Referee 2.

Firstly many, many thanks for the very helpful comments.

1) There really is no difficulty with extending the axioms to multivariate continuous distributions where the slides and squeezes can readily be interpreted respectively in terms of a multivariate distance measure and the joint distribution. The main difficulty in the context of jointly distributed discrete and continuous variables is applying the squeeze and slide transforms to the discrete components of the distributions because one has to contemplate changes in the values of the outcome variables and lumpy swaps of mass between outcome values. I've tried to address this in the discussion.

2) I have extended and hopefully clarified the discussion of what is meant by the identified, partially identified and non identified cases and I've more clearly labeled the distributions as population distributions.

3) I'd not thought of Mahalanobis, it's a good idea, indeed there are other distance measures (Bregman for instance) that would be equally interesting I've noted as much in footnote 8. Following DER I mean standardized each of the variables and used Euclidean distance largely because it is the simplest and I did not want to complicate issues more than necessary (the covariances would have to be estimated and that fact accommodated in the development of the variance of the estimator of P)

4) I've sketched the proof of the variance formula in the appendix and given references as to where it can be obtained.

5) The continuous component of f is estimated using a multivariate kernel which is used in calculating all possible polarization rectangles in the sample. The average value of these rectangles (which is what the estimator of the index is) is simply the average value of all of these. I've changed the first section to emphasize this (it also helped with the development of the variance formula in the appendix incidentally).

6) I've (hopefully!) clarified the notation.

All the minor comments have been accommodated.