Nowcasting of Advance Estimates of Personal Consumption of Services in the U.S. National Accounts: Individual Versus Forecasting Combination Approach

Authors Baoline Chen and Kyle Hood[1], U.S. Bureau of Economic Analysis

Contact Baoline.chen@bea.gov

Date February 2020

Abstract This paper evaluates two individual nowcasting frameworks, the bridge equation and bridging with factors model, in concert with a set of forecast combination techniques for nowcasting the advance estimates of quarterly personal consumption expenditures (PCE) of services at the detailed component level in the U.S. national accounts, using real time data from 2009:Q3 to 2019:Q4. We show that these individual nowcasting frameworks improve the accuracy of advance estimates of PCE services by reducing revisions in 74 percent of the components when quarterly source data become available. We also show that adding model-averaging techniques to nowcasting further improves the accuracy by reducing revisions in 91 percent of the detailed components. The model-averaging techniques considered in this study include simple averaging (mean, median, trimmed means), information-criterion-based averaging (AIC, BIC, log-likelihood averaging), Bates-Granger averaging with leave-one-out cross-validation errors, and covariance-minimization-based Jackknife and Mallows averaging. We evaluate the performances of all methods by comparing their root mean squared revisions (RMSR) in the advance estimates of each detailed component of PCE services. Our study demonstrates that nowcasting models and model-averaging techniques have the potential to be a powerful tool in reducing revisions in the early estimates in national economic account statistics at the detailed level.

Keywords Nowcasting, forecast combination, model averaging, early estimates of GDP components

JEL Code C52, C53, E17, E01

1. We would like to thank the participants at the 2019 World Congress of the International Statistical Institute, the 2020 Joint Statistical Meetings of the American Statistical Association, and the joint meeting of BEA and Statistics Canada in 2019 for their comments. We would also like to thank Dylan Rassier, Kyle Brown, and Harvey Davis from the National Economic Accounts division at BEA for their helpful comments on the paper.

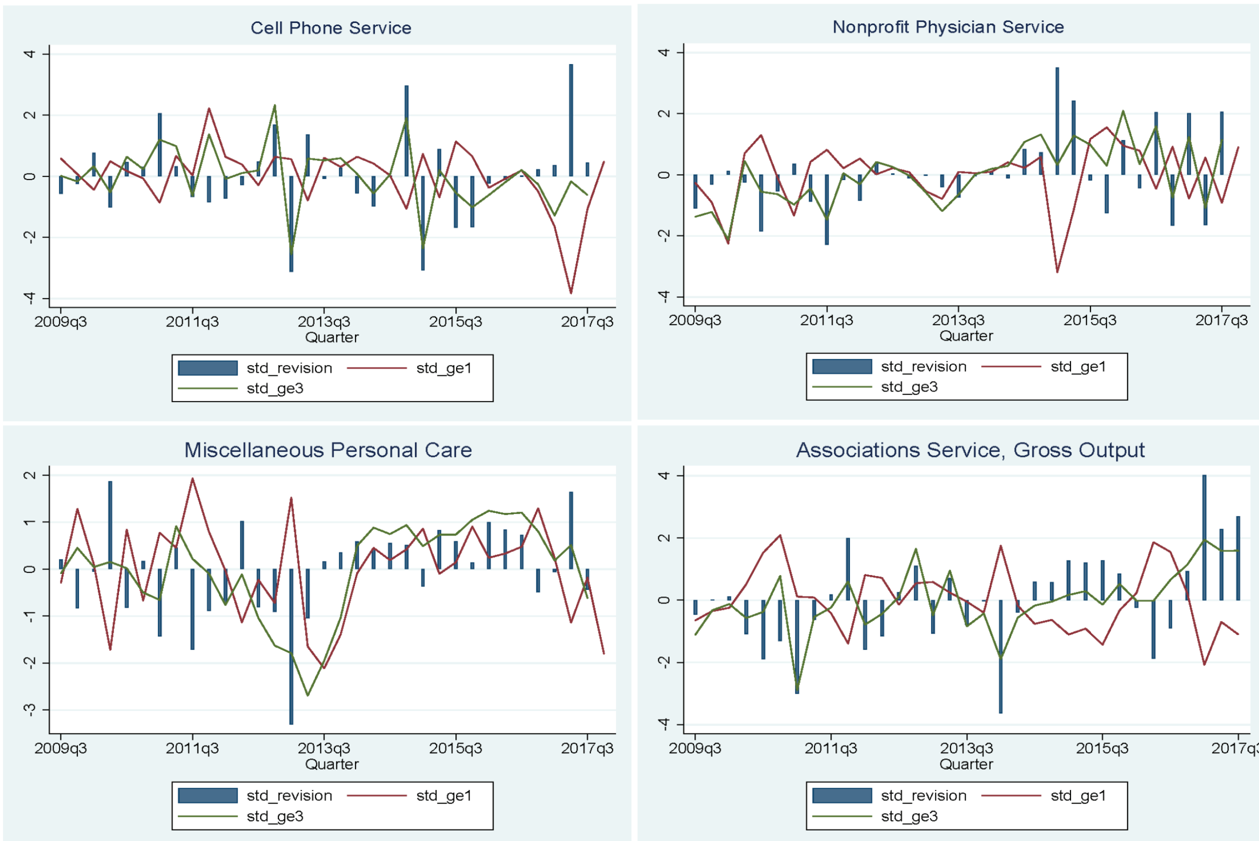
1. Introduction

Gross domestic product (GDP) measures the economic performance of a country and is a vitally important source of information on how macroeconomic conditions are evolving. Thus, timely and accurate estimates of GDP are expected by business cycle analysts and decision makers to monitor economic conditions in order to design economic policies and make business decisions. However, lags in receiving source data needed to compile GDP estimates present practical challenges to the national accounts in producing timely and accurate estimates of GDP.

In this paper, we explore several nowcasting or forecasting methods to overcome these challenges. There are some typical as well as some unique aspects of this application, however, when compared to the typical nowcasting problem. The first aspect of the application we discuss is the release schedule associated with how the United States compiles GDP and its components: Because there are lags in receiving most source data, the Bureau of Economic Analysis compiles GDP several times as new information arrives. We show that the information that is used to compile early releases is not always correlated with the information that is available for later releases, requiring us to look for either new information or for other ways to use the information that we have. The second aspect relates to how GDP and its major components are compiled: In contrast to the usual forecasting problem, national accounts are mostly computed in a "bottom-up" manner, from detailed components which are then summed to produce subaggregates and aggregates. As such, our approach is based on a "bottom-up" method in which we forecast the many detailed components.

We first discuss the GDP release schedule, source data, and revisions. In the United States, GDP is released three times for each quarter with approximately a one-, two- and three-month lag, respectively. The first or "advance" estimate is based on the most limited information, because detailed quarterly source data for most GDP components are not yet available. Some quarterly data become available for the second release; by the time the third release is compiled, more-detailed and less-preliminary quarterly source data have become available. Due to these lags in the quarterly source data, early estimates of GDP may later be revised and revisions for some components can sometimes be significant. Figure 1 shows percentage revisions in four PCEs in standardized data. Revisions are measured as the differences between the advance and the third estimates in growth rates. Lines in red and green are standardized quarterly growth rates of the first and third estimates; blue bars represent the revisions.

Figure 1. Revisions of Selected Detailed Components of Personal Consumption Expenditures Services



Quarterly estimates of GDP are built up from detailed components of its major aggregates— personal consumption expenditures (PCE), government expenditures, private investment, exports, and imports. PCE, which consists of personal consumption of goods and services, constitutes about two-thirds of GDP, and PCE services account for about two-thirds of PCE (or roughly 47 to 48 percent of GDP). Thus, PCE services is the largest major component of GDP. Because of lags in quarterly source data for advance estimates at the most detailed level, PCE services may show large revisions when quarterly source data become available, compared with other components. In this study, we focus on the advance estimates of detailed PCE services, using 2 individual nowcasting models in concert with a set of 10 forecast combination techniques with the goal to reduce the sizes of revisions.

Quarterly source data for the detailed components of PCE services are mostly from the Quarterly Services Survey (QSS) conducted by the U.S. Census Bureau, which are not available for the advance estimates. Moreover, QSS does not yet cover all detailed components of PCE services. For those components that QSS does not cover, quarterly data from other surveys, Federal Reserve banks, and private trade companies are used. A lack of source data for the advance estimate of PCE services means that the national accounts must rely on indicators to compile the advance estimates.

In the sample period in this study, for 45 percent of the detailed components, monthly indicators are percentage changes in wages and salaries (from the Bureau of Labor Statistics) in the relevant industries; for the other 55 percent of the components—those for which information on wages and salaries is not available—percentage changes in population and a component-specific price index or business revenues are used instead as the indicators.[1]

Ideally, these indicators should match as closely as possible the movements in the components being considered. In contrast, these indicators can be weakly or even negatively correlated with the third quarterly estimate of the components that are compiled using QSS estimates or other quarterly source data as indicators. For example, for recreational services, correlations between the indicators and the third quarterly estimates are negative for 5 of the 25 detailed components and are above 0.5 in 12 of the 25 components. For health care services, correlations between the indicators and the third estimates are negative for 9 of the 21 detailed components and are above 0.5 in 3 of the 21 cases. Similar correlation patterns hold for other categories of services.

The absence of better correlation between the monthly indicators and the quarterly target series can be alleviated by addition of other information. However, during the study period, the current extrapolation method uses indicators that match the specific service components only in the current quarter. While analyst judgment is used to inform estimates of consumer spending, additional statistical information that is available at the time of advance estimates is not incorporated. For example, longer-term trends in the quarterly target series and medium-term dynamics in the monthly indicators are not incorporated. In addition, information from short- and medium-term cross-sectional movements in PCE services is also excluded.

Because quarterly source data are unavailable in time for the advance estimate, reducing revisions in the advance estimate hinges on producing them more accurately using all available information for the quarter that has just ended. This amounts to a nowcasting problem, which is defined as the prediction of the present, the very near future, and the very recent past (Giannone, Reichlin, and Small 2008). We follow the basic principle of nowcasting to exploit information published early (and often at higher frequencies than the target variables) in order to obtain more accurate early estimates two months before official estimates based on quarterly source data become available. For the advance estimates of quarterly PCE services, we have access to all monthly indicators for the quarter being estimated, and thus, the "ragged edge" problem, which is sometimes encountered in nowcasting, does not arise in this case.

1. Indicators described here are relevant to the sample period for this study. New indicators, when they become available, could be added or used to replace some existing indicators. Alternative methods, such as several machine learning techniques, and alternative data, have been investigated to predict the economic indicators used for compiling PCE services. See Chen et al. (2019).

Motivated by the performance of the current extrapolation method, we seek to explore statistical techniques to improve advance estimates of PCE services, focusing on the methods that utilize information on the longer-term trend of the quarterly target series, medium- and short-term movements in the quarterly target and the monthly indicator series, and the dynamic cross-sectional behaviors of the indicators. Because, as noted above, PCE services is built up from its detailed components, we focus on the methods that are suited for the "bottom-up" or the "bean-counting" principle of national accounting.

