Federal Reserve Bank of New York

December 2017
Volume 23 Number 2

Economic Policy Review

Economic Policy Review

The *Economic Policy Review* is published by the Research and Statistics Group of the Federal Reserve Bank of New York. The views expressed are those of the individual authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

www.newyorkfed.org/research

Federal Reserve Bank of New York
Economic Policy Review

December 2017
Volume 23 Number 2

Contents

A measure of underlying inflation that uses all relevant information, is available in real time, and forecasts inflation better than traditional underlying inflation measures—such as core inflation measures—would greatly benefit monetary policymakers, market participants, and the public. This article presents the New York Fed Staff Underlying Inflation Gauge (UIG) for the consumer price index and the personal consumption expenditures deflator. Using a dynamic factor model approach, the UIG is derived from a broad data set that extends beyond price series to include a wide range of nominal, real, and financial variables. This modeling approach also makes it possible to combine information simultaneously from the cross-sectional and time dimensions of the sample in a unified framework. In addition, the UIG can be updated on a daily basis to closely monitor changes in underlying inflation—a feature that is especially useful when sudden and large economic fluctuations occur, as was the case during the 2008 global financial crisis. Lastly, the UIG displays greater forecast accuracy than many measures of core inflation.

Marlene Amstad, Simon Potter, and Robert Rich

The New York Fed Staff Underlying Inflation Gauge (UIG)

- Monetary policymakers and others would benefit from a smooth, broad based, real-time measure of underlying inflation.
- The authors introduce the New York Fed Staff Underlying Inflation Gauge (UIG), explain its construction and review the experience of the Federal Reserve Bank of New York with daily, real-time updates of the UIG, made internally since 2005.
- The UIG includes a wide range of nominal, real, and financial variables in addition to prices and focuses on the persistent common component of monthly inflation.
- The UIG proved especially useful in detecting turning points in trend inflation and has shown higher forecast accuracy compared with core inflation measures.

1. Introduction

The two most widely followed measures of consumer price inflation in the United States are the consumer price index (CPI) and the personal consumption expenditures (PCE) deflator, both released monthly. Yet for many observers—including monetary policymakers and market participants—the "headline" readings of both series are too volatile to provide a reliable measure of the trend in inflation even after some averaging of the series. Indeed, the series can fluctuate quite dramatically: the headline twelve-month change in the CPI was 5.6 percent in July 2008, fell to zero in December of the same year, and then reached a low of –2.1 percent in July 2009.

Not surprisingly, the volatility of the two leading measures has prompted a large and ongoing research effort to extract the long-run, or persistent, component of aggregate inflation from the monthly data releases. Approaches to estimating this component—termed "underlying inflation"—have varied, both in their methodology and in the data set used.

Marlene Amstad is professor of practice in economics at the Chinese University of Hong Kong, Shenzhen, and research fellow at the Federal Reserve Bank of New York. Simon Potter is an executive vice president and head of the Markets Group and Robert Rich an assistant vice president at the Federal Reserve Bank of New York.

marleneamstad@cuhk.edu.cn
simon.potter@ny.frb.org
robert.rich@ny.frb.org

Work on the UIG began in 2004-05 when Amstad, on leave from the Swiss National Bank, was a Federal Reserve Bank of New York resident visiting scholar, and it continued during periodic follow-up visits. An earlier version of this article was published as "Real-Time Underlying Inflation Gauges for Monetary Policymakers," Federal Reserve Bank of New York Staff Reports, no. 420 (2009). The authors' work draws from a prior experience developing a similar gauge for Switzerland (Amstad and Fischer 2009a, 2009b) and builds on code developed by Ricardo Cristadoro, Mario Forni, Domenico Giannone, Marc Hallin, Marco Lippi, Lucrezia Reichlin, and Giovanni Veronese (see Cristadoro et al. [2005]).

The authors thank Evan LeFlore, Ariel Zetlin-Jones, Joshua Abel, Christina Patterson, M. Henry Linder, Ravi Bhalla, Matt Cocci, and Linda Wang for excellent research assistance. They are also grateful for comments on versions of this article from Jonathan McCarthy, Stephen Cecchetti, Marvin Goodfriend, Kenneth Rogoff, Domenico Giannone, members of the New York Fed's Economic Advisory Panel, other staff in the Federal Reserve System, and seminar participants at the Bank for International Settlements, European Central Bank, Norges Bank, and the Reserve Bank of New Zealand. The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

To view the authors' disclosure statements, visit https://newyorkfed.org/research/author_disclosure/ad_epr_2017_underlying_inflation_rich.

One well-established approach to estimating underlying inflation is to construct measures of "core" inflation. This approach assumes that transitory changes in the aggregate price index are linked to the volatility of its subcomponents. Consequently, core inflation measures are generally designed to remove the most volatile price changes associated either with the same specific goods and services, or with those goods and services displaying the largest price increases and decreases in a particular month. The former strategy underlies the "ex-food and energy" measure—which removes the impact of food and energy prices on inflation. The latter strategy motivates the trimmed mean and median measures.[1] Although such adjustments may seem reasonable, researchers have identified various limitations in the core inflation measures.[2] One well-known limitation of these measures is that they assume that the source of transitory movements in aggregate inflation remains constant over time. In addition, they focus exclusively on the cross-sectional dimension of the data and therefore neglect potentially useful information in movements of the data over time. Further, core inflation measures can only be updated monthly, which might be too infrequent during periods when there is heightened uncertainty about movements in trend inflation. There are also reasons to question the reliability and timeliness of these measures as a gauge of underlying inflation.[3]

Another common approach to estimating underlying inflation is to use model-based techniques. This approach can involve statistical smoothing methods whose complexity can vary widely. It can also involve the estimation of Phillips curve models and structural vector autoregression (SVAR) models. However, as with the core inflation measures, researchers have raised concerns about the model-based measures—in this case, because of their near-exclusive reliance on price data, sensitivity to particular specifications, or strong model restrictions.

Recognizing the limitations of commonly used measures of underlying inflation, we present the New York Fed Staff Underlying Inflation Gauge (UIG). This measure of underlying inflation for the CPI and PCE deflator provides a complement to existing measures and aims to add value by helping to detect turning points in trend inflation. This article describes the development of the UIG, explains its construction, and reviews the experience of the Federal Reserve Bank of New York with daily, real-time updates of the UIG, made internally since 2005. We note that the New York Fed is preparing to publish monthly updates of the UIG for CPI inflation starting later in 2017.

The design of the UIG is based on the premise that movements in underlying inflation are accompanied by related persistent changes in other economic and financial series. Specifically, the UIG is defined as *the persistent common component of monthly inflation*. Consequently, we examine a large data set and apply modern statistical techniques to extract a small number of "factors" that capture the common fluctuations in the series. The data set includes disaggregated price data as well as a wide range of nominal, real, and financial variables. The statistical techniques, known as dynamic factor models, provide a very tractable framework in which to use large information sets, with the extracted factors serving as the basis to construct the UIG.

The UIG offers several notable features that build on and extend the work done by other researchers on the estimation of underlying inflation. The framework used here combines information simultaneously from the cross-sectional and the time-series dimensions of the sample in a unified framework. In this regard, our modeling strategy follows that of Cristadoro et al. (2005), who derive a measure of underlying inflation for the euro area. In addition, the UIG uses a real-time framework, entailing daily updates of the model, which was introduced by Amstad and Fischer (2009a, 2009b) in the development of an inflation gauge for Switzerland. Our work also finds parallels with that of Stock and Watson (1999, 2016) and Reis and Watson (2010), who use a dynamic factor model to estimate a common component that they associate with trend inflation. The UIG differs from these last studies, however, by moving beyond the common component to extract its persistent element.

Our analysis offers significant evidence of the UIG's effectiveness in monitoring inflation developments in real time and assessing their implications for the inflation outlook of policymakers and market participants. An essential property of a measure of underlying inflation is the ability to look through

[1] There are also strategies that weight inflation subcomponents inversely by their volatility rather than exclude volatile subcomponents. Going forward, we use the terms "traditional underlying inflation measures" and "core inflation measures" interchangeably. With regard to core inflation measures, our study focuses on the ex-food and energy measure, the trimmed mean, and the median.

[2] For example, see Cecchetti (1997), Cecchetti and Moessner (2008), and Bullard (2011) as well as the references therein for further discussion.

[3] During the recent global financial crisis, the twelve-month change in headline CPI inflation fell to 2.1 percent in July 2009—far below the 1.1 percent value that was the lowest reading during the previous recession in 2001. For the CPI ex-food and energy, however, the lowest twelve-month change during the recent global financial crisis was 0.6 percent—a value that was not reached until October 2010 and was not that far from the low of 1.1 percent observed during the 2001 recession.

A similar concern arises in the case of the PCE deflator during these same episodes. The twelve-month change in headline PCE inflation fell to –1.2 percent in July 2009, again far below the low of 0.6 percent seen during the 2001 recession. Meanwhile, PCE inflation ex-food and energy declined to 1 percent in July 2009, which was only slightly below the low of 1.2 percent during the 2001 recession.

the noise—short-term transitory fluctuations—in headline inflation to identify movements in the trend. We show that in past noncrisis periods, during which trend inflation remained fairly stable, the UIG showed little response to noise in headline inflation. However, when the economy was subject to large and persistent shocks, such as in 2008, the UIG was very responsive to the worsening conditions in the economy and offered a daily signal of the speed and scale of changes in underlying inflation. In particular, we find that the addition of nonprice data was especially important for the UIG to quickly signal the sharp and rapid decline in trend inflation during the global financial crisis. Because the UIG was able to generate this signal in real time, this model feature is particularly useful for decision makers, including policymakers and investors.

Last, how do our findings on the performance of the UIG relate to other researchers' assessments of trend inflation measures? Many studies have concluded that no single measure of underlying inflation consistently outperforms other measures across a range of criteria.[4] Other studies have narrowed their analysis to evaluating the relative performance of select measures in *forecasting* inflation. For example, Atkeson and Ohanian (2001) argue that a simple random walk model (that is, the use of the most recently observed change in inflation to forecast future inflation) is just as accurate as Phillips curve models that incorporate nonprice variables in their specification. Stock and Watson (2008) subsequently find that while Phillips curve models remain useful tools for forecasting inflation, their value is "episodic." That is, Phillips curve models do not offer higher forecast accuracy than a random walk model during times of low volatility, but provide additional predictive content around business cycle turning points.[5] We find that the UIG outperforms core inflation measures as well as a simple random walk model in a pseudo out-of-sample forecast exercise that covers subsamples both before and during the recent global financial crisis. Consequently, we conclude that the UIG adds meaningful value compared with alternative measures in forecasting inflation. We attribute the robustness of the UIG's greater accuracy in this regard to its use of a large data panel and its focus on only the persistent part of the common component of inflation.[6]

[4] See, for example, Rich and Steindel (2007) and the references therein. Stock and Watson (2010) and Wynne (2008) give a comprehensive analysis that also supports this assessment for the United States and Vega and Wynne (2001) for the euro area. Cecchetti (1995) shows evidence that this finding is related to structural breaks in the inflation process.

[5] Liu and Rudebusch (2010) confirm the finding of Stock and Watson (2008) including data for the global financial crisis.

[6] The motivation for the found robustness is supported by Gavin and Kliesen's (2008) evidence that data-rich models significantly improve the forecasts for a variety of real output and inflation indicators.

The remainder of this article is organized as follows. Section 2 discusses a suite of measures of underlying inflation, including their strengths and weaknesses. Section 3 motivates our specification of the dynamic factor model, and also describes the data set and estimation procedure used to construct the real-time UIG. In Section 4, we compare the UIG with traditional underlying inflation measures using descriptive statistics as well as forecast performance. Section 5 presents our conclusions.

2. Underlying Inflation: A Review of Approaches and Measures

This section examines various approaches to estimating underlying inflation, and highlights measures included in our analysis. The discussion helps to motivate the modeling strategy adopted for the UIG.

For any observed headline inflation rate π_t, the rate can always be decomposed as:

$$\pi_t = \pi_t^* + c_t \,, \tag{1}$$

where π_t^* denotes the underlying rate of inflation and c_t denotes deviations of inflation from the underlying inflation rate. While the concept of underlying inflation is generally agreed upon, the best method for estimating the underlying inflation rate is not—a wide range of proposed measures of π_t^* exist. One dimension along which the measures differ is the choice of methodology. Another area of difference is the nature of the data set, with some measures only using price data and others including additional variables. We now examine and comment in more detail on some of the more popular approaches and corresponding measures used to estimate underlying inflation.[7]

The term "core inflation" is widely used by practitioners and academics to represent a measure of underlying inflation that is less volatile than headline inflation. Measures of core inflation gained attention in the 1970s when large price movements in food and oil complicated the task of estimating the trend in inflation. This experience highlighted the importance of developing methods that could filter out

[7] There are measures of underlying inflation that are derived from financial markets (for example, breakeven inflation using Treasury Inflation-Protected Securities) or consumer surveys (for example, the University of Michigan Inflation Expectations data). However, these measures provide a forecast of future underlying inflation rather than an estimate of current underlying inflation. Consequently, we exclude them from our analysis.

transitory price movements in order to identify the persistent part of inflation. One strategy suggested by Gordon (1975) and Eckstein (1981) associates the transitory elements with food and energy prices and argues for excluding these items from the price index every month. Another strategy, suggested by Bryan and Cecchetti (1994), associates the transitory elements with those items displaying the largest price movements—both increases and decreases—in a particular month and argues for computing trimmed mean and median measures in which the excluded items are allowed to change each month.[8] In the United States, statistical agencies publish monthly measures of the CPI and the PCE deflator that exclude the food and energy subcomponents, while various Federal Reserve Banks calculate trimmed mean and median measures for the CPI and the PCE deflator.[9]

An attractive feature of core inflation measures is that they are easy to construct and to understand. Further, their forecast performance, as shown by Atkeson and Ohanion (2001), can be very similar to, or even better than, measures of underlying inflation based on more complicated approaches.[10]

There are, however, limitations to core inflation measures and the practice of excluding volatile components. In the case of the ex-food and energy measure, the specific subcomponents to be removed are determined in a strictly backward-looking manner based on the historical behavior of the noise in the inflation release. For example, although in the 1970s it may have been reasonable to exclude temporary oil price increases from core inflation measures, it makes less sense to do so now because oil price changes appear to be more persistent.[11] This discussion illustrates an inherent difficulty in the construction of core inflation measures: What is temporary only becomes apparent in retrospect and not in advance.[12]

[8] See Bryan and Cecchetti (1994, 1999), Bryan, Cecchetti, and Wiggins (1997), Dolmas (2005), and Meyer, Venkatu, and Zaman (2013) for a discussion of methodologies.

[9] The Federal Reserve Bank of Cleveland reports trimmed mean and median measures for the CPI (suggested by Bryan and Pike [1991]), while the Federal Reserve Banks of Dallas and San Francisco report, respectively, trimmed mean and median measures for the PCE deflator.

[10] Although some studies report evidence favorable to the forecast performance of core inflation measures, Crone et al. (2013) have reported that the relative forecast performance of core inflation measures can be sensitive to the choice of the inflation measure and time horizon of the forecast.

[11] James Hamilton and Menzie Chinn have written several blog posts on oil prices that illustrate this point. Furthermore, Cecchetti and Moesnner (2008) points out that the exclusion of energy from this measurement has imparted a bias to medium-term measures of inflation.

[12] In their comprehensive comparison of core inflation measures, Rich and Steindel (2007) conclude that no single core measure outperforms the others over different sample periods owing to the fact that there is considerable variability in the nature and sources of transitory price movements.

In the case of the trimmed mean and median measures, another concern is that excluding components that display large price changes (in either direction) may remove early signals of a change in trend inflation that tend to show up in the tails of the price change distribution. Therefore, even though the trimmed mean or median measures may display a low average forecast error over long-dated episodes, they may be a lagging indicator at important times such as turning points in trend inflation. More generally, the practice of excluding large price changes narrows the range of possible reported outcomes during a given time period. Consequently, core inflation measures can suffer both from being late to recognize changes in underlying inflation and from understating the extent of such changes.[13]

Because of the limitations of core inflation measures, model-based techniques have been used to develop measures of underlying inflation for the United States. Within this approach, one strategy has focused on the application of time-series smoothing methods. Examples include the integrated moving average (IMA) model of Nelson and Schwert (1977), the four-quarter moving average model of Atkeson and Ohanian (2001), the exponential smoothing model of Cogley (2002), and the stochastic volatility model of Stock and Watson (2007). However, these applications involve univariate time-series methods and only examine aggregate inflation for their analyses. More recently, Stock and Watson (2016) have proposed a measure of underlying inflation that is based on the estimation of a multivariate unobserved components-stochastic volatility model using price data for the subcomponents of the PCE deflator. Although Stock and Watson (2016) also associate underlying inflation with the estimated common component of multiple inflation series, they do not include nonprice data.

Another strategy within this approach involves model estimation using additional nonprice data. One prominent example includes Gordon (1982) "triangle"-type models.[14] Gordon estimates a backward-looking Phillips curve model and combines price data along with labor market information and additional covariates to capture exogenous pricing pressures, such as those from energy. Underlying inflation measures can then be derived as the endpoint of the within-sample prediction values from the model, with the estimation period varied either in a recursive manner or through a rolling window. One criticism of the estimated measure of underlying inflation is that there are limitations on the number of variables that can be added to the model as

[13] Footnote 3 in the Introduction touched upon these points.

[14] The triangle model is a common approach to modeling inflation in the Federal Reserve System (Rudd and Whelan 2007).

a result of degrees of freedom issues. Another criticism is that it is very sensitive to the particular model specification (Stock and Watson 2008).

Quah and Vahey (1995) provide another example, in which they propose a slightly different definition of underlying inflation based on the long-run neutrality of inflation. Specifically, they define underlying inflation as the "component of measured inflation" that has no medium- to long-run impact on real output. However, their approach requires the estimation of a SVAR model that has been criticized on the grounds that it is difficult to formulate and imposes tenuous identifying restrictions.

Taken together, the issues we have outlined speak to the limitations associated with various measures of underlying inflation. Given these limitations, we view dynamic factor models as providing an attractive framework in which to develop an improved measure. Among the reasons motivating our choice is the fact that dynamic factor models have received increased attention and gained greater popularity because their specification allows for the use of a broad data set without requiring adherence to strong theoretical guidelines for estimation purposes. The UIG is related to this modeling strategy and is formalized in greater detail in the next section.

3. New York Fed Staff Underlying Inflation Gauge (UIG)

The New York Fed Staff UIG is based on the estimation of a dynamic factor model using price data as well as economic and financial variables. This section motivates our modeling strategy and highlights its important features, including a broad data approach and flexibility to extract information from many indicators. We then describe the specification of the dynamic factor model and illustrate its role in the construction of the UIG. With regard to the dynamic factor model, we also provide a general discussion of issues related to model parameterization and estimation procedure. After describing the data set used for the analysis, we examine the estimated UIG series and their behavior.

The research that corresponds most closely to our work on the UIG is by Amstad and Fischer (2009a, 2009b), who developed a gauge for Switzerland, and by Amstad, Huan, and Ma (2014),[15] who developed one for China—both relying on the methodology of Cristadoro et al. (2005) in a real-time framework. Giannone and Matheson (2007) and Khan, Morel, and Sabourin (2013) adopt a similar approach to construct an inflation gauge for New Zealand and Canada, respectively, but their analyses only use disaggregated price data.[16] Further, related work has employed dynamic factor models of the type used in this study to explore several issues related to inflation dynamics. For example, Altissimo, Mojon, and Zaffaroni (2009) investigate persistence in aggregate inflation in the euro area, while Amstad and Fischer (2009b) explore the impact of macroeconomic announcements on weekly updates of forecasts for Swiss core inflation, and Amstad and Fischer (2010) construct monthly pass-through estimates from import prices to consumer prices in Switzerland.

3.1 Methodology

From a policy perspective as well as a forecasting perspective, there are several reasons why it is beneficial to add rather than exclude information to measure underlying inflation. As argued in Bernanke and Boivin (2003), monetary policymaking operates in a "data-rich environment." Furthermore, Stock and Watson (1999, 2002, 2010) show that broader information sets can improve forecast accuracy in certain time periods. Therefore, several authors (including Galí [2002]) argue that policymakers would benefit from a more comprehensive measure that can cull and encapsulate the relevant information for inflation from a large data set.

By their design, factor models can be applied to a broad data set and therefore offer a particularly attractive framework to summarize price pressures in a formal and systematic way as well as to gauge sustained movements in inflation. The key feature of this class of models is that although the data set contains a large number of variables, a significant amount of their co-movement can be explained using a low number of series—referred to as factors. In addition to the work cited in this article that has used large data factor models to derive measures of underlying inflation, this modeling strategy has been used to construct measures of economic activity.[17]

[15] For an update see People's Bank of China (2016).

[16] The inflation gauge developed by Giannone and Matheson (2007) and Khan, Morel, and Sabourin (2013) is similar to the prices-only version of the UIG discussed later in this article.

[17] With regard to the latter application, Altissimo et al. (2001) use a dynamic factor model to produce EuroCoin, which provides a monthly reading of euro area GDP, while the Chicago Fed National Activity Index offers a monthly gauge of U.S. GDP.

For this study, we follow Cristadoro et al. (2005) and use the generalized dynamic factor model developed by Forni et al. (2000, 2001, 2005) that draws upon the work of Brillinger (1981) and allows for the application to large data sets. The following discussion is intended to provide the reader with a general understanding of the theoretical framework and estimation procedure used to construct the UIG, as well as to preview issues that will receive subsequent attention.

Let X_t represent the time t values of the N series that make up our large data set such that $X_t = [x_{1,t}, x_{2,t}, \ldots, x_{N,t}]$. For convenience, let $x_{1,t}$ denote the monthly inflation rate. We assume that the behavior of $x_{1,t}$ can be described as the sum of two unobserved components using a formulation similar to equation (1):

$$x_{1,t} = x^*_{1,t} + e_{1,t}\,, \tag{2}$$

where $x^*_{1,t}$ denotes our variable of interest, the underlying rate of inflation, and $e_{1,t}$ is a component reflecting movements in inflation related to other factors such as short-run dynamics, seasonality, measurement error, and idiosyncratic shocks. A central element of our analysis is to use the dynamic factor model methodology to estimate $x^*_{1,t}$ using information from present and past values of X.

The dynamic factor model assumes that the variables in X_t can be represented as the sum of two mutually uncorrelated, unobserved components without trend: the common component $\chi_{i,t}$—which is assumed to capture a high degree of co-movement between the variables in X_t—and the idiosyncratic component $\xi_{i,t}$. The premise of a dynamic factor model is that the common component reflects the influence of a few factors that act as a proxy for the fundamental shocks that drive behavior in an economy, while the idiosyncratic component reflects the influence of variable specific shocks. More formally, we can summarize the time-series process for each variable in X_t as

$$x_{i,t} = \chi_{i,t} + \xi_{i,t} = \sum_{h=1}^{q}\sum_{k=0}^{s} \alpha_{i,h,k}\,\mu_{h,t-k} + \xi_{i,t}\,, \tag{3}$$

where the common component $\chi_{i,t}$ is defined by the same q common factors, $\mu_{h,t}$, but which may be associated with different coefficients and lag structures, with maximum lag s. The appeal of the dynamic factor model is that it provides a convenient dimension reduction technique. That is, it enables us to use a small number of factors to summarize the information from a large data set.

Looking at the first time-series variable, $x_{1,t}$, as well as equations (2) and (3) yields

$$x_{1,t} = x^*_{1,t} + e_{1,t} = \chi_{1,t} + \xi_{1,t}\,. \tag{4}$$

Because our notion of the underlying rate of inflation relates to the long-run, or persistent, component of aggregate inflation, we would like this property to carry over to the common component in equation (4). It is important to note that, as proposed by Cristadoro et al. (2005), $\chi_{1,t}$ can be separated into a long-run (persistent) component, $\chi^{LR}_{1,t}$, and a short-run component, $\chi^{SR}_{1,t}$, based on a specified cut-off frequency for the data. Accordingly, we can rewrite equation (4) as

$$x_{1,t} = x^*_{1,t} + e_{1,t} = \chi^{LR}_{1,t} + \chi^{SR}_{1,t} + \xi_{1,t}\,. \tag{5}$$

From equation (5), we can then think of the underlying rate of inflation in terms of the following association:

$$x^*_{1,t} = \chi^{LR}_{1,t}\,. \tag{6}$$

That is, the UIG is defined as the long-run common component of monthly inflation. As previously described, one difference between our approach and that of Stock and Watson (1999, 2016) concerns our additional filtering of the common component to isolate its persistent element. This difference is illustrated and may be best understood by comparing equation (4) with equation (6).

Although our interest focuses on $\chi^{LR}_{1,t}$, neither the common component $\chi_{1,t}$ nor the factors underlying its behavior are observable and therefore they must be estimated. Because some aspects of the estimation and the construction of the UIG are quite technical, we refer readers to Cristadoro et al. (2005) and Forni et al. (2000, 2001, 2005) for more information, rather than explore these issues in further detail here.[18] Instead, we turn our attention to the specification of three key parameters of the model. In particular, we need to select a cut-off horizon to filter out short-run fluctuations in the data as incorporated in equation (5), and select the number of factors q and the number of maximum lags s as described in equation (3).[19]

We select a cut-off frequency of twelve months to extract $\chi^{LR}_{1,t}$ from $\chi_{1,t}$. Lags in the monetary transmission mechanism suggest that inflation at a horizon of one year or less is relatively insensitive to changes in current monetary policy. Therefore there is little that policymakers can do to affect

[18] For example, estimation of the dynamic factor model and smoothing of the UIG are undertaken in the frequency domain.

[19] For New York Fed internal analysis, these settings are evaluated on a regular basis.

these fluctuations in inflation. Consequently, if monetary policy has been achieving its objective of price stability with well-anchored inflation expectations, then the effects of changes in current monetary policy on expected inflation will be at horizons of greater than twelve months. In addition, this choice enables us to remove seasonal effects.

With respect to the number of common factors, our analysis will involve settings of $q = 1$ and $q = 2$. To preview the results discussed in Section 3.3, the difference in the number of specified dynamic factors reflects variations in the nature of the data set. In particular, we find that only one factor is relevant when the price data are considered alone but that two factors provide a proper representation when we include the nonprice variables in the data environment. The q factors are allowed to influence UIG not only contemporaneously but also with a maximum number of lags s. Our choice of $s = 12$ is motivated by several considerations that include consistency with the one-year cut-off band for the common component and the monthly frequency of the data.[20]

Thus, the UIG at time t is then defined as the predicted long-run common component of the monthly inflation rate from estimation of equation (3) with settings of $q = 1$ or $q = 2$, $s = 12$, and a cut-off frequency of twelve months. That is,

$$x^{\star}_{1,t} = \hat{\chi}^{LR}_{1,t} . \tag{7}$$

The previous discussion and formulation in equations (1) through (7) highlight several key properties of the UIG. The definition of the UIG is consistent with the idea that a measure of underlying inflation should reflect a common as well as a persistent element in the component parts of aggregate price indexes. In addition, the presence of multiple factors does not restrict movements in underlying inflation to those driven by a single type of shock. The estimated factors take into account the co-movement of variables in both the cross-sectional and the time-series dimensions, without imposing any restrictions on the sign or magnitude of the correlations.

Moreover, the analysis does not require that the factors either be extracted from a pre-selected partition of the data set or pre-identified as a specific type of shock. Lastly, the UIG is well suited to evaluate whether a large price change is likely to persist over a specified period of time as the UIG's movement is not restricted in either speed or magnitude.[21] Specifically, our inferences about movements in underlying inflation are informed by an empirical framework that allows for a broad representation of economic and financial developments at the same time that it allows information from this large data set to be extracted in a flexible manner and to be summarized in a very parsimonious way.

3.2 Data

There is no objective criterion to judge which data should or should not be included in the large information set. Consequently, we rely on the experience of the New York Fed staff and include the series considered to be the most relevant determinants of inflation. The data set has remained the same since 2005 when we began construction of the UIG.

We use data from the following two broad categories: (1) consumer, producer, and import prices for goods and services and (2) nonprice variables such as labor market measures, money aggregates, producer surveys, and financial variables (short- and long-term government interest rates, corporate and high-yield bonds, consumer credit volumes and real estate loans, stocks, and commodity prices). We refrain from including every available indicator that could have an impact on inflation because research on factor models (Boivin and Ng 2006) shows that doing so does not come without risks.[22] Our approach is to include the variables that were regularly followed by the New York Fed staff in their assessment over several economic cycles. This procedure not only offers the benefit of drawing upon the staff's long-term experience, but also maintains some continuity in the set of variables used to construct the UIG. Such continuity is important because it helps ensure that a change in the UIG is not caused by changes in the data composition through the addition or removal of a data series. The weighting of each series in the UIG changes over time and is determined by the factor model as new observations become available and existing data are revised. Chart 1 provides more information on the current data set used, while the Data Appendix provides a detailed listing of the variables.

[20] Further analysis indicated that the results were not sensitive to variation in the number of these lags.

[21] An additional advantage of our UIG concept compared with traditional underlying inflation measures is that it enables us to focus on a particular horizon of interest that will, in this case, align with that of policymakers. As previously discussed, the horizon of interest for this study is twelve months and longer.

[22] Their results suggest that factors estimated using more data do not necessarily lead to better forecasting results. The quality of the data must be taken into account, with the use of more data increasing the risk of "leakage of noise" into the estimated factors.

Sample Range

Based on substantial evidence of structural breaks in the U.S. inflation process (see Clark [2004] and Stock and Watson [2008] for a comprehensive evaluation), we limit our analysis of the data to the period starting in January 1993. For similar reasons, the OECD (2005) divides the sample for a multicountry study of inflation into the subperiods 1984-95 and 1996-2004. In addition, a tension exists between our large data set and the dynamic factor model—which relies on a balanced data set to start the estimation—requiring us to strike a balance between the length of the time period and the range of indicators for the study. These considerations reinforced the choice of January 1993 as the start date because an earlier time period would have limited significantly the number of time series that could be included in the analysis.

3.3 Estimation Results

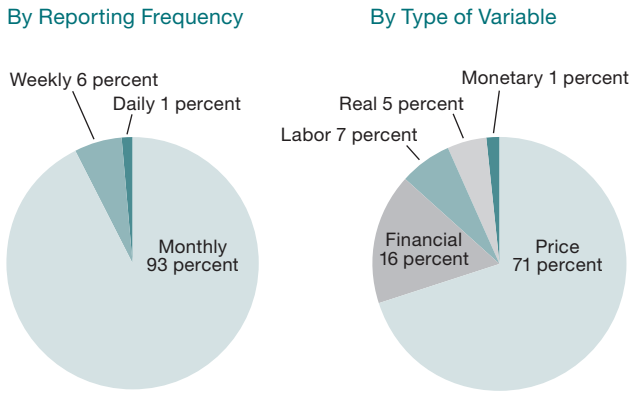
In this section, we discuss some additional details of the estimation procedure, the number of factors used to summarize the information content of our data set, and the behavior of the resulting UIG series. Following conventional practice in the factor model literature, prior to estimation we transformed the data to induce stationarity and standardized each series so that it has zero mean and unit variance.[23] Because of the standardization process, the initially estimated UIG series is driftless and must be re-normalized by assigning an average growth rate to it. We use 2.25 percent for the CPI and 1.75 percent for the PCE. When we began the project at the end of 2004, these numbers were very close to the respective average inflation rates starting from 1993.[24]

[23] Almost all variables were transformed to growth rates to induce stationarity, except for a small number for which no transformation was required. Using the variables listed in the Data Appendix, no transformation was applied to the eighteen variables in the Real Variables group, the first seventeen variables in the Labor group, and the Standard and Poor's 500 Price Earnings Ratio Index in the Financials group.

[24] As noted in the discussion, a value needs to be selected to allow for a nonzero mean of the underlying inflation measure. When we started this analysis, the Federal Reserve Board had not stated a numerical inflation goal. In January 2012, the Federal Open Market Committee agreed to a longer-run goal of a 2 percent PCE inflation rate.

A growing number of countries establish their monetary policy more or less explicitly according to an inflation target. In these countries, information on the inflation target (or the specific point target, if available) can be used to construct the average of the underlying inflation measure.

