Tracing the Effects of Real-Time Data Revisions in Exclusions-from-Core Measures of Inflation

by

Heather L.R. Tierney*

Abstract

Through the use of conditional, local nonparametric estimation, this paper tracks the effects that data revisions have on core and total inflation as measured by Personal Consumption Expenditures and the Consumer Price Index. The purpose of this paper is two-fold. The first is to investigate whether data revisions of PCE does have an impact on total inflation in regards to the exclusions-from-core measures of inflation over a five-period in-sample forecast horizon of one, two, four, eight, and twelve quarts. Previous work has found that the effect of data revisions is difficult to determine when a recursive framework is used since both new data as well as revised data is used in each new vintage of real-time data. Using the same techniques, the second purpose is to examine whether the Consumer Price Index is better able to capture the overall trend of inflation, which is one of the uses of core inflation that could also have policy ramifications as well.

KEY WORDS: Real-Time Data, Local Estimation, Nonparametrics, Inflation Persistence, Monetary Policy

JEL Classification Codes: E40, E52, C14

*Contact author: Heather L.R. Tierney, School of Business and Economics, College of Charleston; 5 Liberty Street, Charleston, SC 29424, email: tierneyh@cofc.edu; phone: (843) 953-7070; fax: (843) 953-5697.
Extended Abstract

The issue of using real-time data could be of significance especially for monetary policy purposes. For instance, since 2004, the Federal Reserve has been using core personal consumption expenditures (PCE) to forecast inflation, which is subject to revision, and since 2007, the Fed has been using total PCE and core PCE for forecasting purposes with total PCE also being subject to revision. In terms of forecasting inflation, expecting the Federal Reserve to make forecasts on information that they did not have available at that time is akin to playing Monday-morning quarterback, which is essentially what making forecasts with revised data does.

The purpose of this paper is two-fold. The first is to investigate whether data revisions of PCE does have an impact on total inflation in regards to the exclusions-from-core measures of inflation over a five-period in-sample forecast horizon of one, two, four, eight, and twelve quarts. Previous work found that the effect of data revisions is difficult to determine when a recursive framework is used since both new data as well as revised data is used in a new vintage of real-time data (Tierney 2008). So, although the estimated parameters of a model are changing, it is difficult to determine where the changes are coming from the newly incorporated data or the revised data. If data revisions do affect the estimates, determining a pattern or patterns that could be useful in terms of policy implementation is yet another one of the end-purposes of this paper. Using the same techniques, the second purpose is to examine whether the Consumer Price Index (CPI) is better able to capture the overall trend of inflation, which is one of the uses of core inflation that could also have policy ramifications as well.

In order to study the first purpose of this paper, real-time data is used to examine whether core inflation is an unbiased estimator of total inflation in the regression model of Laflèche and Armour (2006), which is based upon Cogley (2002). Laflèche and Armour’s (2006) model is a parametric ordinary least squares (OLS) model, which this paper uses as a benchmark comparison to the kernel-weighted least squares method (KWLS) method of nonparametrics. The main reason for using nonparametrics is its ability to provide time-varying local regression estimators. Generally, since the data revisions are typically small, an aggregate parametric framework may not be able to capture the effect of data revisions while a local nonparametric approach is able to do so.
Furthermore, according to Croushore (2007), after the initial release date, which is done mid-quarter, data is revised the first two months following the initial release, in July for the following three years, and the benchmark revision, which generally occurs every five years. Even for the problematic benchmark revisions, this tracking of data revision across vintages is perfectly suited for the flexible nonparametric framework to which there is no parametric equivalent. The benchmark revisions include new information from economic censuses as well as a change of base year, which should not have an effect on chain-weighted data, but the benchmark years could be problematic when a change in methodology has occurred (Croushore 2007). Hence, instead of using just one particular vintage or one particular revision for (in-sample) forecasting, all available vintages and all available revisions are able to be examined simultaneously as is suggested by Elliott (2002).