In this paper, we first evaluate two autoregressive distributed-lag-type models—namely the bridge equation framework and the bridging with factors model—for compiling advance estimates of detailed quarterly PCE services. These two methods can exploit all information on the dynamics of the quarterly target variables and the monthly indicators in the nowcasting of advance estimate. They differ not so much in the formulation but in the type of information used in estimation. In the bridge equation framework, quarterly indicators are derived from monthly indicator data. However, instead of the current method that extrapolates quarterly PCE services from the growth rates of the monthly indicators in the current quarter, we allow a longer-term trend and medium-term dynamics by including lags of the quarterly target variable and lags of the monthly indicator(s). The bridging with factors model replaces the indicator(s) constructed for the bridge equation with common factors derived from available indicators for all categories of PCE services. This collection of indicators yields two common factors that appear to accurately capture the general movement of many of the detailed components of PCE services.

Moving beyond these individual nowcasting models, we note that many empirical studies have shown that an easy way to improve forecast accuracy is to use several different methods on the same time series and to average the resulting forecasts (Bates and Granger 1969, Clemen 1989, Timmermann 2006, and Aiolfi et al. 2010). Thus, in addition to the 2 individual nowcasting frameworks mentioned above, we also evaluate a set of 10 forecast combination or model-averaging techniques to produce combination nowcasts from an array of specifications of the two individual nowcasting models to explore further improvements in the accuracy of advance estimates of PCE services.

In the evaluation of these model-averaging techniques, we aim to (1) identify specific methods that can be used to accurately impute detailed individual components of PCE services for advance estimates of GDP, and (2) look for patterns in the revisions implied by the different classes of model-averaging techniques. These patterns should provide useful evidence for the type(s) of model-averaging techniques that are appropriate in similar situations. The reason that nowcast combination or model averaging is chosen to combine this information rather than using the more traditional techniques of augmenting the model with additional variables is that in this application, we have at most 42 quarterly observations. The short span of the time series that we must work with is not compatible with a large number of explanatory variables. Model-averaging techniques are designed in part to combat this issue.

We estimate and average 12 models, 6 specifications of both the bridge equation and the bridging with factors models, which differ by how many lags of the dependent and independent variables are included. The model-averaging techniques that are considered include: 1) simple averages based on equal-weights, median and two trimmed-means averaging; 2) information-criteria-based averaging (AIC, BIC, and unpenalized); 3) Bates-Granger averaging with leave-one-out cross-validation errors; and 4) quadratic-minimization-based Jackknife and Mallows averaging.

We apply the 2 individual nowcasting models and the 10 selected model-averaging algorithms to compile advance estimates of detailed components of PCE services using real-time data from 2009:Q3 to 2019:Q4 from the U.S. national accounts. The application of these nowcasting methods to real-time data is critical for a proper comparison of the performances of the proposed methods with the current extrapolation methods. Using real-time data allows the nowcast to be computed under the same circumstances that would have been in place in the actual nowcast situation.

To address overspecification and estimation error, we divide the full sample into an estimation sample (sometimes called a "training" sample) and a test sample (sometimes called a "validation" sample). For each period in the test sample, all models are estimated with all data up to but not including the period being considered. Model averages are computed on the same sample that is used for estimation, and the one-step-ahead nowcast is computed using the next pseudo-out-of-sample (POOS) observation from the test sample. We compare improvements in accuracy in terms of reduction in the root mean squared revision (RMSR) for each detailed component. Our results show that the bridge equation and bridging with factors models help reduce RMSR in 74 percent of the detailed components in the study. Moreover, adding combined nowcasts from the model-averaging techniques increases the reductions in RMSR to 91 percent of the components, further improving the accuracy of advance estimates. Our study suggests that individual nowcasting models in concert with model-averaging techniques are potentially powerful tools for reducing revisions in the early estimates in national accounts statistics at the detailed level.

The plan for the paper is as follows: Section 2 describes the current extrapolation method used in the U.S. national accounts and the 2 individual nowcasting models as well as the 10 model-averaging techniques being evaluated in the study. Section 3 describes the empirical application and the strategy for estimation and one-step-ahead POOS predictions. Section 4 reports the results from both individual models and the model-averaging techniques. Section 5 discusses further research and concludes the paper.

2. Methodology for Advance Estimates of Detailed PCE Services

2.1 Current method for advance estimates of PCE services in the U.S. national accounts

Currently, the U.S. national economic accounts compile detailed PCE services using a simple linear extrapolation method.[2] Compilation of advance estimates of quarterly PCE services is done in two steps. First, for each of the three months in a quarter, monthly estimates of detailed PCE services are extrapolated from the previous month using monthly indicators; and second, quarterly averages of the three monthly estimates are computed as the first or the advance quarterly estimates. For detailed components of PCE services that have designated monthly indicators, monthly estimates are extrapolated using the designated monthly indicators. For those detailed components that have no designated monthly indicators, a month-to-month extrapolation is computed using percentage changes in population and price indexes associated with specific services components. Second quarterly estimates are compiled similarly, except that some monthly indicators may have been revised and some quarterly source data may have become available.

The third quarterly estimates of detailed PCE services are directly extrapolated using seasonally unadjusted (NSA) or seasonally adjusted (SA) Quarterly Service Survey (QSS) data as available. For those PCE services for which QSS data are not available, monthly indicators for the advance estimates continue to be used to extrapolate the third quarterly estimates. After the QSS-based quarterly estimates are compiled, final monthly estimates for the quarter being considered are interpolated with Denton's (1971) proportional first difference procedure using the third estimate of the quarter being predicted and the previous quarter as temporal constraints. The interpolated estimate of the third month of the quarter is used as the reference value for the monthly extrapolation of the following quarter.

The current extrapolation method for advance estimates of detailed PCE services employs information on the monthly indicators for the quarter being predicted. It does not utilize information on the longer-term trend of the quarterly PCE services, nor does it systematically utilize information on the dynamics of the monthly indicators. To improve the accuracy of advance estimates of PCE services, we explore estimation methods that allow information on the longer-term trend of the detailed PCE services, as well as information on the current and longer-term dynamics of all the relevant monthly indicators to be incorporated in estimation. Given the limitations of the linear

2. Details on the various linear extrapolation procedures for compiling monthly and quarterly estimates of PCE services are described in Chapter 5 of BEA's National Income and Product Accounts (NIPA) Handbook: Concepts and Methods of the U.S. National Income and Product Accounts .

extrapolation procedure for producing accurate advance estimates, we focus on some nowcasting methods with the objective of reducing revisions in advance estimates of PCE services.

A variety of nowcasting methods has been developed. Examples are univariate and multivariate regression methods such as bridge equations, vector autoregressive (VAR) and Bayesian VAR models, static and dynamic factor models, MIDAS regression models, density forecasting models, forecast combination methods, machine learning, and such. For a survey of related techniques, see Foroni and Marcellino (2013) and Bańbura et al. (2013). Since quarterly estimates of GDP are built up from detailed components of its major aggregates, the "bean-counting" principle of national accounting requires that each detailed component be estimated using indicator data relevant to that component, if quarterly source data are not yet available. Thus, for our purpose, we consider 2 individual autoregressive-distributed-lag types of regression models, namely the bridge equation framework and the bridging with factors model, and we also consider a set of 10 forecast combination techniques to generate combination nowcasts from an array of specifications of the two individual models.

2.2 General bridge equation and bridging with factors models for nowcasting PCE services

We select the bridge equation and bridging with factors models because of their ability to utilize all information on the quarterly target variable and the monthly indicators in the estimation procedure. The general bridge equation framework, first developed by Klein et al. (1989), is perhaps the most widely used method for nowcasting or near-term forecasting of macroeconomic variables. It allows early estimates of low-frequency target variables (quarterly PCE services in this study) to be computed using high-frequency indicators (in this case, monthly), which contain statistically timely information. It is sometimes described as a tracking model, for example, tracking quarterly growth in the quarterly target variables by tracing the arrival of new information via the high-frequency monthly indicators in real time.

For this study, we consider the variables in terms of growth rates. Let y_t^A denote the quarterly growth rate of the advance estimate of a detailed component of PCE service in quarter t; and let $\bar{x}_t = (\bar{x}_{1,t}, \ldots, \bar{x}_{s,t})$ be the quarterly growth rates of the quarterly averages of s monthly indicator variables. Monthly indicator variables in other forms can also be considered in the model (Kitchen and Kitchen 2013). For each detailed component of PCE services, the general specification of the bridge equation framework for quarterly and monthly explanatory variables is

$$y_t^A = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{s} \sum_{i=1}^{k} \delta_{j,i} \bar{x}_{j,t-i} + \varepsilon_t \tag{1}$$

where β_0 is a constant, p is the number of autoregressive parameters, s is the number of monthly indicator variables, k is the number of lags for the indicator variables, and $\varepsilon_t \sim iid\ N(\mu_\varepsilon, \sigma_\varepsilon^2)$. The general bridge equation framework in equation (1) allows multiple monthly indicators with multiple lags.

Although the bridge equation framework allows for multiple high-frequency indicator variables, the number of high-frequency variables that could practically be included in the bridge equation is limited by the length of the time series sample. For PCE services, the use of QSS as quarterly source data for the third quarterly estimate dates back to 2009:Q3. Thus, the maximum number of quarterly observations available for our study, from 2009:Q3 to 2019:Q4, is 42. For each detailed component of PCE services, there are currently one or two designated monthly indicators. Thus, the bridge equation framework is adequate to incorporate lagged quarterly growth rates on the detailed service components and the lagged quarterly growth rates of the designated monthly indicators.

On the other hand, to use information from a much larger set of monthly indicator variables while at the same time limiting the number of explanatory variables, we need a framework that can reduce a large set of monthly indicators to a much smaller number of explanatory variables in the regression equation. In recent years, multiple studies have combined a static or a dynamic factor model with the bridge equation framework for nowcasting GDP (Higgins 2014, Piette 2016). The combined model is referred to as bridging with factors model, in which monthly indicators in the bridge equations are replaced with a small number of common factors to capture the main co-movement of a much larger set of indicators.

Let $F_t = (f_{1,t}, f_{2,t}, \ldots, f_{r,t})'$ be the vector of r common factors from monthly factor models aggregated to the quarterly frequency, where $r \ll \min(n, T)$ with n and T being the number of indicator variables and the number of quarterly observations in the sample. For advance estimate of a PCE services component, the bridging with factor model can be expressed as

(2)
$$y_t^A = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{r} \sum_{i=0}^{k} \gamma_{j,i} f_{j,t-i} + \eta_t$$

where β_0 is a constant, p is the number of autoregressive parameters, k is the number of lags of factor j, and $\eta_t \sim iid\ N(\mu_\eta, \sigma_\eta^2)$.