Chart 1
Breakdown of UIG Series



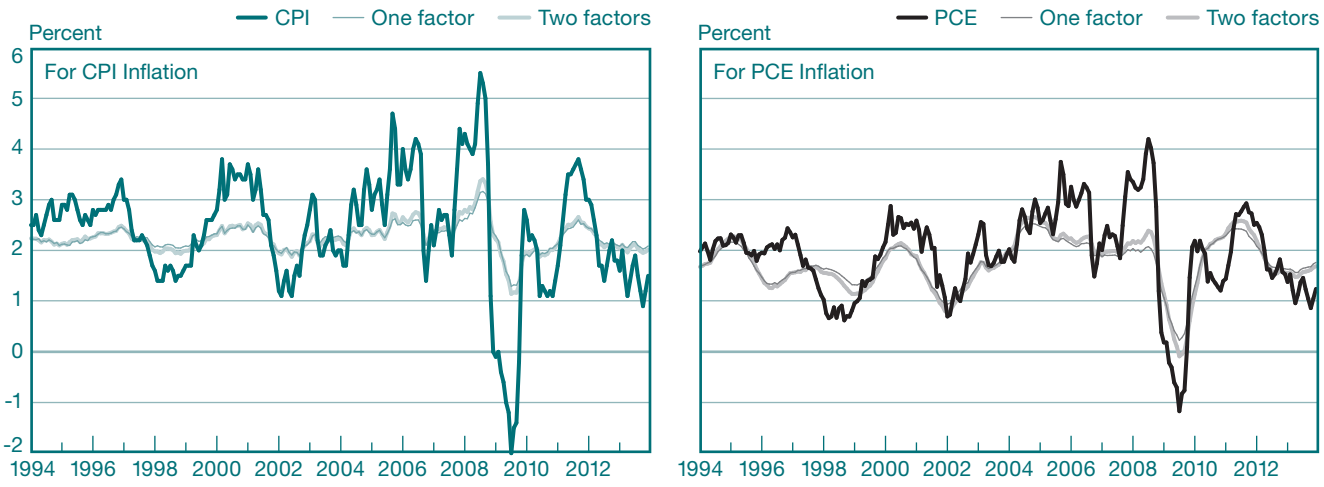
Source: Authors' calculations.

With regard to the number of factors, different articles find that much of the variance in U.S. macroeconomic variables is explained by two factors. Giannone, Reichlin, and Sala (2005) show this result using hundreds of variables for the period 1970-2003, while Sargent and Sims (1977) examine a relatively small set of variables and use frequency domain factor analysis for the period 1950-70. Watson (2004) notes that the two-factor model provides a good fit for U.S. data during the postwar period, and that this finding is quite robust. Hence, in most large data-factor-model applications the number of factors is set to two. Often one factor is associated with real variables (such as GDP or aggregate demand), while the second factor is associated with nominal prices (such as the CPI).

Our choice of the number of factors is not based on the considerations described above. Rather we draw upon the previously cited literature and include the lowest number of factors needed to represent our data environment properly—without labeling the factors (as either real or nominal) or interpreting them. We start our examination of the UIG measure by presenting estimates based only on price data from the CPI and PCE.[25] One

[25] We refer to these as the UIG estimates using prices-only data for the CPI and PCE. References to the "UIG for CPI inflation" and "UIG for PCE inflation" indicate measures derived using additional nonprice variables. The Data Appendix lists the series used in the analysis. In particular, the prices-only model for the UIG for CPI inflation uses the first 222 listed variables in the Prices group, while the prices-only model for the UIG for PCE inflation uses all 254 variables. The former choice facilitates the comparison to a core CPI measure that only uses CPI subcomponents, while the latter choice reflects the earlier release date of the CPI data and their usefulness in predicting PCE inflation. The model for the UIG for CPI inflation uses the first 242 listed variables in the Prices group and the variables from all the other groups (a total of 345), while the model for the UIG for PCE inflation uses all of the listed variables (a total of 357) in the Data Appendix.

Chart 2
UIG Estimates Using Only Price Data



Sources: Bloomberg L.P.; authors' calculations.

Notes: CPI is consumer price index; PCE is personal consumption expenditures deflator.

would expect these series to be driven by a single factor, since the data set comprises nominal variables only. The left and right panels of Chart 2 show the one- and two-factor estimates of the prices-only UIG for CPI inflation and PCE inflation, respectively, along with the twelve-month change in the relevant price index. As shown, there is little difference between the two estimates, offering support for the view that only one factor is relevant when the price data are considered alone.

Chart 3 shows the one- and two-factor estimated UIGs incorporating the nonprice variables in our data set through December 2013, along with the relevant twelve-month inflation rate. Three findings are noteworthy. First, the estimates now show larger cyclical fluctuations and appear to track inflation more closely. Second, starting in 2005 they correctly capture a broadly declining trend despite the temporary large increase in inflation in the first half of 2008. Moreover, when we turn to the period of the global financial crisis, we are immediately struck by how quickly the UIG begins to signal the deceleration in inflation starting in the second half of 2008 as a decline in trend inflation. In particular, a marked downturn in the UIG emerges as early as December 2008. Taken together, these findings suggest that the additional information contained in the nonprice variables is quite important both in terms of trend/cycle decomposition as well as in the timeliness of identifying shifts in underlying inflation. Third, the estimates based on two or more factors for the most part differ little from one another, a result that underlies our adoption of two factors for the dynamic factor model.[26]

Real-Time Updates and Data Revisions

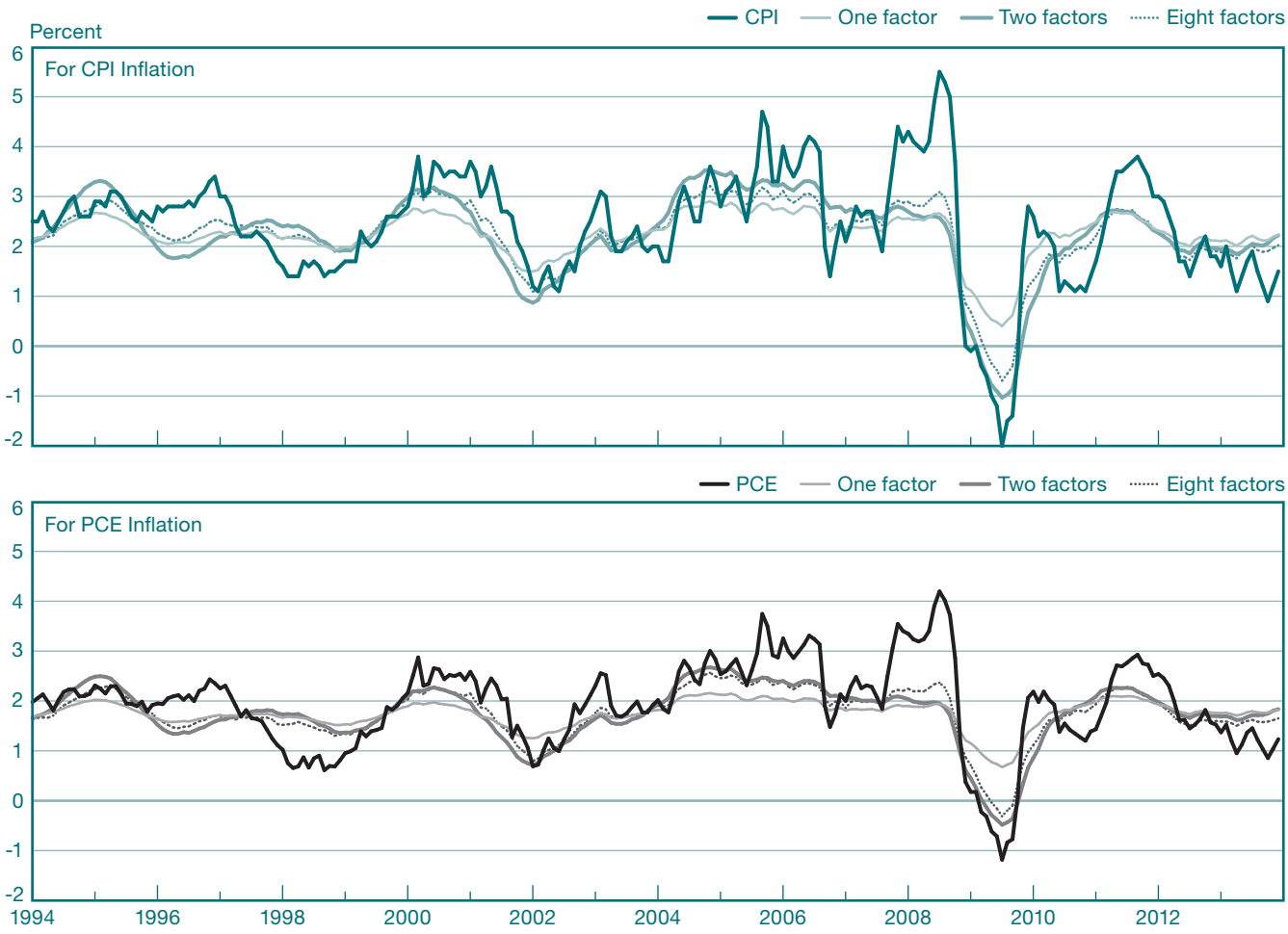
The UIG offers a monthly gauge of underlying inflation but is updated daily, following Amstad and Fischer (2009a, 2010) in their work using Swiss data. The monthly dating of the UIG is motivated by the monthly frequency of inflation reports in the United States. The daily updates allow for a close monitoring of the inflation process and also provide a basis to assess movements in underlying inflation that stem from daily changes in financial markets between monthly inflation reports.[27]

[26] Specifically, we considered estimates of the UIG that included as many as eight factors.

[27] Because our data set includes the most current daily information available, it results in an unbalanced panel structure. Therefore, some series end in month T, while others end in months T-1, T-2, . . . T-j. To address the unbalanced panel structure at the end of the sample, we use the methodology of Altissimo et al. (2001) and Cristadoro et al. (2005), which provides procedures to fill in the missing observations and create a balanced panel for estimation purposes.

Chart 3

UIG Estimates Using Different Numbers of Factors



Sources: Bloomberg L.P.; authors' calculations.

Notes: CPI is consumer price index; PCE is personal consumption expenditures deflator.

The daily UIG updates contrast with the monthly data releases of headline and core inflation measures. More generally, daily UIG updates can also be used to identify the sources of a change in inflation forecasts by determining the impact of a particular economic or financial news release—for example, the unemployment rate or an ISM (Institute for Supply Management) number—on underlying inflation.[28]

One aspect of the UIG updates is particularly important and merits special attention. Specifically, a UIG update not only generates a reassessment of the measure's behavior during the current month, but also for all previous months. This revisionist history occurs because each time the dynamic factor model is re-estimated, the addition of new data and revisions to existing data result in changed parameters as well as a more informed inference about the (estimated) factors throughout time.[29] As shown by equations (3), (5), and (7), changes in the time-series behavior of the factors will result in a different path for the predicted value of the persistent component of monthly inflation and hence the UIG. We explore and quantify the relevance of these revisions in the next section.

[28] Amstad and Fischer (2009b, 2010) provide an example of this type of analysis using an event study approach for Swiss inflation.

[29] Technically, this is referred to as smoothing the state vector in the dynamic factor estimation procedure.

Because of the revisionist nature of the UIG, it is important to limit other sources of variability as much as possible to derive a reliable signal of underlying inflation. Therefore, most of the selected data is either not revised or is subject to limited revisions. This implies that we must rely heavily on survey data for measures of real activity and not use more traditional measures based on National Income and Product Accounts (NIPA) data.[30] Another advantage of survey data is that it is usually released more quickly than expenditure and production data. Additionally, we use data that is not seasonally adjusted and, following Amstad and Fischer (2009a, 2009b), apply filters within the estimation procedure to generate a seasonally adjusted estimate of underlying inflation. We adopt this approach primarily because it prevents revisions in our measure of underlying inflation from being driven by concurrent seasonal adjustment procedures.

4. Comparing Measures of Underlying Inflation

This section compares core inflation measures and the UIG measures for CPI and PCE inflation. We begin by commenting on general features of the measures' behavior. Next we turn to statistical properties of the various underlying inflation measures and compare their ability to track and forecast inflation.

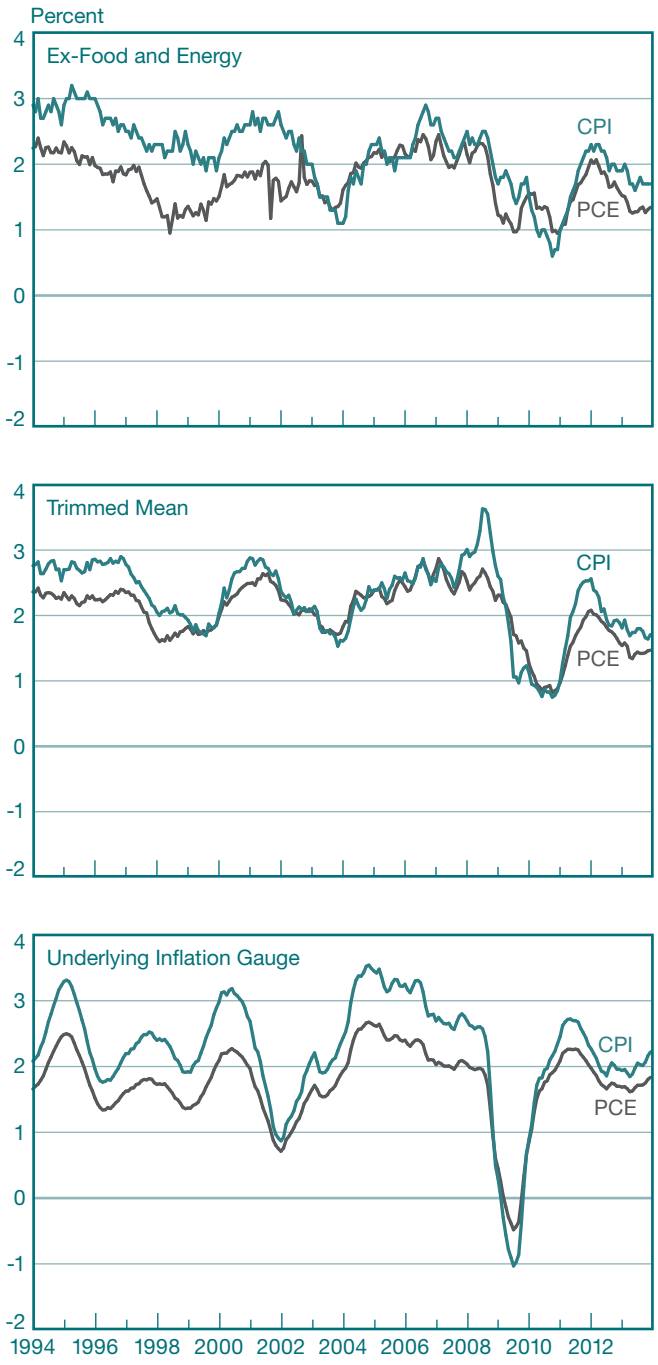
4.1 General Features and Statistical Properties

The underlying inflation measures in this study differ across two dimensions: methodology and price index. We begin the comparison by investigating the relative importance of each of these considerations. Chart 4 plots three underlying inflation measures—ex-food and energy, trimmed mean, and UIG—for the two price indexes, while Chart 5 plots underlying inflation measures for the same price indexes along with the twelve-month inflation rate.[31] As shown, we find that the general behavior of the different measures of underlying inflation is driven mainly by the choice of methodology and less by the choice of the price

[30] The NIPA data provides a detailed snapshot of the production of goods and services in the United States and the income that results. They are produced by the Bureau of Economic Analysis of the Department of Commerce and are an important source of data on U.S. economic activity.

[31] The upper panel of Chart 5 also includes the CPI Median, which is used for the forecast performance evaluation in Section 4.2. There is, however, no measure of the PCE Median that is readily available. The core inflation measures plotted in each panel are constructed as twelve-month changes.

Chart 4
Underlying Inflation Gauges for CPI and PCE Inflation

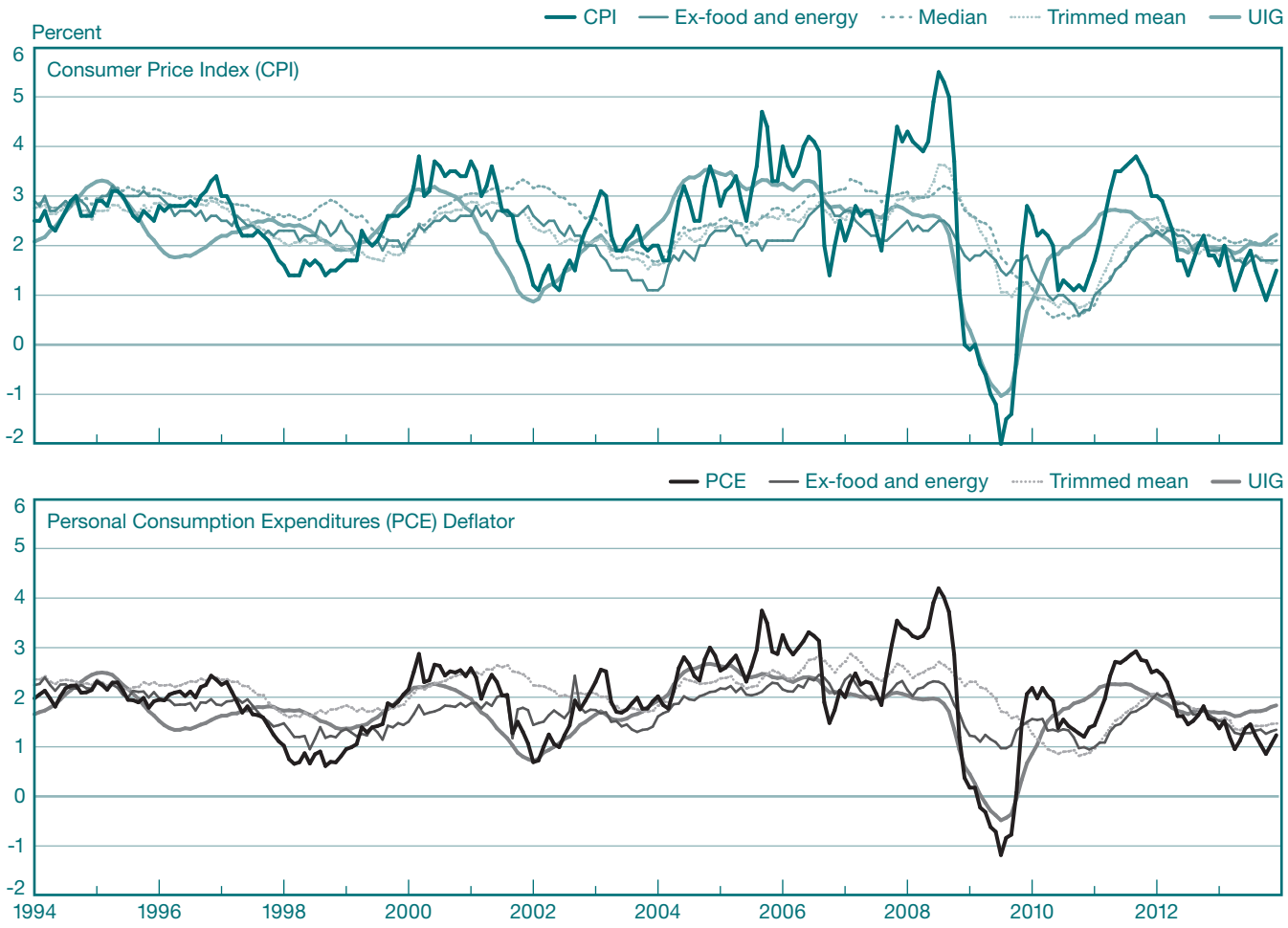


Sources: Bloomberg L.P.; authors' calculations.

Notes: CPI is consumer price index; PCE is personal consumption expenditures deflator.

Chart 5

A Comparison of Underlying Inflation Gauges



Sources: Bloomberg L.P.; authors' calculations.

index. While Chart 4 displays a level shift across the price indexes, there is a strong correlation between the underlying inflation measures within each panel. In Chart 5, however, there is a lower correlation between the underlying inflation measures, which is particularly evident when we look at the core inflation measures relative to the UIG.

We now examine three statistical features of the various underlying inflation measures: smoothness, the correlation with headline CPI inflation and headline PCE inflation, and the correlation between the UIG for CPI inflation and the UIG for PCE inflation.

First, smoothness is typically associated with the volatility of a series—measured using a metric such as a standard deviation—with lower volatility viewed as a favorable criterion in the evaluation of underlying measures of inflation. Our view, however, is that using a conventional measure of volatility for such an evaluation is problematic because it does not distinguish between volatility at high and low frequencies. In particular, the relevant property for a measure of underlying inflation is not its overall volatility, but rather its ability to match the lower-frequency trend of inflation and to produce little high-frequency noise. Consequently, overall volatility is uninformative as a criterion because the same value can be generated from alternative configurations of volatility at high and low frequencies.

With the previous discussion serving as background, we can address the issue of smoothness of the underlying inflation measures by analyzing the nature of their volatility. As shown

Table 1

CPI and PCE Standard Deviation, Sample Period Jan 1994–Dec 2013

	CPI	CPI UIG	CPI UIG Prices Only	CPI Ex-Food and Energy	CPI Trimmed Mean	CPI Median
Standard deviation	1.12	0.85	0.31	0.53	0.57	0.64

	PCE	PCE UIG	PCE UIG Prices Only	PCE Ex-Food and Energy	PCE Trimmed Mean
Standard deviation	0.86	0.59	0.53	0.39	0.44

Sources: Bloomberg L.P.; authors' calculations.

Notes: CPI is consumer price index; PCE is personal consumption expenditures deflator.

in Table 1, the UIG (augmented by the nonprice variables) has a lower standard deviation than CPI/PCE inflation, but a higher standard deviation than the various core measures of inflation. At the same time, Chart 4 shows that the UIG is smoother—that is, has less high-frequency noise—than the various core inflation measures.[32] Thus, the ex-food and energy measure and, to a lesser extent, the trimmed mean retain more high-frequency noise, which makes it more difficult for a policymaker to determine if changes in a core inflation measure merit a policy action. Moreover, it is now evident that the higher standard deviation of the UIG reported in Table 1 is largely driven by its variability around the time of the Great Recession, which likely relates to a shift in trend inflation. Thus, this discussion should make clear the importance of judging the volatility of a measure of underlying inflation in relation to the low-frequency movements in inflation.

Second, the UIG closely tracks headline CPI/PCE inflation and is also able to provide additional information that is not incorporated in core inflation measures. Compared with the core inflation measures, the UIG displays the highest correlation with CPI inflation and PCE inflation, respectively (see Table 2, panels A and B). At the same time, the UIG is less correlated with the core inflation measures, although this finding holds more for the CPI than the PCE deflator. In both cases, however, it is evident that the UIG is providing a different signal than the traditional underlying inflation measures. This conclusion is confirmed by a simple principal components analysis (PCA) on the CPI and underlying inflation measures that include the UIG.[33] As shown by the factor loadings given in Table 3, the traditional underlying inflation measures are grouped in the first principal component, while the UIG and CPI inflation are grouped in the second principal component.

Third, although there are clear differences between the UIG for CPI inflation and the UIG for PCE inflation, the two are highly correlated with one another, as shown in Table 2, panel C. This is also true if we restrict the data set for extracting factors to prices only. Going forward, we will focus more on the CPI-based UIG to streamline the discussion and because the measure has the advantage that the CPI is subject only to very minor and infrequent revisions, whereas the PCE is subject to major revisions, especially in the non-market-based prices.[34]

4.2 Forecast Performance

One rationale for developing underlying measures of inflation is to produce more accurate forecasts of inflation than those generated using only the headline measure. For any evaluation, it is particularly important that the forecast exercise reflects a realistic setting. Following Cogley (2002) and others,

[32] This should not be surprising because we exclude short-run fluctuations in inflation from the construction of the UIG.

[33] Principal component analysis arranges variables in groups (referred to as principal components) based on their statistical behavior. This is done in a way that ensures by construction that variables with similar behavior are grouped in the same principal component, with each of the principal components uncorrelated with the others.

[34] However, both underlying inflation gauges for the CPI and for the PCE are calculated daily by the New York Fed internally.

Table 2

CPI and PCE Correlations

Panel A: CPI Correlations

	CPI UIG	CPI	CPI Ex-Food and Energy	CPI Trimmed Mean	CPI Median
CPI UIG	1.00				
CPI	0.74	1.00			
CPI Ex-food and energy	0.24	0.38	1.00		
CPI Trimmed mean	0.35	0.61	0.83	1.00	
CPI Median	0.20	0.34	0.89	0.89	1.00

Panel B: PCE Correlations

	PCE UIG	PCE	PCE Ex-Food and Energy	PCE Trimmed Mean
PCE UIG	1.00			
PCE	0.74	1.00		
PCE Ex-food and energy	0.53	0.73	1.00	
PCE Trimmed mean	0.21	0.48	0.79	1.00

Panel C: UIG Correlations, Sample Period Jan 1994–Dec 2013

	CPI UIG	CPI UIG Prices Only	PCE UIG	PCE UIG Prices Only
CPI UIG	1.00			
CPI UIG prices only	0.61	1.00		
PCE UIG	0.98	0.59	1.00	
PCE UIG prices only	0.88	0.66	0.93	1.00

Sources: Bloomberg L.P.; authors' calculations.

Notes: CPI is consumer price index; PCE is personal consumption expenditures deflator.

we initially evaluate the within-sample performance of the various measures of underlying inflation by estimating the following regression equation for monthly horizon h:

$$\pi_{t+h} - \pi_t = \alpha_h + \beta_h (\pi_t - \pi_t^m) + \varepsilon_{t+h}, \quad (8)$$

where π_t^m denotes the relevant measure of underlying inflation. Because underlying inflation is intended to measure trend inflation, the term $(\pi_t - \pi_t^m)$ can be interpreted as the transitory component of monthly inflation at time t that is expected to dissipate over time. That is, the term provides a measure of the expected reversal in current inflation.

Two desirable properties of an underlying measure of inflation are unbiasedness ($\alpha_h = 0$ and $\beta_h = -1$) and the capability to explain a substantial amount of the future variation in inflation. If β_h were negative but less than (greater than) one in absolute value, then the deviation between headline inflation and the underlying inflation measure $(\pi_t - \pi_t^m)$ would overstate (understate) the magnitude of subsequent changes in inflation, and thus would also overstate (understate) the magnitude of the current transitory deviation in inflation. This specification also nests the random walk model of Atkeson and Ohanion (2001) when $\alpha_h = \beta_h = 0$.

When equation (8) is estimated within sample, our main interest is testing for unbiasedness and whether the transitory deviation in inflation displays the correct size ($\beta_h = -1$). Using the sample period January 1993 through December 2013, we are unable to reject either hypothesis.[35] However, note that

[35] Using quarterly data from the period 1978-2004 and examining traditional underlying inflation measures, Rich and Steindel (2007) find that the property of unbiasedness can be rejected, but there is less evidence against the hypothesis that the coefficient on the deviation equals –1.

Table 3

Principal Components Analysis (PCA) on Core Inflation Measures and the UIG

	PCA1	PCA2	PCA3	PCA4	PCA5
CPI	0.40	0.53	–0.60	0.28	0.33
UIG	0.31	0.67	0.64	–0.19	–0.12
CPI Ex-food and energy	0.49	–0.32	0.30	0.74	–0.14
CPI Trimmed mean	0.52	–0.15	–0.34	–0.40	–0.65
CPI Median	0.49	–0.37	0.16	–0.42	0.65
Variance Proportion	0.65	0.26	0.06	0.03	0.01
Cumulative Proportion	0.65	0.91	0.96	0.99	1.00

Sources: Bloomberg L.P.; authors' calculations.

Note: CPI is consumer price index.

the test for unbiasedness of the UIG suffers from pre-test bias because the UIG must be centered separately from the estimation of the factors.[36]

Note of Caution for the Forecasting Exercises

We now investigate the relative forecast performance of the underlying inflation measures. It is often argued that such an exercise allows for the identification of a preferred underlying inflation measure. However, this type of comparison raises a number of issues that require careful consideration.

The most difficult issue in the interpretation of forecasting results concerns the appropriate loss function to evaluate forecast accuracy. The standard approach is to use a quadratic loss function for the forecast errors. Consider the following examples:

- Case 1: For total inflation between 1 and 3 percent, the root mean square error (RMSE) at a twelve-month horizon for underlying measure A is 1 percentage point, while for measure B it is 1.1 percentage points.

- Case 2: For total inflation outside the range of 1 to 3 percent, the RMSE at a twelve-month horizon for underlying measure A is 2 percentage points, while for measure B it is 1.2 percentage points.

[36] As mentioned in Section 3.3 and in footnote 24, the standardization of the variables requires us to assign an average value for the underlying inflation gauges for CPI inflation and PCE inflation.

If policymakers use measure A, they will be slower to recognize a change in underlying inflation than they would be if they used measure B. Suppose the policymaker successfully uses measure B to conduct monetary policy so that total inflation is rarely outside a range of 1 to 3 percent; a forecast evaluation would favor measure A if actual inflation was outside the 1 to 3 percent range less than 10 percent of the time. Therefore, forecast accuracy may not be informative about the usefulness of an underlying inflation measure for stabilization purposes.

Another important issue raised by the forecast exercise concerns the choice of the sample period. Long time periods can be problematic because they may cover different inflation regimes. Furthermore, because most industrialized countries successfully stabilized their inflation rates before the global financial crisis, static inflation forecasts (that is, a constant) might be more accurate than model-based forecasts generated from earlier periods when there was greater variability in inflation. The opposite result might hold for measures with greater variability during the global financial crisis. Therefore it is important to conduct our forecasting exercise over a sample displaying significant variation in inflation as well as over different subsample periods. The behavior of inflation in the United States since 2000 displays these features because it is relatively tranquil during the pre-2008 period but extremely volatile during the post-2008 period.

Finally, forecasting exercises are often undertaken in a pseudo-real-time manner in which estimation is conducted using a single vintage data set. In practice, the actual data used might have been revised subsequently. In our case,

the UIG (for CPI inflation) is constructed from data that is either not revised or only revised slightly (some PPI [producer price index] prices) but whose future values may lead to reassessments of the UIG's previous values—a feature not found in more traditional underlying inflation measures.[37]

A "Horse Race": UIG versus Core Inflation Measures

We first consider the results of a forecasting exercise based on equation (8).[38] Using data through period t, we can estimate the following regression equation:

$$\pi_t = \pi_{t-h} + \alpha_h + \beta_h (\pi_{t-h} - \pi^m_{t-h}) + \varepsilon_t \,. \tag{9}$$

The estimated equation can then be iterated forward by h periods to generate a forecast:

$$\hat{\pi}_{t+h} = \pi_t + \hat{\alpha}_{h|t} + \hat{\beta}_{h|t}(\pi_t - \pi^m_t), \tag{10}$$

where $\hat{\alpha}_{h|t}$ and $\hat{\beta}_{h|t}$ are the estimated regression coefficients using data through time t.

Estimation starts in 1994, while the forecasting range spans the period from 2000 through the end of 2013. To account for possible sensitivity of the forecast comparisons to this sample period, we also consider three different subsample periods: first, a pre-crisis subsample from 2000-07, a time range that could be considered a representative inflation cycle because it encompasses moderate cyclical phases in CPI inflation; second, a crisis subsample that captures the period from 2008 until the end of 2013; and third, for comparison purposes, a sample from 2001-07 that is also considered in Stock and Watson (2008). We compare the forecast performance of the UIG with the ex-food and energy, trimmed mean, and median measures. We also include a prices-only version of the UIG as well as the prior twelve-month change in the CPI in the forecast exercise.

The results in Table 4 show that the UIG clearly outperforms the traditional underlying inflation measures in forecasting headline CPI inflation before the crisis,

[37] This feature of the UIG was discussed in Section 3.3 and footnote 29.

[38] To ensure comparability we use the same setting as in Rich and Steindel (2007), which compares forecast performance of traditional core measures. The same regression model has been used in studies such as Clark (2001), Hogan, Johnson, and Laflèche (2001), Cutler (2001), and Cogley (2002).

