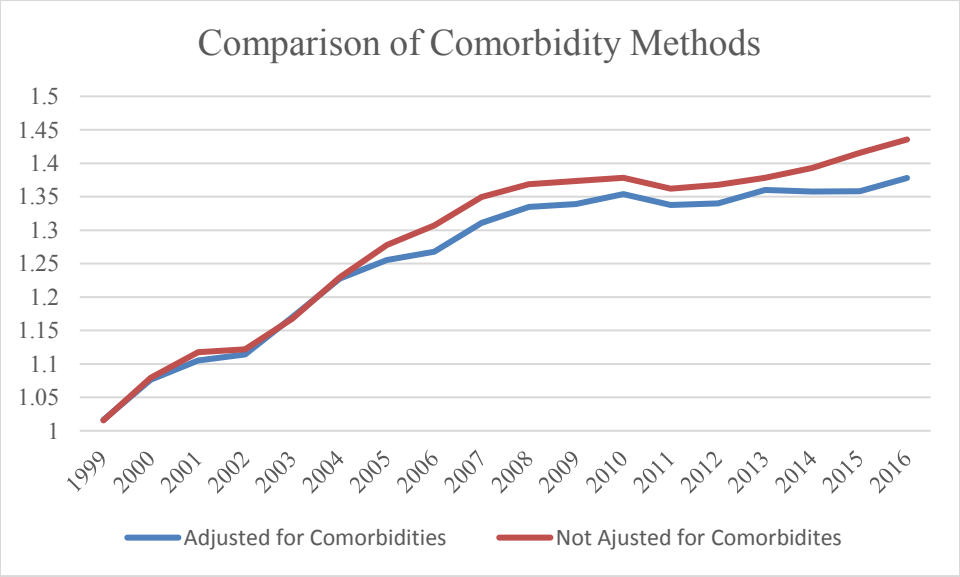
BLSTAC - Future Priorities discussion, November 18th, 2016

Disease Based Price Indexes (OPLC – June 2017)

Federal statistical agencies are being encouraged to publish medical statistics by disease rather than by goods and services. This means that medical expenditures would be reported by disease (i.e. cancer, heart disease, etc.) rather than by service (i.e. physicians, hospitals, etc.). Disease based price indexes are a necessary tool to deflate nominal disease expenditures into real disease expenditures. In 2015, both the Bureau of Economic Analysis (BEA) and the Bureau of Labor Statistics published non-production quality disease based price indexes. BEA's indexes are yearly, while the BLS indexes are timely monthly indexes. BEA indexes are produced as part of BEA work on the Health Satellite Account and are currently reported through 2012. BEA generates price indexes that use a combination of data sources – the Medical Expenditure Panel Survey (MEPS), private claims data, and Medicare Claims data. Both claims data sets need to be purchased. MEPS is freely downloadable. The BLS indexes use a combination of MEPS data along with currently published medical price indexes and incurs no data set costs. Since July 2015, BLS has posted experimental disease based price indexes that have been updated on a monthly basis after that initial release, shortly after the release of CPI indexes each month. Presently, BLS generates a variety of experimental disease based price indexes. Some adjust for comorbidities and some do not. Some combine dental services with diseases of the digestive system and others do not. The chart below compares the differences between adjusting and not adjusting for comorbidities. Over a sixteen year period, there is a five percentage point difference.



Many issues arise when deciding how to construct disease-based price indexes. First, decisions are made about which data set to use. Do we use expensive private insurance claims data that comes from convenience samples or do we use publicly available databases that are free to the public but have much smaller sample sizes? BLS obviously took the second approach with the use of MEPS in combination with PPI and CPI indexes. Second, we must decide how to treat

comorbidities. For example, if an individual visits a physician to treat both her diabetes and asthma, what fraction of the visit gets allocated to diabetes and what fraction gets allocated to asthma? A third issue involves the choice of a disease classification system. The ICD-10 (International Classification of Diseases) system has too many categories and most items would not be populated. However, a classification system such as the Chapter of the ICD-10 are so broad that each category contains a widely heterogeneous group of patients. Finally, there is the quality adjustment issue. Any medical price index needs to be quality adjusted, but at this point in time, we do not have the necessary outcome database to make these quality adjustments. Both BEA and BLS are considering options for approaching quality adjustment.

Questions:

1. Do you think that it is better to report medical statistics by disease than by service? Alternatively, do you think reporting by disease should be treated as another view of medical inflation rather than ultimately as a replacement for current approaches to medical care price indexes?

2. Do you believe that it worth the additional expense of purchasing the larger private insurance claims databases as a source for computing these disease based price indexes, or can we continue to rely on the Medical Expenditure Panel Survey? Should some combination of the BEA and BLS approaches be considered and if so, what would you recommend?