The vector of r common factors is related to the monthly indicators via

$$x_m = \Lambda \tilde{F}_m + \varsigma_m$$

where $x_m = (x_{1,m}, x_{2,m}, \ldots, x_{n,m})'$ is a vector of n standardized monthly indicator series; and $\Lambda \tilde{F}_m$ are common components, with Λ being the $n \times r$ factor loading matrix and $\tilde{F}_m = (\tilde{f}_{1,m}, \tilde{f}_{2,m}, \ldots, \tilde{f}_{r,m})'$

being the vector of r common factors, $r \ll \min(n, T_M)$ and $\mathrm{T_M}$ is the number of monthly observations in the monthly factor model. The principal component estimator of Λ and $\tilde{F}_m$ is obtained from minimizing the sum of squared residuals (SSR) of the regression equation $x_m = \Lambda\tilde{F}_m + \varsigma_m$, where ς_m is a vector of idiosyncratic components with $\mathrm{E}(\varsigma_{i,m,}\ \varsigma_{j,m-s}) = 0$ for all s if $i \neq j$. The number of factors selected is determined by the criteria in Bai and Ng (2002).

The practical advantage of the bridging with factors model is that it extracts common factors from all available monthly indicators for the detailed components from all categories of PCE services and uses them in estimation. Because we have an insufficient number of designated monthly indicators for all detailed PCE services, the bridging with factors model allows us to investigate whether by incorporating information on the general business conditions of the service sector in estimation via common factors could help further improve accuracy. The "best" autoregressive models specified in equations (1) and (2), in terms of the optimal number of lags of the target variables and the indicators, could be determined based on an information criterion selected for estimation.

2.3 Forecast combination methods

Studies have shown that forecast combinations have frequently been found in empirical studies to produce better forecasts than methods based on the "best" individual forecasting models. Moreover, simple combinations often outperform more refined combination algorithms that are aimed at producing "better" combination weights based on model fit (Timmermann 2006, Aiolfi et al. 2010).

In this study, we consider forecast combination or model-averaging methods based on 4 categories of weighting schemes to generate combination nowcasts from a set of 12 candidate models. The four categories of weighting schemes include: 1) simple averaging (equal weight (EW), median (MED), trimmed-mean averaging (TRM) (10 percent, 25 percent)); 2) information-criterion-based averaging (AIC, BIC, and Log-likelihood function); 3) Bates-Granger averaging with leave-one-out cross-validation errors (BG-LOOCV); and 4) quadratic-minimization-based Jackknife averaging (JMA) and Mallows averaging (MA).[3]

3. Each of these model-averaging techniques is discussed by Diks and Vrugt (2017). Our Bates-Granger technique differs slightly in that we are using the LOOCV errors, rather than in-sample residuals.

Model specifications

We select 10 model specifications to be supplied to each of the model-averaging algorithms. They are further grouped into five specifications of the bridge equation framework in equation (1) and five specifications of the bridging with factors model in equation (2). Within each grouping, we consider model specifications with lags from 0 to 4. Models with zero lag means only the contemporaneous indicator or factors are included in the regressions; and models with up to four lags means that the contemporaneous indicator or factors, lags of the indicator or factors, and the lags of the quarterly target variable up to and including lag four are included in the regression. In this study, there is one designated indicator for each service component in the bridge equation model and two common factors in the bridging with factor model.

The bridge equation model with one designated indicator is reduced to

$$y_t^A = \beta_0 + \sum_{i=1}^{4} \beta_i y_{t-i} + \sum_{i=1}^{4} \delta_i \bar{x}_{t-i} + \varepsilon_t.$$

The five specifications are: 1) $\beta_0 \neq 0$, $\beta_i = 0$, $i = 1, \dots, 4$, $\delta_0 \neq 0$, $\delta_i = 0$, $i = 1, \dots, 4$; 2) $\beta_0 \neq 0$, $\beta_1 \neq 0$, $\beta_i = 0$, $i = 2, \dots, 4$, $\delta_i \neq 0$, $i = 0, 1$, $\delta_i = 0$, $i = 2, \dots, 4$; ...; 5) $\beta_0 \neq 0$, $\beta_i \neq 0$, $i = 1, \dots, 4$, $\delta_i \neq 0$, $i = 0, 1, \dots, 4$.

The bridging with factor model with two common factors is written as,

$$y_t^A = \beta_0 + \sum_{i=1}^{4} \beta_i y_{t-i} + \sum_{j=1}^{2} \sum_{i=0}^{4} \gamma_{j,i} f_{j,t-i} + \eta_t.$$

The five specifications are: 1) $\beta_0 \neq 0$, $\beta_i = 0$, $i = 1, \dots, 4$, $\gamma_{j,i} \neq 0$, $\gamma_{j,i} = 0$, $i = 1, \dots, 4$, $j = 1, 2$; 2) $\beta_0 \neq 0$, $\beta_1 \neq 0$, $\beta_i = 0$, $i = 2, \dots, 4$, $\gamma_{j,i} \neq 0$, $i = 0, 1$, $\gamma_{j,i} = 0$, $i = 2, \dots, 4$, $j = 1, 2$; ...; 5) $\beta_0 \neq 0$ $\beta_i \neq 0$, $i = 1, \dots, 4$, $\gamma_{j,i} \neq 0$, $i = 0, 1, \dots, 4$; *and* $j = 1, 2$.

In addition, we also consider a model specification based on model selection from the most general model specification in each grouping and use them as the individual nowcasting models. For the general bridge equation framework, we allow a maximum of four lagged growth rates of the quarterly target variables, and the current and a maximum of four lagged quarterly growth rates of the monthly indicator variables. The maximum number of the lags of the quarterly target variables and the monthly indicators are restricted by the length of the time series samples available for the study. For bridging with factors model, we allow a maximum of four lagged quarterly growth rates of the target variables and the current and four lagged quarterly growth rates of the selected common factors aggregated from the monthly factors. The "best" bridge equation and the "best" bridging with factors models are selected based on a selected information criterion. This provides two additional model specifications, for a total of 12.

We focus on averaging the nowcasts associated with the 12 specified models. However, because the models are linear, model-averaging here is equivalent to averaging the parameters. Table 1 shows the number of parameters by specification of the 12 models.

Table 1. Number of Parameters in Each Specification of the Models

Lag	Bridge equation framework	Bridging with factors framework
(s)	(x_t, x_{t-s}, y_{t-s})	$(f_{1,t}, f_{2,t}, f_{1,t-s}, f_{2,t-s}, y_{t-s})$
0	1	r*
1	3	1+2r
2	5	2+3r
3	7	3+4r
4	9	4+5r
"Best" model (AICC)	≤9	≤ 4+5r

*r= 2 is the number of common factors and s = 0, 1, ..., 4.

Model-averaging techniques

Most of the techniques for combination forecasts are distinguished by the weighting schemes used for the combination. Let y_t denote the variable of interest; let $\hat{y}_{t+h|t}$ denote combined h-step-ahead forecast of y_{t+h} conditional on information up to t, using a set of M forecasts, $\hat{y}_{i,t+h|t}$, $i = 1, ..., M$. Let $w_{i,t}$ denote the weight for period t placed on the i^{th} forecast, $\hat{y}_{i,t+h|t}$, in the combination forecast.

Simple averaging: In the category of simple averaging schemes, we consider two models that are averaged using either equally weighted means or medians of the models, and two trimmed mean or ordered forecasts averaging schemes. Simple averaging is typically considered optimal when weights would be estimated on short samples (Smith and Wallis 2009). Using the equal weights averaging scheme, the combined nowcast is

$$\hat{y}^{EW}_{t+h|t} = \frac{1}{M} \sum_{i=1}^{M} \hat{y}_{i,t+h|t}. \tag{3}$$

Median averaging uses the median of the M forecasts at period t to compute the h-step-ahead combined forecast

$$\hat{y}^{MED}_{t+h|t} = median\{\hat{y}_{i,t+h|t}\}_{i=1}^{M}. \tag{4}$$

Trimmed mean averaging removes a certain designated percentage of the largest and smallest values before calculating the mean. Here we consider trimming the top and bottom $\lambda\%$ = {10%, 25%} of the nowcasts,

(5) $$\hat{y}^{TRM}_{t+h|t} = \frac{1}{M(1-2\lambda)} \sum_{i=\lambda M+1}^{(1-\lambda)M} \hat{y}_{i,t+h|t}.$$

Information-criterion-based averaging: The Bayesian information criterion (BIC) (Schwarz 1978) and Akaike information criterion (AIC) (Akaike 1974) are two of the most popular information criteria used for model selection. Assuming again there are *M* models, one of which is the true model, AIC or BIC is calculated for each candidate model and the model with the smallest AIC or BIC is chosen to be the best model. BIC is consistent in the sense that if the true model is among the candidates, the probability of selecting the true model approaches 1. On the other hand, AIC is minimax-rate optimal for both parametric and nonparametric cases for estimating the regression function. When comparing the BIC and AIC, the penalty for additional parameters is larger in BIC than in AIC. In recent years, BIC has also been more frequently used as a model selection criterion (Hansen 2010). In information-criterion-based averaging, weights are proportional to the exponential of the negative of an information criterion (Burnham and Anderson 2002). In general,

(6) $$w^{I}_{i,t} = \frac{\exp(^{-I_{i,t}}/_{2})}{\sum_{j=1}^{M} \exp(^{-I_{i,t}}/_{2})}$$

where I = {AIC, BIC} and weights are computed from recursive regression using information up to period t. Weights based on BIC are called Bayesian weights (Hansen 2010).

Information criteria based on the likelihood function are also used in model selection and model averaging. In some studies, model averaging using weighted model log-likelihoods function values is considered to incorporate model uncertainty into the profile of heterogeneity across a set of different models in terms of the endogenous and exogenous variables included. Since log-likelihood can be used to assess the uncertainty associated with the specification of the empirical model (Moral-Benito 2011, Nguefack-Tsague 2014), we also include the "non-penalized" information-criterion averaging or log-Likelihood (LL) averaging based on weighted log-likelihoods, with weights being defined to be proportional to the sum of the log likelihood function values from all models,

(7) $$w^{LL}_{i,t} = \frac{\mathcal{L}(\theta_{i,t}|x)}{\sum_{j=1}^{M} \mathcal{L}(\theta_{j,t}|x)}$$

where θ denote the vector of parameters and x denotes the observed explanatory variables.

Bates-Granger-averaging: The original Bates-Granger averaging (Bates and Granger 1969) model weights each of the M models by $1/\sigma_i^2$, where σ_i^2 is the forecast variance. In practice, the variance needs to be estimated. The Bates-Granger weights are computed according to

(8) $$w_{i,t} = \frac{1/\hat{\sigma}_i^2}{\sum_{j=1}^{M} 1/\hat{\sigma}_j^2}.$$

Our Bates-Granger technique differs slightly in that we use the leave-one-out cross-validation (LOOCV) errors rather than in-sample residuals (Hansen 2010, Hansen and Racine 2012) to compute the weights for model averaging or nowcast combination.