Table 4

Out-of-Sample Performance in Root Mean Square Error for CPI

Whole Inflation Cycle: Sample Period, Jan 2000–Dec 2013

	$h = 12$
UIG	**1.35**
UIG prices only	1.54*
CPI Ex-food and energy	1.73**
CPI Trimmed mean	1.80**
CPI Median	1.81**
CPI ($t - h$)	*1.94***

Pre-Crisis: Sample Period, Jan 2000–Dec 2007

	$h = 12$
UIG	**0.93**
UIG prices only	**0.93**
CPI Ex-food and energy	*1.32***
CPI Trimmed mean	1.28**
CPI Median	1.26**
CPI ($t - h$)	1.25***

During the Crisis: Sample Period, Jan 2008–Dec 2012

	$h = 12$
UIG	**1.85**
UIG prices only	2.25**
CPI Ex-food and energy	2.32*
CPI Trimmed mean	2.56**
CPI Median	2.62**
CPI ($t - h$)	*2.88****

Stock and Watson (2008): Sample Period, Jan 2001–Dec 2007

	$h = 12$
UIG	**0.96**
UIG prices only	**0.96**
CPI Ex-food and energy	1.27*
CPI Trimmed mean	1.22**
CPI Median	1.24**
CPI ($t - h$)	*1.28****

Source: Authors' calculations.

Notes: Estimation starts in January 1994. Out-of-sample forecast exercise runs through December 2013. Text in **boldface** signifies the lowest root mean square error (RMSE). Text in *italics* signifies the highest RMSE. Diebold-Mariano test of the null hypothesis of equal RMSE against the alternative hypothesis that the RMSE of UIG is lower. Test statistics use the Newey-West covariance matrix estimator. CPI is consumer price index.

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

during the crisis, and over the whole sample range. This is evident from the lowest reported RMSE over all samples. To analyze the UIG forecast performance further, we apply the Diebold-Mariano (1995) testing procedure.[39] The results show that the forecast errors from the UIG are lower than those from the traditional underlying inflation measures, at a 5 percent statistical significance level during the crisis and mostly at a 1 percent statistical significance level before the crisis and over the whole sample.

When we focus solely on the traditional underlying inflation measures, they do not differ much in their forecasting performance, confirming the previous findings in Rich and Steindel (2007). However, there are three notable observations for the traditional underlying inflation measures. First, all underlying inflation measures outperform the use of the prior twelve-month change in total CPI—the random walk forecast—which, not surprisingly, displays the highest forecast errors among the reported measures and samples during the crisis.[40] Second, the forecasting performance of the CPI trimmed mean and CPI median are remarkably similar over all samples. Third, the forecasting performance of the popular CPI ex-food and energy measure relative to the other measures is better during the crisis than before the crisis.

What makes the forecast accuracy of the UIG superior to that of core inflation measures and the popular random walk model? One consideration is that our methodology combines cross-sectional and time-series smoothing methods to derive a measure of underlying inflation. As noted by Cristadoro et al. (2005), the application of filtering techniques within the dynamic-factor-model structure enables us to move from isolating the $\chi_{1,t}^{LR} + \chi_{1,t}^{SR}$ component in equation (5) to extracting only the $\chi_{1,t}^{LR}$ component in equation (6). Gains to forecast accuracy also seem to arise from including nonprice data in the sample. While the UIG and the prices-only version display equal forecast accuracy in two of the cases in Table 4, the UIG always achieves the lowest RMSE across each time period. Consequently, the results suggest that the combination of the large data panel and filtering techniques has the benefit of offering forecast accuracy that is either comparable to or better than forecasts based solely on prices.

[39] Diebold and Mariano (1995) propose and evaluate explicit tests of the null hypothesis of no difference in the forecast accuracy of two competing models.

[40] The random walk forecast is the current value of the variable, which would be expected to perform poorly during episodes when inflation is particularly volatile.

UIG Revisions Historically and during the Crisis Period

An important consideration in judging the results in Table 4 is that the UIG is derived using the full sample data set that incorporates the latest revised values of the nonprice components. That is, all previous monthly readings of the UIG are informed by future information. Even though equations (9) and (10) are estimated in a recursive manner, this feature of the UIG might be viewed as an advantage in the conduct of the forecast exercises. However, there would appear to be a more general question about the nature of the UIG revisions that extends beyond the significance of using the currently updated values for forecasting purposes.

There are several ways that we can try to qualify and quantify the importance of this issue. One option is that we can examine the magnitude of revisions to past UIG estimates for CPI inflation and determine if they were small. In doing so, we will consider a twenty-six-month period before the crisis from November 2005 to December 2007 and a forty-four-month period during the crisis from January 2008 to August 2011. This first phase covers a time period with economic changes that were very typical when judged on a historical basis, while the second phase covers a time period of historically large economic changes. Given the events in the most recent crisis, we think of the second subsample as a real-world stress test that provides an assessment of the maximal revision that can occur to the UIG.

We examine the daily revisions to each of the monthly UIG estimates over 240 workdays (approximately one year).[41] The results of this exercise are presented in Chart 6 for the absolute size of the change, where we plot the mean and median of the change of the UIG estimate from the x^{th} workday compared with the final estimate. We examine absolute values to ensure that large changes in one direction are not canceled out by large changes in the opposite direction. Although the CPI release for a particular month is not made available until the middle of the following month, estimation of the UIG for that month can proceed without delay.

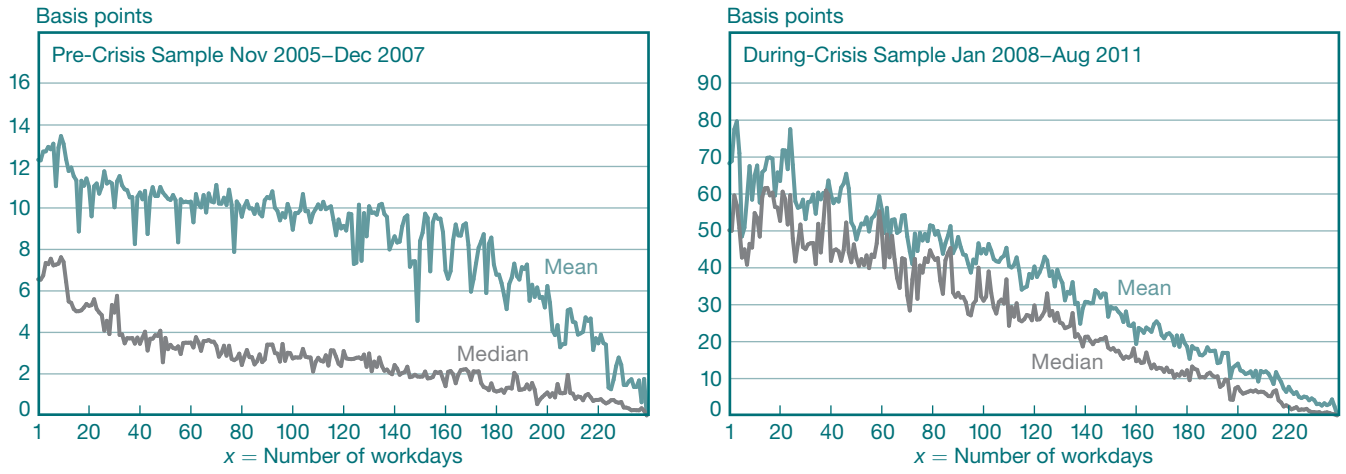
As shown, the largest changes in the estimate of the UIG for a month usually occur within the first one and a half months (thirty workdays). During a normal business cycle (November 2005 to December 2007), the maximal

[41] For the November 2005 to December 2007 sample period, we look at the revisions for each of these months for up to a year. This results in an equal number of observations for each month.

Chart 6

Absolute Change of UIG Estimates

From the x^{th} Workday Compared with the Final Estimate One Year Later



Source: Authors' calculations.

Note: One year equals 240 workdays.

median revision in the UIG peaks at about 7 basis points (0.07 percentage point) before and 4 basis points after the monthly CPI publication (Chart 6, left panel).[42] Given an average CPI inflation rate of around 2.25 percent (twelve-month change) between 1994 and 2014, these maximum changes in the UIG seem relatively minor. After the first thirty days, the median and mean revisions converge to zero.[43] Since 2008, with the large decline in CPI inflation and the deep recession in the United States, revisions in the input variables and consequently the UIG have been considerably larger. During this period of extremely volatile news flows, the maximal median revision in the UIG was around 60 basis points before and 40 basis points after the CPI publication (Chart 6, right panel).

We can also explore the issue of UIG revisions by examining the behavior of the UIG estimated in real-time using different data vintages. The upper and lower panels of Chart 7 depict the estimated UIG series on a quarterly basis from December 2005 to December 2007 and from March 2008 to December 2011, respectively. In each case, the series is estimated through the relevant end-of-month period and provides a value through the previous month.[44] The first set of data vintages again relates to the pre-crisis period, while the second set includes the crisis period. The plots also depict the realized twelve-month change in the CPI.

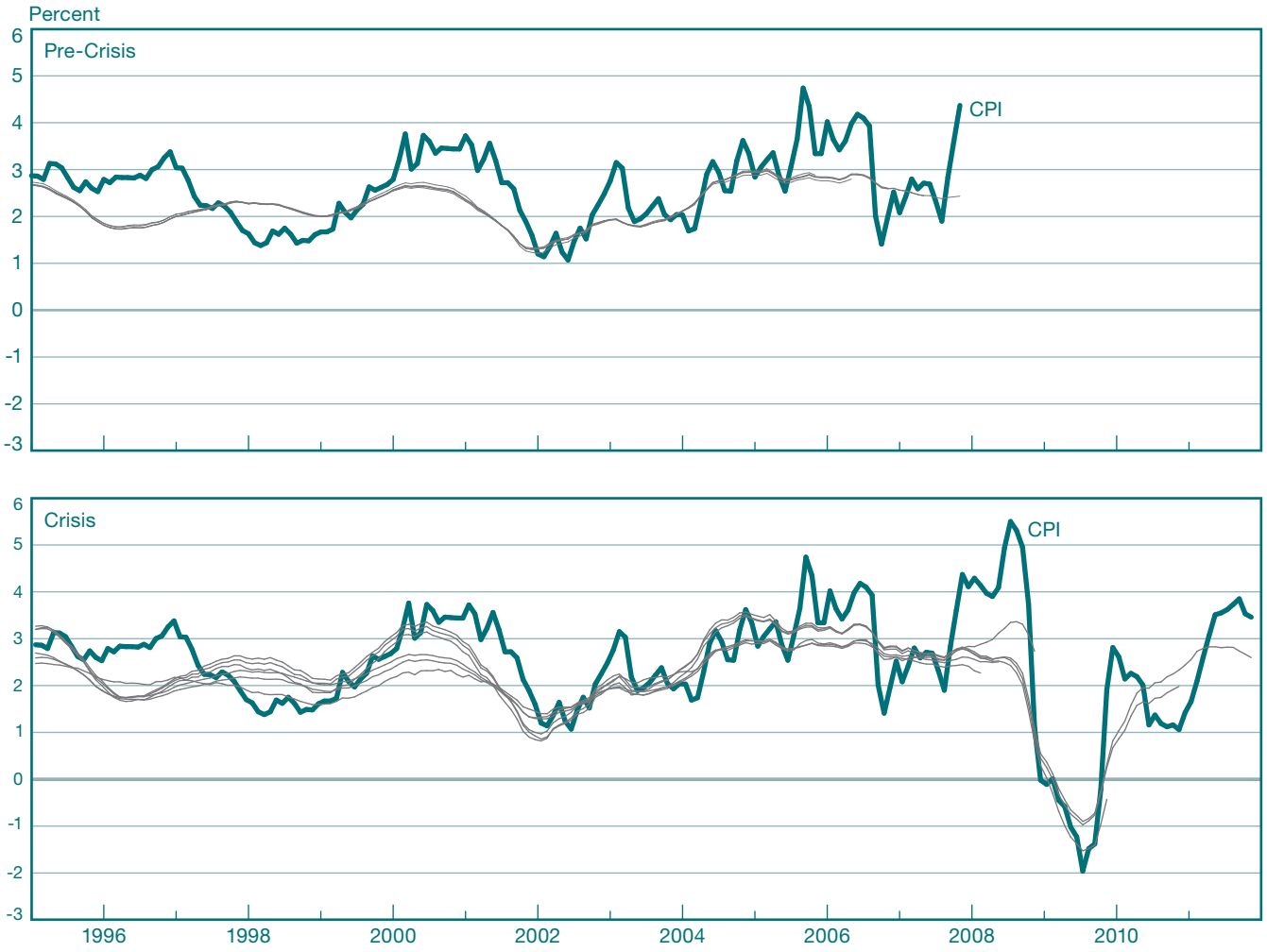
Several interesting findings emerge from the charts. As shown in Chart 7, upper panel, while the CPI inflation rate displays considerable variability, the UIG is more stable. This stability suggests that the UIG viewed the fluctuation in inflation as largely transitory. In addition, the subsequent updates do not yield significant revisions to the historical behavior of the UIG. It is also interesting to note from the lower panel that subsequent updates during the crisis period generated meaningful revisions to the UIG around turning points in inflation. However, the revisions largely exclude the Great Recession episode and focus on the level rather than the timing associated with the other turning points. This latter finding is particularly

[42] For convenience, Chart 6 is plotted in basis points—100 basis points is equivalent to 1 percentage point.

[43] Because the mean is more sensitive to outliers than the median, the slower convergence of the mean to zero likely reflects the sustained period of CPI inflation greater than 3 percent in the evaluation period—an ex ante unlikely event given our re-normalization process that centers the UIG at 2.25 percent and the volatility of the CPI from 1993 to 2005.

[44] This is because of the one-month publication lag of the CPI price series. For example, the UIG estimated using the December 30, 2005, data vintage covers the period January 1995–November 2005.

Chart 7
UIG Revisions during Pre-Crisis and Crisis Periods



Sources: Bloomberg, L.P.; authors' calculations.

Notes: Gray lines depict the estimated UIG series as measured on the following dates: Upper panel: Dec 2005, Jun 2006, Dec 2006, Jun 2007, and Dec 2007. Lower panel: Mar 2008, Jun 2008, Dec 2008, Dec 2009, Dec 2010, and Dec 2011. The UIG series is estimated through the relevant end-of-month period and provides a value through the previous month. CPI is consumer price index.

noteworthy because of the importance and difficulty of identifying turning points in the inflation process.

The preceding evidence suggests several important findings about revisions to the UIG. The revisions converge to zero fairly quickly, particularly after the first month. In addition, while revisions to the UIG have been more notable during the post-2007 period, they have not affected the dating of the turning point during the Great Recession. Rather, revisions have largely changed the level of the UIG associated with earlier turning points, not the timing of these points. Consequently, we view the UIG as providing a strong and reliable signal for an approaching change in trend inflation. Taking all of this evidence together, we consider the impact of revisions on the UIG as limited.

CPI and the Labor Market as Drivers of UIG

As a final step, we examine in more detail the changes in the estimated path of the UIG since 1995 using data through the last

Chart 8

Change in the UIG Following the Release of Various Economic Indicators

Source: Authors' calculations.

Notes: CPI is consumer price index. The shaded areas indicate periods designated recessions by the National Bureau of Economic Research.

two months of 2008 and the first month of 2009—three months during which economic activity was contracting sharply. For each month we show the path of the UIG after the release of the CPI in the two prior months and the release of the U.S. Employment Situation report for the prior month that falls between the two CPI releases. The results are presented in Chart 8. The results for November indicate little response to the CPI releases or the employment report for October 2008. In December 2008, it can be seen that the November CPI had a large effect on the current value of the UIG and the estimates for the previous twenty-four months. Lastly, the December 2008 employment report produced a large change in the UIG estimated during January 2009 and significantly altered its whole history.

5. Conclusions

This article explains the construction of the New York Fed Staff Underlying Inflation Gauge (UIG), highlights several of its attractive features and properties, compares its performance to existing measures of underlying inflation and reviews the experience of the New York Fed with real-time updates of the UIG, made internally since 2005. The article serves as useful background for the publication of monthly updates of the UIG for CPI inflation later in 2017.

Of particular note, the UIG summarizes the information content in a broad data set including asset prices and real variables such as the unemployment rate. Unlike traditional core inflation measures, the UIG does not restrict its scope to price data. Therefore it can incorporate the idea that many economic variables may affect the inflation process. The carefully chosen data set reflects the information that New York Fed staff economists consider to be the most relevant determinants of inflation.

In addition, unlike traditional underlying inflation gauges, the UIG can be updated daily. As shown in the analysis, this property is of particular importance during a crisis period, such as 2007-09. Further, the UIG adds to the literature in that it focuses on the persistent part of the common component in the broad data set. The resulting smooth movements of the UIG provide policymakers with a strong and reliable signal for an approaching turning point in trend inflation—that is, a change in underlying inflation that is likely to persist and therefore warrant a possible policy response.

The UIG is also strongly correlated with headline inflation and contains additional useful information beyond that found in traditional core measures. As a result, the

UIG can be used as a complement to, rather than as a substitute for, other core inflation measures.

Last, the UIG significantly outperforms traditional core measures when forecasting headline inflation. These findings hold for a sample from 2000 through 2013, as well as for a sample focusing on an average economic regime before the crisis and an extremely volatile sample during the crisis.

Data Appendix: UIG Variables

Prices

1. CPI-U: All items (NSA, 1982–84 = 100)
2. CPI-U: All items less energy (NSA, 1982–84 = 100)
3. CPI-U: All items less food (NSA, 1982–84 = 100)
4. CPI-U: All items less food and energy (NSA, 1982–84 = 100)
5. CPI-U: All items less medical care (NSA, 1982–84 = 100)
6. CPI-U: All items less shelter (NSA, 1982–84 = 100)
7. CPI-U: All items less food and shelter (NSA, 1982–84 = 100)
8. CPI-U: All items less food, shelter, and energy (NSA, 1982–84 = 100)
9. CPI-U: All items less food, shelter, energy, used cars and trucks (NSA, 1982–84 = 100)
10. CPI-U: Commodities (NSA, 1982–84 = 100)
11. CPI-U: Durable commodities (NSA, 1982–84 = 100)
12. CPI-U: Nondurable commodities (NSA, 1982–84 = 100)
13. CPI-U: Services (NSA, 1982–84 = 100)
14. CPI-U: Services less rent of shelter (NSA, Dec 82 = 100)
15. CPI-U: Transportation services (NSA, 1982–84 = 100)
16. CPI-U: Other services (NSA, 1982–84 = 100)
17. CPI-U: Services less medical care services (NSA, 1982–84 = 100)
18. CPI-U: Energy (NSA, 1982–84 = 100)
19. CPI-U: Apparel less footwear (NSA, 1982–84 = 100)
20. CPI-U: Energy commodities (NSA, 1982–84 = 100)
21. CPI-U: Utilities and public transportation (NSA, 1982–84 = 100)
22. CPI-U: Food and beverages (NSA, 1982–84 = 100)
23. CPI-U: Food (NSA, 1982–84 = 100)
24. CPI-U: Food at home (NSA, 1982–84 = 100)
25. CPI-U: Domestically produced farm food (NSA, 1982–84 = 100)
26. CPI-U: Cereals and bakery products (NSA, 1982–84 = 100)
27. CPI-U: Cereals and cereal products (NSA, 1982–84 = 100)
28. CPI-U: Flour and prepared flour mixes (NSA, 1982–84 = 100)
29. CPI-U: Breakfast cereal (NSA, 1982–84 = 100)
30. CPI-U: Rice, pasta, and cornmeal (NSA, 1982–84 = 100)
31. CPI-U: Bakery products (NSA, 1982–84 = 100)
32. CPI-U: White bread (NSA, 1982–84 = 100)
33. CPI-U: Bread other than white (NSA, 1982–84 = 100)
34. CPI-U: Cakes, cupcakes, and cookies (NSA, 1982–84 = 100)
35. CPI-U: Fresh cakes and cupcakes (NSA, 1982–84 = 100)
36. CPI-U: Cookies (NSA, 1982–84 = 100)
37. CPI-U: Other bakery products (NSA, 1982–84 = 100)
38. CPI-U: Fresh sweetrolls, coffeecakes, and doughnuts (NSA, 1982–84 = 100)
39. CPI-U: Crackers, bread, and cracker products (NSA, 1982–84 = 100)
40. CPI-U: Frozen and refrigerated bakery products, pies, tarts, etc. (NSA, 1982–84 = 100)
41. CPI-U: Meats, poultry, fish, and eggs (NSA, 1982–84 = 100)
42. CPI-U: Meats, poultry, and fish (NSA, 1982–84 = 100)
43. CPI-U: Meats (NSA, 1982–84 = 100)
44. CPI-U: Beef and veal (NSA, 1982–84 = 100)
45. CPI-U: Uncooked ground beef (NSA, 1982–84 = 100)
46. CPI-U: Pork (NSA, 1982–84 = 100)

Data Appendix: UIG Variables (Continued)

Prices (*continued*)

47. CPI-U: Bacon and related products (NSA, 1982–84 = 100)
48. CPI-U: Ham (NSA, 1982–84 = 100)
49. CPI-U: Ham excluding canned (NSA, 1982–84 = 100)
50. CPI-U: Pork chops (NSA, 1982–84 = 100)
51. CPI-U: Other meats (NSA, 1982–84 = 100)
52. CPI-U: Frankfurters (NSA, 1982–84 = 100)
53. CPI-U: Lamb and organ meats (NSA, 1982–84 = 100)
54. CPI-U: Poultry (NSA, 1982–84 = 100)
55. CPI-U: Fresh whole chicken (NSA, 1982–84 = 100)
56. CPI-U: Fresh and frozen chicken parts (NSA, 1982–84 = 100)
57. CPI-U: Fish and seafood (NSA, 1982–84 = 100)
58. CPI-U: Canned fish and seafood (NSA, 1982–84 = 100)
59. CPI-U: Frozen fish and seafood (NSA, 1982–84 = 100)
60. CPI-U: Eggs (NSA, 1982–84 = 100)
61. CPI-U: Dairy and related products (NSA, 1982–84 = 100)
62. CPI-U: Fresh whole milk (NSA, 1982–84 = 100)
63. CPI-U: Cheese and related products (NSA, 1982–84 = 100)
64. CPI-U: Ice cream and related products (NSA, 1982–84 = 100)
65. CPI-U: Fruits and vegetables (NSA, 1982–84 = 100)
66. CPI-U: Fresh fruits and vegetables (NSA, 1982–84 = 100)
67. CPI-U: Fresh fruits (NSA, 1982–84 = 100)
68. CPI-U: Apples (NSA, 1982–84 = 100)
69. CPI-U: Bananas (NSA, 1982–84 = 100)
70. CPI-U: Oranges, including tangerines (NSA, 1982–84 = 100)
71. CPI-U: Fresh vegetables (NSA, 1982–84 = 100)
72. CPI-U: Potatoes (NSA, 1982–84 = 100)
73. CPI-U: Lettuce (NSA, 1982–84 = 100)
74. CPI-U: Tomatoes (NSA, 1982–84 = 100)
75. CPI-U: Other fresh vegetables (NSA, 1982–84 = 100)
76. CPI-U: Frozen vegetables (NSA, 1982–84 = 100)
77. CPI-U: Nonalcoholic beverages and beverage materials (NSA, 1982–84 = 100)
78. CPI-U: Carbonated drinks (NSA, 1982–84 = 100)
79. CPI-U: Coffee (NSA, 1982–84 = 100)
80. CPI-U: Roasted coffee (NSA, 1982–84 = 100)
81. CPI-U: Instant freeze-dried coffee (NSA, 1982–84 = 100)
82. CPI-U: Other food at home (NSA, 1982–84 = 100)
83. CPI-U: Sugar and sweets (NSA, 1982–84 = 100)
84. CPI-U: Sugar and artificial sweeteners (NSA, 1982–84 = 100)
85. CPI-U: Fats and oils (NSA, 1982–84 = 100)
86. CPI-U: Butter (NSA, 1982–84 = 100)
87. CPI-U: Margarine (NSA, 1982–84 = 100)
88. CPI-U: Other foods at home (NSA, 1982–84 = 100)
89. CPI-U: Soups (NSA, 1982–84 = 100)
90. CPI-U: Frozen and freeze dried prepared food (NSA, 1982–84 = 100)
91. CPI-U: Snacks (NSA, 1982–84 = 100)
92. CPI-U: Seasonings, condiments, sauces, spices (NSA, 1982–84 = 100)

Prices (*continued*)

93. CPI-U: Other condiments (NSA, 1982–84 = 100)
94. CPI-U: Food away from home (NSA, 1982–84 = 100)
95. CPI-U: Alcoholic beverages (NSA, 1982–84 = 100)
96. CPI-U: Alcoholic beverages at home (NSA, 1982–84 = 100)
97. CPI-U: Beer, ale and malt beverages at home (NSA, 1982–84 = 100)
98. CPI-U: Distilled spirits at home (NSA, 1982–84 = 100)
99. CPI-U: Whiskey at home (NSA, 1982–84 = 100)
100. CPI-U: Distilled spirits excluding whiskey at home (NSA, 1982–84 = 100)
101. CPI-U: Wine at home (NSA, 1982–84 = 100)
102. CPI-U: Alcoholic beverages away from home (NSA, 1982–84 = 100)
103. CPI-U: Housing (NSA, 1982–84 = 100)
104. CPI-U: Shelter (NSA, 1982–84 = 100)
105. CPI-U: Rent of primary residence (NSA, 1982–84 = 100)
106. CPI-U: Rent of shelter (NSA, 1982–84 = 100)
107. CPI-U: Housing at school excluding board (NSA, Dec 82 = 100)
108. CPI-U: Other lodging away from home including hotels/motels (NSA, 1982–84 = 100)
109. CPI-U: Owners' equivalent rent of primary residence (NSA, Dec 82 = 100)
110. CPI-U: Fuels and utilities (NSA, 1982–84 = 100)
111. CPI-U: Fuels (NSA, 1982–84 = 100)
112. CPI-U: Fuel oil and other fuels (NSA, 1982–84 = 100)
113. CPI-U: Fuel oil (NSA, 1982–84 = 100)
114. CPI-U: Other [than fuel oil] household fuels (NSA, Dec 86 = 100)
115. CPI-U: Household piped gas and electricity (NSA, 1982–84 = 100)
116. CPI-U: Household electricity (NSA, 1982–84 = 100)
117. CPI-U: Utility [piped] gas service (NSA, 1982–84 = 100)
118. CPI-U: Water and sewerage maintenance (NSA, 1982–84 = 100)
119. CPI-U: Garbage and trash collection (NSA, Dec 83 = 100)
120. CPI-U: Household furnishings and operation (NSA, 1982–84 = 100)
121. CPI-U: Household furniture and bedding (NSA, 1982–84 = 100)
122. CPI-U: Bedroom furniture (NSA, 1982–84 = 100)
123. CPI-U: Household laundry equipment (NSA, 1982–84 = 100)
124. CPI-U: Clocks, lamps, and decorator items (NSA, 1982–84 = 100)
125. CPI-U: Indoor plants and flowers (NSA, Dec 90 = 100)
126. CPI-U: Housekeeping supplies (NSA, 1982–84 = 100)
127. CPI-U: Apparel (NSA, 1982–84 = 100)
128. CPI-U: Men's and boys' apparel (NSA, 1982–84 = 100)
129. CPI-U: Men's apparel (NSA, 1982–84 = 100)
130. CPI-U: Men's suits, sport coats, and outerwear (NSA, 1982–84 = 100)
131. CPI-U: Men's furnishings (NSA, 1982–84 = 100)
132. CPI-U: Men's pants and shorts (NSA, 1982–84 = 100)
133. CPI-U: Boys' apparel (NSA, 1982–84 = 100)
134. CPI-U: Women's and girls' apparel (NSA, 1982–84 = 100)
135. CPI-U: Women's apparel (NSA, 1982–84 = 100)
136. CPI-U: Women's outerwear (NSA, 1982–84 = 100)
137. CPI-U: Women's dresses (NSA, 1982–84 = 100)
138. CPI-U: Girls' apparel (NSA, 1982–84 = 100)

Data Appendix: UIG Variables (Continued)

Prices (*continued*)

139. CPI-U: Footwear (NSA, 1982–84 = 100)
140. CPI-U: Men's footwear (NSA, 1982–84 = 100)
141. CPI-U: Boys' and girls' footwear (NSA, 1982–84 = 100)
142. CPI-U: Women's footwear (NSA, 1982–84 = 100)
143. CPI-U: Infants' and toddlers' apparel (NSA, 1982–84 = 100)
144. CPI-U: Watches and jewelry (NSA, Dec 86 = 100)
145. CPI-U: Watches (NSA, Dec 86 = 100)
146. CPI-U: Jewelry (NSA, Dec 86 = 100)
147. CPI-U: Transportation (NSA, 1982–84 = 100)
148. CPI-U: Private transportation (NSA, 1982–84 = 100)
149. CPI-U: New and used vehicles (NSA, Dec 97 = 100)
150. CPI-U: New vehicles (NSA, 1982–84 = 100)
151. CPI-U: New cars (NSA, 1982–84 = 100)
152. CPI-U: New trucks (NSA, Dec 83 = 100)
153. CPI-U: Used cars and trucks (NSA, 1982–84 = 100)
154. CPI-U: Motor fuel (NSA, 1982–84 = 100)
155. CPI-U: Gasoline (NSA, 1982–84 = 100)
156. CPI-U: Unleaded regular gasoline (NSA, 1982–84 = 100)
157. CPI-U: Unleaded premium gasoline (NSA, 1982–84 = 100)
158. CPI-U: Motor vehicle parts and equipment (NSA, 1982–84 = 100)
159. CPI-U: Tires (NSA, 1982–84 = 100)
160. CPI-U: Vehicle parts and equipment excluding tires (NSA, 1982–84 = 100)
161. CPI-U: Motor oil, coolants, and fluids (NSA, 1982–84 = 100)
162. CPI-U: Motor vehicle maintenance and repair (NSA, 1982–84 = 100)
163. CPI-U: Motor vehicle body work (NSA, 1982–84 = 100)
164. CPI-U: Motor vehicle maintenance and servicing (NSA, 1982–84 = 100)
165. CPI-U: Motor vehicle insurance (NSA, 1982–84 = 100)
166. CPI-U: Public transportation (NSA, 1982–84 = 100)
167. CPI-U: Airline fare (NSA, 1982–84 = 100)
168. CPI-U: Other intercity transportation (NSA, 1982–84 = 100)
169. CPI-U: Intracity public transportation (NSA, 1982–84 = 100)
170. CPI-U: Medical care (NSA, 1982–84 = 100)
171. CPI-U: Medical care commodities (NSA, 1982–84 = 100)
172. CPI-U: Prescription drugs and medical supplies (NSA, 1982–84 = 100)
173. CPI-U: Nonprescription drugs and medical supplies (NSA, 1982–84 = 100)
174. CPI-U: Internal/respiratory over-the-counter drugs (NSA, 1982–84 = 100)
175. CPI-U: Nonprescription medical equipment and supplies (NSA, 1982–84 = 100)
176. CPI-U: Medical care services (NSA, 1982–84 = 100)
177. CPI-U: Professional medical care services (NSA, 1982–84 = 100)
178. CPI-U: Physicians' services (NSA, 1982–84 = 100)
179. CPI-U: Dental services (NSA, 1982–84 = 100)
180. CPI-U: Eyeglasses and eye care (NSA, Dec 86 = 100)
181. CPI-U: Services by other medical professionals (NSA, Dec 86 = 100)
182. CPI-U: Hospital and related services (NSA, 1982–84 = 100)
183. CPI-U: Outpatient hospital services (NSA, Dec 86 = 100)
184. CPI-U: Recreation (NSA, Dec 97 = 100)

Prices (*continued*)

