Towards a new paradigm for scanner data price indices: applying big data techniques to big data
(*my preferred paper for oral presentation should I need to make a ranking*)

Jens Mehrhoff, European Commission

The advent of scanner data brought new opportunities but also new challenges to statistical offices. One of these challenges is the treatment of product churn, i.e. a dynamic universe due to new and disappearing products. Chaining price indices at monthly frequency, say, can then lead to significant drift. In order to overcome chain drift, multilateral methods have been proposed that are by construction drift-free. These methods are borrowed from the literature on international purchasing power parity comparisons and may not be tailored to the problem in intertemporal comparisons; after all countries and months might deserve different treatments in their measurement.

The present paper proposes a shift towards a new paradigm: dealing with the question at hand in the time domain. To this end, a model-based procedure is derived that yields figures, which do no longer possess the classical formula interpretation. Rather the index numbers follow a structural time series model – the starting point for the development being the time-product dummy approach. The paper continues by estimating price index numbers using the new method and comparing the results with the traditional approaches. The new index series convey a similar information content in terms of the statistical signal but come with much lower noise than the classical concepts.