Quadratic-minimization-based averaging: We consider Jackknife averaging (JMA) and Mallows averaging (MA) in this category (Hansen and Racine 2007, 2012; Hansen 2010). Both methods derive weights by minimizing a quadratic criterion function. JMA selects the weights by minimizing a cross-validation (CV) criterion. In models that are linear in the parameters, the cross-validation criterion is a simple quadratic function of the weights. The JAM-selected weights vector minimizes the quadratic function $\varepsilon_T(w)$ over $w \in H_T$, the unit simplex,

(9) $$\hat{w} = argmin_{w \in H_T} \varepsilon_T(w),$$

where $\varepsilon_T(w) = \frac{1}{n} ||\tilde{e}(w)||^2 = w'S_T w$, $S_T = \frac{1}{n} \tilde{e}'\tilde{e}$ and $\tilde{e} = (\tilde{e}^1, \ldots, \tilde{e}^n)'$ contains the vectors of Jackknife residuals. The i^{th} vector, $\tilde{e}^i = (\tilde{e}_1^i, \ldots, \tilde{e}_T^i)$, contains the residuals of the i^{th} model where each $\tilde{e}_t^i$ is the residual of the estimation with the t^{th} observation deleted. Although $\varepsilon_T(w)$ is a quadratic function of w, the solution to it is not typically available in closed form due to the inequality constraints on w. The minimization problem falls in the class of quadratic programming problems, for which numerical solutions have been thoroughly studied and algorithms are widely available (Hansen and Racine 2007). In this study, we solve this quadratic function using the *qprog* function in MATLAB.

The Mallows-selected weight vector minimizes the criterion $C(w)$ over $w \in H_n$, the unit simplex,

(10) $$\hat{w} = argmin_{w \in H_n} C(w),$$

where the Mallows criterion $C(w) = w'\hat{e}'\hat{e}w + 2\sigma^2 w'K$, and $K = (k_1, \ldots, k_M)'$ contains M vectors, each of which contains the number of coefficients in the i^{th} model. This is again a quadratic programing problem with inequality constraints (Hansen 2010). In this study, we solve this quadratic programming model also using the *qprog* function in MATLAB.

3. Estimation and nowcasting of detailed components of PCE services

We apply the bridge equation framework and the bridging with factors model specified in equations (1) and (2) to compile advance estimates of detailed components of PCE services in the U.S. national accounts. We also apply the 10 model-averaging techniques discussed above to compile combination nowcasts over the 12 models specified above. Detailed components from nine PCE service categories are included in the application. (See the appendix for the nine PCE service categories.) Of the 121 detailed components in the 9 service categories, 33 use identical indicator data for the first and the third estimates. Thus, for these components, revisions are zero or close to zero. Any minor revisions in the advance estimates are attributable to the revisions in the source data for the indicators. The remaining 88 components use distinct source data to compile the first and the third estimates. Our evaluation of improvements in accuracy is, therefore, based on percentage reductions in RMSR for these 88 components.

To evaluate improvements in accuracy in terms of reductions in revisions relative to the current extrapolation method, we use the real-time data that were used to compile the official advance estimates of detailed PCE services from 2009:Q3 to 2019:Q4. Data used in the application include the quarterly growth rates of the first and the third estimates of the detailed PCE services from the 2009:Q3 to 2019:Q4 vintages, the quarterly growth rates of the QSS series from the vintages of the same sample period, and the monthly indicators series from the 2009M7 to 2019M12 vintages. Monthly indicators, mostly from the Bureau of Labor Statistics, include monthly percentage changes in population, average wage earnings for relevant service industries, and Consumer and Producer Price Indices (CPI and PPI) of the corresponding services. Each vintage of the quarterly data includes the current and four lagged quarterly growth rates of the target variables, that is, the third quarterly estimate of the PCE services, and each vintage of the monthly indicators includes current and the lagged values sufficient to compute four lagged quarterly growth rates of the monthly indicators.

To estimate the bridging with factors models, a monthly factor model is first estimated using a total of all 92 monthly indicators designated for the components in the 9 categories of PCE services. Two common factors are extracted according to the criterion from Bai and Ng (2002) for all specifications of the bridging with factors models.

To provide an unbiased picture of the improvements in revision anticipated for the 2 individual nowcasting models and the 10 proposed model-averaging schemes, the full sample is split into an estimation sample and a test sample. The estimation sample starts with 75 percent of the total quarterly observations and the test sample contains 25 percent of the total observations at the end of the full sample. The nowcasting exercise proceeds in three steps: 1) in-sample estimation using the estimation sample; 2) recursive one-step-head POOS predictions computed from the estimated bridge equation and bridging with factors models, and from the weighted averages of the 12 estimated models; and 3) performance evaluation of all methods using the test sample for validation. We measure improvements in accuracy brought by all proposed methods by comparing root mean squared prediction errors (RMSE) relative to the actual third estimate, or RMSR, with those from the current extrapolation method.

To fully utilize the information from our short time spans, we choose recursive estimation based on an expanding data window to compute the POOS predictions, as this approach makes most efficient use of the data. This means that each one-step-ahead prediction is computed based on the estimated model or the weighted average of the estimated models in which estimation was performed using all observations prior to the quarter being nowcasted.

The programs for in-sample estimation and out-of-sample predictions are written in Stata. For the bridge equation and bridging with factors models, in-sample estimation is conducted using the VSELECT (Lindsey 2014) command in Stata, and the "best" models are selected using Akaike Information Criterion corrected for small samples (AICC).

4. Results from estimation and nowcasting

This section presents the estimation results from the bridge equation and bridging with factors models using data with and without the outliers being removed, and it compares the improvements in accuracy brought by the two individual models and the model-averaging techniques. The improvements in accuracy is measured by the RMSR from the one-step-ahead POOS predictions based on each method relative to the RMSR computed from the current extrapolation method.

4.1 Results from bridge equation and bridging with factors models

Of all the component series in our study, about one-third have outliers in the estimation sample. Outliers are defined as the quarterly growth rates, in absolute values, greater than or equal to 3 standard deviations of the series. To examine the impact of the outliers on estimation results, we also estimate the models with the outliers removed. Thus, we evaluate four in-sample estimations in our application: bridge equation regression (M1), bridge equation regression with outliers removed (M2), bridging with factors model (M3), and bridging with factors model with outliers removed (M4).

Results from the in-sample estimation exercise validate the choices of using bridge equation and bridging with factors models to compile advance estimates of detailed PCE services. Table 2.1 compares the results from in-sample estimation of selected components of PCE services from the bridge equations using data with and without outliers.

Table 2.1. In-Sample Estimation of General Bridge Equations With or Without Outliers—Selected Components of PCE Services

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Model	**PCE category**	**HLC code**	β_0	β_1	β_2	β_3	β_4	α_0	α_1	α_2	α_3	α_4	**Adj R^2**	**Prob > F**
M1	HLC	PHH	.0064 (3.69)				.5225 (3.12)		–.4982 (–2.30)				0.27	0.01
M2	HLC	PHH	.0054 (2.23)			.3027 (1.81)	.3054 (2.04)		–.4779 (–1.52)				0.28	0.01
M1	TRS	MVR	.0078 (1.92)	.7236 (4.04)	–.6900 (–4.17)	.4204 (2.22)				–.1802 (–2.12)		–.1861 (–2.32)	0.53	0.00
M2	TRS	MVR	.0068 (2.44)				.3752 (3.06)	.2136 (3.23)	.1537 (2.05)				0.36	0.00
M1	RCA	SPE	–.0015 (–.13)	–.3140 (–1.97)	–.7338 (–4.38)					1.4271 (2.34)	1.3627 (1.91)		0.40	0.00
M2	RCA	SPE	.0154 (1.05)		–.4807 (–2.12)		.5205 (2.44)				1.2240 (1.67)	–1.8631 (–1.99)	0.40	0.00
M1	COM	ODS	.0015 (4.31)	.5170 (2.88)		–.2553 (–2.78)	–.2939 (–2.83)				.2148 (3.34)	–.1950 (–2.84)	0.72	0.00
M2	COM	ODS	.0064 (3.46)	.2411 (2.03)					.2240 (3.63)	–.0847 (–1.73)	.2803 (5.18)	–.1446 (–2.78)	0.50	0.00
M1	PRS	AXS	.0160 (3.19)				–.4771 (–2.52)		–.4966 (–1.90)				0.23	0.01
M2	PRS	AXS	.0065 (1.11)				–.5611 (–2.91)	.6316 (1.65)					0.21	0.01
M1	PER	BBB	–.0017 (–.52)		.5400 (3.90)						1.4683 (3.05)		0.41	0.00
M2	PER	BBB	.0026 (.60)	.3729 (2.25)	.3395 (2.05)				–1.0837 (–1.75)		1.4541 (2.39)		0.43	0.00
M1	SOC	CIO	0.0092 (3.28)				.3018 (1.78)						0.07	0.09
M2	SOC	CIO	.0098 (4.75)					.2661 (2.38)				.2237 (2.18)	0.30	0.00

Note: HLC-PHH: Health care-physician services; TRS-MVR: Transportation–motor vehicle rental; RCA-SPE: Recreational services-spectator sport; COM-ODS: Communication services–other delivery; PRS-AXS: Professional services–professional association dues; PER-BBB: Personal care-hair salons and personal grooming; SOC-CIO: Social services–civil and organization services.

Column 1 in table 2.1 identifies the model used for in-sample estimation (M1, M2). Columns 2 and 3 identify the selected components (PCE_Code) and the corresponding categories of PCE services (PCE_Category) they belong to. Columns 4 to 8 display the estimated coefficients of the constant and the lagged quarterly growth rates of the target variables. Columns 9 to 13 show the estimated coefficients of the current and lagged quarterly growth rates of the monthly indicators. Numbers in parentheses are t-values for the estimated coefficients. Columns 14 and 15 display adjusted R^2 values and p-values for the F-tests.

Similarly, table 2.2 shows the in-sample estimation results for selected PCE services detailed components from the bridging with factors models. Column 1 identifies the model used for the in-sample estimation (M3, M4). Columns 2 and 3 identify the selected PCE services components and the services categories they belong to. Columns 4 to 8 display the estimated coefficients of the constant and the lagged quarterly growth rates of the selected PCE services. Columns 9 to 18 display the estimated coefficients for the current and lagged quarterly growth rates of the two common factors. Numbers in parenthesis are t-values for the estimated coefficients, and adjusted R^2 values and p-values for the F-tests are shown in Column 19 and 20.