185. CPI-U: Video and audio (NSA, Dec 97 = 100)
186. CPI-U: TV sets (NSA, 1982–84 = 100)
187. CPI-U: Cable and satellite TV and radio service (NSA, Dec 83 = 100)
188. CPI-U: Audio equipment (NSA, 1982–84 = 100)
189. CPI-U: Pets and pet products (NSA, 1982–84 = 100)
190. CPI-U: Sporting goods (NSA, 1982–84 = 100)
191. CPI-U: Sport vehicles including bicycles (NSA, 1982–84 = 100)
192. CPI-U: Sports equipment (NSA, 1982–84 = 100)
193. CPI-U: Photographic equipment and supplies (NSA, 1982–84 = 100)
194. CPI-U: Toys (NSA, 1982–84 = 100)
195. CPI-U: Admissions (NSA, 1982–84 = 100)
196. CPI-U: Fees for recreational lessons/instructions (NSA, Dec 86 = 100)
197. CPI-U: Recreational reading materials (NSA, 1982–84 = 100)
198. CPI-U: Education and communication (NSA, Dec 97 = 100)
199. CPI-U: Education (NSA, Dec 97 = 100)
200. CPI-U: Educational books and supplies (NSA, 1982–84 = 100)
201. CPI-U: Tuition, other school fees, and child care (NSA, 1982–84 = 100)
202. CPI-U: College tuition and fees (NSA, 1982–84 = 100)
203. CPI-U: Elementary and high school tuition and fees (NSA, 1982–84 = 100)
204. CPI-U: Child care and nursery school (NSA, Dec 90 = 100)
205. CPI-U: Communication (NSA, Dec 97 = 100)
206. CPI-U: Postage services (NSA, 1982–84 = 100)
207. CPI-U: Information and information processing (NSA, Dec 97 = 100)
208. CPI-U: Land-line telephone services, local charges (NSA,1982–84 = 100)
209. CPI-U: Land-line interstate toll calls (NSA, 1982–84 = 100)
210. CPI-U: Land-line intrastate toll calls (NSA, 1982–84 = 100)
211. CPI-U: Information technology, hardware, and services (NSA, Dec 1988 = 100)
212. CPI-U: Other goods and services (NSA, 1982–84 = 100)
213. CPI-U: Tobacco and smoking products (NSA, 1982–84 = 100)
214. CPI-U: Personal care (NSA, 1982–84 = 100)
215. CPI-U: Personal care products (NSA, 1982–84 = 100)
216. CPI-U: Cosmetics, perfumes, bath, nail preparations and implements (NSA, 1982–84 = 100)
217. CPI-U: Personal care services (NSA, 1982–84 = 100)
218. CPI-U: Miscellaneous personal services (NSA, 1982–84 = 100)
219. CPI-U: Legal services (NSA, Dec 86 = 100)
220. CPI-U: Funeral expenses (NSA, Dec 86 = 100)
221. CPI-U: Financial services (NSA, Dec 86 = 100)
222. CPI-U: Stationery, stationery supplies, gift wrap (NSA, 1982–84 = 100)
223. PPI: Finished consumer goods (NSA, 1982 = 100)
224. PPI: Finished consumer foods (NSA, 1982 = 100)
225. PPI: Finished consumer foods: Unprocessed (NSA, 1982 = 100)
226. PPI: Finished consumer foods: Processed (NSA, 1982 = 100)
227. PPI: Finished consumer goods excluding foods (NSA, 1982 = 100)
228. PPI: Consumer nondurable goods less food (NSA,1982 = 100)
229. PPI: Consumer durable goods (NSA, 1982 = 100)
230. PPI: Finished capital equipment (NSA, 1982 = 100)

Data Appendix: UIG Variables (Continued)

Prices (*continued*)

231. PPI: Capital equipment: Manufacturing industries (NSA, 1982 = 100)
232. PPI: Capital equipment: Nonmanufacturing industries (NSA, 1982 = 100)
233. PPI: Finished goods [including foods and fuel] (NSA, 1982 = 100)
234. PPI: Intermediate materials, supplies, and components (NSA, 1982 = 100)
235. PPI: Crude materials for further processing (NSA, 1982 = 100)
236. PPI: Finished goods excluding foods (NSA, 1982 = 100)
237. PPI: Offices of physicians (Dec 96 = 100)
238. PPI: Home health care services (Dec 96 = 100)
239. PPI: Commercial natural gas (NSA, Dec 90 = 100)
240. Import Price Index: All imports (NSA, 2000 = 100)
241. Export Price Index: All exports (NSA, 2000 = 100)
242. FRB Dallas: Trimmed-mean 12-month PCE inflation rate (%)
243. PCE: Chain Price Index (SA, 2000 = 100)
244. PCE less food and energy: Chain Price Index (SA, 2000 = 100)
245. PCE: Durable goods: Chain Price Index (SA, 2000 = 100)
246. PCE: Nondurable goods: Chain Price Index (SA, 2000 = 100)
247. PCE: Services: Chain Price Index (SA, 2000 = 100)
248. Real PCE: Durable goods: Motor vehicles and parts (SAAR, Mil.Chn.2000$)
249. Import Price Index: Foods, feeds and beverages (NSA, 2000 = 100)
250. Import Price Index: Industrial supplies and materials (NSA, 2000 = 100)
251. Import Price Index: Capital goods (NSA, 2000 = 100)
252. Export Price Index: Foods, feeds, and beverages (NSA, 2000 = 100)
253. Export Price Index: Industrial supplies and materials (NSA, 2000 = 100)
254. Export Price Index: Capital goods (NSA, 2000 = 100)

Real Variables

1. ISM: Mfg: New Orders Index (NSA, 50+ = Econ Expand)
2. ISM: Mfg: Production Index (NSA, 50+ = Econ Expand)
3. ISM: Mfg: Employment Index (NSA, 50+ = Econ Expand)
4. ISM: Mfg: Vendor Deliveries Index (NSA, 50+ = Econ Expand)
5. ISM: Mfg: Inventories Index (NSA, 50+ = Econ Expand)
6. ISM: Mfg: Prices Index (NSA, 50+ = Econ Expand)
7. ISM: Mfg: Backlog of Orders Index (NSA, 50+ = Econ Expand)
8. ISM: Mfg: New Export Orders Index(NSA, 50+ = Econ Expand)
9. ISM: Mfg: Imports Index (NSA, 50+ = Econ Expand)
10. ISM: Nonmfg: New Orders Index (NSA, 50+ = Econ Expand)
11. ISM: Nonmfg: Business Activity Index (NSA, 50+ = Econ Expand)
12. ISM: Nonmfg: Employment Index (NSA, 50+ = Econ Expand)
13. ISM: Nonmfg: Supplier Deliveries Index (NSA, 50+ = Econ Expand)
14. ISM: Nonmfg: Inventory Change Index (NSA, 50+ = Econ Expand)
15. ISM: Nonmfg: Prices Index (NSA, 50+ = Econ Expand)
16. ISM: Nonmfg: Orders Backlog Index (NSA, 50+ = Econ Expand)
17. ISM: Nonmfg: New Export Orders Index (NSA, 50+ = Econ Expand)
18. ISM: Nonmfg: Imports Index (NSA, 50+ = Econ Expand)

Data Appendix: UIG Variables (Continued)

Labor

1. Unemployment rate: 16–24 years (NSA, %)
2. Unemployment rate: 25–34 years (NSA, %)
3. Unemployment rate: 35–44 years (NSA, %)
4. Unemployment rate: 45–54 years (NSA, %)
5. Unemployment rate: 55 years and over (NSA, %)
6. Civilian employment-population ratio: 16–24 years (NSA, ratio)
7. Civilian employment-population ratio: 25–34 years (NSA, ratio)
8. Civilian employment-population ratio: 35–44 years (NSA, ratio)
9. Civilian employment-population ratio: 45–54 years (NSA, ratio)
10. Civilian employment-population ratio: 55 years and over (NSA, ratio)
11. Average weeks unemployed: 16–19 years (NSA)
12. Average weeks unemployed: 20–24 years (NSA)
13. Average weeks unemployed: 25–34 years (NSA)
14. Average weeks unemployed: 35–44 years (NSA)
15. Average weeks unemployed: 45–54 years (NSA)
16. Average weeks unemployed: 55–64 years (NSA)
17. Average weeks unemployed: 65 years and over (NSA)
18. Unemployment (NSA, thousands)
19. Number unemployed for less than 5 weeks (NSA, thousands)
20. Number unemployed for 5–14 weeks (NSA, thousands)
21. Number unemployed for 15–26 weeks (NSA, thousands)
22. Number unemployed for 15 weeks and over (NSA, thousands)
23. Unemployment insurance: Initial claims (Number, NSA)

Money

1. Money stock: M1 (NSA, billions $)
2. Money stock: M2 (NSA, billions $)
3. Adjusted monetary base (NSA, millions $)
4. Adjusted reserves of depository institutions (NSA, millions $)
5. Adjusted nonborrowed reserves of depository institutions (NSA, millions $)

Financials

1. Cash price: gold bullion, London commodity price, PM Fix (US$/troy oz)
2. Gold: London PM Fix (US$/troy oz)
3. Gold spot ($/oz) NSA
4. Spot commodity price—West Texas Intermediate crude oil, Cushing OK
5. Federal funds effective rate
6. 3-month Treasury bill rate coupon equivalent
7. 6-month Treasury bill rate coupon equivalent
8. 1-year Treasury bill yield at constant maturity (% p.a.)
9. 5-year Treasury note yield at constant maturity (% p.a.)
10. 7-year Treasury note yield at constant maturity (% p.a.)
11. 10-year Treasury note yield at constant maturity (% p.a.)
12. LIBOR Eurodollar 11 A.M. Fixing 1 month
13. LIBOR Eurodollar 11 A.M. Fixing 3 month
14. LIBOR Eurodollar 11 A.M. Fixing 6 month

Data Appendix: UIG Variables (Continued)

Financials (*continued*)

15. LIBOR Eurodollar 11 A.M. Fixing 9 month
16. LIBOR Eurodollar 11 A.M. Fixing 1 year
17. Spot price (euro/$) (Revised backwards)
18. Spot price (GBP/$)
19. Spot price (yen/$)
20. Spot Price (Swiss franc/$)
21. Board Narrow Nominal Effective Exchange Rate Index: United States (2000 = 100)
22. Board Broad Nominal Effective Exchange Rate: United States (2000 = 100)
23. Bank credit: all commercial banks (NSA, billions $)
24. Total revolving U.S. consumer credit outstanding
25. Total non-revolving U.S. consumer credit outstanding
26. Securities in bank credit: all commercial banks (NSA, billions $)
27. U.S. government securities in bank credit: all commercial banks (NSA, billions $)
28. Real estate loans in bank credit: all commercial banks (NSA, billions $)
29. Commercial and Industrial loans in bank credit: All commercial banks (NSA, billions $)
30. Consumer loans in bank credit: All commercial banks (NSA, billions $)
31. Moody's seasoned Aaa corporate bond yield (% p.a.)
32. Moody's seasoned Baa corporate bond yield (% p.a.)
33. Merrill Lynch High Yield Master II yield
34. New York Stock Exchange Composite Index
35. New York Stock Exchange total volume
36. Standard and Poor's 500 Price Earnings Ratio Index
37. Dow Jones Industrial Average
38. Dow Jones Wilshire 5000 Composite Index Full Cap
39. Light Sweet Crude Oil Futures Price: 1st exp contract nearby settlement (EOP, $/bbl)
40. Light Sweet Crude Oil Futures Price: 3 month contract settlement (EOP, $/bbl)
41. Light Sweet Crude Oil Futures Price: 6 month contract settlement (EOP, $/bbl)
42. No 2 Heating Oil Futures Price: 1st exp contract nearby settlement (EOP, $/gal)
43. No 2 Heating Oil Futures Price: 3 month contract settlement (EOP, $/gal)
44. No 2 Heating Oil Futures Price: 6 month contract settlement (EOP, $/gal)
45. Unleaded gasoline futures price: 1st exp contract nearby settlement (EOP, $/gal)
46. Unleaded Gasoline Futures Price: 3 month contract settlement (EOP, $/gal)
47. New York Harbor Conventional Gasoline Regular Spot Price FOB (EOP cents/gal)
48. Gas Oil Futures Price: 1st exp contract nearby settlement (EOP, $/metric ton)
49. Unleaded Premium Gasoline Price, NY gal (EOP, $/gal)
50. Unleaded Gas, Regular, Non-Oxygenated: NY (EOP, $/gal)
51. Natural Gas Price, Henry Hub, LA ($/mmbtu)
52. Dow Jones AIG Futures Price Index (Jan 2, 1991 = 100)
53. Dow Jones AIG Spot Price Index (Jan 7, 1991 = 100)
54. FIBER Industrial Materials Index: All Items (1990 = 100)
55. Goldman Sachs Commodity Nearby Index (EOP, Dec 31, 1969 = 100)
56. S&P 500 Futures Price: 1st exp contract nearby settlement (EOP, Index)
57. S&P 400 Midcap Futures Price: 1st exp contract nearby settlement (EOP, Index)

Editor's note:
This data appendix has been updated to reflect the removal of a duplicate price series (CPI-U: Other fresh vegetables). The article's conclusions remain the same.

References

Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Lippi, L. Reichlin, and G. Veronese. 2001. "EuroCOIN: A Real-Time Coincident Indicator for the Euro Area Business Cycle." Centre for Economic Policy Research Discussion Paper, no. 3108.

Altissimo, F., B. Mojon, and P. Zaffaroni. 2009. "Can Aggregation Explain the Persistence of Inflation?" Journal of Monetary Economics 56, no. 2 (March): 231-41.

Amstad, M., and S. Potter. 2009. "Real-Time Underlying Inflation Gauges for Monetary Policymakers." Federal Reserve Bank of New York Staff Reports, no. 420, December.

Amstad, M., and A. M. Fischer. 2009a. "Are Weekly Inflation Forecasts Informative?" Oxford Bulletin of Economics and Statistics 71, no. 2 (April): 237-52.

———. 2009b. "Do Macroeconomic Announcements Move Inflation Forecasts?" Federal Reserve Bank of St. Louis Review 91, no. 5, part 2 (September-October): 507-18.

———. 2010. "Monthly Pass-Through Ratios." Journal of Economic Dynamics and Control 34, no. 7 (July): 1202-13.

Amstad, M., Y. Huan, and G. Ma. 2014. "Developing an Underlying Inflation Gauge for China." BIS Working Papers, no. 465.

Atkeson, A., and L. E. Ohanian. 2001 "Are Phillips Curves Useful for Forecasting Inflation?" Federal Reserve Bank of Minneapolis Quarterly Review 25, no. 1 (Winter): 2-11.

Bernanke, B. S., and Boivin, J. 2003. "Monetary Policy in a Data-Rich Environment." Journal of Monetary Economics 50, no. 3 (April): 525-46.

Boivin, J., and S. Ng. 2006. "Are More Data Always Better for Factor Analysis?" Journal of Econometrics 132, no. 1 (May): 169-94.

Brillinger, D. R. 1981. Time Series: Data Analysis and Theory. San Francisco: Holden-Day.

Bryan, M. F., and C. J. Pike. 1991. "Median Price Changes: An Alternative Approach to Measuring Current Monetary Inflation." Federal Reserve Bank of Cleveland Economic Commentary, December 1.

Bryan, M. F., and S. G. Cecchetti. 1994. "Measuring Core Inflation." In N. G. Mankiw, ed., Monetary Policy. Chicago: University of Chicago Press, 195-215.

———. 1999. "Inflation and the Distribution of Price Changes." The Review of Economics and Statistics 81, no. 2 (May): 188–96.

Bryan, M. F., S. G. Cecchetti, and R. L. Wiggins II. 1997. "Efficient Inflation Estimation." NBER Working Paper no. 6183, September.

Bullard, J. 2011. "Measuring Inflation: The Core is Rotten." Federal Reserve Bank of St. Louis Review 93, no. 4 (July-August): 223-33.

Cecchetti, S. G. 1995. "Inflation Indicators and Inflation Policy." In B. S. Bernanke and J.J. Rotemberg, eds., NBER Macroeconomics Annual 1995, Volume 10. Cambridge: MIT Press, 189-236.

———. 1997. "Measuring Short-Run Inflation for Central Bankers." Federal Reserve Bank of St. Louis Review (May-June): 143-55.

Cecchetti, S. G. and R. Moessner. 2008. "Commodity Prices and Inflation Dynamics." BIS Quarterly Review (December): 55-66.

Clark, T. E. 2001. "Comparing Measures of Core Inflation." Federal Reserve Bank of Kansas City Economic Review 86, no. 2 (Second Quarter): 5-31.

———. 2004. "An Evaluation of the Decline in Goods Inflation." Federal Reserve Bank of Kansas City Economic Review 89, no. 2 (Second Quarter): 19-51.

Cogley, T. 2002. "A Simple Adaptive Measure of Core Inflation." Journal of Money, Credit, and Banking 34, no. 1 (February): 94-113.

Cristadoro, R., M. Forni, L. Reichlin, and G. Veronese. 2005. "A Core Inflation Indicator for the Euro Area." Journal of Money, Credit, and Banking 37, no. 3 (June): 539-60.

Crone, T. M., N. N. K. Khettry, L. J. Mester, and J. A. Novak. 2013. "Core Measures of Inflation as Predictors of Total Inflation." Journal of Money, Credit, and Banking 45, no. 2-3 (March-April): 505-19.

References (Continued)

Cutler, J. 2001. "Core Inflation in the U.K." Monetary Policy Committee Unit, Bank of England, Discussion Papers, no. 3 (March).

Diebold, F. X. and R. S. Mariano. 1995. "Comparing Predictive Accuracy." Journal of Business and Economic Statistics 13, no. 3 (July): 253-63.

Dolmas, J. 2005. "Trimmed Mean PCE Inflation." Federal Reserve Bank of Dallas Working Paper, no. 0506.

Eckstein, O. 1981. Core Inflation. Englewood Cliffs, N.J.: Prentice Hall.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin. 2000. "The Generalized Dynamic-Factor Model: Identification and Estimation." The Review of Economics and Statistics 82, no. 4 (November): 540-54.

———. 2001. "Coincident and Leading Indicators for the Euro Area." The Economic Journal 111, no. 471 (May): C62-C85.

———. 2005. "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting." Journal of the American Statistical Association 100, no. 471 (September): 830-40.

Galí, J. 2002. "New Perspectives on Monetary Policy, Inflation, and the Business Cycle." Centre for Economic Policy Research Discussion Paper Series, no. 3210 (February).

Gavin, W. T., and K. L. Kliesen. 2008. "Forecasting Inflation and Output: Comparing Data-Rich Models with Simple Rules." Federal Reserve Bank of St. Louis Review 90, no. 3 pt. 1 (May-June): 175-92.

Giannone, D., and T. D. Matheson. 2007. "A New Core Inflation Indicator for New Zealand." International Journal of Central Banking 3, no. 4 (December): 145-80.

Giannone, D., L. Reichlin, and L. Sala. 2005. "Monetary Policy in Real Time." In M. Gertler and K. Rogoff, eds., NBER Macroeconomics Annual 2004, Volume 19. Cambridge: MIT Press, 161-224.

Gordon, R. J. 1975. "Alternative Responses of Policy to External Supply Shocks." Brookings Papers on Economic Activity 6, no. 1: 183-206.

Gordon, R. J. 1982. "Price Inertia and Policy Ineffectiveness in the United States, 1890-1980, Journal of Political Economy 90, no. 6 (December): 1087-117.

Hogan, S., M. Johnson, and T. Laflèche. 2001. "Core Inflation." Bank of Canada Technical Report No. 89, January.

Khan, M., L. Morel, and P. Sabourin. 2013. "The Common Component of CPI: An Alternative Measure of Underlying Inflation for Canada." Bank of Canada Working Paper 2013-35, October.

Liu, Z., and G. Rudebusch. 2010. "Inflation: Mind the Gap." Federal Reserve Bank of San Francisco FRBSF Economic Letter, no. 2010-02 (January 19).

Meyer, B., G. Venkatu, and S. Zaman. 2013. "Forecasting Inflation? Target the Middle." Federal Reserve Bank of Cleveland Economic Commentary, April.

Nelson, C. R. and G. W. Schwert. 1977. "Short-Term Interest Rates as Predictors of Inflation: On Testing the Hypothesis That the Real Rate of Interest Is Constant," American Economic Review 67, no. 3 (June): 478-86.

Organisation for Economic Co-operation and Development. 2005. "Measuring and Assessing Underlying Inflation." OECD Economic Outlook 2005, no. 1 (June).

People's Bank of China. 2016. "An Underlying Inflation Gauge (UIG) for China." In Inflation Mechanisms, Expectations, and Monetary Policy, BIS Papers, no. 89 (November): 117-21.

Quah, D., and S. P. Vahey. 1995. "Measuring Core Inflation." The Economic Journal 105, no. 432: (September): 1130-44.

Reis, R., and M. W. Watson. 2010. "Relative Goods' Prices, Pure Inflation, and the Phillips Correlation." American Economic Journal: Macroeconomics 2, no. 3 (July): 128-57.

Rich, R., and C. Steindel. 2007. "A Comparison of Measures of Core Inflation." Federal Reserve Bank of New York Economic Policy Review 13, no. 3 (December): 19-38.

Rudd, J., and K. Whelan. 2007. "Modeling Inflation Dynamics: A Critical Review of Recent Research." Journal of Money, Credit, and Banking 39, no. s1 (February): 155-70.

References (Continued)

Sargent, T. J., and C. A. Sims. 1977. "Business Cycle Modeling without Pretending to Have Too Much *A Priori* Economic Theory." In New Methods in Business Cycle Research: Proceedings from a Conference, 45-109. Federal Reserve Bank of Minneapolis.

Stock, J. H., and M. W. Watson. 1999. "Forecasting Inflation." Journal of Monetary Economics 44, no. 2 (October): 293-335.

———. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." Journal of the American Statistical Association 97, no. 460 (December): 1167-79.

———. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" Journal of Money, Credit, and Banking 39, no. s1 (February): 3-33.

———. 2008. "Phillips Curve Inflation Forecasts." In Understanding Inflation and the Consequences for Monetary Policy: A Phillips Curve Retrospective, Federal Reserve Bank of Boston Conference Series, no. 53.

———. 2010. "Modeling Inflation after the Crisis." NBER Working Paper no. 16488, October.

———. 2016. "Core Inflation and Trend Inflation." The Review of Economics and Statistics 98, no. 4 (October): 770-84.

Vega, J. L., and M. A. Wynne. 2001. "An Evaluation of Some Measures of Core Inflation for the Euro Area." European Central Bank Working Paper Series, no. 53 (April).

Watson, M. W. 2004. "Comment" on Giannone, Reichlin, and Sala's "Monetary Policy in Real-Time" In M. Gertler and K. Rogoff, eds., NBER Macroeconomics Annual 2004, Volume 19. Cambridge: MIT Press, 216-221.

Wynne, M. A. 2008. "Core Inflation: A Review of Some Conceptual Issues." Federal Reserve Bank of St. Louis Review 90, no. 3, pt. 2 (May-June): 205-28.

Jeffrey F. Ingber

The Development of the Government Securities Clearing Corporation

- Despite its global importance, the U.S. government securities market was late in adopting centralized clearance and settlement services.
- The path toward provision of such services—and the accompanying boost in market efficiency and reduction in risk—began with the 1986 launch of the Government Securities Clearing Corporation (GSCC), now a part of the Depository Trust & Clearing Corporation.
- This history of the formation and development of the GSCC describes the state of the market in the 1980s; the establishment of the GSCC and its adoption of an automated comparison and netting system; the expansion of the system to include Treasury auction awards, and later, repos and reverse repos; the addition of services for brokered repos; and the launch of the General Collateral Finance Repo service (GCF Repo®).

1. Introduction

The U.S. government securities market is one of the largest and most liquid securities markets in the world, and arguably the single most important financial marketplace. Yields on Treasury securities are benchmarks for other interest rates globally. U.S. government repurchase agreements represent the most important short-term credit market in the country. The U.S. government securities market overall is the market others look to for safety and risk mitigation, particularly in times of trouble. And the market provides the federal government the ability to conduct monetary policy and, even more essentially, to fund itself.

Surprisingly, the government securities market was one of the last major securities markets to receive the benefits of centralized clearance and settlement services. This article reviews the development of such services through the formation of the Government Securities Clearing Corporation (GSCC) in 1986. The GSCC—which is now the Government Securities Division of the Fixed Income Clearing Corporation, a subsidiary of the Depository Trust & Clearing Corporation—is considered by many to be the largest and most significant clearing corporation in the world. (For a description of the processes involved in clearance and settlement services, see Box 1 on the next page.)

Jeffrey F. Ingber is a former managing director of the Depository Trust & Clearing Corporation (DTCC) and general counsel of the Fixed Income Clearing Corporation (earlier the Government Securities Clearing Corporation), a subsidiary of DTCC. At the time this article was written, he was a senior vice president in the Federal Reserve Bank of New York's Supervision Group.

jingber@gmail.com

The author thanks David Buckmaster and Andrew Weg for their assistance with this article. The views expressed are those of the author and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

To view the author's disclosure statement, visit https://www.newyorkfed.org/research/author_disclosure/ad_epr_2017_gscc_ingber.html.

Box 1
Clearing and Settlement Functions

The process of clearing and settling trades includes three main functions.

- *Comparison*: The process of matching the terms of each side of a transaction to identify differences in reported trades. The ability to correct or resolve those differences is usually attendant on a comparison service.
- *Clearance*: The process of preparing compared trades for settlement. This preparation can take several forms, ranging from the most basic (producing individual receive and deliver instructions for each matched trade) to the most sophisticated (netting all deliver and receive obligations in each security on a continuous basis).
- *Settlement*: The actual exchange of securities and funds.

We begin our look at the history of this corporation by describing the government securities market in the 1980s and the events that led to the formation of the GSCC, including the impact of a market scandal. We then show how the adoption of a centralized, automated system of comparing the buy and sell sides of securities transactions increased efficiency and reduced risk in GSCC operations. Next, we outline the effect of process enhancements and broadened access on the growth of the netting system, and how the addition of proprietary Treasury auction awards to the system further increased efficiency and reduced risk for the GSCC, its members, and the U.S. Treasury. We also explore the extension of comparison and netting services to repurchases (repos) and reverse repurchases (reverse repos) of government securities, which provided the repo market with cost and efficiency benefits similar to those provided for buy-sell trades. Finally, we review the introduction of netting and settlement services for brokered repos—which lowered brokers' costs and eliminated counterparty risk—and the launch of the General Collateral Finance Repo service (GCF Repo®) for Treasury securities collateral.

2. The Government Securities Marketplace in the 1980s

Before the 1980s, the government securities market was thought to be transparent, efficient, and safe, even though it was essentially unregulated.[1] During the first half of that decade, however, a number of government securities dealers failed, including Bevill, Bresler, and Schulman Asset Management Corporation, Drysdale Government Securities, E.S.M. Government Securities Inc., and Lombard-Wall Inc. These failures led to congressional hearings in 1985 on the safety of the marketplace, and ultimately, to the signing by President Reagan of the Government Securities Act of 1986 (GSA), which, among other things, provided for the registration of government securities brokers and dealers. The enactment of the GSA, together with the Treasury Department's move to end issuance of certificated debt in 1986 (by then, all new Treasury securities were issued in book-entry form), prompted the establishment of GSCC.

The enactment of the GSA laid an important legal foundation for the formation of GSCC in at least two respects. First, the GSA amended the definition of "exempted security" in the Securities Exchange Act of 1934 (SEA) to include a new section stating that, "government securities shall not be deemed to be 'exempted securities' for the purposes of section 17A."[2] As a result, the GSA required the registration of any entity seeking to act as a clearing agency for government securities and granted the U.S. Securities and Exchange Commission (SEC) jurisdiction over such clearing agency activity. The motivation for this measure was to ensure appropriate supervisory oversight of entities performing clearance and settlement functions for the government securities marketplace, and to encourage the development of a clearing agency akin to those existing for other markets, such as the National Securities Clearing Corporation for corporate equities and municipal debt securities.

Another important legal foundation for the successful operation of a government securities clearing agency was Congress' decision through the GSA to make many of the prospective participants in such a clearing agency, such as government securities brokers, government securities dealers, and depository institutions, subject to federal regulation.

[1] The Securities Exchange Act of 1934 (SEA) specifically exempted U.S. government securities from its key provisions.

[2] Section 17A of the SEA (National System for Clearance and Settlement of Securities Transactions), among other things, directs the Securities and Exchange Commission to use its authority to facilitate the establishment of a national system for the prompt and accurate clearance and settlement of transactions in securities other than exempted securities.

This broad extension of federal regulation greatly facilitated the ability of any prospective government securities clearing agency to build and maintain the comprehensive risk management systems that would be essential for the agency's successful operation.

The idea for GSCC arose in the mid-1980s as government securities trading volume increased and the Federal Reserve, the Public Securities Association (PSA),[3] and several large primary dealers became concerned about the safety and soundness of the existing processes for clearing and settling government securities.[4] Their concerns included the risks associated with the failure of a major firm, the inefficiencies of manual paper processing of trade confirmations, and bilateral trade-for-trade settlement.

The Fed's concerns also included the large and increasing levels of intraday credit extensions, or "daylight overdrafts," on Fedwire (the Federal Reserve's system for transferring money and securities between banks and certain other financial institutions), as well as the frequent delays in the closing of Fedwire and the bunching of deliveries within peak "traffic periods."[5] It had become standard practice at many government dealer firms to hold large deliveries for which the firms had only part of the required securities (for example, $150,000,000 to make a $200,000,000 delivery) and at least some smaller deliveries until five minutes before the scheduled close of Fedwire. The Federal Reserve Bank of New York would announce at least five minutes prior to the scheduled close whether it was going to extend the wire. If the wire was not extended, the firms would try to minimize their failures to deliver either by borrowing securities to make their larger deliveries or by making their smaller deliveries. This practice contributed to late-day congestion on Fedwire as dealers made last-minute deliveries.[6]

The Fed also had long-standing concerns, dating back at least to the failure of Drysdale Government Securities in May 1982, about the risks arising from government securities trading.[7] One concern in particular was the potential for insolvency of a major firm and the consequences for the marketplace if that occurred. What was needed was a central guarantor. Not wanting that role, the Fed asked the private sector to devise a solution: a clearing corporation that risk-managed and guaranteed the settlement of government securities trades. The Fed was also concerned that the bilateral netting relied on by dealers and interdealer brokers did not have a solid legal underpinning and might unravel if a firm became insolvent.

3. Formation of GSCC

In May 1986, at the suggestion of several primary dealers, the board of directors of the National Securities Clearing Corporation (NSCC) established a Government Securities Committee to consider applying NSCC's expertise in automated comparison and netting to the government securities

[3] In 1997, the name of the PSA was changed to the Bond Markets Association (TBMA). In 2006, TBMA merged with the Securities Industry Association to form the Securities Industry Financial Markets Association (SIFMA).

[4] "In recent years tremendous investor losses have occurred in the government securities market due to dealer failures. Although the . . . SEC . . . generally has regulatory authority over broker-dealers in corporate and municipal securities, and the [Federal Reserve] has regulatory authority over dealer banks, dealers who trade only in government securities have operated outside the federal system of financial supervision. Most of these dealer failures occurred among dealers operating outside the federal regulatory structure. Alarmed by these losses, Congress enacted the [GSA], placing the government securities market under complete federal regulation. . . . Many of the failed dealers operated outside the federal regulatory structure because they dealt solely in exempt government securities. As a result of these failures, many savings and loans, municipalities, and other public institutions lost millions of dollars. In response to these dealer failures, Congress sought to provide for a formal system of regulation of government securities dealers and brokers by enacting the [GSA]." Joseph G. Fallon, "The Government Securities Act of 1986: Balancing Investor Protection with Market Liquidity," *Catholic University Law Review* 36, no. 4 (Summer 1987) [citing from the Act's legislative history].

[5] On December 12, 1986, at a PSA conference on regulation of the U.S. government securities markets, Cathy Minehan, vice president of the Electronic Payments Function at the Federal Reserve Bank of New York (later to become the president of the Federal Reserve Bank of Boston), reported that the settlement of Treasury securities was a significant portion of total transfers on Fedwire, and that, in 1986, the average closing time of Fedwire was 4:30 P.M., two hours later than the established closing time.

[6] In January 1988, in an attempt to reduce risk and daylight overdrafts and to even out the flow of traffic on Fedwire, the Fed instituted a maximum par value limit of $50 million per government securities transaction. However, this measure did not significantly alleviate the problem and market participants became concerned that the Fed would further react by imposing strict debit cap requirements.

