat, the threshold value for which a manager prefers to fix a violation. That is, a manager of type $s \in \{0,1\}$ prefers to fix a given violation v iff v-hat(s) ≤ v. Specifically, v-hat(1)=0 and v-hat(0) = $\alpha/q_p(\alpha+\delta)$. These two results do not depend on the paper's claim that $q_p > q_\phi$, so remain consistent for the reverse assumption. However, v-dot[1], the cutpoint where a manager prefers to be privately informed of a violation (derived in Lemma 1b), partially relies on the relationship between q_p and q_ϕ, so we review each manager's preferences given the reverse assumption $q_p < q_\phi$.

A high-type manager (s=1) derives zero utility from fixing a violation and always derives negative utility from not fixing a violation or not being informed of a violation, so trivially we have v-hat(1)=v-dot(1)=0. Thus, the high type manager always fixes a violation that is privately reported, and a high-type manager always prefers an employee to privately report a violation rather than stay silent. This result does not rely on the assumption $q_p < q_\phi$ and so remains consistent from the original version of the paper.

To derive the preferences of a low-type manager (s=0), we compare the payoffs when an employee stays silent to the payoffs when an employee privately reports. When an employee stays silent, the utility of the low-type manager is $-q_\phi v(\alpha+\delta)v$. Conversely, when an employee privately reports, we consider two sub-cases: the low-type manager either does not fix or does fix the violation. If v<v-hat(0), the manager does not fix the violation, and receives payoff $-q_p v(\alpha+\delta)v$. Given $q_p < q_\phi$, we observe that the low-type manager is now strictly better off if the employee privately reports when v < v-hat(0).

Turning to the other case, if v≥v-hat(0), the manager fixes the violation and receives payoff $-\alpha v$. So with v≥v-hat(0), the manager is better off when the employee reports privately iff his payoffs are higher, that is $-\alpha v \geq q_\phi v(\alpha+\delta)v$. Rearranging terms and solving for v yields the condition: $v \geq \alpha/q_\phi(\alpha+\delta)$. Recall that v-hat(0) = $\alpha/q_p(\alpha+\delta)$. So if v≥v-hat(0) and $q_p < q_\phi$, then the above condition always holds with certainty.

Combining these two cases for the low-type manager, we have that when $q_p < q_\phi$, v-dot(0)=0. In words, a low-type manager now prefers to be privately informed of all violations. But he continues to fix the violation iff v≥ v-hat(0) = $\alpha/q_p(\alpha+\delta)$. Briefly, there are three implications for the paper's later results. First, we observe that both managers now share an interest in preventing employee silence and that neither will deter private reporting by employees. So both manager types levy a maximum penalty for employee silence and no penalty for private reporting. Second, if the manager's type is publicly known, then a low-type manager will still impose a maximum whistle-blowing penalty when v < v-hat(0), as

[1] Formally, a manager of type $s \in \{0,1\}$ prefers to be privately informed of a violation v iff v-dot(s) ≤ v.

ethical employees know these violations will not be fixed if privately reported. And third, a separating equilibrium will still exist for a sufficiently large share of ethical employee types in the population.

Question 5b) What happens if the types (i.e. s and t) are correlated? What results would this affect?

We assume a positive correlation, so ethical employees are more likely to work with ethical managers, and the reverse holds for non-ethical managers. We also assume this correlation is known by all types, and thus affects expectations about the manager type β and the employee type distribution η. We restrict our attention to the no penalties model for ease of analysis.

For low type violations with high-type managers, β is higher and therefore $T_{\phi p}$ is lower. In expectation, more employees will private report because the value of t is higher and the threshold $T_{\phi p}$ has dropped. Naturally, the reverse holds for low-type managers, with β lower and therefore $T_{\phi p}$ higher, and thus more employees choose to stay silent rather than privately report.

For medium type violations, $T_{\phi p}$ is now independent of β, and so does not change for either manager type. But the distribution of employees above and below $T_{\phi p}$ will now change depending on manager type. We expect more employees above $T_{\phi p}$ for high type managers (below $T_{\phi p}$ for low-type managers) and thus more private reporting of medium type violations (less private reporting of medium violations).

Finally, for high violations, changes in β have no impact on $T_{\phi p}= 0$ or $T_{pw}=1$, so all high violations continue to be privately reported independent of the manager's type, and are fixed with certainty by both types.