Table 2.2. In-Sample Estimation of Bridging with Factors Model With or Without Outliers—Selected Components of PCE Services

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Model	**PCE category**	**PCE code**	β_0	β_1	β_2	β_3	β_4	γ_{10}	γ_{11}	γ_{12}	γ_{13}	γ_{14}	γ_{20}	γ_{21}	γ_{22}	γ_{23}	γ_{24}	**Adj R^2**	**Prob > F**
M3	HLC	OPO	0.0076 (4.42)	.3580 (2.98)									.0406 (3.62)			.0383 (3.53)		0.75	0.00
M4	HLC	OPO	.0127 (8.28)		.4260 (4.39)	−.3967 (−3.77)				.0199 (2.98)			.0330 (3.96)			.0324 (3.83)		0.80	0.00
M3	TRS	ORT	0.0078 (2.54)	.4213 (2.62)														0.16	0.01
M4	TRS	ORT	.0077 (2.30)	.5839 (4.11)		.5533 (3.03)	−.7423 (−3.55)	.0548 (2.79)		.0705 (3.46)						−.0619 (−3.28)		0.44	0.00
M3	RCA	PAR	0.0045 (.98)				.6688 (4.10)		.0885 (2.39)	.0938 (2.35)		.0911 (2.13)						0.46	0.00
M4	RCA	PAR	.0033 (.77)		−.2652 (−1.67)		.4191 (2.88)			−.0842 (−1.87)			−.1212 (−2.26)					0.26	0.00
M3	COM	LDT	.0001 (.03)			.2539 (1.59)							−.0878 (−3.15)					0.21	0.01
M4	COM	LDT	−.0010 (−.48)									−.0291 (−1.60)	−.0398 (01.76)					0.14	0.05
M3	PRS	AXO	.0118 (5.53)					.0453 (2.51)		.0662 (3.44)								0.27	0.00
M4	PRS	AXO	.0133 (7.83)					.0265 (1.83)		.0497 (2.22)	.0236 (1.66)							0.23	0.02
M3	PER	FRE	0.0046 (1.29)					0.0706 (2.52)	0.0459 (1.58)									0.17	0.03
M4	PER	FRE	0.0025 (.80)									.0486 (1.78)						0.07	0.09
M3	SOC	FAH	.0147 (2.42)	1.0211 (5.23)			−.6058 (−2.88)											0.49	0.00
M4	SOC	FAH	.0105 (3.16)	.7375 (6.89)			−.2661 (−2.31)	−.0340 (−2.17)		−.0637 (−3.82)								0.68	0.00

Note: HLC-OPO: Health care-nonprofit all other health services; TRS-ORT: Transportation-other road services; RCA-PAR: pari-mutuel net receipts; COM-LDT: Communication-land line telephone services; PRS-AXO: Professional services-professional association services; PER-FRE: Personal Services-furniture repair and floor servicing; SOC-FAH: Social services-proprietary and public individual and family services.

A few general observations that stand out from the in-sample estimation results are: 1) dynamic characteristics of each component of PCE services and of its indicator (or common factors in the bridging with factors model) determine the "best" model for its advance estimate; 2) lagged quarterly growth of PCE services and lagged monthly indicators (or lagged common factors) provide useful information in the estimation of advance estimates for most components of PCE services; 3) common factors extracted from all available monthly indicators for the nine PCE services categories are shown to provide useful information in compiling advance estimates using the bridging with factors model; and 4) removing outliers affects the selection of the best models.

Lagged growth rates of PCE services are shown to be significant in the estimated models for most of the detailed components. However, although the maximum number of lagged growth rates is set to four for both target variables and the indicators or common factors, the Stata VSELECT command selected fewer than four lags of the target variables and fewer than four lags of the indicators or common factors for most components, based on AICC.

The main observations from the one-step-ahead POOS predictions are that: 1) one-step-ahead POOS predictions from both bridge equation and bridging with factors models produce reductions in RMSR in 64 (or 74 percent) of the 88 components of PCE services included in the application; 2) the bridge equation framework outperforms the bridging with factors model by producing larger reduction in RMSR in 42 of the 64 components (that is, 66 percent); and 3) degrees of reduction in RMSR brought by either of the models vary over time and across services categories.

To further illustrate what these results mean, table 3 shows the percentage changes in the RMSR from the one-step-ahead POOS predictions for the HLC services. The percentage changes are relative to the RMSR from the current extrapolation method. The first column contains the 20 components of the HLC services. Columns 2 and 3 show percentage changes in the RMSR from the one-step-ahead POOS predictions computed from the estimated bridge equation model with and without outliers (M1, M2). Columns 4 and 5 show the percentage changes in the RMSR computed from the estimated bridging with factors models with and without outliers (M3, M4).

Numbers with negative values indicate the reductions in the RMSR from the one-step-ahead predictions. Reductions in RMSR are seen in 18 out of the 20 detailed HLC services components. Degrees of reduction in RMSR vary across the four estimated models. The largest reduction in RMSR for each service component among the four models is indicated with a blue shade. Of the 18 components that show reductions in the RMSR, the 10 largest reductions are from the estimated bridge equation models. The better performance of the POOS predictions from the estimated bridge equation models is also observed in other categories of PCE services.

Table 3. Reduction in Root Mean Squared Revision from One-Step-Ahead POOS Predictions of the Estimated Bridge Equation and Bridging with Factors Models with or without Outliers for Health Care Services

	1	2	3	4	5
No	Health care services (HLC)	Δ% RMSR BE (M1)	Δ% RMSR BE_No OL (M2)	Δ% RMSR BF (M3)	Δ% RMSR BF_No OL (M4)
1	Proprietary specialty outpatient (SOH)	−10.51	−11.92	−13.11	−19.86
2	Non-profit specialty outpatient (SPS)	−8.10	−8.10	−4.77	−19.67
3	Nonprofit specialty out-patient, gross output (SPO)	−27.84	−27.84	−11.08	−10.56
4	Proprietary home health (HHH)	−27.13	−38.65	1.13	−38.75
5	Non-profit home health care (HIX)	−11.19	−8.11	−3.74	−6.92
6	Nonprofit home health care, gross output (HOX)	−24.43	−24.43	−29.62	−30.43
7	Non-profit hospital services (NPH)	14.99	14.99	10.86	9.37
8	Proprietary hospital services (FPH)	8.64	8.64	22.37	22.37
9	Non-profit hospital, gross output (HSO)	−5.16	5.07	1.47	1.25
10	Non-profit nursing homes, gross output (NXO)	1.10	1.10	−9.76	−4.20
11	Non-profit nursing home services (NPN)	−10.74	−10.74	−8.93	−12.00
12	Proprietary nursing home services (FPN)	−30.63	−30.63	2.56	16.21
13	Proprietary physician services (PHH)	−32.33	−41.61	−28.20	−28.20
14	Non-profit physician services (PSX)	−28.81	−33.92	−35.25	−40.17
15	Dental services (DEN)	5.35	3.25	−3.50	−3.50
16	Medical laboratories (MLB)	−24.35	−27.22	−22.56	−18.92
17	Proprietary all other med. services (OHH)	−11.41	−15.56	1.01	−8.40
18	Non-profit all other med. services (OPS)	−11.43	−10.78	−24.39	−21.24
19	Nonprofit physicians, gross output (POX)	7.52	−13.85	1.47	−7.07
20	Nonprofit other med. services, gross output (OPO)	−18.84	−4.88	−5.14	−3.66

/ — The blue-shaded areas indicate the largest reduction in RMSR for each service component among the four models.

/ — The grey-shaded areas reflect no reduction in RMSR for the components from any models.

POOS pseudo-out-of-sample
RMSR root mean squared revisions

The results from the POOS predictions allow us to compare the performance of the current extrapolation method against the performances of both models. Table 4 summarizes the percentage changes in the RMSR from the "best" bridge equation and the "best" bridging with factors models (M1 and M3). Since the RMSRs from the current extrapolation method are computed without removing the outliers from the data, we report summary results for the two models without removing the outliers from the data for comparability. Columns 1 and 2 show the services categories and the number of components in each category. Column 3 shows the number of components in each category with no reduction in RMSR from the two models. Columns 4 and 5 show the largest reductions in RMSR in each services category from each individual model; Column 6 indicates the range of reductions in RMSR from the "best" models in each category. In sum, the two individual models produce reductions in RMSR in 65 (or 74 percent), of the 88 components, ranging from 0.5 percent to 53 percent.

Table 4. Summary of Reduction in RMSR from Bridge Equation or Bridging with Factor Models

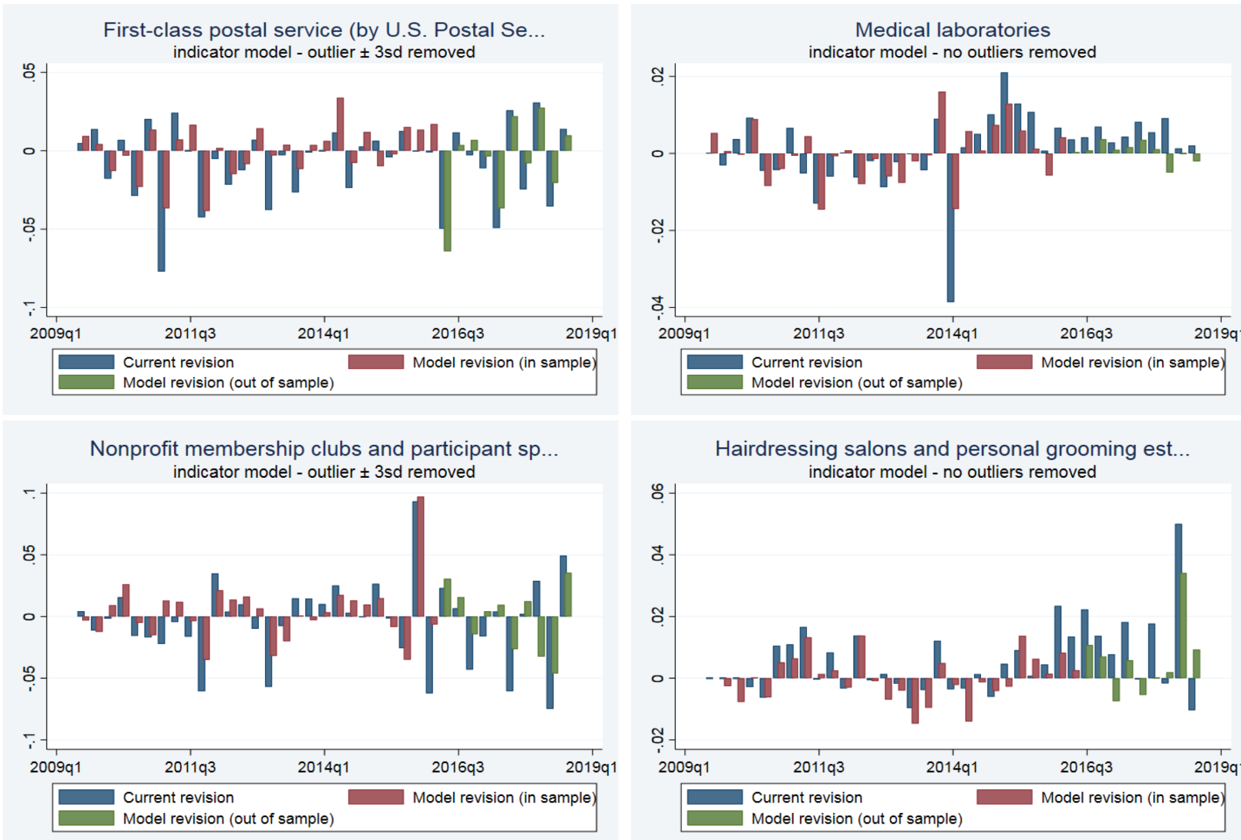
1	2	3	4	5	6
Service group	No. of components	No. of largest reduction in RMSR from current method	No. of largest reduction in RMSR from GB	No. of largest reduction in RMSR from BF	Range of % reduction in RMSR
Communication service	6	4	2	0	(29.5, 50.0)
Health care service	20	2	10	8	(3.5, 35.3)
Personal service	7	1	1	5	(2.5, 46.7)
Professional service	5	0	3	2	(6.85, 33.9)
Recreational service	17	6	10	1	(1.0, 46.2)
Social service	19	7	6	6	(5.5, 29.6)
Transportation service	14	3	9	2	(0.5, 53.1)
SUM	**88**	**23**	**41**	**24**	**(0.5, 53.1)**
% of total components	**100%**	**26%**	**47%**	**27%**	

Note: Personal service includes categories of personal and home maintenance services; transportation services include categories of domestic transportation and foreign travel services.

