

DISCUSSION OF “COAUTHORSHIP AND CITATION NETWORKS FOR STATISTICIANS”

BY FORREST W. CRAWFORD

Yale School of Public Health

I congratulate the authors on this valuable contribution to the statistics profession, and for their diligent work collecting the coauthorship and citation data sets upon which their analysis is based. The paper provides a valuable perspective on an important aspect of relationships between papers and between individual statisticians in a few of the most prominent journals in the field. The authors’ analyses of these relationships yields new insights, paves the way for future data collection efforts, and provides a valuable data set for further analytical exploration. In this brief discussion, I will give a few general comments and questions, and give suggestions for future work. I focus on three areas: author, paper and journal attributes, selection of authors and journals, and the role of time in the process of research collaboration and citation.

1. The role of author, paper and journal attributes. Collection of additional metadata, including paper content and author characteristics (current institution, department, Ph.D. institution, time since Ph.D., dissertation advisor, etc.), could potentially yield additional insight into the complex process of coauthorship and citation. The characteristics of authors themselves may be important. For example, what proportion of authors are students versus professors? Often the order of authorship matters: usually the first author did most of the work, and the last author is in charge. Middle authors have moderate contributions. How does author order arise from coauthorship arrangements? Is the more senior author usually the last author?

One might expect that coauthorship relationships are most common among authors who have been physically proximate in the past or currently. Coauthorship within the same institution might be most common. Is the same true for coauthorship in the same department? Can we learn about the collaborative character of academic statistics and biostatistics departments by studying the pattern of collaborations within and between them? Online access to scholarly publications, blogs and researcher websites have made it increasingly easy to identify potential collaborators all around the world. We might expect the prevalence of collaboration across large geographic distances to become more common as information barriers become less pronounced.

Received August 2016.