3. Do you have any thoughts/suggestions regarding approaches to quality adjustment for these indexes?

Current Employment Statistics (CES) Establishment Age/Size Class Estimates (OEUS)

For the Current Employment Statistics (CES) survey, experimental establishment size class estimates are expected to be published in 2015. An earlier iteration of CES firm size class data was released that utilized a base-period sizing for monthly estimates and an end-period resizing for generating a benchmarked series. This benchmarking methodology provided information about change to employment levels from one end-period size classification to the next. However, most interest in size class data is in answering where growth is occurring. For this, a consistent base-period sizing methodology appears to be a better choice. As a result, CES has been researching methods to do just that. The methodology chosen has led to choices about the data that we can publish (changes but no levels). Additional choices include the size classes to use, and the industries to publish.

The methodology chosen for developing estimates of employment change by business size may also be applied to develop data based on business age. This work is in an earlier stage of development.

Questions:

1. Does the Committee have any concerns with or suggestions about the methodology chosen to produce firm-size class estimates for CES?

2. Should BLS pursue specific additional research before considering the firm size class benchmarking methodology final?

3. Does the Committee have any concerns with or suggestions about applying the methodology to produce data based on firm-age class for CES?

Occupational Requirements Survey (OCWC – June 2017)

In 2015, the Occupational Requirements Survey implemented production of a nation-wide sample of establishments to support production of occupation specific data for most of the O*NET 8-digit SOC codes. Under the current design, establishments in all industries and sectors of the economy are sampled each year and data from three (or more) years of data collection will be used to produce the ORS estimates. This sample design was carefully tested to ensure that it would meet the goals for this survey but has some potential drawbacks.

The current design does not provide the opportunity to control the number of quotes collected for any specific occupations or even for specific occupational groups resulting in the potential of collecting too much data for occupations in which there is a lot of homogeneity for the ORS data elements and not enough data for occupations where there is wide variation in the ORS data elements. Although the design will provide data for the vast majority of occupations, it is possible that it will not support release of estimates for all the SOC codes that appear frequently in the disability claims processes. In addition, initial research into the frequency with which ORS data elements change over time indicates that ORS values will vary more often for some jobs or types of jobs than for other jobs or types of jobs. This suggests that it could be necessary to update ORS data for some types of jobs more often than for other jobs. But our current sample design only supports updating data for all jobs on a single schedule (i.e. all at once).

So research is being conducted into the potential for using other sample designs that would permit targeting specific occupations with the selection of each sample. Among the options being researched are the potential of selecting samples of occupations that have been collected by the Occupational Employment Statistics (OES) program and/or using prior OES data to adjust sample allocations by industry. Using OES data poses many challenges including acceptance of OES non-response rates and potential occupation coding errors which vary by industry and occupation and may have a negative impact on future OES response rates. This presentation for TAC will include an overview of the challenges with the current design and a summary of the current research efforts to target occupations during sampling.

Questions:

1. What are additional options for targeting occupations within an industry-based sample?

2. Should we move completely away from an industry-based sample and use an occupation-based sample? If so, do you have suggestions aside from recontacting OES respondents?

Incorporating OSHA collected data with the Survey of Occupational Injuries and Illnesses (OCWC - November 2017)

In 2016, the Occupational Safety and Health Administration (OSHA) issued a new rule mandating certain employers to electronically submit injury and illness data directly to OSHA. Employers in non-exempt industries had previously only been required to keep a log of their workplace injuries and illnesses on forms that the employer maintained. These forms are the basis for the data collected by BLS as part of the Survey of Occupational Injuries and Illnesses (SOII). In principle the new OSHA reporting requirements could provide complete administrative records – precisely the forms SOII seeks – for a large part of the SOII scope.

The Office of Management and Budget (OMB) has requested that BLS and OSHA work together in order to minimize any increased burden on employers due to this new rule. In accordance with this request, BLS has begun researching ways of combining the data electronically collected by OSHA with data collected through the SOII. There are a number of challenges in combining these two data sources. First, the data collected by OSHA will contain employer name and address but will likely not contain additional fields such as EIN or NAICS codes to facilitate matching of individual records from OSHA to either the SOII or to the BLS establishment frame. Second, OSHA does not have the resources to ensure high levels of compliance with their recordkeeping initiative. Non-compliance is likely to be nonrandom. Finally, in order to elicit complete information, respondents to BLS surveys are assured confidentiality under CIPSEA. OSHA is a regulatory agency and plans to post the electronic records publicly. The lack of confidentiality may have an impact on the quality of responses.

Matching individual records between OSHA and BLS is likely to be time consuming and error prone, and this may preclude record linkage as a part of regular SOII estimation processes. Therefore, BLS has been focused on methods of utilizing both datasets that do not necessarily rely on this level of match. One approach is outlined in Raghunathan (2015).[1] This method relies on explicitly modeling the likelihood of selection into the "big data" (here, the OSHA electronic records). The advantage of this method is that it makes use of the representative survey data in combination with the power of the number of records from the big data in order to draw statistical inferences.

Questions:

1. What additional methodologies should BLS be considering for combining OSHA electronic records with the SOII?
2. What caveats should be considered with the Raghunathan method?
3. Are there advances in record linkage that should be considered in matching the establishment data from OSHA to that from BLS given the tight time frame to produce estimates?

[1] https://deepblue.lib.umich.edu/bitstream/handle/2027.42/120417/NAS-Paper.pdf?sequence=1&isAllowed=y