RMSR root mean squared revisions

We also graphically compare revisions in quarterly growth rates from the in-sample estimation and POOS predictions from the bridge equation and bridging with factors models against those from the current extrapolation method. In Figure 2.1, blue bars show revisions from the current extrapolation method; red bars show revisions computed from the in-sample estimation of the bridge equation model; and green bars indicate revisions from the one-step-ahead predictions computed from the estimated bridge equation model. Not surprisingly, revisions from in-sample estimation are generally smaller than those from the POOS predictions. Moreover, revisions from in-sample estimation and POOS predictions are noticeably dampened in the periods in which the current extrapolation method produced large spikes of revisions. However, a reduction in revision does not necessarily occur in every quarter in the sample.

Figure 2.1. Reduction in RMSR in Selected PCE Services Based on Bridge Equation Model

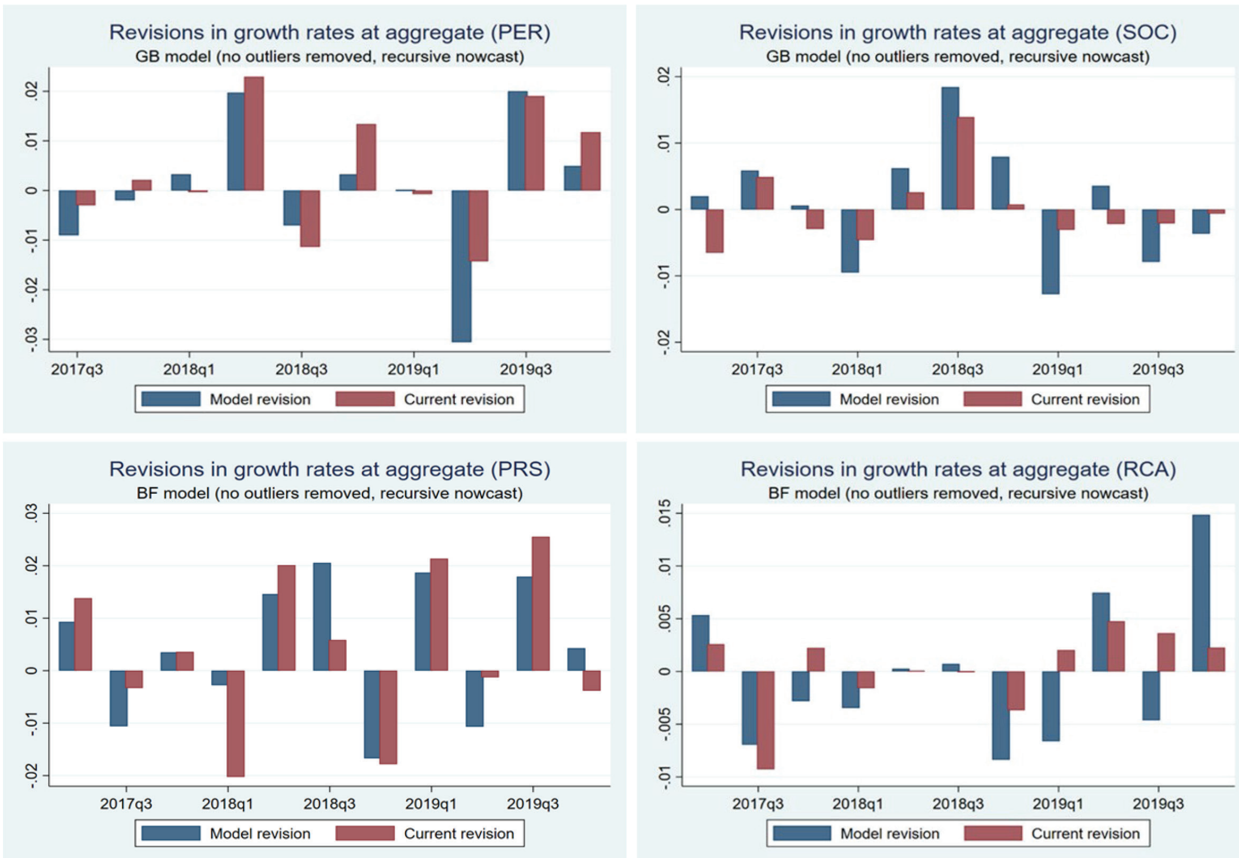


Similarly, Figure 2.2 compares revisions in selected components of PCE services from the current extrapolation method, the in-sample estimation exercise, and the one-step-ahead POOS predictions computed from the estimated bridging with factors models. Observations on the sizes and smoothness of revisions from in-sample estimation and POOS predictions are similar to those shown in Figure 2.1.

Figure 2.2. Reduction in RMSR in Selected PCE Services Based on Bridging with Factors Models

Since changes in revisions in the detailed components of services are not identical, it is useful to compare revisions by category of PCE services. Figure 2.3 compares revisions at the aggregate level for selected categories of PCE services. The graphs in the top panel compare revisions from the current method with those from the one-step-ahead predictions based on estimated bridge equations; the graphs in the bottom panel compare revisions from the current method with those from the one-step-ahead predictions based on the estimated bridging with factors model. The red bars represent revisions from the current extrapolation method and the blue bars show revisions from the one-step-ahead predictions of the estimated models. Again, we observe that revisions from the one-step-ahead POOS predictions computed from the two models dampen the large revision spikes produced by the current extrapolation method. We also observe that even though revisions are greatly reduced in some periods, reductions in revisions are not seen in every period.

Figure 2.3. Revisions at Aggregated PCE Services—Selected PCE Services Categories



Note: PCE: Personal consumption expenditures; PER: Personal services; PRS: Professional services; RCA: Recreational services; SOC: Social services.

4.2 Results from the proposed model-averaging schemes for combination nowcasts

To evaluate the performance of each model-averaging scheme, the full sample is split into an estimation and a test sample, and model-averaging weights are computed only on the estimation sample. To compute the complete set of revisions for the test sample, we estimate each model and compute the averaging weights on an "expanding window" basis, meaning that for each period in the test sample, the model-averaging weights are computed based on the estimated models using all observations prior to the period under consideration. The estimation results share similar features as those from the individual models. However, because we consider 12 models and 10 model-averaging schemes, it would be overwhelming to present results from in-sample estimation of all models for weights calculation. Thus, we focus on reporting reductions in revision for detailed PCE services from combined one-step-ahead POOS predictions based on each model-averaging scheme.

For each detailed component of PCE services, we compute RMSR from the one-step-ahead POOS combination nowcasts using the test sample and compare it with the RMSR based on the current extrapolation method. The main observations are that: 1) characteristics of the target and indicator (or factor) series of each component determine the model-averaging scheme that produces the largest reduction in RMSR; 2) by adding model-averaging techniques to nowcasting, reductions in RMSR are seen in 80 detailed components of PCE services, an increase of 16 components (or 18 percent) compared with nowcasting with only the 2 "best" individual models; and 3) the weighted log likelihood averaging scheme outperforms the other methods by producing the largest reductions in revision in 40 (or 45 percent) of the 88 components.

Table 5 shows, as an example, percentage changes in RMSR for detailed components in HLC services based on the 10 model-averaging schemes and the 2 "best" individual models, relative to the RMSR from the current method. Column 1 displays the 20 components in health care services; columns 2 to 11 contain the percentage change in RMSR brought by each model-averaging scheme; and columns 12 to 13 show reductions in RMSR produced by the "best" bridge equation and the "best" bridging with factors models. The shaded blue entry in each row indicates the largest reduction in RMSR for a given component from one of the twelve methods. Column 14 and 15 display the largest reduction in RMSR and the corresponding best method for each component, ranging from 1.5 percent to 43.4 percent.

For health care services, reduction in RMSR is seen in 19 out of the 20 components. One-step-ahead predictions from each model show a reduction in RMSR in more than 50 percent of the 20 components. It is worth noting that each simple averaging scheme yields reductions in RMSR in 18 out of the 20 components, more than any other method. It is also worth noting that weighted log-likelihood averaging technique outperforms all other models, producing the largest reduction in RMSR in 13 of the 20 components.

To evaluate the performance of each proposed method in all categories of PCE services, table 6 compares the number of largest reductions in RMSR brought by each method among the 88 detailed components in the study.

