2020

36th IARIW General Conference

Prediction Intervals for Estimates of Quarterly Labor Productivity

Jay Stewart

John Glaser

Kendra Hathaway

Peter Meyer

Jerin Varghese

Background: The US Bureau of Labor Statistics (BLS) produces estimates of quarterly labor productivity growth that combine output data from the US Bureau of Economic Analysis (BEA) with labor hours data from the BLS. BLS produces a preliminary estimate, a first revision, and a second revision for each quarter. Revisions to output are due to the replacement of projections with survey or other data, revision of seasonal factors, and occasional comprehensive revisions that redefine the output concept. Revisions to hours occur due to receipt of additional data, annual benchmarking, and revision of seasonal factors. Thus productivity estimates continue to be revised long after the second revision, and our maintained assumption is each revision moves the estimate closer to "truth." (For this project, we ignore sampling and non-sampling error, and focus on what Manski (2015) refers to as transitory uncertainty.)

Currently, BLS provides reliability estimates for the labor productivity index, but not for estimates of labor productivity growth. The goal of this project is to develop prediction intervals for initial estimates of labor productivity growth so that we can communicate to data users (policymakers, forecasters, etc.) how close subsequent estimates are likely to be from initial estimates.

Phase I of this project examines revisions of preliminary and first revision estimates relative to the second revision estimate. We find that the distributions of both types of revisions are left skewed and fairly compact. There is no evidence of a trend in either of the revisions, and the size of the revisions are not correlated with initial values. Next, we examine revisions to output and hours separately. Decomposing the revisions, we find that the revisions to labor productivity are primarily driven by revisions to output. There is some variation in the revisions by quarter, mainly due to the timing of regularly scheduled revisions.

Finally, we construct prediction intervals based on historical revisions. We consider two alternative methodologies for constructing 70-percent and 90-percent intervals—conventional prediction intervals and percentile-based intervals. To compare these methods, we use the leave-one-out cross validation methodology. We calculate intervals and determine whether the left-out observations fall within the prediction interval and then calculate the accuracy to determine the fraction of revised estimates that fall within those intervals. Both methods performed similarly in our initial tests, although the conventional prediction intervals appear to be more sensitive to outliers.

Phase II examines longer-run revisions to the growth rates of labor productivity, output and hours. We find that the timing of revisions to output and hours is different. Most of the revisions to hours (mainly due to revisions to employment) occur in the first 2-3 years after the reference quarter. In contrast, output is never really final, although the revisions become much smaller about 5 years after the reference quarter.

We are in the process of developing prediction intervals for the three initial estimates (preliminary, first revision, and second revision) relative to a "final" estimate. Determining which estimate is final is critical. Estimates based on data from, say, 10 years after the reference quarter (CV10 – current values 10 years later) will be more accurate than those based on data 5 years hence (CV5). But using CV10 values results in fewer observations to calculate prediction intervals. Given that revisions become small after 5 years, we use the estimates based on the latest information 5 years after their respective reference quarters (CV5) as our "final" estimates. As expected, these prediction intervals are larger than those developed in Phase I. We also find that prediction intervals for year-over-year estimates are much narrower than those for quarter-to-quarter estimates.

Phase I of the project is complete, and much of the data work for Phase II has been done. We anticipate completing Phase II by early Spring 2020.

Reference

Manski, Charles F. (2015) "Communicating Uncertainty in Official Economic Statistics: An Appraisal Fifty Years after Morgenstern." Journal of Economic Literature 53(3), pp. 631-653.