Main contribution: We make significant contributions in this work by completely removing the independence number $\alpha(G)$ from known regret bounds and replace it by the domination number $\beta(G)$. We note that there have been other recent works which also aim at replacing other graph-theoretic quantities like clique partition (e.g. Stochastic Bandits with Side Observations on Networks by Buccapatnam et al. and Graph regret bounds for Thompson Sampling and UCB by Lykouris et al.) by the domination number. These works are motivated by real-world social network graphs, where the clique partition number is much larger compared to the domination number. Note that for star-graphs or star-like graphs, the domination number is substantially smaller than the independence number. It is also important to note that we fix an error in the lower bound of Rangi and Franceschetti.

Main algorithmic novelty: Let us emphasize that the main algorithmic novelty in our work is the star graph algorithm (Algorithm 1). We note that the adaptive mini-batch idea is novel and that even though the algorithm seems natural it was non-trivial to prove regret guarantees. For example another natural algorithm, which decides to remain at the current arm with some probability (as in Rangi et al.) can be shown not to achieve a bound depending on the domination number. We also note that applying the corralling algorithm was not straightforward.

Suboptimality of the bound: Even though our intuition suggests that our upper bound might be suboptimal, this is only an intuition. It might be the case that our lower bounds are actually not tight and that the optimal regret bound is some interpolating quantity between $\sqrt{\beta(G)}$ and $\alpha(G)^{1/3}$. We wish to stress that we tried other, maybe more natural algorithms based on the star-graph algorithm, which did not require corralling. However, our attempts were unsuccessful and it is very likely that the most natural extension of the star-graph algorithm, which just constructs adaptive mini-batches based on the revealing arms can provably not work. The conjectured suboptimality likely stems from the corralling algorithm and either a more careful analysis or an improved corralling algorithm taking into account the additional disjoint structure of arms and the homogeneous nature of the corralled algorithms are the key to improving the bound. Both tasks, however, are very non-trivial and are left as future work.

Motivation: Our motivating examples are stated on lines 28-38 and 198-205. Both the stock market example and the investor example (lines 28-38) are important and do actually occur in real world. Another example is the treatment of patients. In most cases it is harmful to switch between different treatments for the same patient and additional information can be recovered from other patients who were treated differently but have the same disease sub-type. This example also motivates the feedback graph structure arising in the policy regret minimizing section. Indeed we expect to see counterfactual feedback only about the same type of treatment administered for longer periods of time and it is less likely that we get to observe counterfactual feedback about switching treatments often. We would like to note that technically the graph does not need to be restricted in Section 4, however, our algorithm only needs to see the described feedback. This is because of the comparator class against which we compete. We can also give other examples both for the policy regret minimizing setting and switching costs setting.

Reviewer 1: Please see paragraphs about motivation, main algorithmic novelty and suboptimality of the bound.

Reviewer 4: Please see paragraphs about main contributions, main algorithmic novelty, suboptimality of the bound and motivation. We can add experiments in the extended version of the paper. We would like to emphasize that $\sqrt{\beta(G)} \ll \alpha(G)^{1/3}$ for many graphs. Take for example the star graph or union of star graphs, where $\alpha(G)$ is order number of vertices, while $\beta(G)$ is small or constant!

Reviewer 2: We will polish the Appendix more. Regarding the improvement of the regret upper bound, please see the paragraph above on the suboptimality of the bound. We thank the reviewer for pointing out the typos. *Regarding domination number in prior work:* We refer the reviewer to Alon et al. 2015, Buccapatnam et al. 2014 and Lykouris et al. 2019 which we will reference. In general, when there are no switching costs there are two settings in which either $\sqrt{\alpha(G)T}$ is optimal or $\beta(G)^{1/3}T^{2/3}$ is optimal, depending on the available feedback. Please see Appendix A for additional discussion. *Regarding mini-batch sizes:* Mini-batch sizes are always integers. We apologize for the confusion and will add a floor function in the definition. This will not change the final result or the proofs for Section 3 since the important term to control for the regret bound only decreases when we take the floor of τ_t. *Regarding exploration parameter, high probability bounds and discarding feedback:* We thank the reviewer for taking the time to carefully read the paper and notice the extra exploration parameter. This parameter is needed to ensure that $\lfloor \tau_t \rfloor \geq 1$, so that the algorithm terminates. We will fix the typo in Algorithm 1 and elaborate more about the role of the parameter. We also agree that this parameter can help when trying to show high probability regret bounds, however, we expect these to be very non-trivial to show. We also agree that discarding feedback seems prohibitive but we expect that there are adversarial sequences in which the feedback does not really contribute much. For example consider the star graph case, where the leaf arms all have the same loss of 1 and the revealing arm is the best arm. In this case we gain almost no information by playing a leaf arm. In practice it might happen that an algorithm which takes into account the additional information performs better in some settings. We will run experiments to try and verify the above claim. *"- p16, l514: I did not understand why Thm 3.1 implies the condition of Thm C.5 with alpha=1/2 and not 1":* The ρ_t in Theorem C.5 is defined differently than the ρ in Theorem 3.1. In particular the second can be thought of as a squared version of the first and hence we get $\alpha = 1/2$. We will fix this discrepancy in the final version of our work.