Table 5. Percentage Changes in RMSR from Nowcasts Produced by Alternative Model-Averaging Schemes and the "Best" Individual Models—Health Care Services

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	PCE_HLC	LL AVERAGING	IC_BASED AVERAGING		BG-AVERAGING	SIMPLE AVERAGING				COV. MIN BASED AVERAGING		MODEL SELECTED BY AICC			
No.	CODE	LL	AIC	BIC	BATES-GRANGER	EW	MED	TRIM MEAN 10%	TRIM MEAN 25%	JMA	MA	BEST BE	BEST BF	ΔRMSRss%	Best method
1	SOH	-29.10	-17.13	-20.05	-18.28	-19.12	-21.10	-19.82	-20.21	-12.58	-12.58	-10.51	-13.11	-29.10	LL
2	SPS	-19.37	-4.78	-4.83	-20.60	-21.21	-24.69	-23.28	-23.42	-15.68	-12.88	-8.10	-4.77	-24.69	MED
3	SPO	-25.47	-11.40	-10.46	-26.81	-27.85	-29.27	-28.30	-28.80	-17.93	-17.18	-27.84	-11.08	-29.27	MED
4	HHH	-30.41	1.01	-2.95	-24.65	-25.16	-27.37	-25.57	-26.35	-9.69	-0.57	-27.13	1.13	-30.41	LL
5	HIX	-14.96	-5.46	-3.64	-12.01	-12.19	-10.38	-11.88	-11.29	-7.28	-10.74	-11.19	-3.74	-14.96	LL
6	HOX	-30.55	-29.93	-28.78	-29.13	-29.02	-28.53	-28.83	-28.77	-29.33	-30.07	-24.43	-29.62	-30.55	LL
7	NPH	0.74	23.04	17.81	0.83	0.74	-1.50	-0.76	-1.28	14.09	11.50	14.99	10.86	-1.50	MED
8	FPH	23.63	22.24	21.40	24.75	27.21	25.18	25.05	24.04	17.35	20.09	8.64	22.37	0.00	CUR
9	HSO	-7.30	0.67	0.17	-2.35	-2.09	-6.14	-3.90	-5.89	-2.28	-1.96	-5.16	1.47	-7.30	LL
10	NXO	-20.32	0.93	0.18	-1.78	-1.92	-2.38	-1.23	-1.81	-0.89	-2.82	1.10	-9.76	-20.32	LL
11	NPN	-2.85	-10.78	-9.99	-12.28	-11.82	-13.05	-12.04	-12.24	-9.11	-11.14	-10.74	-8.93	-13.05	MED
12	FPN	-11.50	-13.94	-21.35	-14.35	-12.45	-12.38	-12.11	-12.46	-17.54	-13.47	-30.63	2.56	-30.63	BE
13	PHH	-34.36	-28.20	-28.21	-34.91	-34.55	-32.59	-33.57	-33.11	-33.27	-32.41	-32.33	-28.20	-34.91	ISESQ
14	PSX	-43.40	-31.96	-24.14	-39.96	-40.79	-41.04	-41.00	-41.31	-35.21	-40.83	-28.81	-35.25	-43.40	LL
15	DEN	-13.13	0.82	-1.35	-10.16	-10.57	-5.77	-9.73	-8.93	-1.98	-2.78	5.35	-3.50	-13.13	LL
16	MLB	-29.29	-19.16	-20.51	-22.62	-23.48	-23.23	-23.14	-23.63	-21.53	-23.35	-24.35	-22.56	-29.29	LL
17	OHH	-18.99	0.44	-4.68	-15.14	-15.72	-14.64	-14.89	-14.63	-6.95	-2.49	-11.41	1.01	-18.99	LL
18	OPS	-28.67	-22.58	-22.46	-26.05	-26.48	-26.42	-26.92	-26.87	-23.17	-20.89	-11.43	-24.39	-28.67	LL
19	POX	-18.76	1.48	2.18	1.62	-0.41	4.57	2.46	3.59	-2.65	-1.51	7.52	1.47	-18.76	LL
20	OPO	-21.57	4.82	-10.61	-21.53	-19.95	-17.63	-18.05	-17.85	-18.86	-15.01	-18.84	-5.14	-21.57	LL
		13	0	0	1	0	4	0	0	0	0	1	0		

/ — The shaded blue entry in each row indicates the largest reduction in RMSR for a given component from one of the twelve methods.

Note: Columns 2-13 display changes in RMSR from nowcasting techniques relative to the extrapolation method. A negative value indicates a reduction in the RMSR. Column 14 show the maximum reduction in RMSR and the corresponding nowcasting technique for the component in Column 1.

RMSR root mean squared revisions

Table 6. Best Method for Reducing RMSR for PCE Services Components by Category

				IC based averaging				Simple averaging				Cov. minimization averaging		Model selection based on AICC	
No.	Service category	No. of components	Current method	Log likelihood	AIC	BIC	Bates-Grange LOOCV	Equal weight	Median weight	Trimmed mean (10%)	Trimmed mean (25%)	Jackknife averaging	Mallows average	Bridge equation (VSELECT)	Factor model (VSELEC)
1	COM	6	3	2										1	
2	HLC	20	1	13					5					1	
3	PER	7	1		1	1								1	3
4	PRS	5		5											
5	RCA	17	2	4	1			1	1	1	2		1	4	
6	SOC	19	0	12							1		1	1	4
7	TRSFTR	14	1	4		1	2		2			2		1	1
	SUM	**88**	**8**	**40**	**2**	**2**	**2**	**1**	**8**	**1**	**3**	**2**	**2**	**9**	**8**
	%SUM	**100%**	**9.1**	**45.5**	**2.3**	**2.3**	**2.3**	**1.1**	**9.1**	**1.1**	**3.4**	**2.3**	**2.3**	**10.2**	**9.1**

Note: Personal services include personal and home services; transportation includes domestic transportation and foreign travel.

PCE personal consumption expenditures

RMSR root mean squared revisions

The weighted log-likelihood averaging scheme stands out, outperforming all other methods by producing the largest reduction in RMSR in 40 components; the "best" bridge equation and the "best" bridging with factors model produced the largest reduction in RMSR in 9 and 8 components, respectively. Four simple averaging schemes combined produced largest reduction in RMSR in 13 components. In contrast, the AIC, BIC, Bates-Granger, JMA and MA model-averaging techniques performed less impressively in this sample of data, each producing the largest reduction in RMSR in only 2 components. There are 8 components for which the proposed nowcasting models did not yield a reduction in RMSR.

To examine additional improvements in revisions from adding the proposed model-averaging techniques, we compare the largest reductions in RMSR in each category produced by the two individual nowcasting models and the 10 model-averaging techniques. Table 7 shows that model-averaging techniques would help further improve the accuracy of advance estimates of PCE services. By including the model-averaging schemes in the study, the number of components showing reduction in RMSR increases from 64 to 80, ranging from 1.5 percent and 63 percent.

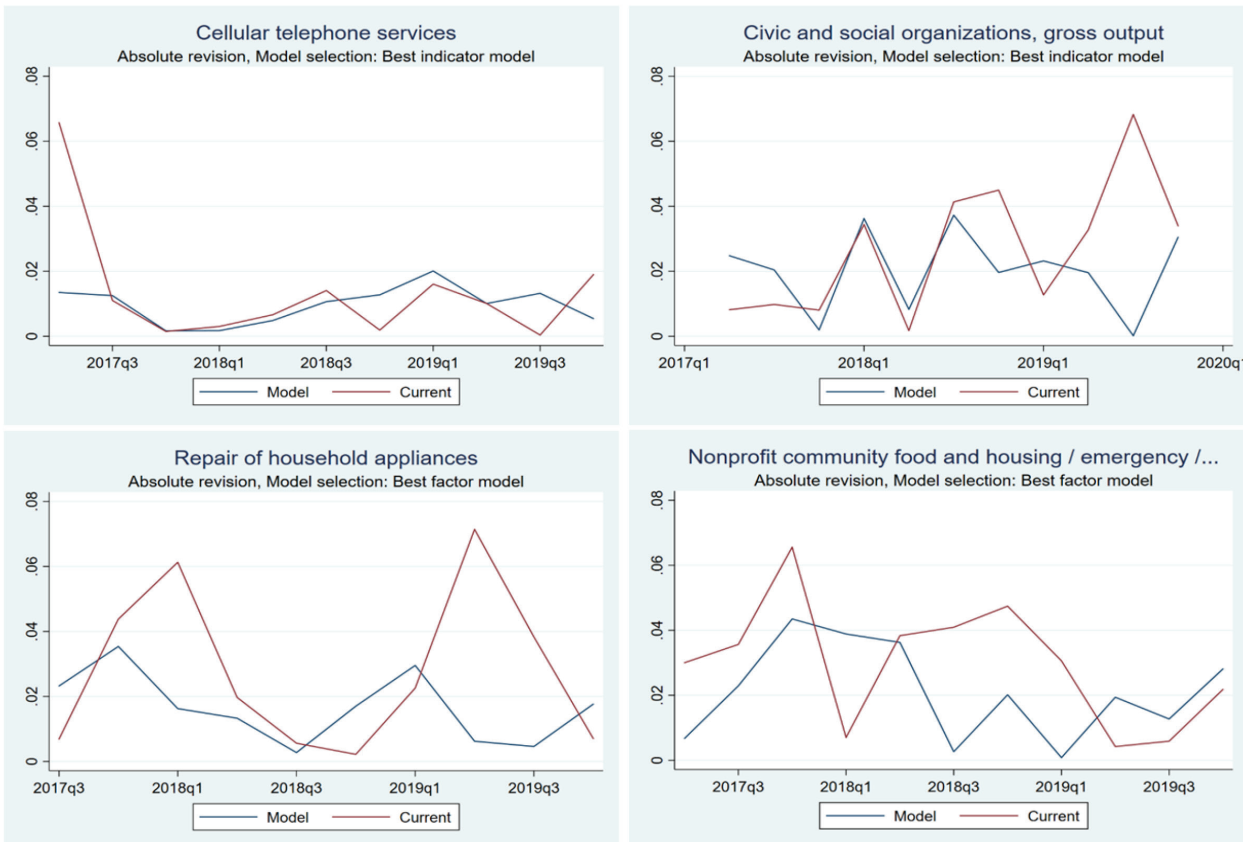
Table 7. Summary of Reduction in RMSR from All Models in Evaluation

Service category	No. of components	No. of largest reduction in RMSR from current methods	No. of largest reduction in RMSR from GB	No. of largest reduction in RMSR from BF	No. of largest reduction in RMSR from Model-averaging methods	Range of % reduction in RMSR
Communication service	6	3	1	0	2	(29.3, 50.2)
Health care service	20	1	1	0	18	(1.5, 43.4)
Personal service	7	1	1	3	2	(2.5, 49.3)
Professional service	5	0	0	0	5	(11.5, 36.5)
Recreational service	17	2	4	0	11	(2.3, 46.1)
Social service	19	0	1	4	14	(5.8, 35.5)
Transportation service	14	1	1	1	11	(5.4, 63.0)
SUM	**88**	**8**	**9**	**8**	**63**	**(2.3, 63.0)**
% of SUM	**100%**	**9.1%**	**10.2%**	**9.1%**	**71.6%**	

Note: Models in evaluation include general bridge equation, bridging with factor model, and the 10 model-averaging schemes.

In general, revisions from the bridge equation and bridging with factors models as well as from the model-averaging schemes tend to be smaller and smoother compared with revisions from the current extrapolation method. Figure 3.1 uses four detailed components (proprietary and government physician services, medical laboratories, legal services, and proprietary and public community food and housing services) as examples to compare revisions, in absolute values, from the "best" bridge equation and "best" bridging with factors models with revisions from the current extrapolation method. Lines in red represent revisions from the current method and blue lines show the smaller and smoother revisions from the "best" estimated bridge equation and "best" bridging with factors models.

