lar treatment and the points at which the curve bifurcates are precisely those at which the treatments begin to differ. Again, it would be hard to estimate such a bifurcating curve using kernel methods, but a way of thinking based on penalized least squares and spline smoothing gives a natural way to proceed.

In conclusion, it would be facile to suggest that any particular method will yield answers to all the problems one might encounter. But I do believe that if we are making detailed comparisons between different approaches, the spline smoothing method as a general approach has a great deal to commend it.

Rejoinder

C.-K. Chu and J. S. Marron

We are very grateful to the discussants for their interesting and thoughtful additions to the points we made in this paper. We also thank the editor for many helpful comments, and for the nontrivial task of organizing the discussion.

Our response is organized into sections, with the first three concerning topics raised by more than one discussant, followed by some individual responses in alphabetical order.

1. PHILOSOPHICAL ISSUES

Useful elaboration of our P1-P2 formulation of the viewpoints that have been adopted to consider smoothing, has been provided by Grund and Härdle and by Silverman.

We agree with Grund and Härdle that computational issues are very important and welcome the addition of their P3 as a general principle. However, in the present context, we do not view proper incorporation of this factor as having a major impact on the ideas indicated here. The reason is that both $\hat{m}_C$ and $\hat{m}_E$, when properly implemented, for example, as described in Section 3 of Grund and Härdle, have roughly comparable computation time. On the other hand, this P3 could easily become vital in, for example, a comparison of splines versus kernels as suggested by Silverman.

We also find Silverman's P4 and the surrounding discussion very useful. This is a very nice extension of the points we were trying to make. One small point we would like to clear up is that when we attached the phrase "nonparametric regression estimation" to P2, we were referring to the phrase, not the methodology. Our intention was to convey the point that most people who have used this phrase in the literature tend to lie in the P2 camp. However, we wholeheartedly endorse Silverman's main point (also expressed well by Grund and Härdle) that there needs to be more combined use of P1 and P2.

2. ADDITIONAL COMPARISON

Some interesting and carefully considered alternative ways of comparing $\hat{m}_C$ and $\hat{m}_E$ are presented by Grund and Härdle and by Hart. We welcome these deeper analyses and are happy that the main conclusions are not much different from what we saw by simpler methods.

The figures of Grund and Härdle are very informative and provide excellent visual quantification of the points we were driving at in the paper. However, we caution against trying to infer too much from these examples. We are hesitant to make a recommendation as to which estimator is better on the basis of the size of the region where the ratio is bigger than one, because this is only one example. Even if one looked at several such examples, there are doubtless other examples that give the opposite conclusion. Furthermore, even in the presented example caution is indicated, because these sizes of regions are also dependent on the parameterization that has been chosen for the example. For example, the regions seen in Figure 4 could look quite different if this picture were based on the logarithms of these two parameters. But, of course, the main point here is that $\hat{m}_C$ and $\hat{m}_E$ are not really comparable in terms of one being always preferable, and the figure illustrates this in a compelling fashion.

Hart's idea of looking at the joint probability structure of $\hat{m}_C$ and $\hat{m}_E$ is excellent. He admirably illustrates that this is an important issue, and things are not as one might expect at first guess. This clearly needs to be borne in mind in future comparisons of estimators.

3. OTHER KERNEL ESTIMATORS

Hart and Jones discuss alternative kernel smoothers to $\hat{m}_C$ and $\hat{m}_E$ and make some interesting cases for their serious consideration. We were