[7] The failure of Drysdale had enormous implications for the marketplace. Prior to that, it was common practice in the repo market to ignore the value of accrued interest in pricing repos using coupon-bearing securities. This practice enabled Drysdale to acquire a substantial amount of undervalued securities, despite its limited capital base. Drysdale used the securities that it had reversed in to settle short sales for an amount that included the accrued interest. Using the surplus cash generated, Drysdale was able to raise working capital and make interest payments to its repo counterparties. This strategy worked until May 17, 1982, when cumulative losses on Drysdale's interest rate bets caused it to be unable to pay the coupon interest on securities it had borrowed. As a result of the weaknesses exposed by the Drysdale matter, full accrual pricing, in which accrued interest was included in the initial purchase and resale prices, was adopted as standard repo market practice. See Stephen A. Lumpkin, "Repurchase and Reverse Repurchase Agreements," Federal Reserve Bank of Richmond, *Monograph*, no. 1998rarr, 1998.

market.[8] In doing so, the NSCC board noted the fact that many of the concerns expressed about the government securities marketplace were not dissimilar from those faced and resolved by NSCC in the corporate and municipal securities markets. Subsequently, the NSCC board created a more broadly representative Ad Hoc Committee on Clearance of Treasury Securities, which included representatives from each of the major industry groups in the government securities market—primary dealers, interdealer brokers, and clearing banks.[9]

In September 1986, in contemplation of the passage of the GSA, the NSCC board approved the establishment of Government Securities Clearing Corporation as a wholly owned subsidiary of NSCC, capitalized with $1 million. On October 28, 1986, the Government Securities Act was signed into law. Three weeks later, on November 18, GSCC was incorporated under the New York Business Corporation Law. At its December 11, 1986, meeting, the NSCC board selected the first board of directors and officers of GSCC.[10]

In December 1987, a private placement of GSCC shares began. About 81 percent of GSCC's shares were sold to forty-four participant firms, including a majority but not all of the primary dealers. (The rest of the shares were retained by NSCC.) By May of the following year, $10.4 million had been raised.[11] Also that May, the SEC granted GSCC temporary registration as a clearing agency.[12] On August 31, 1988, the first participant shareholder board of directors of GSCC was elected. The board was made up of representatives from primary dealers, interdealer brokers, and clearing banks, plus a management director (GSCC's president) and two directors designated by NSCC.[13]

4. Launch of the Comparison System

GSCC's operations began on August 26, 1988, with the implementation of its Comparison System, which provided for the reporting, validating, and matching of the buy and sell sides of securities transactions. GSCC began to match, in an automated fashion, the next-day and future-settling Treasury and agency trades of thirty primary dealers and interdealer brokers. The comparison of trade data was deemed to have occurred when GSCC made a report of the comparison of such trades available to its members.[14] By a rule filing

[8] NSCC, an SEC-registered clearing agency, was formed in 1977 as a result of the merger of the clearing corporations of the American Stock Exchange, the New York Stock Exchange, and the National Association of Securities Dealers. Now a subsidiary of the DTCC, NSCC was the nation's largest provider of post-trade processing, clearance, and settlement services for equity, corporate and municipal securities, unit investment trusts, and mutual fund transactions to the broker-dealer, bank, and mutual fund communities.

[9] William Tierney of Salomon Brothers Inc. was chairman of the Ad Hoc Committee.

[10] The members of the first GSCC board were taken from the Ad Hoc Committee.

[11] On March 14, 1988, the Legal Advisory Services Division of the Office of the Comptroller of the Currency issued a letter holding that the proposed acquisition of GSCC shares by a national bank was permissible. On April 18, 1988, the Board of Governors of the Federal Reserve System ruled that bank holding companies could invest in the voting shares of GSCC without filing an application under the Bank Holding Company Act, provided that no bank holding company acquired more than 5 percent of such shares. On June 8, 1988, the New York State Banking Department issued a letter approving the application filed on behalf of several state-chartered banks requesting the authorization to invest up to $250,000 in the capital stock of GSCC.

[12] The SEC took note of the fundamental change in the government securities market regulatory environment in its May 24, 1988, order granting GSCC temporary registration as a clearing agency (Release No. 34-25740). The SEC observed that, pursuant to the GSA, all government securities brokers and dealers were subject to registration, examination, and financial regulatory requirements. In footnote 21 of the order, the SEC linked this regulatory expansion to the newly required registration of government securities clearing agencies as follows:

> The [GSA], among other things, authorizes and directs the Secretary of the Treasury to issue financial responsibility, recordkeeping, and financial reporting and audit rules. The Secretary also must regulate the possession and control of customer securities and funds. The law requires clearing agencies that provide centralized clearance and settlement services in Government Securities to register with the Commission under Section 17A of the [Exchange] Act and requires dealers and brokers that were previously unregulated to register with the Commission and to join either an exchange or a registered securities association.

[13] The initial directors were Jorge Brathwaite of the Bank of New York; Allen B. Clark of Manufacturers Hanover Trust Company; Frank Cuoco of Garban Ltd.; Herbert Friedman of Salomon Brothers; Peter Gall of Discount Corporation; Edward Geng of Fundamental Brokers; David Kelly (president of NSCC and the first NSCC-designated director); Bruce Lakefield of Lehman Brothers; Charles Moran (president of GSCC); Alexander Neamtu of Morgan Stanley; Howard Shallcross of Merrill Lynch; and Ronald Upton of Irving Trust Company. The second NSCC-designated director—Andrew Threadgold of JPMorgan Securities Inc.—was named at the December 7, 1988, board meeting. At the January 1989 GSCC board meeting, Mr. Lakefield was elected chairman of GSCC, a position he held until 1994. By January 1989, senior management consisted of Mr. Moran, Thomas Costa (chief operating officer), and Jeffrey Ingber (general counsel).

[14] In 2000, when GSCC moved to a real-time trade-matching environment, the time of comparison effectively was moved from end of day to minutes after the execution of a trade.

approved by the SEC on February 22, 1989, GSCC comparison output was established as constituting the final and binding evidence of a correctly matched trade.

The Comparison System supported CPU-to-CPU (computer-to-computer) transmission and machine-readable input and output. While the system was built to support real-time interactive comparison, these capabilities would not be utilized for more than a decade.

The Comparison System, as well as GSCC's subsequent Netting System, was developed by SPC Software Services, a subsidiary of Security Pacific. The software, which was based on the "SPEED" system used by Security Pacific National Trust Company (SPNTCO) to clear government securities on a book-entry basis, was dubbed "IONS," for industry-owned netting system. (The term "netting" was used in anticipation of the software being used in the future to provide netting services.) The Securities Industry Automation Corporation (SIAC) was retained to manage GSCC's hardware and communication facilities.

The introduction of a centralized, automated comparison system was of critical importance to the industry. Prior to 1988, trades in Treasury securities among dealers and brokers (which were largely done for settlement on the next business day) were verbally confirmed between the parties on the trade date, with written confirmations to follow the next day.[15] The Comparison System eliminated the need for such physical confirmation, bringing more certainty, greater efficiencies, and lower costs to the comparison process. Use of the system also eliminated risk by providing for the easy (and early) resolution of trade data differences. As one participant indicated:

> The GSCC comparison [system] has reduced our operational cost by an estimated $100,000 a year and has increased our efficiency. We now process the computer information you make available and distribute reports to our traders before 7 A.M. These reports outline the compared and uncompared trades at our internal trading account level. As a result, the traders resolve previous day's differences before the opening of trading.[16]

[15] Representatives of the various brokers and dealers would physically exchange paper confirmations at a facility provided by Bankers Trust Co. (the "Bankers' drop") and bring back counterparty confirmations for comparison with their records in the hope of identifying any problems before the opening of Fedwire or, more importantly, before any market-moving news affected a trading decision based on a faulty position.

[16] Letter of February 21, 1989, from Alexander Neamtu, principal, Morgan Stanley & Co., to Charles Moran, president of GSCC.

The Comparison System was successful from the start. On average, more than 16,000 sides (a "side" being one-half of a trade—either the buy side or the sell side) were being compared each day by December 1988 and more than 24,000 sides were compared on November 9, 1988, the record number per day for that year. By December, the average comparison rate was 94 percent of all submitted sides. Comparison System participation grew rapidly in 1989, to fifty-six primary dealer and interdealer broker members by year-end, with a record volume on August 11 of $258.5 billion, representing more than 34,000 sides.

The capabilities of the Comparison System also grew rapidly. "As-of" trades (trades compared on or after their scheduled settlement date, commonly done for audit trail purposes) were made eligible for comparison in March 1989. Trade cancellation and replacement features were added in May 1989.

5. Commencement of Netting

The Comparison System was a prelude to a more ambitious initiative: the Netting System. On July 7, 1989, after months of planning, programming, testing, and training on the part of GSCC staff and member firms,[17] GSCC's Netting System was implemented.[18] The system aggregated and matched offsetting deliver and receive obligations resulting from netting members' trades, in order to establish a single net settlement position for a member's activity in each CUSIP.

As explained in Box 2, netting, in its essence, is math. For a member's activity in a particular CUSIP, all the buy activity par amounts that contributed to the creation of a long obligation were added, and then all the sell activity par amounts that contributed to the creation of a short obligation were added; the difference between the two totals was the member's net-long or net-short position for the CUSIP. (Buy activity later also included Treasury auction awards, reverse repo start-leg activity, and repo close-leg activity, while sell

[17] GSCC began distributing test output to participants in April and, in conjunction, held training classes for participant operations and system personnel.

[18] GSCC had announced on June 22 that July 7, 1989, would be the implementation date for the Netting System, but the necessary regulatory approvals were obtained virtually at the last minute. On the morning of July 7, at a meeting of the Board of Governors of the Federal Reserve System, and after a presentation by Ernest Patrikis, general counsel of the New York Fed, the Board determined that GSCC's proposed operating rules for the Netting System were consistent with its risk reduction policies. That afternoon, the SEC issued an order approving the extensive rules governing netting, settlement, and related risk management.

Box 2

Net Settlement

The easiest way to explain net settlement is with an example.

Suppose four market participants are members of a net settlement system, trade a security among themselves as shown in Table A1, and submit their purchases and sales to the system. (Note that the four participants may also have traded with other, nonmember market participants, but any resulting purchases and sales are not submitted to the net settlement system.)

The first step in settling the trades of the four participants is to mark all of the transactions to a common system settlement price—which we assume is $99 per unit.

Marking to the common price results in credits and debits for the accounts of the four participants, as shown in Table A2. For example, marking the sale by A to B of 1 unit of the security at $97 per unit to a sale by A to B of 1 unit at $99 per unit results in a $2 debit to A (because A is due $99 instead of $97 following the new mark) and a $2 credit to B (because B will now have to pay $99 instead of $97).

Adding up the credits and debits for each of the participants shows that A has a net credit of $1, B and C have net credits of $2 each, and D has a net debit of $5.

The next step is to net the purchases and sales of each participant. As shown in Table A3, A purchased a total of 7 units of the security and sold a total of 9 units and is, therefore, a net seller of 2 units. Similarly, B is a net buyer of 4 units, C is a net seller of 3 units, and D is a net buyer of 1 unit.

On the settlement day, A delivers 2 units of the security to the settlement system against payment of $198 ($198 = $99 system settlement price per unit, times 2 units) and C delivers 3 units (against payment of $297). The net settlement system redelivers 4 units to B (against payment of $396) and 1 unit to D (against payment of $99).

Additionally, but as a separate matter, D transfers $5 to the net settlement system to clear the debit balance that appeared when D's transactions were marked to the system settlement price of $99 per unit, and the net settlement system transfers $1 to A, $2 to B, and $2 to C to clear their respective credit balances.

Significantly, the net settlement arrangement results in transfers of 10 units of the security, including 5 from the two net sellers to the system and 5 from the system to the two net buyers. Bilateral settlement would have required transfers of 33 units of the security.

Table A1

Four Market Participants' Hypothetical Transactions in a Security

Number of Units at Specified Prices

	Buyer			
Seller	A	B	C	D
A	—	1 at $ 97 3 at $100	2 at $100 1 at $ 99	1 at $101 1 at $100
B	1 at $ 98	—	3 at $101	1 at $ 99 1 at $100
C	3 at $100 2 at $ 99	1 at $ 99	—	2 at $101 2 at $100
D	1 at $102	2 at $100 3 at $ 99	1 at $ 98	—

activity later also included repo start-leg activity and reverse repo close-leg activity.) Thus, there was complete fungibility between buys, sells, auction awards, and repos for settlement netting purposes. Once the net settlement position was established by GSCC, the identity of the underlying trade activity was lost for clearance and settlement purposes.

GSCC netted on a multilateral basis, meaning that netting members were fungible and indistinguishable for netting purposes. The netting was done with a full guarantee of settlement for each net settlement position established; the buyer was guaranteed to receive the specific securities it purchased and the seller was guaranteed to receive the precise dollar amount it agreed to. Once a trade entered the net, GSCC became, through "novation," the effective counterparty to each of the original parties for credit and settlement purposes. In other words, all of a member's obligations to pay or receive money and to deliver or receive securities arising from its trades with counterparty members were terminated and replaced by similar obligations to and from GSCC.

After net settlement positions were determined, on the night before a scheduled settlement date, the resulting receive and deliver obligations were established. Each business day, GSCC established and reported by CUSIP, in a manner that preserved anonymity, net settlement positions and deliver and

Table A2

Credits and Debits Consequent upon Marking Transactions to a Common Settlement Price of $99

Transaction		Consequence for the Seller	Consequence for the Buyer
A sells	1 unit to B at $ 97	A debited $2	B credited $2
	3 units to B at $100	A credited $3	B debited $3
	2 units to C at $100	A credited $2	C debited $2
	1 unit to C at $ 99	—	—
	1 unit to D at $101	A credited $2	D debited $2
	1 unit to D at $100	A credited $1	D debited $1
B sells	1 unit to A at $ 98	B debited $1	A credited $1
	3 units to C at $101	B credited $6	C debited $6
	1 unit to D at $ 99	—	—
	1 unit to D at $100	B credited $1	D debited $1
C sells	3 units to A at $100	C credited $3	A debited $3
	2 units to A at $ 99	—	—
	1 unit to B at $ 99	—	—
	2 units to D at $101	C credited $4	D debited $4
	2 units to D at $100	C credited $2	D debited $2
D sells	1 unit to A at $102	D credited $3	A debited $3
	2 units to B at $100	D credited $2	B debited $2
	3 units to B at $ 99	—	—
	1 unit to C at $ 98	D debited $1	C credited $1

Note: In the table, A sells one unit for $97 and is entitled to receive $97. However, A will deliver the unit to the Government Securities Clearing Corporation (GSCC) at the common price of $99, requiring GSCC to recover the $2 difference.

Table A3

Total and Net Purchases and Sales of a Security by Four Market Participants

	Total Purchases	Total Sales	Net Purchases	Net Sales
A	7	9	—	2
B	10	6	4	—
C	7	10	—	3
D	8	7	1	—
Total	32	32	5	5

receive obligations. GSCC interposed itself between all receive and deliver obligations so that a long position represented securities that the member would receive from GSCC, while a short position represented securities that were due to be delivered to GSCC by the member.

Settlement of netted positions was done through GSCC's settlement process, which had three basic underpinnings: 1) every securities delivery, whether to or from GSCC, was made against full payment; 2) GSCC did not "build a box" during the day; rather, deliveries that came in to GSCC were instantaneously redelivered to another netting member (in other words, GSCC specified to each dealer the exact par and dollar amounts that were to be received or delivered, so that all movements were in a pre-bound status); and 3) all settlements were made over Fedwire (or intrabank, meaning that both the buying and selling members cleared at the same bank and thus payment needed to be made only on the books of that bank); therefore, finality of settlement was obtained at the time of the securities movement. Deliver and receive obligations were satisfied through delivery to and receipt from clearing banks designated by GSCC. All deliveries were made against simultaneous payment at that day's system value for the obligations.

Net settlement positions (including "fail" and "forward" positions[19]) and any resultant deliver and receive obligations of a netting member were fixed and guaranteed by GSCC at the time GSCC made available to the member the reports of such positions and obligations. (At the outset of the Netting System, that time was around 2:30 A.M. the following day.) At that point, all deliver, receive, and related payment obligations between members that had been created by the trades that determined the net settlement positions were terminated and replaced by the GSCC-issued settlement obligations. GSCC did not unwind positions that it had netted and guaranteed, meaning that it would not return a buyer and seller to their original positions.

Twenty firms, including three interdealer brokers and seventeen primary dealers, participated in the first net settlement.[20] The then-current four-year note was the first security netted. As the list of eligible securities expanded, the necessary clearance, settlement, custody (for margin), and financing services were provided to GSCC by SPNTCO (for notes)[21] and the Bank of New York (for all other products). It is interesting to note that in the early years of GSCC, the Fed, the Treasury, and GSCC had ongoing discussions about taking GSCC out of the business of moving securities, which requires the use of an intermediary clearing bank. GSCC argued that the most efficient and least risky scenario for settling government securities would involve GSCC issuing settlement balance orders directly to the Fed through a GSCC account at the Fed, which would have acted like a transfer agent. No Fed intraday credit would be required by GSCC. Apparently, this idea was never seriously considered by the Fed for at least two reasons: 1) the Fed would have had to develop or acquire all of the intricate operational capabilities developed by the clearing banks over decades, and 2) the Fed would have taken on counterparty credit risk (in other words, the risk of default by a dealer) directly.

The introduction of the GSCC Netting System had enormous implications for the government securities marketplace. Operational savings to members, particularly brokers, were quite significant. The benefits of GSCC were highlighted in an internal Chemical Bank publication called the *Data Bank*:

> Everybody wins. GSCC members were happy because netting eliminates the labor-intensive, time-consuming, and error-prone process of manual comparison. Accuracy levels have greatly improved, lowering a major risk factor. . . . Here at Chemical, Operations staff have witnessed a drop in the volume of buy and sell "deliveries," which was "particularly dramatic on our most active day," says Kyle Conselyea, who managed the GSCC project. . . . The common practice of holding onto outgoing deliveries until closing time—and the ripple effect that had on the subsequent task of "proving"—meant that overtime was a fact of life. Now his staff gets to go home at 5:00. "I don't know when I last paid overtime," Conselyea reports.

The Netting System also ensured the safety and soundness of the overall settlement process, and, for the first time, brought to the government securities market the significant risk protections that stem from multilateral netting of obligations by novation (with GSCC assuming the position of counterparty on all trades for settlement purposes) and daily margining and marking-to-market (taking into account accrued interest) of the net settlement positions of each netting member.

GSCC imposed on its members the discipline of having to meet various financial, operational, and other standards for admission to and continued participation in the system. Moreover, GSCC put in place a centralized loss allocation procedure for handling the insolvency of a member.

During the initial weeks of the netting operation, eligible securities were limited to newly issued Treasury notes, allowing participants time to acclimate to the process. The first Treasury bond—the most recent one at the time—was added in September 1989. Thereafter, product eligibility grew rapidly. By January 1990, all Treasury securities other than STRIPS (Separate Trading of Registered Interest and Principal of Securities) were eligible for the net, and there were thirty-four netting participants. In April 1990, the scope of the Netting

[19] A fail net settlement position is one that is past its scheduled settlement date and has not yet settled. A forward net settlement position is one that is scheduled to settle one or more days in the future.

[20] The three brokers were RMJ Securities Corporation, Garban Ltd., and Cantor Fitzgerald Securities Corporation. The seventeen primary dealers were Carroll, McEntee & McGinley; Daiwa Securities America; Discount Corporation of New York; Dillon, Read; First Boston Corporation; First National Bank of Chicago; Goldman Sachs; Kidder, Peabody; Kleinwort Benson Government Securities; Merrill Lynch Government Securities; Morgan Stanley; Nikko Securities; Nomura Securities International; Prudential-Bache Securities; Salomon Brothers; Smith, Barney, Harris Upham; and Yamaichi International (America).

[21] In 1992, SPNTCO's parent was merged into Bank of America National Trust and Savings Association (BOA). BOA subsequently determined, for risk reasons, to exit the clearing business. Its Security Pacific National Bank subsidiary continued to operate and provide clearance services for two years after BOA announced that it was planning to exit the business, in order to facilitate the smooth, seamless conversions of its customers to other clearing banks. Interestingly, NSCC/GSCC considered buying SPNTCO at that time, in order to allow nonbanks to have direct access to Federal Reserve services. Myriad issues prevented this, including NSCC and GSCC becoming subject to the Bank Holding Company Act, lack of sufficient capital, and intraday overdraft and overnight credit considerations.

System was expanded to encompass all forward-settling trades (those scheduled to settle within fifteen days of execution) of netting members in Treasury securities. By the following month, the securities constituting the quarterly Treasury refunding were eligible for the net for the first time, and in July 1990, agency securities became eligible for netting. On February 1, 1991, STRIPS were made eligible for netting. By the end of 1991, GSCC was processing an average of more than 20,000 sides, representing over $150 billion, in net settlements each business day.

6. Impact of the Salomon Brothers Scandal

In August 1991, Salomon Brothers Inc. admitted that it had seriously violated U.S. Treasury auction rules by submitting fraudulent bids. The firm managed to avoid an indictment but paid $290 million in fines. Investor Warren Buffett took control of the company and a number of senior executives resigned, including Chairman and CEO John Gutfreund.

The Salomon Brothers scandal had significant implications for GSCC and the government securities market. By the early 1990s, the interdealer brokers, who formed the core of this over-the-counter market, found themselves in an untenable economic position because of severe cuts in commission rates in previous years. The cuts were caused by primary dealer actions, including the formation of Liberty Brokerage, which was owned by several primary dealers and established to bring down commissions through competition. To make up for lost commission income, the interdealer brokers sought to "go national" and expand their customer base beyond primary and "aspiring primary" dealers.[22]

The scandal focused public attention on the activities and role of the primary dealers and raised questions about their integrity and that of the entire Treasury marketplace. This, in turn, constrained the ability of those dealers to continue taking actions that might be perceived as hindering the transparency and fairness of the market. As a result, in late 1991, four interdealer brokers—Fundamental Brokers, Garban, Liberty, and RMJ—announced that they were expanding access to their screens to non-primary dealers that were GSCC netting members. This action had not been tenable earlier because a primary dealer would not do business with a broker that might match it against someone other than a primary dealer on a blind basis, even if that dealer was a GSCC netting member. The interdealer brokers, taking advantage of the post-scandal climate (and the intensified scrutiny of primary dealer actions), adopted the standard of status as a GSCC netting member as an objective means of expanding their customer base beyond the primary dealer community. The first non-primary dealer to receive broker screens was the Chicago Corporation (on October 28, 1991), followed by Continental Illinois Bank.[23]

7. Growth of the Netting Process

In January 1992, the Federal Reserve, the SEC, and the Treasury Department issued the *Joint Report on the Government Securities Market*, which stated that the three agencies did not believe the government securities market was "flawed or broken in any fundamental economic sense."[24] The report said that GSCC had made the market "even more efficient," and that GSCC's netting process "substantially reduces counterparty risk" for GSCC members. The report further noted the benefits of a GSCC proposal to include Treasury auction awards in its Netting System, and it encouraged GSCC to 1) "develop efficient processing systems for market participants' repo activity," 2) "expand to a greater universe of trades the benefits of netting," and 3) accelerate its efforts to expand membership.

GSCC moved ahead quickly to provide additional enhancements. On February 21, 1992, it introduced an enhancement to the executing firm information field in the Comparison System that offered members improved comparison results through identification of the true executing parties to a trade (see Box 3). It also allowed nonmember firms that cleared through GSCC members to more readily

[22] An "aspiring primary dealer" category existed at the time. It was treated the same as primary dealer for purposes of access to broker screens.

[23] Certain primary dealers were vocal in their opposition to this development, voicing credit concerns. GSCC's position was that this was a positive development for a number of reasons, including enhanced market liquidity and transparency. GSCC also emphasized that, if a non-primary dealer with interdealer broker screen access failed, the primary dealers were far better off having that failed dealer be a netting member because GSCC would have guaranteed the transactions of, and collected appropriate margin and mark from, the insolvent member.

[24] In conjunction with the release of the *Joint Report*, the New York Fed issued a revised set of criteria for designation of a firm as a primary dealer and for the administration of its relationship with primary dealers. The Fed eliminated certain market-making requirements and replaced them with criteria including making "reasonably good markets" in trading with the Fed's trading desk, participating meaningfully in Treasury auctions, and providing the Fed's trading desk with useful market information and analysis.

Box 3

Identifying the True Executing Parties to a Trade

When participants were aware of the correspondent clearing relationships of other members and had information on the names and accounts of the nonmember parties they actually traded with, trade data could successfully be submitted for comparison against the member that was acting for the correspondents. Difficulties arose, however, when the clearing relationships were not well understood. A GSCC member, unaware that its trading partner was the correspondent of another member, very often failed to submit the matching side of such trade for GSCC processing.

For example, if Primary Dealer A traded with XYZ Small Firm, and XYZ cleared its activity through Primary Dealer B, Primary Dealer A might not have submitted the trade to GSCC because it assumed that it had traded with a nonmember and, thus, that the trade was not GSCC-eligible. In addition to causing an uncompared trade for the submitting party, this lack of awareness made it difficult for the counterparty, who got an advisory, to determine the cause of and resolve that advisory. Generally speaking, these unmatched trades pended in the system until they were deleted by GSCC.

The situation was complicated by the fact that some members guaranteed the trading activity of their correspondents but others did not (a situation that remains to this day). Trading partners that were netting members often were hesitant to submit the counterparty side of nonguaranteed correspondent trades to GSCC, because this activity was subject to netting and margin and mark-to-market requirements. To avoid these problems, GSCC provided new fields for identifying correspondents, and netting members were allowed to indicate whether the activity of a given correspondent with another netting member was eligible for netting.

access GSCC's comparison and netting services. By the end of the year, more than 250 executing firms were taking advantage of this feature.

On October 16, 1992, an automatic yield-to-price conversion feature was implemented, eliminating the two-step pricing process for when-issued trades[25] and allowing when-issued trades to be netted and novated on the trade date. As the coupon of the new issues was not determined until the auction date, members' when-issued trading activity between the announcement date and the auction date was submitted to GSCC with a yield. Those transactions then had to be resubmitted to GSCC for comparison with a dollar price during the *auction date* + 1 processing cycle. Only then were trades, if compared, eligible for GSCC's netting and novation services and the resultant credit protections. With the launch of the yield-to-price facility, GSCC automatically converted yield trades into price trades following the announcement of auction results. The service thus reduced both risk (because the guarantee of settlement occurred as soon as a yield trade compared) and costs (because of the elimination of the double submission of when-issued trades).[26]

8. Treasury Auction Takedown Process

Membership and trade data submissions continued to grow rapidly. On March 31, 1994, GSCC had its first $1 trillion netting day. More than $950 billion was eliminated from settlement.

A month before that, GSCC had filed a rule change proposal with the SEC to permit it to extend its comparison, netting, settlement, and risk management services to U.S. Treasury securities purchased at auction and issued through Federal Reserve Banks. The proposal allowed the inclusion of all proprietary (or "house") purchases of Treasury bills, notes, and bonds by GSCC netting members, whether done on a competitive or noncompetitive basis.[27]

Three years earlier, GSCC had opened discussions with the Treasury and Fed on its proposal to expand its Netting System to include auction awards, or "takedowns." The 1992 *Joint Report* encouraged the effort, noting that "the benefits of netting were greater as more trades were included in the net, because a greater number of receive and deliver obligations were reduced to as small a number as possible."

[25] Trades in securities that are about to be issued. The when-issued market allowed dealers to presell to customers ahead of the auction date and then cover the sale in the auction.

[26] Participation in the service was initially voluntary. Those members that did not participate had to submit final money on the auction date or their trades would be rejected. Also, they needed to submit final money for all trades executed between the auction date and the settlement date in order for those trades to compare (whereas GSCC would calculate final money for participating members).

[27] Auction awards resulting from bids made by netting members on behalf of customers that had been named on the bidding member's tender form were not eligible. This was a requirement imposed by the Treasury Department to ensure that customer awards would always be filled (and not netted out against short sales in the secondary market). Owing to system limitations, securities that were auctioned and issued on the same date also were (and remain) ineligible.

GSCC initiated its auction takedown process in September 1994. Prior to that time, the Treasury settled an auction award to a dealer that was a GSCC member in the same way that it settled an auction award to any other institutional investor: by delivering securities to the book-entry account of the dealer's custodian. This process was inefficient because in many cases the dealer had already sold some or all of its auction award in when-issued transactions.

The key idea of the new process was that auction awards and when-issued purchases were equivalent for purposes of netting and settlement. Additionally, if a new issue reopened an outstanding security, auction awards were also equivalent to conventional secondary market purchases of securities with the same CUSIP. Pursuant to the auction takedown process, on the issue date, the Fed delivered to GSCC securities equal to the aggregate awards of its members. GSCC then redelivered those securities, along with securities received from members with net-short positions, to members with net-long positions. Thus, the auction awards lost their separate identities and became part of a consolidated net settlement process.

On September 12, 1994, the first day that members' proprietary Treasury auction awards were encompassed within GSCC's net, three Federal Reserve Banks (New York, Chicago, and San Francisco) submitted the auction award details of the three-month and six-month bill auctions to GSCC, which in turn generated locked-in confirmations for the thirty-nine participating GSCC members. The size of each auction was $11.6 billion, nearly half of which reflected proprietary auction awards to eligible GSCC participants. Auction purchases were then netted with when-issued and other secondary market trades in the same securities submitted by GSCC netting participants. Combined, auction purchases and secondary market trades totaled more than $120 billion. However, through netting, the resulting receive and deliver obligations generated on September 15, when the securities were issued, totaled approximately $28 billion. The Treasury soon began to make the large majority of all auction deliveries to dealers indirectly through GSCC.

The implementation of auction processing enabled GSCC to 1) accept eligible auction award details from Federal Reserve Banks and generate comparison output based on those details, 2) net auction purchases with when-issued and other secondary-market trades in the same security submitted by netting members, and 3) take direct delivery of purchased securities from Federal Reserve Banks at one of GSCC's designated clearing banks for prompt redelivery to members with net-long positions through GSCC's settlement process. The prompt redelivery of auction awards to participants with long positions (within minutes of receipt of the securities from the Fed), among other things, reduced the daylight overdraft exposures associated with new issue distribution.[28]

The auction takedown service also resolved several risk management problems associated with gross settlement of auction awards. First, the unnecessary deliveries to dealers that were not ultimate buyers created risk for GSCC because of its guaranteed settlement of the redeliveries by those dealers. Second, because GSCC did not have knowledge of auction awards made to its netting members, it could not guarantee settlement of those awards (as it would for secondary market trades), thus leaving the Treasury exposed to credit risk. Finally, GSCC was unable to assess proper performance guarantees, or margin on purchasers and sellers, and it could not mark their positions to market accurately. The auction takedown service allowed GSCC to margin and mark dealer positions on a true net basis.

9. Repo Netting

In 1990, with the basic netting engine built, GSCC started to analyze the possibility of applying its comparison, netting, and risk management processes to repos—repurchase agreements and reverse repurchase agreements involving government securities as collateral. Successful application would provide the repo market with efficiencies and risk protections akin to those provided for buy-sell trades.

At the time, there were a variety of risks and inefficiencies in the government securities repo market. To begin with, repo transactions were confirmed on a nonautomated basis by telephone or fax. Also, while the bulk of repo activity was conducted through interdealer brokers, such brokered transactions were not transacted anonymously but rather were done on a "give-up" basis, meaning that the broker matched the two parties and then stepped out of the trade after revealing the identity of each party to the other. This withdrawal was done because repos, given their average size and potential time to settlement, presented more settlement risk than most dealers wanted to take on with a broker counterparty. However, the consequent lack of an intermediary resulted in a greater flow of information that exposed a dealer's trading strategies to competitors. Moreover, market participants faced the risk (which might not be sufficiently, if at all, covered by

[28] Before the implementation of auction processing by GSCC, Treasury securities delivered in settlement of auction awards would be sent to the purchasing dealer's clearing bank account in the morning, and often would sit in that account until redelivered to dealers with long positions later in the day, thus requiring the bank to provide intraday credit to the purchasing dealer.

margin) that a counterparty would fail to pay back principal plus interest owed or fail to deliver back collateral. Furthermore, repos required settlement of both their start and close legs on a trade-by-trade basis.

Key to the feasibility of including repos in the net was that, from a clearing perspective, receives and delivers generated by repos did not differ significantly from receives and delivers generated from netting outright buys and sells. In view of this, GSCC management realized that it could offer a service that would keep track of and net out offsetting securities movements, whether arising from other repos or from non-repo trading involving the same CUSIP.