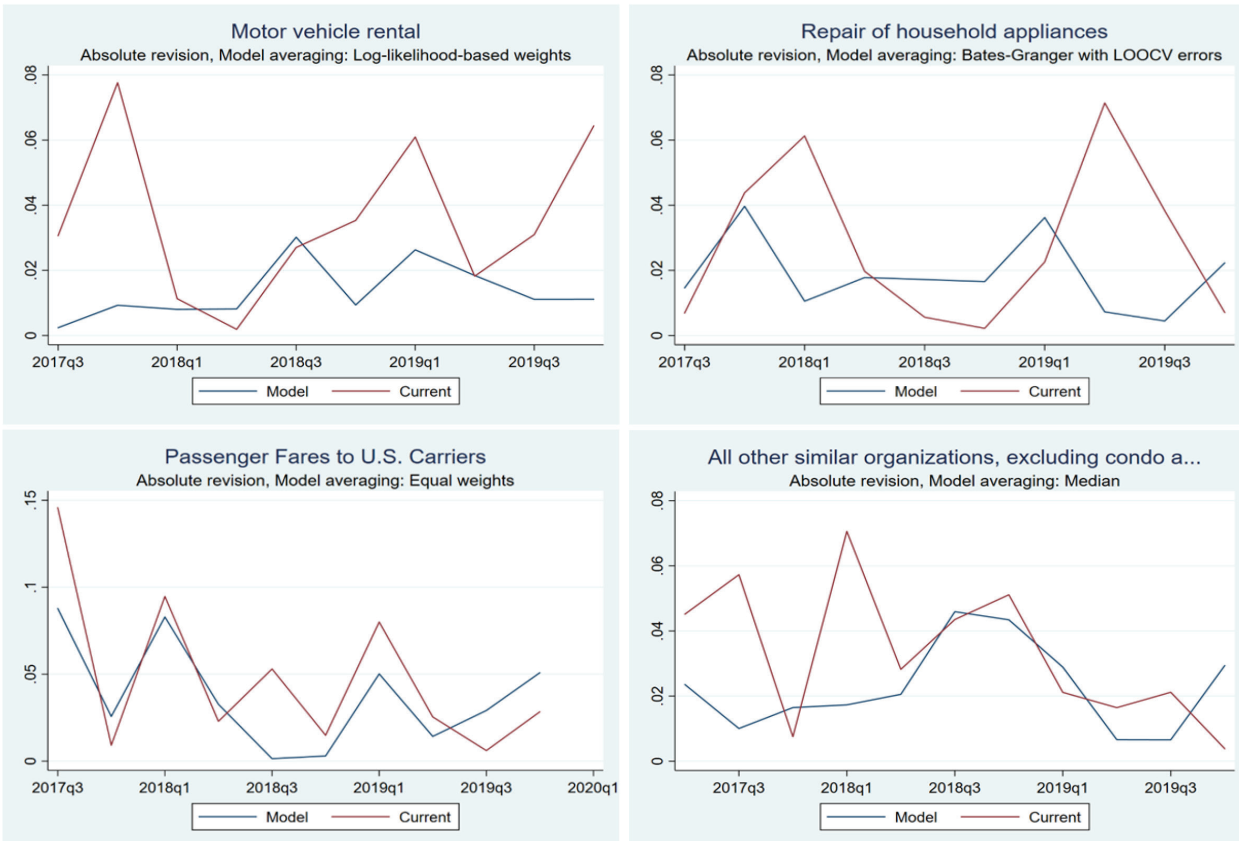
Figure 3.1. Comparison of Revisions in Advance Estimates from Bridge Equation or Bridging with Factors Models and the Current Extrapolation Method in Selected PCE Services



Note: Best indicator model refers to the best bridge equation model.

Similarly, in Figure 3.2, we compare revisions from selected model-averaging techniques with revisions from the current extrapolation method in four components of PCE services (physician services, medical laboratories, motor vehicle rental, and other postal delivery services). In all four cases, revisions from the model-averaging techniques are smaller and smoother. The big revision spikes from the current method are dampened and smoothed out compared to when the combination nowcasts are used.

Figure 3.2. Comparison of Revisions in Advance Estimates from Model-Averaging Schemes and Current Extrapolation Method in Selected PCE Services



5. Conclusion

In this study, we apply two individual nowcasting models and a collection of 10 model-averaging techniques to compile advance estimates of detailed components of PCE services using real-time data from the U.S. national accounts. The results demonstrate that nowcasting methods are useful for improving the accuracy of advance estimates of detailed PCE services categories, as these methods allow more information on the quarterly target variables and the monthly indicators to be used in estimation.

Comparing the two individual nowcasting methods, the bridge equation framework frequently outperforms the bridging with factors model when a single model is considered, suggesting that many of the designated indicators that are already used to produce estimates of the detailed PCE services contain sufficient relevant information on the movements of the target series. The two individual nowcasting methods considered here produced reductions in RMSR in 73 percent of the components included in the study. However, degrees of reduction vary over time and across services categories.

Model-averaging techniques are shown to further improve the accuracy of advance estimates of detailed PCE services. In fact, model-averaging techniques are shown to have the potential to be a powerful tool in reducing revisions in national accounts statistics at the detailed component level. For some detailed components, model-averaging techniques reduce revisions by more than 50 percent. Adding the 10 proposed model-averaging techniques in the study, reduction in RMSR is achieved in 80 (or 91 percent) of the 88 components series.

Each of the 10 model-averaging techniques produced the largest reduction in RMSR in some components of PCE services. However, the non-penalized information-criterion-based averaging based on the log-likelihood weights outperformed all other techniques, producing the largest reduction in RMSR in 40 (or 45 percent) of the 88 components series. The “best” bridge equation and “best” bridging with factors models combined stand out as the second-best performer, producing the largest reduction in RMSR in 17 (or 19 percent) of the 88 series. The four simple averaging schemes combined produced the largest reduction in RMSR in 13 (or 14 percent) of all components. The other model-averaging schemes performed less impressively in this study, each producing the largest reduction in RMSR in two components. Thus, the results seem to suggest that to achieve computational efficacy, the two individual models and a small set of the model-averaging techniques could be considered for implementation in the national accounts to produce more accurate advance estimates of PCE services at the detailed component level.

There are some component series for which revision reductions are small or nonexistent. Perhaps these series require additional information to properly forecast. We also continue to try to fully understand the reasons the forecast combination puzzle (that is, simple averaging outperforms other sophisticated techniques) did not arise in more cases in our study, and the reasons that the weighted log-likelihood averaging outperformed all other methods.

Finally, the recent COVID-19 pandemic caused unexpected delays in the collection and production of monthly indicator data, which inevitably led to incomplete information for compiling advance estimates of PCE services. One remedy is to explore information from high frequency electronic transactions data, which are especially useful for capturing market activities during the sharp fluctuations in the market caused by unexpected events. To try to capture the sudden sharp fluctuations in the market activities, in further studies we will extend our models to investigate whether high-frequency credit card payments data provide additional information on market activities that is not already covered in the traditional indicator data.

References

Aiolfi, M., C. Capistrán, and A. Timmermann. 2010. “Forecast Combination,” Bank of Mexico, Working Paper 2010-4.

Bai, J. and S. Ng. 2002. “Determining the number of factors in approximate factor models,” *Econometrica*, 70(1): 191–221.

Bańbura M., D. Giannone, M. Modugno, and L. Reichlin. 2013. “*Now-casting and the real-time data flow*, European Central Bank, Working Paper Series no. 1564, 1–54.

Bates, J. M. and C. W. J. Granger. 1969. “The combination of forecasts,” *Journal of the Operational Research Society*, 20(4): 451–468.

Burnham, K. P. and D.R. Anderson. 2003. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. (Springer: December 2003)

Chen, J., A. Dunn, K. Hood, A. Driessen, and A. Batch. 2019. “Off to the Races: A Comparison of Machine Learning and Alternative Data for Predicting Economic Indicators,” Forthcoming in *Big Data for 21st Century Economic Statistics*, edited by K. G. Abraham, R. S. Jarmin, B. Moyer, and M. D. Shapiro. Chicago: University of Chicago Press, for the National Bureau of Economic Research.

Clement, Robert T. 1989. “Combining forecasts: A review and annotated bibliography,” *International Journal of Forecasting*, vol. 5(4): 559–583.

Denton, Frank T. 1971. “Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization,” *Journal of the American Statistical Association*, vol. 66, 99–102.

Diks, C. G. H. and J. A. Vrugt (2010), “Comparison of point forecast accuracy of model averaging methods in hydrologic applications,” *Stochastic Environmental Research and Risk Assessment*, 24(6): 809–820.

Foroni, C. and M. J. Marcellino. 2013. “A survey of econometric methods for mixed-frequency data,” Norges Bank Research Working Paper 2013–6.

Giannone, D., L. Reichlin, and D. Small. 2008. “Nowcasting: The real-time informational content of macroeconomic data,” *Journal of Monetary Economics*, 55(4): 665–676.

Hansen, Bruce E. 2010. “Averaging estimators for autoregressions with a near unit root,” *Journal of Econometrics*, Elsevier, vol. 158(September): 142–155.

Hansen, B. E. and Racine, J. S. 2012. "Jackknife model averaging," *Journal of Econometrics*, Elsevier, vol. 167(1): 38–46.

Higgins, Patrick. 2014. "GDP Now: A Model for GDP 'Now-casting,'" Working Paper 2014-7, Federal Reserve Bank of Atlanta.

Kitchen, J. and R. Monaco. 2003. "Real-Time Forecasting in Practice: The U.S. Treasury Staff's Real-time GDP Forecast System," MPRA Paper no. 21068, U.S. Department of Treasury.

Klein, L. R. and E. Sojo. 1989. "Combinations of high and low frequency data in macroeconometric models," 1–13 in *Economics in Theory and Practice: An Eclectic Approach*, L.R. Klein and Marquez (eds.), Kluwer Academic Publisher.

Moral-Benito, Enrique. 2011. "Model averaging in economics," Working Paper no. 1123, Bank of Spain, 1–47.

Lindsey, Charles. 2014. "VSELECT: Stata module to perform linear regression variable selection," Statistical Software Components S457808, Boston College Department of Economics, revised 15 Dec 2014.

Nguefack-Tsague, Georges. 2014. "On optimal weighting scheme in model averaging," *American Journal of Applied Mathematics and Statistics*, vol. 2(3): 150–156.

Piette, Christopher. 2016. "Predicting Belgium's GDP using targeted bridge models," Working Paper Research No. 290, National Bank of Belgium.

Smith, J. and K. F. Wallis. 2009. "A simple explanation of the forecast combination puzzle," *Oxford Bulletin of Economics and Statistics*, 71(3): 331–355.

Timmermann, Allan. 2006. "Forecast Combinations." In *Handbook of Economic Forecasting*, edited by G. Elliott, C. Granger, and A. Timmermann, 1st ed., vol. 1, ch. 4, 135–196, Elsevier.

Appendix: Additional Tables

Table A.1. PCE Services by Category

No.	PCE service category	No. of components
1	Health care (HLC)	21
2	Transportation (TRS)	12
3	Net foreign travel (FTR)	6
4	Recreation (RCA)	25
5	Communication (COM)	6
6	Professional and other services (PRS)	12
7	Personal care and clothing (PER)	5
8	Household maintenance (HHM)	7
9	Social and religious activities (SOC)	27
	SUM	**121**