

CURRENT ISSUES IN CANADIAN AGRI-FOOD POLICY

Eighth Annual Canadian Agri-Food Policy Conference

JANUARY 24-26, 2018 • CHÂTEAU LAURIER • OTTAWA

POSTER 11

Is There Too Much History in Historical Yield Data

Presenter: Yong Liu, PhD Candidate

Supervisor: Alan Ker

Food, Agricultural and Resource Economics

Institute for the Advanced Study of Food and Agricultural Policy

University of Guelph

County crop yield data from United States Department of Agriculture - National Agricultural Statistics Service (USDA-NASS) has and continues to be extensively used in the literature as well as practice. The most notable example is crop insurance; RMA uses this data to rate and conduct claims for their area yield and revenue programs. Examples from the literature include investigation of rating methodologies, issues related to land use, modeling the climate-yield relationship, and productivity analysis. In many of these applications --- and certainly with respect to RMA and the crop insurance literature --- yield data are detrended and adjusted for possible heteroskedasticity and then assumed to be independent and identically distributed. For most major crop-region combinations, county yield data exist from 1955 onwards and reflect very significant innovations in both seed and farm management technologies. Despite correcting for changes in the first two moments of the yield data generating process (DGP), these innovations raise doubt regarding the identically distributed assumption. This manuscript considers the question of how much historical yield data should be used in empirical analyses. The answer is obviously dependent on the empirical application, crop-region combination, econometric methodology, and chosen loss function. Nonetheless, we attempt to tackle this question in three ways using county-level yield data for corn, soybean, and winter wheat. First we use distributional tests to assess if and when the adjusted yield data may result from different DGPs. Second, we consider the application to crop insurance by using an out-of-sample rating game --- commonly employed in the literature --- to compare rates from the full versus restricted data sets. Third, we estimate flexible time-varying DGPs and then simulate to quantify the additional error when the identically distribution assumption is imposed. Overall, the results indicate that despite accounting for time-varying movements in the first two moments, using yield data more than 30 years old can substantially increase estimation error. Given that discarding data is unappetizing --- particularly so in applications with relatively small T --- we investigate three methodologies that can re-incorporate the discarded data while both explicitly acknowledging the unknown DGPs are different and not requiring knowledge about the extent or form of those differences. Our results suggest gains in efficiency can be realized by using these methodologies. While our results are most applicable to the crop insurance literature, we certainly feel they suggest proceeding with caution when using historical yield data in other applications as well.