Initially, there was industry resistance to the GSCC proposal for providing services for repos, primarily owing to the perception among larger primary dealers that GSCC's netting process would "level the playing field" to their disadvantage. Among the factors that helped to overcome this resistance was the focus on the government securities marketplace brought about by the Salomon Brothers scandal. In January 1992, in the wake of the scandal, the *Joint Report* stated that "GSCC could benefit the repo market by offering a system that clearly defines which stage of a transaction was occurring . . . and that automatically generates a comparison of the transaction."

9.1 Initial Repo Netting Proposal

By August 1992, GSCC had designed a repo netting proposal that would provide the following benefits to the repo market:

1. Automated comparison of the start and close legs of a repo, including the capture of all key elements of the transaction, which would help members monitor repos and maintain appropriate recordkeeping and audit trail information.[29]
2. Netting and settlement of underlying collateral movements, which would offer significant cost savings and alleviate operational burdens.
3. Pass-through of coupon interest, which would provide for coupon payment protection.
4. Guaranteed settlement of repo transactions, with GSCC assuming the role of counterparty to each side (as it did for buy-sells).
5. Centralized and standardized daily margin and mark-to-market for each repo position.
6. Favorable accounting treatment that would facilitate members' ability to offset, for balance sheet purposes, repos and reverse repos netted and guaranteed by GSCC.[30]
7. Net capital relief: Under the SEC's net capital rule, broker-dealers must deduct from their net worth certain repo agreement deficits when computing net capital. But when computing the deductions, broker-dealers may net obligations due under repo agreements entered into with the same party. Having the clearing corporation as the common counterparty to repo dealers would provide substantial net capital relief.[31]

GSCC staff then sought guidance from the Repo Committee of the Public Securities Association (PSA), which established a working group comprising dealers and repo brokers to focus on the proposal. After meeting from September through December of 1992, the working group wrote to GSCC management encouraging GSCC to provide a comparison service for repos "as expeditiously as possible," noting that comparison would have a number of benefits for the repo market, such as "helping counterparties detect errors and creating an audit trail." The working group also asked for a more detailed "blueprint" for netting repo transactions. It noted that three general principles should govern the development of the blueprint: 1) implementation should be designed so as not to require conversion costs that might exceed savings from future operational efficiencies, 2) novation through GSCC should achieve counterparty netting for accounting purposes, and 3) netting should reduce the cost to GSCC members associated with daylight overdrafts.

[29] At that time, settlement of the close leg of an overnight repo often occurred before the counterparty had the opportunity to check the information contained in its confirmation (which it received only on the morning of the close date). GSCC proposed that members be allowed to compare their repos on a same-day basis and thus be able to properly monitor overnight repo transactions and reconcile incorrect information.

[30] In this regard, an important development occurred in March 1992 when the Financial Accounting Standards Board (FASB) issued an interpretation (No. 39 – Offsetting of Amounts Related to Certain Contracts) stating that fair market amounts recognized for forward and other conditional or exchange contracts executed with the same counterparty under a master netting agreement could be offset. GSCC received an opinion on December 16, 1993, from Michael Passarella of Price Waterhouse essentially providing that GSCC repo participants would be able to satisfy all of the criteria specified in FASB Interpretation No. 39 and thus would be able to offset, for balance sheet purposes, the asset and liability amounts that arose from netted repo transactions that had the same close date.

[31] By letter dated March 13, 1996, from Michael A. Macchiaroli, associate director of the SEC's Division of Market Regulation, to the author, GSCC obtained no-action relief to the effect that a broker-dealer, for net capital computation purposes under paragraph (c)(2)(iv)(F) of Rule 15c3-1 (17 C.F.R. § 240.15c3-1(c)(2)(iv)(F)), could treat GSCC as its counterparty for repo transactions entered into GSCC's netting system.

In April 1993, GSCC established its own Repo Implementation Committee. The committee's initial focus was on the implementation of a comparison service for repos, designed to provide benefits such as the elimination of physical confirmations, more timely comparison of repo trade data, easier monitoring of the status of repos, ready ability to link and monitor the start and close legs of a repo, enhanced ability to identify and correct errors, easier recordkeeping, and improved access to audit trail information. More than a year of business specification design and technological development followed.

9.2 Repo Comparison Service

In January 1995, GSCC launched a Repo Comparison Pilot with more than twenty participants. Taking part in the pilot enabled firms to identify the system and operational changes needed to provide GSCC with accurate repo comparison data, and to identify unforeseen problems.

On May 12, 1995, GSCC went live with its repo comparison service, with twenty-six participants. Overnight and fixed-term repos on government collateral were eligible for comparison. GSCC did not make any characteristics pertaining to rights of substitution a required match item because it felt that doing so would greatly hinder the comparison rate for repos, given the difficulty firms had in submitting substitution data.

9.3 Development of the Repo Netting Service

GSCC next turned its focus to a repo netting service. It struggled with the resolution of several key issues, including whether GSCC should guarantee settlement of a forward starting repo (a repo where the start leg was scheduled to settle one or more days after the trade date) prior to the actual start of the repo. One option was to net and fully guarantee the repo at the time of its comparison, one or more days before the scheduled settlement of the start leg. The problem with this approach was that it created an unacceptable level of exposure for GSCC, obligating it to conduct two settlements when the underlying repo might never be initiated by the parties to the repo.

Another alternative was to neither net nor guarantee the repo until the start leg actually settled. GSCC management rejected this approach because it left participants with no protection during the forward start period.

GSCC ultimately chose the middle ground of providing rate protection—guaranteeing the payment of interest due but not guaranteeing actual settlement of the start and close legs of a repo that had not in fact started—during the period between trade execution and the start of the repo. This protection reflected the difference between the contract repo rate and the current rate for a repo of like term and underlying collateral. The approach made sense because no securities would have moved between the parties before the start date and there was, therefore, no risk other than interest rate risk. (Once the start leg settled, there would be full guarantee of settlement of the close leg.)

Another major issue involved guaranteeing settlement of repos of an extended length, where the daily financing mark to the parties, and the interest rate exposure to GSCC, could be quite large. In view of the risk involved, GSCC chose to limit the number of business days between the submission date and the settlement date for the close leg of an eligible repo to a half-year.[32]

Finally, because of potential operational difficulties, GSCC ultimately chose to make open repos (repos with no fixed end date) ineligible. Other eligibility requirements established for netting were as follows: The data on a repo had to be submitted by netting members that had agreed to adhere to the heightened mark payment and margin deposit requirements and other aspects of the repo service; such data had to be compared; and the underlying securities had to be Treasury or book-entry federal agency securities.

[32] The term of an eligible repo initially was limited to 180 days in order to evaluate GSCC's risk management measures. On September 23, 1996, the term was extended to 360 days. On June 2, 1997, GSCC made eligible those repos having closing leg settlement dates of up to two years after submission. In taking this action, GSCC was cognizant that the Bankruptcy Code's automatic stay exception for the liquidation of a "repurchase agreement" as defined by the Code, and for the setoff by a repo participant of a debt or claim arising in connection with such a defined "repurchase agreement," applied only to a repo transaction with a term of not more than one year. GSCC noted to its members that, "[i]n this regard, there was a comparable automatic stay exception in the Code for the liquidation of a repo transaction as a 'securities contract' (even if it has a term longer than one year) on which GSCC, as a clearing corporation, could rely in the exercise of its netting rights in respect of such transactions." The enforceability of GSCC's netting rights also were supplemented and made clear by the application of other federal legislation. (GSCC Important Notice 42.97, May 28, 1997)

9.4 Implementation of Repo Netting

On November 17, 1995, the initial phase of the repo netting service—involving the netting and settlement of the close leg of overnight and term repos (and the start leg of forward-settling repos)—began, with thirteen participants.[33] Repo transactions were netted with conventional buy-sell activity and Treasury auction purchases in the same CUSIP to arrive at a single net position in each security. As it did for cash transactions, when GSCC netted repos it interposed itself between the two parties for settlement purposes. GSCC's guarantee included the return of repo collateral to the repo participant, the return of principal (the repo start amount) to the reverse repo participant, and payment of repo interest to the full term of the repo to the reverse repo participant. The guarantee also included coupon interest protection, meaning that, once the repo started, GSCC would automatically pass a coupon payment from the reverse repo participant to the repo participant on the coupon payment date, crediting the repo participant and charging the reverse repo participant in the process.

GSCC had to significantly revise its risk protections to accommodate the greater risk arising from repo activity. Repos, on average, were much larger in size than buy-sells (averaging $38 million at the time, compared with roughly $9 million for buy-sell transactions) and many were long term and carried a financing component that buy-sells did not. With regard to its mark-to-market process for repos (also known as "forward margin"), GSCC, in addition to applying a mark to the underlying securities on forward trades, began to use a new, separate financing mark, which took into account the potential financing cost GSCC would earn or incur if it had to finance the repo position of a failed participant between the date of failure and the settlement date for the close leg of a long-term repo.

In determining the repo rate used in these calculations, GSCC decided that the rate would need to be tailored to each individual repo transaction. For general collateral repos, GSCC would use the remaining term of the repo to determine the appropriate market repo rate. For special collateral repos, GSCC would determine the special repo rate on the basis of the CUSIP and the remaining term of the special. In order to determine the various rates, GSCC was given full access to the broker's repo screens.

[33] The initial participants were Cantor Fitzgerald Securities, Deutsche Morgan Grenfell/C.J. Lawrence Inc., Dillon, Read, Eastbridge Capital, Goldman Sachs, HSBC Securities Inc., Lehman Brothers, Merrill Lynch Government Securities Inc., NatWest Bank NA, Oppenheimer & Co., Spear, Leeds & Kellogg LP, UBS Securities Inc., and Zion First National Bank. Repo transactions worth more than $6 billion went into the net on the first day.

GSCC's mark-to-market process also changed fundamentally in that GSCC had previously collected debit marks (which could be satisfied by cash or eligible collateral) but held credit marks, and it had not used a credit mark in one CUSIP to offset a debit mark in a different CUSIP. To facilitate repo netting, it shifted to offsetting credit marks against debit marks across CUSIPs (thus providing for cross-margining of the cash and repo markets) with full pass-through of collected marks. In starting to pay out credit marks, GSCC had to convert its mark-to-market process into a cash-only process.

GSCC also sought to provide its full guarantees without adversely affecting the economics of a repo. Thus, GSCC determined to pay interest on debit mark amounts collected and charge interest on credit margin amounts paid on a daily basis using an effective federal funds rate.

Meanwhile, GSCC added a new "repo volatility" component to its clearing fund to guard against risk associated with guaranteeing the payment of repo interest to the term of the repo. GSCC requires that its netting members maintain deposits in the GSCC clearing fund account to provide adequate risk protection and liquidity in the event of a participant failure. Clearing fund margin is the pool of margin collateral collected from netting members and held by the clearing corporation to help manage the risk of a netting member defaulting on its payment and delivery obligations. The need for the repo volatility aspect of the clearing fund arose from the liquidation process that GSCC would conduct in the event that a member with an outstanding term repo failed. Under that process, if the member was a funds borrower and had defaulted on its obligation to repurchase the underlying securities and pay interest at the end of the repo term, GSCC would immediately sell the same securities and reverse them in (for the same remaining term as the original repo transaction). GSCC would thus have the opportunity to earn interest income; however, this amount could be less than the interest payment that would have to be made to the funds lender on the scheduled settlement date. Conversely, if the member was a funds lender and had defaulted on its obligation to redeliver the underlying securities at the end of the repo term, GSCC would immediately buy the same securities and put them out on repo (for the same remaining term as the original repo transaction). GSCC would thus incur interest expense, which may be greater than the interest payment to be received from the funds borrower at the end of the repo. The immediate open market purchase and sale transactions were necessary to mitigate the market risk of the underlying securities. GSCC marked-to-market and required margin each day up to the firm's insolvency, and it needed to eliminate the risk that any future market moves would create.

Clearing fund margin was calculated in a manner designed to protect GSCC from fluctuations in the value of a net settlement position from the latest marking-to-market until liquidation. The repo volatility amount, which corresponds to the volatility of repo rates, was used to provide GSCC with protection from the portion of that fluctuation in value that represented interest exposure.[34]

GSCC's margining and repricing services provided, for the first time, a standardized approach for the coordinated, risk-managed movement of both the collateral underlying a repo and the interest owed on the repo. The services fundamentally changed the marketplace in that participants no longer needed to build margin (or a "haircut," typically 2 percent) into the original value of a repo, but could instead price the repo at the current market value of the collateral.

The repo netting service was a major success, and participants and volumes grew steadily.[35] Critical to this success was the Financial Accounting Standards Board's adoption, in December 1994, of Interpretation No. 41, which permitted financial entities to offset, on their financial statements, repos and reverse repos when the transactions met certain criteria, such as having the same counterparty and settlement date, being executed in accordance with a master netting agreement, involving securities in book-entry form, and settling via an appropriate securities transfer system.

The "same counterparty" requirement was critical; that was where GSCC could provide great value. By becoming the common counterparty to each side of a repo upon novation, GSCC could maximize the ability of participants to take balance sheet offset. This role was crystallized in a May 30, 1995, opinion from Price Waterhouse LLP to GSCC stating that "members would be allowed to offset, for financial statement purposes, a short and long net settlement position . . . in a particular CUSIP comprised in whole or part of repo transactions against a long or short net settlement position with the same scheduled settlement date in another CUSIP comprised in whole or part of repo transactions."[36]

10. Netting and Settlement of Brokered Repos

The establishment of the repo netting system transformed the Treasury repo market. But even more fundamental change was yet to come. GSCC's next major effort was to arrange for netting and guaranteed settlement services for *same-day start-leg brokered* repos (which represented the majority of repos and were done on a "give-up" of identity basis), including the automation of start- and close-leg processing, which are integrally related.

A critical issue in this regard was that, as noted earlier, most repos started on the day that they were executed, but GSCC was not equipped to handle same-day start legs. GSCC's Repo Implementation Committee and its broker membership proposed a solution that would have brokers assume responsibility for the movement of securities between dealers for same-day start-leg settlement. Brokers and dealers would send transaction details to GSCC for comparison, netting, and guaranteed settlement of repo close legs.

[34] The formula provided that the gross amount of margin would be calculated by multiplying the system value of the repo position by the repo volatility factor (expressed in basis points) and then by a fraction, the numerator being the number of days to the scheduled settlement date of the close leg and the denominator being 360. The repo volatility factor for general collateral repos (defined as all repos other than special repos) was set at 50 basis points. For special repos (defined as any repo with a system rate that was more than 100 basis points less than the system rate for general collateral repos), a distinction would be made between those expected to come off special on a certain date (such as an upcoming issue date) and all others. The factor for those expected to come off special would be the same as the factor for general collateral repos, while the factor for all other specials would be equal to the spread between the system rate for the repo and the system rate for general collateral repos (but in no event less than 50 basis points). Repo volatility amounts on long and short net positions were allowed to offset each other.

[35] Still, there were issues. For example, the SEC filing made by GSCC for authority to implement the repo netting service was challenged by Delta Government Options Corporation, a competitor clearing corporation at that time, which asserted that the proposed repo netting system would not afford participants adequate financial protections. The SEC ultimately rejected Delta's objections. (See SEC Release No. 34-36491, 60 S.E.C. Docket 1814, 1995 WL 704170 [November 17, 1995]; File No. SR-GSCC-95-02, Order Approving a Proposed Rule Change Relating to Netting Services for the Non-Same-Day-Settling Aspects of Next-Day and Term Repurchase and Reverse Repurchase Transactions.)

[36] In June 1996, the FASB issued Financial Accounting Standards Statement No. 125 (FAS 125), which provided accounting and reporting standards for transfers and the servicing of financial assets and extinguishment of liabilities, and established new criteria for determining whether a transfer of financial assets in exchange for cash or other consideration should be accounted for as a sale or as a pledge of collateral in a secured borrowing. After GSCC members raised questions about the impact of FAS 125 on repo netting members' relationship with GSCC and, particularly, whether a right of substitution that was included in the underlying agreement continued to exist after novation, GSCC made a filing in 1997 (1997-3) in which it amended its rules to explicitly provide that: 1) GSCC would recognize that a right of substitution exists with regard to a repo if either of the parties submitted matching data indicating such a right, or if GSCC, in its sole discretion, determined that the parties intended that such a right exist; 2) if the parties to a repo entered into the transaction with a right of substitution, that right would continue once the repo was netted by GSCC, and GSCC would facilitate the parties' ability to conduct such a substitution or termination; and 3) a right of substitution continued once the repo was netted by GSCC. The GSCC filing made clear that, if a GSCC repo netting participant provided for a right of substitution in the underlying repo agreement, there was no need for that participant to seek to enter into an additional agreement with GSCC regarding that right of substitution.

On August 5, 1996, interdealer broker netting members became eligible to participate in the new brokered repo service.[37] Starting on that date, brokers meeting GSCC's financial and operational requirements began submitting repos to GSCC in lieu of give-up (where the counterparties would reveal or "give up" their names to one another when submitting repo trades to GSCC through brokers). Settlement of same-day start legs occurred directly between the broker and the repo and reverse repo dealers, while compared close legs and forward start legs were netted, guaranteed, and settled through GSCC.

The implementation of the brokered repo service revolutionized the marketplace. As GSCC guaranteed settlement of the repo close leg, it eliminated counterparty risk. The service turned a market that was entirely give-up into a largely anonymous blind brokered market. Bringing more transactions into the net also dramatically reduced the number of total daily settlements for broker participants and made their back-office settlement process much more cost-efficient.

Using brokers as counterparties was not without controversy, since it was perceived by some market participants as introducing a new risk by allowing somewhat thinly capitalized brokers to act as principals. One means by which GSCC limited this risk was the imposition of enhanced minimum capital requirements. At the time, give-up repo brokers operated with a low level of capital, based on the principle that they did not participate in settlements or take market risk. GSCC changed that by requiring each interdealer broker netting member engaged in repo activity to have a minimum of $10 million in excess of SEC-required net or liquid capital (an increase from the then-current $4.2 million excess net or liquid capital standard). Among other things, this requirement helped to mitigate the risk of a broker failing between the time a transaction was executed and the submission of data to GSCC.

Another way that GSCC limited risk was by imposing a scope-of-business requirement. Each repo transaction submitted to GSCC by a broker was required to have an actual Netting System participant as the counterparty and had to be bound to a corresponding reverse repo transaction. This rule ensured that the broker would net out of the settlement process for the close leg.

From an operational perspective, the brokers already had in place the systems needed to submit data to, and receive output from, GSCC; thus, they only had to make minor changes to their regular buy-sell input and output specifications to accommodate repos and then test with GSCC. However, additional operational requirements were imposed on the participating brokers:

1. Each broker had to establish a separate account, with a separate Fedwire address, at a clearing bank for use exclusively for repo start-leg intraday settlements. This account would be subject to review by GSCC.
2. Each eligible repo transaction had to be submitted to GSCC as soon as possible after execution and, in any event, no later than fifteen minutes thereafter.
3. If a counterparty netting member indicated that an error had been made by a broker, the broker was obligated to take steps to promptly resolve the error or dispute.
4. Brokers would be involved in collateral substitutions on long-term repos.[38]

Moreover, GSCC required that data on all "regular-way" repos (which start on the trade date), forward-starting repos, and repos in which the start leg had failed be submitted to GSCC, in order to preserve the integrity of the netting process. Finally, GSCC reserved the right, for risk management purposes, to compare repos based on data submitted by only one side.[39]

Soon after implementation of the brokered repo service, a concern arose among certain dealers that, because a dealer or broker counterparty could potentially fail to submit data, a brokered repo would not be compared and, therefore, would not enter GSCC's netting process. This issue was of particular concern for repos that started on the same day they were

[37] Eight brokers (Exco RMJ Securities, Liberty Brokerage, Garban, Tullett & Tokyo Securities, C.F. Kross, GFI Group, Prebon Securities, and Euro Brokers Maxcor) and twenty-six dealers participated. On the launch date, more than $5 billion in repo activity was processed. By week's end, over $31 billion was processed.

[38] The repo dealer initiating the substitution would contact the broker and provide it with all relevant information regarding the substitution. The broker would then contact the reverse repo dealer to arrange for the substitution, providing it with all pertinent information. The broker would also provide GSCC with the terms of the substitution. GSCC, acting as an "honest" third party, would hold the collateral received from one dealer until it could be passed through against payment. Once GSCC had the collateral from both parties, the substitution would be made, with GSCC automatically reversing any previous mark-to-market and clearing fund amounts calculated for the old collateral.

[39] This right proved extremely beneficial on the evening of September 11, 2001, when, in order to facilitate an orderly settlement process and mitigate the potential for the enormous systemic risk associated with thousands of unmatched trades, GSCC made the decision to create and administratively compare 2,178 broker trades valued at more than $71 billion based on the presumption that the dealer counterparty submission was accurate. In taking this action, GSCC moved a massive reconciliation effort that would have been conducted between its dealer and broker members into one central location within GSCC.

entered into and closed the next day because there would be relatively little time to correct errors or omissions. To address this concern, GSCC established a policy under which it would effectively guarantee settlement of the start and close legs of every netting-eligible blind-brokered repo transaction that had been entered into in good faith by a member.[40]

Various enhancements were quickly made to the brokered repo service. In October 1996, GSCC implemented a collateral substitution facility, allowing members to submit substitution details on their comparison input and to process collateral substitutions online.[41] In November 1996, GSCC began offering repo-to-maturity processing services for repos on collateral that matured on the repo close date. In January 1997, the repo netting service was enhanced to provide services for repos on collateral that matured prior to the repo close date. These enhancements provided members with great flexibility when selecting collateral for repo transactions. A participant simply had to substitute acceptable new collateral no later than the business date prior to the maturity date of the existing collateral. Thus, a participant could substitute collateral as many times as it wanted over the term of the repo (subject to the terms of the repo), and substituted collateral could mature before the repo close date, as long as appropriate substitutions were made on a timely basis.

11. GCF Repo

In November 1998, after approval by the SEC,[42] GSCC revolutionized the financing marketplace by introducing a new product, the General Collateral Finance Repo service (GCF Repo) for Treasury securities collateral.[43] The GCF Repo service enabled GSCC's non-interdealer-broker netting members (dealers, for the purposes of this discussion) to trade general collateral repos—based on rate, term, and underlying product—throughout the day without requiring intraday, trade-for-trade settlement on a delivery-versus-payment (DVP) basis. Dealers executed GCF Repos with GSCC interdealer broker netting members (brokers, for the purposes of this discussion) on an anonymous, or blind, basis. The brokers were required to submit data on GCF Repos to GSCC shortly after trade execution.

Brokers could submit GCF Repo transactions in amounts of up to $2 billion, compared with the delivery maximum of $50 million for each non-GCF, or DVP, repo trade (although there was no limit on the number of non-GCF trades that could be submitted). In addition, brokers had the ability to submit data for both the repo and the reverse repo sides of a trade using a single screen. The dealer counterparties would automatically be locked in to the trades submitted by the brokers unless they specifically said, within set time limits, that they had not conducted the trade. Standardized, generic CUSIP numbers requested from Standard & Poor's Corporation, to be utilized by GSCC exclusively for GCF Repo processing, were used to specify the acceptable type of underlying eligible collateral. (Initially, GCF Repo service participants were limited to trading in a single generic CUSIP, encompassing only Treasury securities with not more than ten years remaining to maturity. In September 1999, GSCC added a second generic CUSIP for all Treasury securities.)

Soon after a predetermined trading cutoff, GSCC conducted an afternoon net exclusively for GCF Repo activity, combining each dealer's carryover activity and new GCF Repo activity to establish a single net receive or net deliver position in each generic CUSIP. For each such CUSIP, a dealer member was either a net securities borrower (money lender) or a net securities lender (money borrower), or it netted flat. The Bank

[40] See letter dated February 14, 1997, from the author to each repo netting participant. In a letter dated July 18, 1997, the author wrote again to each repo netting participant to make clear that GSCC's guarantee would hold even if the broker executed the transaction with one dealer counterparty but did not, by the end of the day, have a matching, offsetting transaction with another dealer netting member. In 2000, real-time trade matching for government securities transactions was introduced. After that, GSCC changed its policy to guarantee a repo transaction upon its comparison.

[41] Initially, participants were not permitted to submit repos with right of substitution. With the new facility, participants could specify rights of substitution using a new screen input facility (rather than via telephone or fax), thereby providing an automated audit trail for those rights. GSCC placed itself in the middle of all substitutions, with all collateral deliveries (of both old and replacement collateral) being due to or due from GSCC. GSCC also kept track of the final money for each repo throughout its life, regardless of the number of substitutions.

[42] See Self-Regulatory Organization, Government Securities Clearing Corporation, Securities Exchange Act Release Nos. 40057 (June 2, 1998, notice of filing) and 40623 (October 30, 1998, approval order). In its approval order, the SEC stated its belief that "the use of the GCF Repo service should reduce exposure to counterparty default, increase payment netting, and apply advanced clearing and risk management practices to the market in general collateral repos."

[43] After GSCC filed in 1998 for approval to commence the GCF Repo service, the Chicago Board of Trade submitted a comment letter to the SEC stating that the GCF Repo service might be subject to the exclusive jurisdiction of the Commodity Futures Trading Commission (CFTC) and that if any of the repos that would be cleared through the service were futures, then the CFTC would have exclusive jurisdiction over the service even if the repos also were securities. The CFTC's Division of Trading and Markets ultimately advised the SEC that it had completed its review of the GCF Repo service and had determined that it had no further comment on the service or GCF Repo transactions.

of New York and J.P. Morgan Chase provided the mechanism for allowing a chain of simultaneous collateral and cash movements to occur between GSCC and its dealer members and for allowing those securities to be available for various purposes, including tri-party repo processing and bank loans.[44] All positions were reversed on the morning of the next business day prior to the opening of the securities Fedwire. Term repos thus were collateralized by cash between a morning reversal and an afternoon settlement.

GSCC became a counterparty for settlement purposes to each dealer that was party to a GCF Repo transaction, and guaranteed the settlement of GCF Repos upon receipt of trade data. In order to do so prudently, GSCC used the same risk management protections for the GCF Repo service that it had in place for its non-GCF activity, including the collection of margin and the receipt and pass-through of mark-to-market amounts.

[44] When the GCF Repo service was introduced in 1998, participants were limited to intrabank trading; in other words, dealers could engage in GCF Repo trading only with other dealers that used the same clearing bank. This allowed each bank to transfer collateral without the need to involve the other bank or use Fedwire. In June 1999, GSCC broadened the service to allow for the trading of GCF Repos on an interbank basis, meaning that a participating dealer could engage in GCF Repo trading with any other participating dealer, even if it used a different clearing bank. See Self-Regulatory Organizations, Government Securities Clearing Corporation, Securities Exchange Act Release Nos. 41002 (February 5, 1999, notice of filing) and 41303 (April 16, 1999, approval order).

The GCF Repo service provided important benefits to participants in the U.S. government securities market. It gave dealers an additional borrowing source (other than tri-party and DVP repo), bringing greater depth to the general collateral marketplace. It also led to increased liquidity, lower costs, more efficient collateral allocation, reduced operational costs, and improved safety—in large part because GCF Repos, like DVP repos, were guaranteed when compared (generally within minutes of a trade), thus eliminating intraday counterparty credit risk.

Unlike the quick success seen with the basic repo netting service, participation in the GCF Repo service was initially disappointing. Indeed, by 2000 GSCC began to consider closing the service. However, some market participants suggested that the service would grow once other products were made eligible. Ultimately, that is what happened. On January 24, 2000, agency securities became eligible for the GCF Repo service. On March 20, 2000, the first mortgage-backed securities (MBS)—Fannie Mae and Freddie Mac fixed-rates—were made eligible. With the introduction of these new products (particularly MBS, which were much more readily available to dealers for allocating to GCF Repo lending than Treasury securities or agencies), the service took off.

In 2002, GSCC merged with the MBS Clearing Corporation to form the Fixed Income Clearing Corporation, a subsidiary of the Depository Trust & Clearing Corporation. In the decade and a half since, the combined business has evolved into what is arguably the largest and most significant clearing corporation in the world.

Olivier Armantier, Giorgio Topa, Wilbert van der Klaauw, and Basit Zafar

An Overview of the Survey of Consumer Expectations

- The New York Fed's Survey of Consumer Expectations (SCE) gathers information on consumer expectations regarding inflation, household finance, the labor and housing markets, and other economic issues.
- Launched in 2013 and fielded monthly, the survey aims to help researchers and policymakers understand how expectations are formed and how they affect consumer behavior—behavior that, in aggregate, helps drive macroeconomic activity and hence has implications for monetary policy.
- This article explores the survey's history, format, and question construction; details the procedure for selecting the rotating panel of survey participants; and describes the methods used to calculate statistics and disseminate results.
- The authors also outline the benefits of soliciting "density forecasts," which measure uncertainty, in addition to simple "point forecasts," and explain the rationale for seeking expectations on "inflation" rather than "prices."

1. Introduction

The importance of compiling high-quality data on the expectations held by economic agents has been increasingly recognized in both academic research and policymaking. Most economic decisions involve uncertainty and should therefore take into account not only preferences but also expectations about the future. Expectations should drive a variety of economic choices made by households, including those related to saving, investment, purchases of durable goods, wage negotiations, and so on. The aggregation of these choices in turn determines macroeconomic outcomes, including realized inflation, in equilibrium. Given the role of households in aggregate as an important driver of economic activity, the monitoring and management of consumers' expectations have become primary goals of policymakers and central components of modern monetary policy (Woodford 2004; Bernanke 2007; Gali 2008; Sims 2009).[1] The effectiveness of monetary policy and central bank communication relies on longer-run inflation expectations being well-anchored, making the measurement of these expectations important for policymakers. More generally, consumers' expectations about a number of personal and

[1] In particular, Bernanke (2004) argued that "an essential prerequisite to controlling inflation is to control inflation expectations."

Olivier Armantier is an assistant vice president, Giorgio Topa a vice president, Wilbert van der Klaauw a senior vice president, and Basit Zafar a former research officer at the Federal Reserve Bank of New York.

olivier.armantier@ny.frb.org
giorgio.topa@ny.frb.org
wilbert.vanderklaauw@ny.frb.org
basit.zafar@asu.edu

The views expressed in this article are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

To view the authors' disclosure statements, visit https://www.newyorkfed.org/research/author_disclosure/ad_epr_2017_overview-of-sce_armantier.

macroeconomic outcomes are increasingly useful inputs into a variety of forecasting models.

Effective monitoring and management of expectations requires the measurement of consumer expectations and an understanding of how these expectations are formed. The Federal Reserve Bank of New York's Survey of Consumer Expectations (SCE) was developed with precisely these goals in mind. The SCE collects timely, high-frequency information about consumer expectations and decisions on a broad variety of topics. Its overall goal is to fill the gaps in existing data sources (such as the University of Michigan Survey of Consumers, the Federal Reserve Board's Survey of Consumer Finances, and the Bureau of Labor Statistics' Consumer Expenditure Survey) pertaining to household expectations and behavior by providing a more integrated data approach. The SCE aims to cover a broad range of economic outcomes, including inflation, household finance, the labor market, and the housing market, as well as special topics as the need arises for policy or research analysis.[2]

The SCE is designed as a rotating panel, which enables researchers and policymakers to follow the same individuals over time, reducing changes in the sample's composition and thus minimizing sampling volatility in survey responses from month to month. The panel structure of the SCE is designed to provide valuable input into the evaluation of the economic outlook and policy formulation, and to offer an important resource for the research community to both increase its understanding of how consumers form and update their expectations and better assess the links between expectations and behavior. For instance, the data allow one to study how expectations about house prices and interest rates affect consumers' choices regarding buying or renting a home or regarding the type of mortgage used to purchase a home. Data on expectations about the likelihood of finding a job and about future wage earnings may be used to analyze workers' job search behavior or retirement decisions. Researchers may also study how inflation expectations shape consumers' spending and saving behavior. Collecting such data for the same households over time enables researchers to study potential interactions between decisions and expectations across many different domains of consumer behavior. Finally, the data can be used to study the evolution of a diverse but related set of expectations and the way they co-vary over time at the individual level, and to identify any (structural) breaks in this relationship.

