We would like to thank the reviewers for their interesting comments and questions.

Reviewer 1: It is great to hear that you really like the conceptual point of our paper and that you appreciate our strong technical contribution.

Q1: *Does chaotic behavior persist if agents use different values of epsilon?* Great question! In Appendix I, pages 33-36 of the full paper, we examine the case of two learning rates ϵ_1 and ϵ_2 with both theoretical and experimental results (see Figs 11, 12). These results can be extended to any finite number of learning rates by paying mostly a notational overhead. Rev. 1 has an excellent idea about a continuum of learning rates distributed (without atoms) over the population. This is a very interesting open question. Novel ideas and techniques would be required. Unfortunately, this setting is not really conducive to simulations as the full system state is a continuous measure. A discretized version with a finite large set of learning rates would result in a more complex version of our ϵ_1 and ϵ_2 figures. We will be happy to expand upon this in a discussion section for future work.

Q2: *Scaling of costs in* $[0, 1]$*:"If an agent is repeatedly interacting with a system at a fixed load, I'd expect them to adapt their learning to that environment, which (arguably) corresponds to scaling to* $[0, 1]$*."* There is a subtle but important distinction to be made here. The normalization you describe here does not result in games with costs in $[0, 1]$, not even approximately. Let's assume when you travel to work it takes somewhere between 10 to 30 minutes and hence you normalize your costs by dividing by 30 or even 60 to be safe. Your worst case cost can still be arbitrarily larger than 1 especially in games with large populations as it corresponds to the case where you choose the longest possible route and *everyone else* in your city chooses that route too. If one was to work on a formal model of the suggested normalization two things should be clear: 1) whether chaos emerges or not still has to be carefully explored using our techniques 2) this is not the normalization in PoA literature that assumes that not just *your typical operating costs* but the much more demanding condition that even *worst case costs* lie in $[0, 1]$. *" if . . . the load of a system is actually changing (slowly) over time, I think it's reasonable that agents wouldn't update their scaling in response."* We agree and we want to point out that in (Lykouris et al. Learning and efficiency in games with dynamic population, SODA 2016) although the authors explicitly focus on a time-dependent population size, they assume that *all* stage game (worst case) costs are bounded in [0,1] (e.g. Theorem 3.1). Since this modelling assumption is not always easily applicable and as we show it can totally dictate system performance, we need to study it carefully.

Reviewer 4: Thank you for enumerating such a long list of strengths for our paper.

In regards to the background on dynamical systems, as you point out, we included materials both in the main part as well as the full paper. Although the theoretical analysis is technical, the message itself is easily understood even by non-experts. We believe that there is important value and insights both for theoretical minded people as well as experimentalists, especially given our very thorough experimental study with numerous figures. We agree that NeurIPS space constraints are tight and have forced us to push the presentation of some important results, as you point out, into the appendix. Nevertheless, all of our results are at least described completely in the introduction and we hope that the interested reader would benefit from our expansive supplement and directly jump in the section that they find more interesting. Finally, we respectfully disagree that our paper is only tangentially related to NeurIPS. The Palaiopanos et al. paper was published at NeurIPS in 2017 with a spotlight distinction. Since then it has received 46 citations including from at least 8 other NeurIPS papers. The ideas relating to period 3 inducing Li-Yorke chaos in dynamical systems have since then found new applications related to the representational power of Deep Neural Networks: Chatziafratis et al. Depth-Width Trade-offs for ReLU Networks via Sharkovsky's Theorem ICLR '20 (spotlight); Chatziafratis et al. Better Depth-Width Trade-offs for Neural Networks through the lens of Dynamical Systems ICML '20.

We are very grateful for your detailed comments. We will incorporate them in an updated version of our paper.

Reviewer 7: Thank you for your support of our paper.

Are there implications for MWU in practice? In applications Exp3, the bandit version of MWU, is typically used. Recent work on wireless network selection, which is a congestion game, has shown that Exp3 is unstable and has bad performance necessitating new domain-specialized algs (Oh et al. Periodic Bandits and Wireless Network Selection ICALP '19, Appavoo et al. Shrewd Selection Speeds Surfing: Use Smart EXP3! Int. Conf. on Distributed Computing Systems '18). These issues are important and largely unsolved from a practical perspective.

The studied games are relatively simple, even those games which the author claims the results can be extended to. The emergence of chaos is clearly a hardness/complexity type of result. Such results only increase in strength the simpler the class of examples is. Even a single instance of a bad example suffices as, e.g., in Palaiopanos et al. We instead, show that *all* simple networks, no matter their costs, number of links, (except for the case of two links with equal costs) will *all* exhibit chaos, *every single instance of them.* The simplicity, robustness of our games is a major strength of our paper.

Our regret formula is correct as $\min\{\sum_{n=1}^{T} \alpha N x_n, \sum_{n=1}^{T} \beta N\,(1 - x_n)\}$ captures the minimum aggregate cost over time of the two paths. We need the knowledge of x_n to compute that.