Reply to Anonymous Reviewer #1

Runoff ratio did change during the experiment (see new graph in Supplement, when uploaded). The primary change occurred after the first run of the replicates, or after run 2 in the case of the 5% slope. Later runs were not significantly different form one another. We recognized this, which is why we decided to do our analyses primarily in terms of cumulative runoff rather than time or cumulative rainfall. It is an interesting result, but we don't think that it alters the message of the paper.

Mixing and uniformity of the initial soil box was a challenge. The only way that we could come up with for doing a better job in uniformity would have been to sieve the material into size fractions, recombine and mix the material in common proportions, and then place in the box. This would have involved sorting and sieving tons of material. But the problem even with that is then getting the material into the box in a uniform manner. One can see in Figure 1 that initial rock cover did vary for the 6 experiments. However, based on what we saw happen this was due to the way the material was placed into the box rather than the rock content of the bulk soil in each experiment. In other words, the fine material pretty quickly was washed off the surface of the rocks and later rock cover amounts did not correlate with the initial rock cover values. We did not notice differences in the spatial variabilities of the rock cover. We were trying to get at that by our measurements of the rock cover on the three slope positions, but we did not see differences (as reported in the paper) that we might have expected. Yes, there was significant change in rock cover even at the top of the plot, which actually we were not expecting. We expected that it would take significantly longer for the rock cover to form on the upper plot, but as we can see in Figure 2 this did not happen. We expect that was due to splash erosion, but we have no way to know for certain what the relative mechanisms of splash vs. flow detachment were occurring. We probably need to do more experiments to flesh that out.

We added a couple of photos of the experimental setup and rock cover in the supplement as examples. Good idea.

Text added: "*Values of Chezy C and Manning's n were calculated using the standard equations for each and the measured velocities and known slopes. Hydraulic radii were calculated from the measured average discharge and flow velocities.*" These are straightforward calculations. The tricky part is, as we discuss in the paper, the interpolations needed to match up rock cover and roughness measurements with the velocity measurements.

We are not sure what more to say about the differences in timing of the different slope experiments. It is difficult, if not impossible to know a priori when to stop, or how long to run, the different experiments at the different slopes to have comparable end states. For example, it might make sense to stop when sediment loads reduce to a certain value, or when the total sediment lost is of a certain value, but sediment is only measured after the experiment. There is no (we have no) way to instantaneously measure the soil loss as it occurs. We had to base the decision on what the plots looked like as they evolved. There is no doubt that there was variation in the degree of evolution of the various experiments, but

we do argue that the state in all cases was "well developed" based both on our observations at the time of the experiments and from the data analyses after the fact.

Regarding the comment on the "shifts in rock fragment cover with slope," we are not exactly sure what is meant here. We can say that we did sort of expect that there would be different final rock covers for the different slopes, and were surprised when it did not pan out in that way. I guess we learned something new here. We plan to follow up on this in the field under natural conditions to try understand what processes are controlling the rock cover, as well as the roughness and velocity, and we are formulating experiments to do so for this summer. The point of this experiment, and doing it prior to the field experiments, was to try to see what happens when we can control the conditions as much as possible. I think that we did approach maximum rock cover. It was certainly above or on the very high end in all cases compared to our field measurements.

No, we do not mention erosion in this paper. Initially we had planned to combine the erosion and runoff results into a single paper, but we felt that inclusion of the sediment data here would dilute the results of the runoff and velocity measurements. The focus here was really to test the slope-velocity-equilibrium hypothesis. Erosion data will be a follow-up to this paper, where we can give it the focus requisite to that data.

Regarding the characteristics of the flow: we'd like to know more also: tortuosity, flow depths compared to roughness elements, and local hydraulic roughness included. The challenge as we see it is to quantify the spatial variability of the flow depth, and how and if flow depths correlate to the rock cover or roughness elements. What is also not included in this paper (along with the erosion data) is our terrestrial LiDAR data. We hope to use this data to try to get at some of the more thorny questions about the variability of the flow depths, variability of velocity (not sure yet how to do that, but would involve models), and why we ended up in these experiments with different hydraulic and physical roughness, yet essentially the same rock cover amounts. Deeper flows – different conclusions? Perhaps, which would have implications for downslope erosion variability. We are setting up field experiments this summer to try to get at some if this, where we will have virtual plots along the slope and us LiDAR and photo imaging with tracers to try to get roughness and velocities.

Good comments. Thanks.