A key feature of the SCE is its reliance on a probabilistic question format to elicit the likelihood respondents assign to different future events. In addition to questions asking respondents for *point forecasts*—for example, in the case of year-ahead inflation, *What do you expect the rate of [inflation/deflation] to be over the next twelve months?*—for several continuous outcomes, we ask for *density forecasts*—that is, the percent chance the respondent assigns to different future possible values of that variable. In the case of future inflation, for example, respondents are asked for the likelihood that future inflation will fall within different prespecified intervals. These density forecasts allow respondents to express uncertainty regarding their expectations. For binary outcomes (where the event either occurs or does not occur), eliciting the percent chance associated with the event fully identifies the underlying subjective distribution. By obtaining density forecasts, the SCE extends a practice with a longer tradition in the field of psychology and in surveys of professional forecasters, economists, and other financial experts. Our approach also builds on a large and growing body of economic research, led by Charles Manski, that has demonstrated survey respondents' willingness and ability to answer questions expressed in this way.

Finally, the SCE is conducted as a monthly internet survey in order to provide more flexibility in terms of question design and more real-time capabilities for data collection. An internet platform enables the researcher to design more user-friendly questions, with the help of visual aids and other tools that make it easier for respondents to understand and answer a specific question. An internet platform also makes it easier to develop and field new questions on special topics at short notice. For example, we designed and fielded special surveys to help assess how consumers' spending and inflation expectations responded to sudden large declines in gas prices and to elicit beliefs regarding the early impact of the Affordable Care Act on future healthcare spending, prices, and coverage.

[2] National surveys of public (inflation) expectations are now conducted in multiple countries. In the United States, these include the University of Michigan Survey of Consumers, the Livingston Survey, the Conference Board's Consumer Confidence Survey, and the Survey of Professional Forecasters. Other central banks that survey consumers about their inflation expectations include the Bank of England, the European Central Bank, the Bank of Japan, the Reserve Bank of India, and the Sveriges Riksbank. Since 2015, the Bank of Canada has conducted a largely comparable version of the SCE, fielded at a quarterly frequency. Since 2013, the Federal Reserve Board has conducted the Survey of Household Economics and Decisionmaking (SHED), which elicits some expectations about the economic well-being of U.S. households.

1.1 Survey Overview

The SCE started in June 2013, after a six-month initial testing phase.[3] It is a nationally representative, internet-based survey of a rotating panel of about 1,300 household heads, where household head is defined as the person in the household who owns, is buying, or rents the home. The survey is conducted monthly. New respondents are drawn each month to match various demographic targets from the American Community Survey (ACS), and they stay on the panel for up to twelve months before rotating out. The survey instrument is fielded on an internet platform designed by the Demand Institute, a nonprofit organization jointly operated by the Conference Board and Nielsen. The respondents for the SCE come from the sample of respondents to the Consumer Confidence Survey (CCS), a mail survey conducted by the Conference Board. In turn, the respondents for the CCS are selected from the universe of U.S. Postal Service addresses. From that universe, a new random sample is drawn each month, stratified only by Census division.

The SCE has several components. First, it includes a core monthly module on expectations about a number of macroeconomic and household-level variables.[4] In this module, respondents are asked about their inflation expectations, as well as their expectations regarding changes in home prices and the prices of various specific spending items, such as gasoline, food, rent, medical care, and college education. The core survey also asks for expectations about unemployment, interest rates, the stock market, credit availability, taxes, and government debt. In addition, respondents are asked to report their expectations about several labor market outcomes that pertain to them, including changes in their earnings, the perceived probability of losing their current job (or leaving their job voluntarily), and the perceived probability of finding a job. Finally, the core survey asks about the expected change in respondent households' overall income and spending. As described in more detail below, these questions about expectations are fielded at various time horizons and with various formats, including both point and density forecasts.

Second, each month, the SCE contains a supplementary "ad hoc" module on special topics. Three such modules are repeated every four months, leaving three "floating" supplements per year on topics that are determined as the need arises. The three repeating supplements are on credit access, labor market, and spending. Topics covered so far in the "floating" supplement include (but are not limited to) the Affordable Care Act, student loans, workplace benefits such as childcare and family leave, and the use of insurance products. Together, the core monthly module and the monthly supplement take about fifteen minutes to complete.

Finally, SCE respondents also fill out longer surveys (up to thirty minutes in length, and separate from the monthly survey) each quarter on various topics. Most of these surveys are repeated at a yearly frequency. Since each SCE panelist stays in the panel for up to twelve months, these annual surveys can be used as independent repeated cross sections, although they obviously can be linked to the monthly core survey panel responses. The SCE currently contains quarterly surveys on the housing market, the labor market, informal work participation, and consumption, saving, and assets. A subset of these surveys is designed in part or wholly by other Federal Reserve Banks.

2. Questionnaire Design

The 1990s represented a period of significant change in the way economists elicit expectations through surveys. Traditionally, researchers have measured expectations through verbal and qualitative questions, asking respondents whether they expect that an event will occur or not, or asking whether they think it is "very likely," "fairly likely," "not too likely," or "not at all likely" that a specific event will occur. In addition to the limited information captured owing to the coarseness of choice options or to the lack of means for expressing uncertainty altogether, a major drawback of this traditional approach concerns the lack of inter- and intrapersonal comparability of responses.

Led and inspired by the work of Charles Manski, who in turn built on the early work of Juster (1966) and a longer tradition of collecting such data in cognitive psychology, economists began to elicit probabilistic expectations during the 1990s. It quickly became clear that, with some guidance, survey respondents are able and willing to answer probabilistic expectations questions.[5] Moreover, with a fixed numerical scale, responses are interpersonally comparable and have been found to be better predictors of outcomes. A growing number of large-scale surveys in the United States and abroad now use probabilistic formats to elicit expectations for a wide

[3] As discussed below, the SCE was preceded by an extensive feasibility study over the period 2006-13, using experimental surveys on the RAND Corporation's American Life Panel. Early findings from that study are described in van der Klaauw et al. (2008).

[4] Each entering cohort is also administered a module with demographic questions about the respondent and the respondent's household.

[5] Usually before such questions are asked, respondents are provided with a brief introduction or explanation of basic probabilistic ideas through examples.

range of events, and these formats have also been successfully implemented in surveys associated with laboratory and field experiments, including several conducted in less developed countries. Manski (2004) reviews research eliciting probabilistic expectations in surveys and assesses the state of the art at that time. Delavande (2014) and Delavande, Giné, and McKenzie (2011) provide a more recent review.

The questionnaire design of the SCE builds on this previous research and was informed, in large part, by our experiences with the Household Inflation Expectations Project (HIEP), wherein we fielded surveys every six weeks on RAND's American Life Panel (ALP) going back to 2007 (Bruine de Bruin et al. 2010a). The HIEP questionnaires were developed in collaboration with RAND, other Federal Reserve Banks, academic economists, psychologists, and survey design experts. During the period 2006 to 2012, the HIEP conducted in-depth cognitive interviews, fielded psychometric surveys on the ALP as well as on various "convenience" (that is, nonrandom) samples, and administered a number of experimental consumer surveys in the ALP. In addition to testing probabilistic question formats, we also experimented with alternative wording of questions, especially those related to inflation (discussed in more detail below). The findings from this project formed the basis for the creation of the SCE (see van der Klaauw et al. 2008; Armantier et al. 2013).

During the SCE's experimental or development phase, between December 2012 and June 2013, we further sharpened and tested the questionnaire. The process involved conducting additional cognitive interviews with a small auxiliary sample to identify potential interpretations of the questions. After interviewees read the questions out loud, they were instructed to think out loud while generating their answers. This step allowed us to gauge whether, in fact, interviewees interpreted the questions the way we had intended them to. When necessary, we modified the wording of questions accordingly. Pilot surveys were also conducted and we analyzed the resulting data to make sure the questions were eliciting meaningful variation.

Of particular importance were our findings in the SCE questionnaire development phase regarding the elicitation of expectations. In both the HIEP and the SCE experimental phase, we tested a large set of probabilistic questions. Respondents in our surveys showed a consistent ability and willingness to assign a probability (or "percent chance") to future events. Unlike simple point forecasts, probabilistic expectations of binary outcomes as well as density forecasts for continuous outcomes provide a valuable measure of individual uncertainty. Moreover, we found these responses to be largely internally consistent in terms of simple laws of probabilities. Finally, we conducted an experiment that confirmed that the densities elicited were informative (Armantier et al. 2013). For a review of this work, see van der Klaauw et al. (2008) and Bruine de Bruin et al. (2010a, 2011).

Having discussed the SCE questionnaire development process, we now turn to the actual questions, focusing on the SCE core survey. The quantitative questions can be broadly divided into three categories: (1) questions that elicit expectations of binary outcomes (such as the likelihood that the U.S. stock market will be higher in twelve months), (2) questions that elicit pointwise expectations for continuous outcomes (such as the rate of inflation over the next twelve months), and (3) questions that elicit respondents' probability densities for forecasts of continuous outcomes. Besides these questions, the survey also includes some qualitative questions, including those in which respondents are asked to answer using a (for example) seven-point rating scale. The full questionnaire is available on the New York Fed's website.[6]

The bulk of the survey elicits near-term expectations, that is, expectations regarding outcomes over the next twelve months, the next three months, or both. The ad hoc supplements also contain questions at the four-month horizon. However, in certain instances, such as inflation or home prices, expectations are also elicited for the medium term—that is, three years out.

We next describe the rationale for how we elicit some of these expectations.

2.1 Eliciting Expectations of Continuous Outcomes

For some continuous outcomes, we elicit both point and density forecasts. We begin here by illustrating the format of our point forecast questions, and we describe the density forecast questions in Section 2.3.

Inflation Expectations

Inflation expectations are elicited using the following two-stage format. Respondents are first asked: *Over the next twelve months, do you think that there will be inflation or deflation? (Note: Deflation is the opposite of inflation).*

[6] https://www.newyorkfed.org/medialibrary/interactives/sce/sce/downloads/data/FRBNY-SCE-Survey-Core-Module-Public-Questionnaire.pdf.

Depending on their answer to this question, respondents are next asked for a point estimate: *What do you expect the rate of [inflation/deflation] to be over the next twelve months? Please give your best guess.*

Note that we directly ask respondents for the rate of "inflation." This differs from the widely used approach of avoiding the term "inflation" in consumer surveys. Most existing inflation expectations questions, such as those posed in the Michigan Survey (Curtin 2006), ask for expected changes in "prices" as follows: *During the next twelve months, do you think that prices in general will go up, or go down, or stay where they are now?,* with the response options "Go up," "Stay the same," and "Go down." Those who respond "go up" or "go down" are then asked to give a specific point estimate.

We prefer asking respondents for their expectations about "inflation" because our prior research in the HIEP suggests that the way the Michigan Survey question is phrased induces mixed interpretations, with some respondents thinking about specific prices they pay and others thinking about the overall rate of inflation. As the former tend to think more of salient price changes, they are more likely to provide extreme responses (Bruine de Bruin, van der Klaauw, and Topa 2011). We found that asking directly about the rate of inflation reduces ambiguity in interpreting questions, and yields, we believe, more reliable, more interpersonally comparable responses. Moreover, this approach is consistent with the concept of forward inflation expectations, which are of interest to central banks.

One source of hesitation in asking consumers about inflation may be the concern that inflation is a relatively complex concept. However, evidence suggests that consumers tend to have a basic understanding of what the term means (Leiser and Drori 2005; Svenson and Nilsson 1986). Cognitive interviews conducted during the HIEP similarly indicate that the vast majority of consumers have a good understanding of the concept of inflation (van der Klaauw et al. 2008; Bruine de Bruin et al. 2012). Furthermore, our research indicates that consumers act on their reported inflation expectations in sensible ways (Armantier et al. 2015) and that they update their inflation expectations meaningfully when provided with arguably inflation-relevant information (Armantier et al. 2016).

In addition, during the period May 2013 to September 2015, we asked SCE survey respondents the following question: *On a scale of 1 to 7, how well would you say you understand what "inflation" means? (where 1 means "I don't know what 'inflation' means," and 7 is labeled as "I know exactly what 'inflation' means").* Of the 5,182 first-time respondents to whom this question was posed, only 44 (less than 0.9 percent) chose 1 (no understanding) on the scale. In fact, 82 percent of respondents chose 5 or higher, suggesting that inflation is a fairly well-understood concept. This, of course, does not mean that households find it easy to express inflation in quantitative terms. To investigate this, we asked our survey respondents the following: *On a scale of 1 to 7, how easy is it for you to express the rate of inflation as a number? (where 1 is "Very easy" and 7 is "Very difficult").* Of the 5,179 respondents, 82 percent reported 5 or lower, with just 5.6 percent choosing 7 on the scale. Overall, these results offer convincing evidence that the vast majority of consumers understand the concept of inflation and are able to express it numerically.

Expectations for Other Continuous Outcomes

We next present the wording for one other representative set of questions, pertaining to earnings expectations for those who are employed. Respondents are asked:

> *Please think ahead to twelve months from now. Suppose that you are working in the exact same job at the same place you currently work, and working the exact same number of hours. What do you expect to have happened to your earnings on this job, before taxes and deductions?*
>
> ○ *increase by 0 percent or more*
> ○ *decrease by 0 percent or more*

This first part is then followed by: *By about what percent do you expect your earnings to have [increased/decreased]?*

Note that respondents are not presented with a "stay the same" option in the first part of the question. Instead, the instructions specify that they can enter a zero change by picking either the increase or the decrease option and then entering zero in the second part of the question. When experimenting with the wording of questions, we found that a substantial proportion of respondents would choose "stay the same" when that option was available. Upon questioning respondents further, we found that a substantial proportion of those who had chosen "stay the same" had changes in mind that were bigger than 0.5 percent in magnitude, suggesting that the mere availability of the "stay the same" option was leading some respondents to select it. In our analysis, we found that the distribution of responses to a given question with and without the "stay the same" option was noticeably different. More research is clearly

needed to understand which elicitation method is best at recovering the true underlying subjective belief. However, given that respondents can always enter zero and hence we arguably elicit more information without the "stay the same" option, we chose to ask for changes without the "stay the same" option.

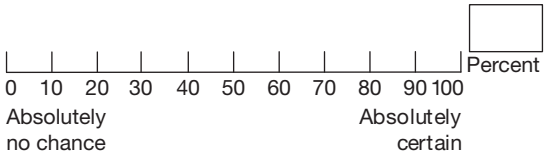
In addition to the questions we just discussed that ask for point forecasts for inflation and earnings, we use this same question format for a wide range of other outcomes.

2.2 Eliciting Expectations of Binary Outcomes

The survey includes several probabilistic questions that elicit the likelihood (or percent chance) of a certain event. These questions are preceded by some instructions regarding the use of percentages:

> *In some of the following questions, we will ask you to think about the percent chance of something happening in the future. Your answers can range from 0 to 100, where 0 means there is absolutely no chance, and 100 means that it is absolutely certain. For example, numbers like 2 and 5 percent may indicate "almost no chance"; 18 percent or so may mean "not much chance"; 47 or 52 percent may be a "pretty even chance"; 83 percent or so may mean a "very good chance"; 95 or 98 percent may be "almost certain."*

For example, those who are unemployed and actively looking for work are asked: *What do you think is the percent chance that within the coming twelve months, you will find a job that you will accept, considering the pay and type of work?*



Respondents can either enter a number (on a scale of 0 to 100) directly into the box (see above) or click anywhere along the sliding scale. To prevent respondents from anchoring their response, no marker appears on the scale until the respondent clicks somewhere on it.

2.3 Eliciting Forecast Densities

Relative to existing surveys of consumer expectations, one of the SCE's innovations is that it also elicits consumers' subjective probability distributions for certain continuous outcomes, such as future inflation, earnings, and home prices. These density data allow us to construct individual measures of central tendency (for example, the density mean or median), uncertainty, and perceived tail risks (such as the probability of extreme positive or negative outcomes).

Our density questions follow a format similar to that of the Survey of Professional Forecasters and the Bank of Italy's Survey of Household Income and Wealth. Respondents are presented with various predefined, non-overlapping bins that exhaust the entire range of values that the random variable may take, and are then asked for the percent chance that the variable would take values in each of those intervals, with the reminder that numbers need to add up to 100 percent. The density forecast for year-ahead national home price changes, for example, is elicited as follows:

> *And in your view, what would you say is the percent chance that,* ***over the next twelve months****, the average home price nationwide will . . .*

Increase by 12 percent or more	_____	*percent chance*
Increase by 8 to 12 percent	_____	*percent chance*
Increase by 4 to 8 percent	_____	*percent chance*
Increase by 2 to 4 percent	_____	*percent chance*
Increase by 0 to 2 percent	_____	*percent chance*
Decrease by 0 to 2 percent	_____	*percent chance*
Decrease by 2 to 4 percent	_____	*percent chance*
Decrease by 4 to 8 percent	_____	*percent chance*
Decrease by 8 to 12 percent	_____	*percent chance*
Decrease by 12 percent or more	_____	*percent chance*
Total	**XXX**	

As respondents enter their answers, they can see the running total. Respondents who nevertheless give answers that do not add up to 100 percent receive the notice *Please change the numbers in the table so they add up to 100.*

We use each individual's responses to the probabilistic questions to parametrically estimate the underlying forecast density function, following Engelberg, Manski, and Williams (2009). We describe this estimation in more detail in Section 5.2. Using the probability density function for each respondent, we compute corresponding density means and medians. Further, we use the density interquartile range (IQR)—the difference

between the third and the first quartiles—as a measure of individual forecast uncertainty. We choose this measure because the IQR is less sensitive than, say, the standard deviation to small variations in the tails of the estimated density.

3. Implementation

3.1 Sample Design

The SCE sample design is based on that of the Conference Board's monthly Consumer Confidence Survey (CCS). The CCS is a mail survey that uses an address-based probability sample design to select a new random sample each month based on the universe of U.S. Postal Service addresses.[7] The universe of addresses is derived from the files created by the U.S. Postal Service and represents near-universal coverage of all residential households in the United States. It is updated monthly to ensure up-to-date coverage of U.S. households. The targeted responding CCS sample size is approximately 3,000 completed questionnaires each month from household heads. Questionnaire instructions define household head as *the person in your household who owns, is buying, or rents this home.*[8]

The SCE sampling frame (or sampling population) consists of CCS respondents who expressed an ability and a willingness to participate in the SCE based on their answers to two questions included at the end of the CCS questionnaire. The first asks: *Do you have access to the internet and an email address?* Those who answer "Yes" are then asked:

> *You may be eligible to participate in a survey about your perceptions of the economy, employment, finances, and related topics. This is a paid survey that would be conducted monthly for up to twelve months. You would receive $15 for each completed survey.*[9] *If selected, we would email you a web address where you could respond to the online questionnaire. Would you be interested in participating in this monthly, paid survey?*
>
> *Yes, my email address is:_____*
> *No, I am not interested in participating*

As discussed in more detail below, on average, 53 percent of CCS respondents in a given month express a willingness to participate in a new online survey.[10] Of those who are interested in participating, approximately 300 to 320 are invited within the following two months to join the SCE internet panel, of whom about 150 to 180 actually end up joining. A stratified random sampling approach is used to draw new CCS respondents into the SCE, with strata based on income, gender, age, race/ethnicity, and Census division,[11] and weights are chosen to maximize the representativeness of the SCE panel.

3.2 Data Collection

The goal of the survey is to capture consumers' expectations over a given month. To do so, the survey is sent to respondents in three batches throughout the month. Specifically, each month, the pool of respondents is partitioned into three batches of roughly equal size. In general, the first, second, and third batches receive an email invitation to fill out the survey on the second, eleventh, and twentieth of the month, respectively. On occasion, this schedule is amended by a day or two to reflect holidays or shorter months (that is, February). If they have not yet completed the survey, respondents in each batch receive two reminders by email, three and seven days after their initial invitation. On rare occasions, a third reminder is sent to the first and second batches on an ad hoc basis (for example, if the response rate is perceived to be lower than usual). Survey responses for all three batches are collected until the last day of the month.

In 2014, the median respondent in each batch completed the survey three days after receiving the initial invitation. In Chart 1, we plot the number of surveys completed each day

[7] The CCS random sample of household addresses is drawn after first stratifying geographically by nine Census divisions to provide a proportionate geographic distribution. To ensure proportional representation in the sample of respondents, the CCS uses weights based on gender, income, geography, and age.

[8] This definition is similar to that used in the Current Population Survey, in the ACS, and by the Census more generally: there, the "reference person" in the household (or the "householder") is the person who owns or rents the unit of residence. Note that the instructions state that if that person is not available or is unable, an adult aged eighteen or older who lives in the household should complete the survey. In a representative month (March 2013), 96 percent (95 percent weighted) of CCS respondents were household heads, with very little variation across gender, age, income, and race/ethnicity. Note also that this definition does not exclude the possibility that a household may have multiple "co-household heads."

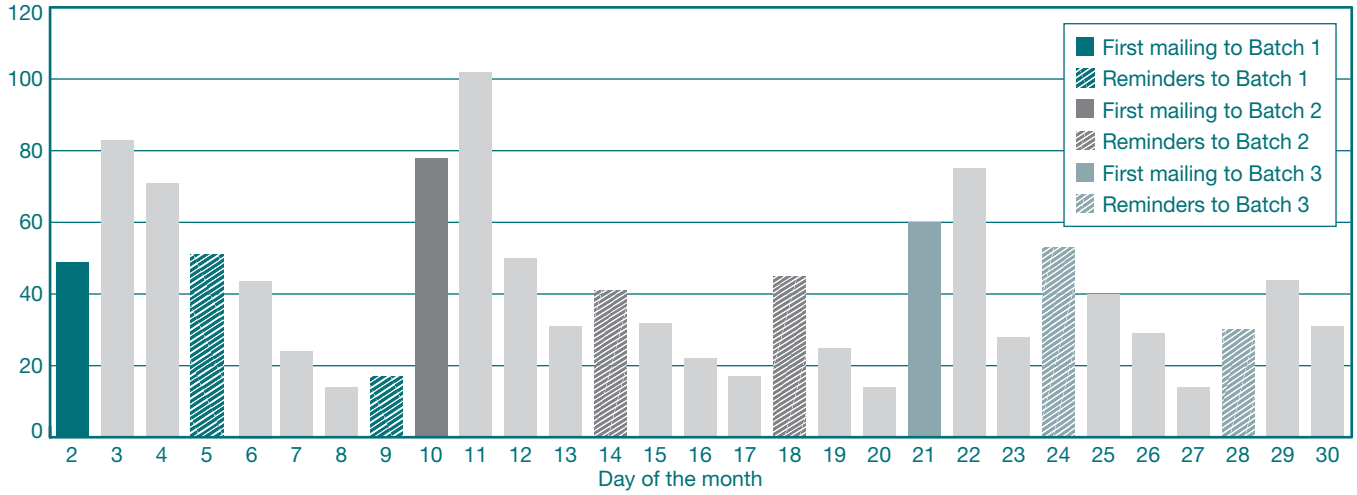
[9] As explained later in this section, before July 2013, some respondents received a letter stating a different amount.

[10] The monthly average of 53 percent was based on CCS responses during the period December 2012 to September 2015.

[11] We distinguish among eight household income groups, five age groups, five race and ethnicity groups, and nine Census divisions.

Chart 1

Number of Surveys Completed during April 2016, by Day



Source: New York Fed Survey of Consumer Expectations.

Notes: The full bars in dark blue, gray, and light blue represent the day on which respondents from batches 1, 2, and 3, respectively, are invited to fill out the survey. The shaded bars in dark blue, gray, and light blue represent the day on which respondents from batches 1, 2, and 3, respectively, receive a reminder to complete the survey.

of a typical month (April 2016).[12] Although not uniformly distributed, the completion of surveys is spread out throughout the month, with three major peaks on the days after each batch receives its invitation to fill out the survey, and smaller peaks on the days on or after which each batch receives a reminder to fill out the survey.

Each month, the panel of household heads invited to answer the survey consists of roughly 300 new respondents and 1,100 "repeat" respondents (that is, respondents who have completed at least one survey within the past eleven months). The new respondents invited to answer the survey for the first time are randomly allocated to one of the three batches. A few days before they are to receive the invitation to fill out the survey, new respondents are contacted by mail and by email to welcome them to the panel. These letters inform the respondent about the nature, the number, the duration, and the timing of the surveys they will be asked to complete over the next twelve months.[13] The new respondents are also told about the payment they will receive for each survey completed, and they are given access to a website where they can find additional information and ask questions of the help desk.

At the beginning of each month, repeat respondents (that is, respondents who have already completed at least one survey in the past) are partitioned into two groups: the "skippers" (those who failed to complete the survey in the previous month) and the "nonskippers." The wide majority of repeat respondents are nonskippers (93 percent in 2014). Skippers are assigned randomly to one of the three batches. The assignment procedure for nonskippers is designed so that (1) there are an equal number of nonskippers in each batch, and (2) nonskippers in each batch have (roughly) the same average number of days between the completion of two consecutive surveys. On the first of each month, nonskippers are ranked according to the number of days since they completed the survey in the previous month and are partitioned into terciles. The first tercile (that is, the respondents who completed

[12] The chart in Appendix A shows the mean and 25th, 50th, and 75th percentiles of the daily frequency of responses over all months from December 2012 to September 2015, thus combining months with different dates for the invitations and reminders.

[13] We experimented with sending a welcome email only (and no postal mail) to the new respondents. However, that approach led to a noticeable decline in the response rate, suggesting that the welcome mail lent greater credibility to the survey. Thus, we reverted to new respondents receiving both a welcome mail and a welcome email.

the survey most recently) is assigned to batch 3, the second to batch 2, and the third to batch 1.[14]

Any respondent invited after July 2013 has been paid $15 for each monthly survey completed. We settled on this amount after testing whether the amount paid for each completed survey affected the response rate. Specifically, we had three groups of respondents between December 2012 and July 2013. During their twelve-month tenure, each group was randomly assigned to be paid $10, $15, or $20 for each survey completed. The response rate in the first month was 61 percent, 66 percent, and 56 percent in the $10, $15, and $20 group, respectively. Further, 28 percent, 37 percent, and 32 percent of the respondents in the $10, $15, and $20 group (respectively) completed all twelve surveys.[15] Thus, we concluded that a payment of $15 per survey was the most cost-effective.

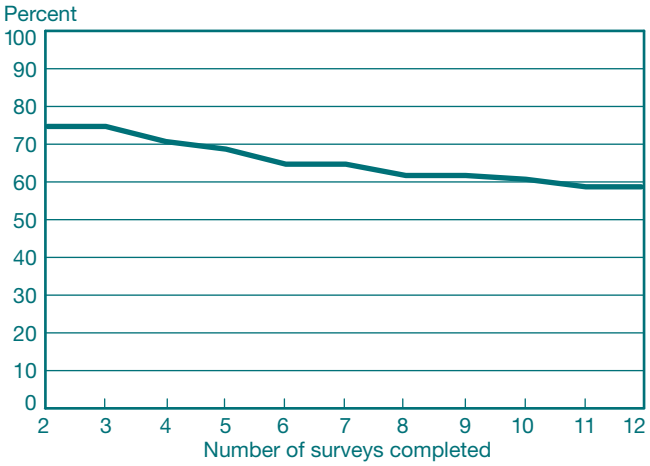
Respondents can be removed from the panel if they fail to respond to the monthly survey invitations. This is the case in particular for respondents who do not complete the first survey they are invited to fill out. Otherwise, if a respondent does not complete the monthly survey in three consecutive months, the respondent is dropped from the panel and no longer invited to fill out any additional surveys. Twelve months after completion of their first survey, every respondent is rotated out of the panel.

We now turn to the issue of survey participation. Most of the nonresponse occurs in the first month. Out of the 3,582 household heads we invited to participate in the survey in 2014, 1,647 (or 46 percent) failed to complete the first survey and were therefore not invited again. Once a respondent is in the panel, however, attrition drops rapidly. Indeed, we can see in Chart 2 that while 26 percent of first-time respondents failed to complete a second survey, the response rate after the second month is essentially flat. In particular, observe in Chart 2 that 58 percent of the respondents who entered the panel in 2014 completed all twelve surveys.

[14] Prior to February 2016, the allocation procedure for nonskippers was also applied to skippers. As a result, skippers were found predominantly in batch 1 (because skippers had completed their last survey more than thirty days earlier). Because skippers may have specific unobserved characteristics, we were concerned that the response rate and the survey responses from batch 1 would be different from those of batches 2 and 3. Thus, we decided to allocate skippers randomly across the three batches.

[15] The lower response rate for the $20 group may stem from the fact that these respondents were pulled from an older CCS sample.

Chart 2
Response Rate for Respondents who Entered the Panel in 2014



Source: New York Fed Survey of Consumer Expectations.

4. Panel Representativeness

The representativeness of our panel of respondents depends on a number of factors, including the composition of: (1) the sample of CCS respondents who report having access to the internet and email and who are willing to participate in our survey; (2) the sample of invited and interested CCS respondents who actually choose to enter our panel by completing their first SCE survey; and (3) the sample of SCE participants who continue to participate in our panel after entry.

As discussed earlier, the CCS target population is the U.S. population of household heads, with household head defined as the person who owns, is buying, or rents the home. As shown in the table in Appendix A, average characteristics of household heads who participated in the CCS during the period from October 2013 to September 2015 are largely comparable to those in the 2013 and 2014 American Community Surveys. The main difference in sample composition between the CCS and ACS concerns the age distribution, with younger household heads being somewhat underrepresented in the CCS and older household heads being overrepresented—a common feature of mail surveys.

The SCE sampling frame consists of CCS respondents who reported having access to the internet and email and who expressed a willingness to join a new online survey. Columns 1, 2, and 3 of Table 1 report the characteristics of,

Table 1

Sample Comparisons—CCS and SCE Survey Respondents

	Full CCS Sample ($N = 64,133$)	CCS Respondents with Internet and Email ($N = 50,089$)	CCS Respondents Who Consented ($N = 26,439$)	SCE Respondents ($N = 3,853$)
	(1)	(2)	(3)	(4)
	Percent			
Age				
Under 30	3.7	4.2	5.9	11.7
30–39	11.3	13.2	17.0	19.0
40–49	15.7	17.7	20.1	18.8
50–59	23.9	25.3	25.0	20.6
60 or over	45.5	39.5	31.9	29.9
Gender				
Female	47.7	47.0	47.9	48.1
Male	52.3	53.0	52.1	51.9
Income				
Less than $15,000	8.3	4.7	5.6	8.5
$15,000–$24,999	9.9	7.0	7.3	11.3
$25,000–$34,999	10.3	8.7	8.6	9.9
$35,000–$49,999	15.2	14.7	13.8	13.1
$50,000–$74,999	19.7	21.3	20.6	21.0
$75,000–$99,999	13.1	15.3	15.2	13.5
$100,000–$124,999	9.4	11.1	11.6	7.3
$125,000 or more	14.0	17.2	17.2	15.4
U.S. Census Division				
New England	5.0	5.2	4.7	4.3
Middle Atlantic	13.6	13.6	13.4	12.9
East North Central	17.6	17.1	17.3	14.4
West North Central	7.7	7.5	7.1	7.6
South Atlantic	19.9	20.1	20.8	20.4
East South Central	5.8	5.2	4.9	5.1
West South Central	9.2	9.3	9.2	11.4
Mountain	6.9	7.3	7.1	8.8
Pacific	14.5	15.2	15.4	15.1
U.S. Census Region				
Northeast	18.5	18.5	17.9	17.2
Midwest	25.2	24.6	24.6	22.0
South	34.9	34.5	35.2	36.9
West	21.4	22.3	22.5	23.9
Mean household size	2.4	2.5	2.6	2.5
Any child under 12	16.4	18.6	23.0	23.0
Race/Ethnicity				
Asian	3.3	3.6	3.5	3.5
Black	9.1	8.2	9.7	10.4
White	82.1	83.5	82.1	81.8
Other	4.1	4.1	4.9	4.4
Has internet	78.1	100.0	100.0	100.0
Is interested	41.2	52.7	100.0	100.0

Sources: New York Fed Survey of Consumer Expectations (SCE); Conference Board, Consumer Confidence Survey (CCS).

Note: Each number in the table is the percentage of the sample that falls into that category.

respectively, all CCS respondents, CCS respondents who reported having access to the internet and an email address, and the subset of those who indicated an interest in joining an online panel survey, during the period October 2013 to September 2015. As shown in the table, relative to CCS respondents overall, those with internet access were somewhat younger—with those over age 60 especially underrepresented—and marginally more likely to be male and white or Asian. Those with internet access also were more likely to have family incomes exceeding $50,000 and were more likely to have young children, to have slightly higher household incomes, and to reside in the western United States. Those who expressed an interest in joining an online survey had average characteristics very similar to those of the CCS respondents with internet access, but, compared with CCS respondents overall, were even more likely to be younger and to have a child under age twelve in the household.