In order to study the effect that the deviations-from-core inflation has on total inflation, the following regression model is used:

\[
(\pi_{t+h} - \pi_t) = \alpha + \beta(\pi^\text{core}_t - \pi_t) + u_t
\]  

(1.1)

where \( \pi_{t+h} \) is the \( h \)-period-ahead total inflation at time \( t \), \( \pi_t \) is total inflation at time \( t \), \( \pi^\text{core}_t \) is core inflation at time \( t \) with \( u_t \sim (\alpha, \sigma^2) \) being the random error term with \( h \) representing the in-sample forecast horizon.

In particular, the estimated slope coefficient from Equation (1.1) is examined for all five in-sample forecast horizons for the regressions involving PCE and CPI, respectively, in a parametric and nonparametric framework. Regarding the data, the measures of core PCE, PCE, and CPI are obtained in real-time from the Philadelphia Fed. The seasonally-adjusted core CPI is obtained from the St. Louis Federal Reserve Economic Data (F.R.E.D) since it is not provided in real-time. Although CPI is not a revised time series, seasonally adjusted CPI does contain some small adjustments due to seasonality, which is why the real time data of seasonally adjusted CPI is used.

The real-time dataset begins with first vintage being V_1996:Q1 and the last vintage being V_2008:Q2.\(^1\) Only 50 vintages are examined since these are the only available

\(^1\) To make it easier to determine when a particular vintage of a real-time dataset as opposed to a given observation is being discussed the notation of “V_” will appear before the vintage of the real-time dataset. For instance, V_1996:Q1 refers to the vintage of the real-time dataset released in the middle of the first quarter of 1996 with the observable data ranging from 1959:Q4 to 1995:Q4.
vintages for both core PCE and PCE. Vintages of CPI go farther back, but in order to keep the real-time data analysis as symmetric as possible in regards to the vintages between the PCE and CPI, the shorter available time span of vintages of the PCE are used. Each of the 50 vintages begins with 1959:Q4 before the calculation of inflation. In the first vintage of V_1996:Q1, that ends with the measure of inflation in 1995:Q4, there are 145 observations before the calculation of inflation, and in V_2008:Q2, that ends with the measure of inflation in 2008:Q1, there are 194 observations before the calculation of inflation.

In estimating the regression model, the methodology is as follows:

(i.) Beginning with V_1996:Q1 to V_2008:Q2, the dataset for each of the fifty vintages starts from 1959:Q4, before the calculation of inflation, and ends in 1995:Q4. The OLS estimates for each of the fifty vintages as well as the average of the local nonparametric estimates are obtained for each of the fifty vintages. In order to trace the local effects, only the conditional, local KWLS nonparametric slope coefficients of the very last observation dataset which in this case is 1995:Q4, will be investigated across vintages. This will enable the conditional, local nonparametric estimates to be compared against both the parametric and average nonparametric estimates for all fifty vintages. This will be done for all five in-sample forecast horizons.

(ii.) Similar to Item (i.), the regression models are again estimated but only for forty-nine vintages from V_1996:Q2 to V_2008:Q2 with the data sample being from 1959:Q4 to 1996:Q1. The local conditional effects on 1996:Q1 are traced across vintages and compared to the parametric and averaged local nonparametric estimates.

(iii.) This iterative method is done for all remaining vintages while holding each of the datasets constant.

In order to examine whether the local nonparametric estimates are statistically significant, a local conditional LR test statistic, which is referred to as the Generalized Method of Moments (GMM) Distance Statistic in Tierney (2008), for only the jth conditional nonparametric estimates based upon the Newey and West (1987) GMM Distance Statistic is used to test whether core inflation is locally unbiased in regards to total inflation. In a recursive framework, Tierney (2008) found the conditional, local nonparametric estimates to be time-varying as well highly nonlinear and statistically significant.
References