Instead of demographic characteristics, Table 2 shows average responses to the standard set of CCS consumer sentiment questions. While differences are generally remarkably small compared with CCS respondents overall, those with internet access and email, on average, are slightly more positive and optimistic about current and future business conditions, job availability, and income, and expect slightly lower inflation. We find the same pattern for those interested in joining an online survey, except that the differences are slightly larger in magnitude. In terms of consumer sentiment, we find the pool of interested CCS respondents to be quite similar to CCS respondents overall.

Turning now to the SCE sample, as discussed earlier, in drawing a sample of new panel members each month from among those who expressed an interest in joining an online panel, we use a stratified sampling procedure that attempts to account for differential SCE survey participation and attrition rates across different demographic groups, in terms of income, gender, age, race/ethnicity, and Census division.[16]

Of those SCE volunteers newly invited to participate in the SCE, on average, 53 percent actually participate, with this proportion ranging between 48 percent and 60 percent during the period October 2013 to September 2015.[17] As shown in the fourth column of Tables 1 and 2, CCS respondents who end up participating in the SCE have (unweighted) demographic characteristics and consumer sentiment that are very similar to those of CCS volunteers (those who consent to being contacted for online surveys) and CCS respondents overall. Given that the pool of CCS respondents already is highly representative of the U.S. population of household heads (as shown in the table in Appendix 1), the similarity between SCE and CCS respondents indicates that our stratified sampling procedure in inviting CCS respondents is largely effective. This is further exemplified by the notable difference between the CCS and SCE samples in the age distribution of respondents. Reflecting the efficacy of our pre-stratification approach to inviting CCS consenting respondents, SCE participants are somewhat younger than CCS respondents and, in fact, have an age distribution of household heads that is very comparable to that in the ACS.

While the previous comparison is concerned with how SCE *entrants* compare with CCS participants, we finally assess the representativeness of SCE respondents overall. That is, how representative are SCE respondents in a typical cross section? The sample of SCE respondents each month, of course, reflects not only their initial recruitment into the panel but also their continued participation over time. Table 3 reports the means and standard deviations of the monthly average sample characteristics of SCE respondents during the period October 2014 to September 2015. The first column of the table shows that the average (unweighted) characteristics of respondents in the SCE are very similar to those of SCE entrants (shown earlier in column 4 of Table 1), but SCE respondents are slightly older and have slightly higher incomes, on average, reflecting differences in survey participation rates after entering the SCE panel. The relatively small standard deviations reported in the first column further indicate that the sample composition of SCE participants each month is highly stable over time. This, of course, is not surprising given that SCE respondents constitute a panel, with approximately 90 percent of respondents in a given month participating again in the following month.

As mentioned earlier, to account for any remaining differences between the SCE and ACS (for example, because of differential sample attrition or skipping behavior), we apply

[16] That is, in inviting SCE volunteers, we oversample not only those less likely to consent but also those less likely to accept our invitation and those who, once entered, are more likely to leave our panel through attrition or to occasionally skip surveys, before completing the twelve-month survey period.

[17] Newly invited SCE volunteers are only provided a one-time opportunity to join the SCE panel in the month for which they are first invited. Those who do not participate in the first month are no longer considered for future participation in the SCE. Note that the first-time participation rates listed include respondents with invalid or inactive email addresses.

Table 2

Sample Comparisons—CCS and SCE Survey Respondents

	Full CCS Sample ($N = 64{,}133$)	CCS Respondents with Internet ($N = 50{,}089$)	CCS Respondents Who Consented ($N = 26{,}439$)	SCE Respondents ($N = 3{,}853$)
	(1)	(2)	(3)	(4)
	Percent			
General business conditions in the area				
Good	23.2	24.8	25.1	24.9
Normal	54.5	54.2	53.7	53.7
Bad	21.9	20.5	20.7	21.0
General business conditions in the area in six months[a]				
Better	17.2	18.3	19.8	19.4
Same	71.1	70.6	69.1	70.0
Worse	11.2	10.7	11.0	10.4
Job availabilty in the area[a]				
Plenty	16.3	18.0	18.7	19.1
Not so many	54.8	55.4	53.9	53.6
Hard to get	27.9	25.6	26.6	26.6
Job availability in the area in six months[a]				
More	15.4	16.1	17.1	17.2
Same	67.2	67.8	66.1	67.1
Fewer	16.5	15.6	16.4	15.4
Family income in six months[a]				
Higher	14.3	16.6	19.9	22.3
Same	73.5	72.2	68.6	67.1
Lower	11.8	10.8	11.2	10.5
Increase in prices over the next twelve months[a]				
2 percent or lower	21.1	22.2	22.7	23.8
3–4 percent	31.0	32.4	31.7	31.7
5–6 percent	21.5	21.5	21.8	21.8
7 percent or more	25.6	23.4	23.2	22.4
Expected change in interest rates[b]				
Mean (percent)	3.7	3.8	3.8	3.8
Expected change in stock prices[b]				
Mean (percent)	3.1	3.1	3.1	3.1

Sources: New York Fed Survey of Consumer Expectations (SCE); Conference Board, Consumer Confidence Survey (CCS).

Note: Each number in the table is the percentage of the sample that falls into that category.

[a] Remainder category is the small proportion of missing or invalid responses.

[b] Averages for responses are based on a Likert scale, ranging from 1 (increase) to 5 (decrease). All statistics are based on CCS surveys from October 2013 to September 2015.

Table 3
SCE Sample Composition

	SCE Average (STDEV) Monthly Unweighted Sample Proportions	SCE Average Monthly Weighted Sample Proportions	2013 ACS Proportions
	(1)	(2)	(3)
		Percent	
Sample Size	24	24	—
Age			
Under 30	10.3 (1.2)	10.9	10.8
30–39	17.5 (0.5)	16.9	16.9
40–49	18.2 (0.6)	19.3	19.2
50–59	21.4 (1.3)	20.6	20.8
60 and over	32.6 (0.8)	32.3	32.4
Gender			
Female	47.4 (1.0)	50.0	49.9
Male	52.6 (1.0)	50.0	50.1
Education			
Up to high school	12.3 (1.0)	37.2	36.7
Some college	33.9 (1.8)	31.2	31.3
College graduate	53.8 (1.9)	31.5	32.0
Income			
Under $50,000	37.7 (1.2)	48.3	47.9
$50,000–$99,999	36.4 (1.3)	30.6	29.7
$100,000 or more	26.0 (1.4)	21.1	22.5
U.S. Census Division			
New England	4.4 (0.7)	4.4	4.8
Middle Atlantic	13.6 (1.5)	13.6	13.2
East North Central	15.7 (2.1)	16.1	15.5
West North Central	6.4 (1.3)	6.1	7.0
South Atlantic	20.0 (1.5)	20.6	19.7
East South Central	5.1 (0.4)	6.2	6.1
West South Central	10.2 (1.6)	10.7	11.5
Mountain	8.9 (0.6)	7.9	7.1
Pacific	15.7 (0.9)	14.5	15.1
U.S. Census Region			
Northeast	18.0 (2.1)	18.0	18.0
Midwest	22.1 (1.4)	22.2	22.5
South	35.3 (2.7)	37.4	37.3
West	24.6 (1.2)	22.4	22.2

Sources: New York Fed Survey of Consumer Expectations (SCE); U.S. Census Bureau, American Community Survey (ACS).

Notes: All SCE statistics are based on surveys from October 2013 to September 2015. Mean proportions are reported in the cells. Standard deviations of sample proportions across months are reported in parentheses in the first column.

weights to make our sample representative of the population of U.S. household heads. The weights are based on four individual characteristics (income, education, region, and age), with targets based on the Census population estimates derived from the American Community Survey for that calendar year.[18]

Column 2 of Table 3 shows the means of the monthly weighted average demographic characteristics of SCE respondents, while column 3 shows the distribution of these characteristics in the 2013 ACS. A comparison indicates that weighting is highly successful in making the SCE samples comparable to the population of household heads in the U.S. overall.

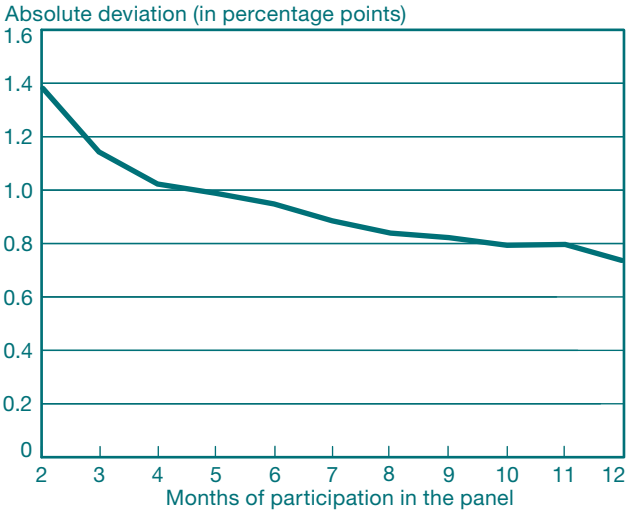
4.1 Learning and Experience

A common feature of survey panels that warrants some discussion is learning. As respondents continue to participate in the survey and answer the same questions over time, their participation in taking the survey may potentially affect their responses through learning. For example, after seeing a question covering a certain topic for the first time, a respondent may pay more attention to that topic in the media or may simply think more about the topic, perhaps in anticipation of receiving the question again in a future survey. Alternatively, the respondent may become more familiar and comfortable with the question formats. If such learning effects exist and influence responses in a systematic way, then changes over time in a respondent's answers may not capture true changes in beliefs.

In analyses we conducted, we find, at best, modest evidence of such effects in our panel. For instance, Chart 3 shows the density mean of short-term inflation expectations elicited for the respondents who entered the panel in 2014.[19] Specifically, we first take the absolute difference from one month to the next in each respondent's density mean. Then we plot the median of these absolute differences across respondents for each month of participation in the panel. We can see that the short-term inflation expectations density mean elicited for the median respondent changes by nearly 1.4 percentage points between the first and second survey she completes. After that, the median respondent reports different beliefs from one month to the next (as should be expected), but the magnitude of the month-to-month change remains relatively stable. Thus, most of the learning occurs within the first few months of the respondent's participation in the panel.

Perhaps most important, the design of the panel, with a constant in- and outflow of respondents each month, ensures a stable survey tenure distribution, so the extent of learning and experience (and any associated impact on responses) is constant over time. As a result, month-to-month changes in median responses should capture real changes in population beliefs.

Chart 3
Individual Changes in Inflation Expectations



Source: New York Fed Survey of Consumer Expectations.

Note: The chart shows the median month-to-month absolute deviation in the density mean for short-term inflation expectations.

[18] The weights applied to the survey responses are obtained using "RIM" (random iterative method) weighting (Sharot 1986). This method essentially uses minimum least squares to find the set of weights that minimize the distance between the marginal distribution in the sample and that in the population, given by the demographic targets. The weights are constructed through an iterative procedure that minimizes the distance between sample frequencies and population proportions sequentially along each dimension (demographic characteristic) separately, then iterates until the weights converge. Target statistics from the American Community Survey are updated each year based on the most recent ACS release. We distinguish among four income groups (up to $30,000, between $30,000 and $50,000, between $50,000 and $100,000, and above $100,000), three education groups (up to high school, some college, and college graduate and above), four Census regions (Northeast, Midwest, South, and West), and five age groups (under 30, 30 to 39, 40 to 49, 50 to 59, and 60 and above).

[19] To avoid selection effects arising from respondents who repeatedly fail to complete the survey and rotate out of the panel quickly, we focus here exclusively on respondents who stay in the panel for at least six months.

5. Computation and Reporting of SCE Statistics

To summarize and present our survey findings, we report median responses overall and by demographic characteristics. The median is a robust measure of central tendency that is less sensitive to the presence of outliers than the mean.[20] Using a robust summary measure is important, since we do not delete or recode outliers in the SCE. In addition to the median, for some survey questions we also report the 25th and 75th percentiles of the distribution of responses, with the difference between the two quartiles, the IQR, representing a measure of dispersion or disagreement among respondents.

5.1 Quantile Interpolation

A common feature of response behavior when a survey question asks for a numerical response is the use of rounding. When asked about past or expected future changes in percentage terms, almost all respondents appear to round to the nearest integer value. Accordingly, when changes in survey responses are tracked over time, it is common to see either no change in the computed raw median or a sudden abrupt change of one or more percentage points. In the case of grouped or rounded responses, it is therefore more informative to compute instead the median (and other quantiles of the distribution of responses) using an interpolation method. Interpolated medians will better capture shifts in the frequencies of responses around the median.[21] The same issue applies to other quantiles of the underlying distribution, including the first and third quartiles.

To compute interpolated quantiles, we use the symmetric linear interpolation approach proposed by Cox (2009).[22] We provide details about the procedure in Appendix B. We have compared Cox's procedure with other interpolation methods, including simple linear interpolation of the cumulative distribution function (asymmetric) and the Harrell-David procedure.[23] Computed quantiles and month-to-month changes in quantiles are generally very similar.

5.2 Density Estimation

In addition to point forecasts and probabilities of binary events, we ask respondents in the SCE for their density forecasts of various continuous variables. As discussed in Section 2.3, we elicit these by asking individuals to assign probabilities to ranges or intervals of possible future realizations. In addition to future inflation (at the one- and three-year horizons), we elicit density forecasts for year-ahead national home price growth and, for those who are employed, year-ahead earnings growth (holding the job and the number of hours fixed).

In reporting and analyzing such density forecasts, we focus on two summary measures: the density mean and the density IQR, defined as the difference between the third and first quartile. To compute the density mean and density quartiles of each individual's reported density, we use the reported bin probabilities to fit an underlying parametric density following the approach adopted by Engelberg, Manski, and Williams (2009). This approach is explained in detail in Appendix C.

Once fitted, the estimated density parameters are used to compute each individual respondent's "density mean" and "density quartiles." The mean represents the expected value, so in the case of the inflation density forecast, we refer to the computed density mean as the respondent's "expected inflation rate." Similarly, we use the estimated parameters to compute density quartiles, with the difference between a respondent's 75th and 25th percentiles (the IQR) measuring the respondent's "uncertainty." When we aggregate across respondents, we obtain the *median density mean* (and the *median density quartiles*), which we use predominantly in our reports (as discussed in the next section).

An important and unique strength of the SCE is its ability to provide quantitative measures of overall uncertainty among respondents and changes therein over time. In our SCE releases, we report the (non-interpolated) median of the respondents' IQRs as a summary measure of overall uncertainty in expectations. This statistic should not be confused with our measures of disagreement of expectations

[20] See Huber (1981) on robust statistics and estimation. Robust methods provide automatic ways of detecting, down-weighting (or removing), and flagging outliers, largely removing the need for manual screening and deletion of outliers.

[21] For example, consider two points x and y (with $x < y$) and two different *empirical* cumulative frequency distributions. The first empirical distribution attains the values 0.4 and 0.51 in x and y, while the second empirical distribution attains the values 0.49 and 0.6 in x and y. When the raw median is defined as the first value at which the cumulative distribution reaches or exceeds 0.5, the two empirical distributions both have the median of y. However, one may expect the median of the *underlying continuous* distribution to be closer to y for the first distribution and closer to x for the second distribution.

[22] In Stata, the procedure is implemented using the *iquantile* module. See Cox (2009).

[23] In Stata, the procedure is implemented using the *hdquantile* module (Xiao 2006).

among respondents. The latter are measured by the IQR of respondents' point forecasts or the IQR of respondents' density means, with both assessing dispersion in beliefs across respondents, while our uncertainty measure captures average forecast uncertainty among respondents.

5.3 Reporting of Multiple Medians

For several expectation questions, we solicit both point forecasts and density forecasts. For example, respondents are asked how much they expect the average home price to change nationwide over the next twelve months. They are also asked for the percent chance that, over the next twelve months, the average home price nationwide will increase (decrease) by: 12 percent or more; between 8 percent and 12 percent; between 4 percent and 8 percent; between 2 percent and 4 percent; and between 0 and 2 percent. As explained earlier, the latter bin probabilities are then used to fit the respondent's underlying density of beliefs about year-ahead changes in home prices.

One would expect the respondent's point forecast to represent some summary statistic of the central tendency of his or her density, such as the density mean or median. While this often appears to be the case, with point forecasts largely tracking density means (as well as density medians), for a nontrivial subset of respondents the reported point forecasts correspond to values in the tails of the respondent's density forecast. Similar findings were reported by Engelberg, Manski, and Williams (2009) for professional forecasters. An important advantage of using the density mean is that it captures the same measure across respondents. This might not be the case for point forecasts, which, for some respondents, may represent the density mean, while, for others, may represent the density median or mode or some other moment of the respondent's forecast distribution. For this reason, in our monthly reporting of SCE findings, we place more emphasis on the *median density mean,* although we include both medians (of point forecasts and density means) in our interactive charts.

6. Dissemination of the Data

The monthly SCE findings are released on the second Monday of each month. The release takes the form of a press release[24] as well as a set of interactive charts[25] that show the trends in the different variables, both for the overall sample as well as various subgroups (such as by age or Census region). The underlying chart data are made available at the same time.

To facilitate the use of these data by researchers and policymakers, the micro data for the monthly survey are also released on the SCE web page with a nine-month lag. Open-ended responses and sensitive information (such as the respondent's zip code) are not released.

The SCE project is still in its infancy, and the process of setting up web pages for the other data collected under the SCE umbrella (either as part of the ad hoc modules or the quarterly surveys) is ongoing. The SCE Credit Access Survey, which is conducted every four months and provides information on consumers' experiences and expectations regarding credit demand and credit access, is available at https://www.newyorkfed.org/microeconomics/sce/credit-access#main; as in the case of the monthly survey, the micro data are made public with a nine-month lag. The annual SCE Housing Survey, which provides rich and high-quality information on consumers' experiences, behaviors, and expectations related to housing, can be accessed at https://www.newyorkfed.org/microeconomics/sce/housing.html#main; the corresponding micro data are released with an eighteen-month lag. Interested readers should check the data page of the New York Fed's Center for Microeconomic Data for the latest products related to the SCE.[26]

[24] Press releases are available at https://www.newyorkfed.org/press/index.html#press-releases.

[25] Charts can be viewed at https://www.newyorkfed.org/microeconomics/sce.

[26] The Center for Microeconomic Data's data page is https://www.newyorkfed.org/microeconomics/databank.html.

Appendix A

Comparison of Consumer Confidence Survey (CCS) and American Community Survey (ACS) Samples

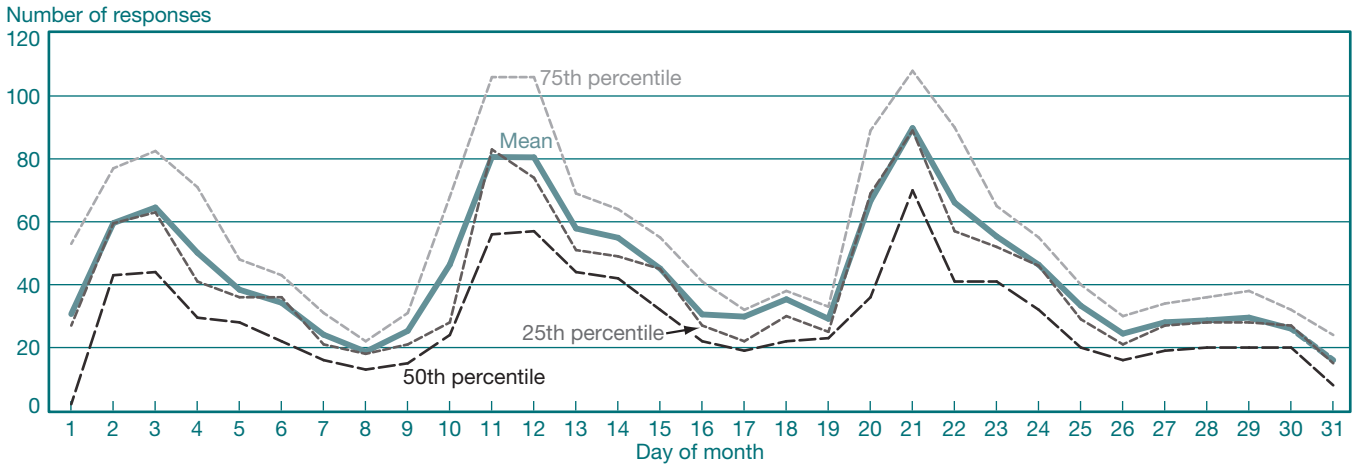
	Full CCS Sample[a]	2013 ACS	2014 ACS
	Percent		
Age			
Under 30	3.7	10.8	10.6
30–39	11.3	16.9	16.9
40–49	15.7	19.2	18.7
50–59	23.9	20.8	20.6
60 and over	45.5	32.4	33.2
Gender			
Female	47.7	49.9	49.9
Male	52.3	50.1	50.1
Education			
Up to high school	NA	36.7	36.3
Some college	NA	31.3	31.2
College graduate	NA	32.0	32.5
Income			
Under $15,000	8.3	13.0	12.6
$15,000–$24,999	9.9	10.9	10.6
$25,000–$34,999	10.3	10.3	10.1
$35,000–$49,999	15.2	13.7	13.4
$50,000–$74,999	19.7	17.9	17.8
$75,000–$99,999	13.1	11.8	12.0
$100,000–$124,999	9.4	7.9	8.1
$125,000 or more	14.0	14.6	15.5
U.S. Census Division			
New England	5.0	4.8	4.8
Middle Atlantic	13.6	13.2	13.2
East North Central	17.6	15.5	15.4
West North Central	7.7	7.0	7.0
South Atlantic	19.9	19.7	19.7
East South Central	5.8	6.1	6.1
West South Central	9.2	11.5	11.6
Mountain	6.9	7.1	7.2
Pacific	14.5	15.1	15.1
U.S. Census Region			
Northeast	18.5	18.0	18.0
Midwest	25.2	22.5	22.4
South	34.9	37.3	37.4
West	21.4	22.2	22.2

Sources: U.S. Census Bureau, American Community Survey (ACS); Conference Board, Consumer Confidence Survey (CCS).

Note: Each number in the table is the proportion of the sample that falls into that category.

[a] CCS averages are unweighted averages, based on 64,133 CCS respondents during the October 2013 to September 2015 period.

Survey Responses by Day of Month



Source: New York Fed Survey of Consumer Expectations.

Note: The chart shows the mean and 25th, 50th, and 75th percentiles of the daily frequency of responses over all months from December 2012 to September 2015, thus combining months with different dates for the invitations and reminders.

Appendix B: Quartile Interpolation

The main idea behind the approach proposed by Cox (2009) to interpolate the cumulative distribution (or quantile) function is the following: rather than linearly interpolating $Pr(X < x)$ or $Pr(X \leq x)$, the average of the two, the *mid-distribution function,* $Pr(X < x) + 0.5\ Pr(X = x)$, is interpolated. More specifically, a brief description of the approach is as follows. First, for all observed values of x, compute the cumulative proportions, symmetrically considered, as $CDFS(x) = Pr(X \leq x) - 0.5Pr(X = x)$. Then to compute the median, determine the values of x observed with positive frequency with cumulative frequency *CDFS* that surround 0.5, defined as L (the smaller of the two) and H, and compute $CDFS(L)$ and $CDFS(H)$. Then the linearly interpolated median m is calculated as follows:

$$m = L + (H - L) \times [0.5 - CDFS(L)] \,/\, [CDFS(H) - CDFS(L)].$$

Similarly for other quantiles, for example the third quartile, we identify the values of x observed with positive frequency with mid-distribution function values closest around 0.75, and in the equation above, replace 0.5 with 0.75. When applying sample weights, the *CDFS* values are computed by calculating frequencies $Pr(X \leq x)$ as sums of the relative weights (normalized to have mean 1) corresponding to all observations below or at x.

Appendix C: Density Estimation

We follow the approach proposed by Engelberg, Manski, and Williams (2009) to fit a parametric distribution for each respondent based on the probabilities the respondent reported for each possible density interval. We assume the underlying distribution to have a generalized beta distribution when the respondent assigns positive probability to three or more outcome intervals. We assume an isosceles triangular distribution when the respondent puts all probability mass in two intervals and a uniform distribution when the respondent puts all probability mass in one interval.

The generalized beta distribution is a flexible four-parameter unimodal distribution that allows different values for its mean, median, and mode and has the following functional form:

$$f(x) = \begin{cases} 0 & \text{if } x < l \\ (x-l)^{\alpha-1}(r-x)^{\beta-1}/B(\alpha,\beta)(r-l)^{\alpha+\beta-1} & \text{if } l \leq < x \leq r \\ 0 & \text{if } x > r \end{cases}$$

where $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha,\beta)$.

It uses two parameters (α and β) to describe the shape of the distribution and two more (l and r) to fix the support of the distribution. Fitting a unique beta distribution requires a respondent to have assigned positive probability mass to at least three (not necessarily adjacent) intervals.[27]

The triangular distribution, for cases where a respondent assigns positive probability to exactly two adjacent bins, has the shape of an isosceles triangle whose base includes the interval with the highest probability mass and part of the adjacent interval. Thus, the triangle is anchored at the outer bound of the interval with probability mass above 50 percent.[28] Its density has the functional form:

$$f(x) = \begin{cases} \frac{4}{(r-l)^2}(x-l), & l \leq x \leq \frac{(l+r)}{2} \\ \frac{4}{(r-l)^2}(r-x), & \frac{(l+r)}{2} \leq x \leq r \\ 0 & \textit{elsewhere.} \end{cases}$$

With the triangle being anchored at one of the outer bounds (l or r), there is only one parameter (either l or r) to fit, which fixes the center and height of the triangle.[29] Note that an isosceles triangle is symmetric, so the mean, median, and mode are identical to each other.

Densities are not fitted for respondents who put positive probability in only two bins that are nonadjacent or for whom the probabilities do not sum to 100. Such respondents make up less than 2 percent of our sample.

[27] In fitting a generalized beta distribution to a respondent's bin probabilities, we use a minimum distance procedure that minimizes the distance between the empirical and estimated parametric distribution. We fix l and r to be the minimum and maximum bound of the positive-probability intervals, unless the corresponding bin is open-ended, in which case l and/or r are estimated together with α and β. In the latter case, we restrict l to be greater than or equal to -38 and restrict r to be at most 38. The sample statistics that we report are generally not sensitive to the choice of the imposed lower and upper bound.

[28] This rule applies only to the case of two adjacent intervals of equal width where neither interval is open-ended. In the case of two adjacent intervals with unequal width, the support of the triangle is assumed to include the smaller-width bin in its entirety if its probability exceeds 40 percent and includes the larger-width bin entirely otherwise, with the triangle covering only part of the adjacent bin. In the former case, the triangle would be anchored at the outer bound of the narrower bin and, in the latter, at the outer bound of the wider bin. In all cases where one of the two bins represents an open-ended interval (the left or right tail of the distribution), the base always includes the inner closed-end bin, with the triangle anchored by the innermost bound of the two intervals.

[29] In the case of two adjacent bins with equal width, no estimation is required, since the support of the triangle now fully includes both intervals, with the triangle anchored at the left-most and right-most interval bounds.

References

Armantier, O., W. Bruine de Bruin, S. Potter, G. Topa, W. van der Klaauw, and B. Zafar. 2013. "Measuring Inflation Expectations." Annual Review of Economics 5: 273-301.

Armantier, O., W. Bruine de Bruin, G. Topa, W. van der Klaauw, and B. Zafar. 2015. "Inflation Expectations and Behavior: Do Survey Respondents Act on Their Beliefs?" International Economic Review 56, no. 2 (May): 505-36.

Armantier, O., W. van der Klaauw, S. Nelson, G. Topa, and B. Zafar. 2016. "The Price Is Right: Updating of Inflation Expectations in a Randomized Price Information Experiment." Review of Economics and Statistics 98, no. 3 (July): 503-23.

Bernanke, B. 2004. "The Economic Outlook and Monetary Policy." Speech at the Bond Market Association Annual Meeting, New York, N.Y.

———. 2007. "Inflation Expectations and Inflation Forecasting." Speech at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Mass.

Bruine de Bruin, W., C. F. Manski, G. Topa, and W. van der Klaauw. 2011. "Measuring Consumer Uncertainty about Future Inflation." Journal of Applied Econometrics 26, no. 3 (April-May): 454-78.

Bruine de Bruin, W., S. Potter, R. Rich, G. Topa, and W. van der Klaauw. 2010a. "Improving Survey Measures of Household Inflation Expectations." Current Issues in Economics and Finance 16, no. 7.

Bruine de Bruin, W., W. van der Klaauw, J. S. Downs, B. Fischhoff, G. Topa, and O. Armantier. 2010b. "Expectations of Inflation: The Role of Demographic Variables, Expectation Formation, and Financial Literacy." Journal of Consumer Affairs 44, no. 2 (Summer): 381-402.

Bruine de Bruin, W., W. van der Klaauw, and G. Topa. 2011. "Expectations of Inflation: The Biasing Effect of Thoughts about Specific Prices." Journal of Economic Psychology 32, no. 5 (October): 834-45.

Bruine de Bruin, W., W. van der Klaauw, G. Topa, J. S. Downs, B. Fischhoff, and O. Armantier. 2012. "The Effect of Question Wording on Consumers' Reported Inflation Expectations." Journal of Economic Psychology 33, no. 4: 749-57.

Cox, N. 2009. "IQUANTILE: Stata Module to Calculate Interpolated Quantiles," http://EconPapers.repec.org/RePEc:boc:bocode:s456992.

Curtin, R. 2006. "Inflation Expectations: Theoretical Models and Empirical Tests." Paper presented at the National Bank of Poland Workshop on The Role of Inflation Expectations in Modeling and Monetary Policy Making, Warsaw, February 9-10.

Delavande, A. 2014. "Probabilistic Expectations in Developing Countries." Annual Review of Economics 6 (August): 1-20.

Delavande, A., X. Giné, and D. McKenzie. 2011. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." Journal of Development Economics 94, no. 2 (March): 151-63.

Engelberg J., C. Manski, and J. Williams. 2009. "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters." Journal of Business and Economic Statistics 27, no.1: 30–41.

*Gali, J. (*2008). Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework. Princeton: Princeton University Press.

Huber, P. J. 1981. Robust Statistics. New York: Wiley.

Juster, T. 1966. "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design." Journal of the American Statistical Association 61: 658–96.

Leiser, D., and S. Drori. 2005. "Naïve Understanding of Inflation." Journal of Socio-Economics 34: 179-98.

Manski, C. 2004. "Measuring Expectations." Econometrica 72, no. 5 (September): 1329-76.

Sharot, T. 1986. "Weighting Survey Results." Journal of the Market Research Society 28: 269-84.

References

Sims C. 2009. "Inflation Expectations, Uncertainty, and Monetary Policy." BIS Working Paper no. 275.

Svenson, O., and G. Nilsson. 1986. "Mental Economics: Subjective Representations of Factors Related to Expected Inflation." Journal of Economic Psychology 7: 327-49.

Van der Klaauw, W., W. Bruine de Bruin, G. Topa, S. Potter, and M. Bryan. 2008. "Rethinking the Measurement of Household Inflation Expectations: Preliminary Findings." Federal Reserve Bank of New York Staff Reports, no. 359, December.

Woodford, M. 2004: "Inflation Targeting and Optimal Monetary Policy." Federal Reserve Bank of St. Louis Economic Review 86, no. 4 (July/August): 15-41.

Xiao, B. 2006. "The Use of the Interpolated Median in Institutional Research." Paper presented at the Association for Institutional Research Annual Forum, Chicago, Ill.

Federal Reserve Bank of New York
Economic Policy Review

Library of Congress Card Number: 77-646559

Federal Reserve Bank of New York research publications—the *Economic Policy Review*, *Staff Reports*, and the *Liberty Street Economics* blog—are available at the Research Group's website. To view these publications online or to sign up for alerts, visit www.newyorkfed.org/research.

Follow us on Twitter: *@NYFedResearch*

Federal Reserve Bank of New York
33 Liberty Street, New York, N.Y. 10045-0001
www.newyorkfed.org/research