The decline in teen labor force participation

Variations in consumer sentiment across demographic groups

Earnings announcements, private information, and liquidity

An alternative measure of inflation
## Contents

First Quarter 2006, Volume XXX, Issue 1

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
<th>Authors</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>The decline in teen labor force participation</td>
<td>Daniel Aaronson, Kyung-Hong Park, and Daniel Sullivan</td>
<td>The authors examine the recent decline in teen work activity, offering explanations for both the long secular decline since the late 1970s and the recent acceleration in this decline since 2000. They argue that much of this pattern is due to a significant increase in the rewards to formal education. They also explore the importance of changes to labor demand, crowding out by substitutable workers, the increased work activity of mothers, and increases in wealth.</td>
</tr>
<tr>
<td>19</td>
<td>Variations in consumer sentiment across demographic groups</td>
<td>Maude Toussaint-Comeau and Leslie McGranahan</td>
<td>Consumer sentiment is one of the many macroeconomic indicators tracked by policymakers and the public. The aggregate numbers in consumer sentiment indexes, such as the University of Michigan’s Index of Consumer Sentiment, conceal a wealth of demographic-specific information. The authors’ findings suggest that index disaggregation by group matters because consumer sentiment varies systematically by demographic group.</td>
</tr>
<tr>
<td>39</td>
<td>Earnings announcements, private information, and liquidity</td>
<td>Craig H. Furfine</td>
<td>In this article, the author examines how the price impact of a trade varies throughout the days surrounding public earnings announcements. The results indicate that public news releases correlate with a reduction in the price impact of a trade on the day of the announcement.</td>
</tr>
<tr>
<td>55</td>
<td>An alternative measure of inflation</td>
<td>François R. Velde</td>
<td>The author proposes an alternative measure of inflation that captures the intuition behind the use of “core” measures. Inflation is modeled as an unobserved factor affecting the components of an aggregate price index (including food and energy). The common component, estimated using Kalman filtering, resembles usual measures of core inflation; its extrapolation can be used to improve performance in forecasting core inflation.</td>
</tr>
<tr>
<td>66</td>
<td>Conference on Bank Structure and Competition</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The decline in teen labor force participation

Introduction and summary
By the middle of 2005, the U.S. civilian unemployment rate had fallen to 5 percent, a level many analysts consider consistent with essentially full employment. However, individuals who have become discouraged over their prospects of finding suitable employment and, as a result, have given up looking are not counted among the unemployed. Thus, analysts often look to the labor force participation (LFP) rate, the fraction of the population that is either employed or unemployed as an additional indicator of labor market conditions. In fact, the participation rate declined significantly during and after the 2001 recession and remains well below its 2000 level. This could imply more labor market slack than the unemployment rate suggests.

The decline in LFP has been especially great for teenagers. As figure 1 shows, teens’ participation rates had been trending down since the late 1970s. However, from 2000 to 2003, teen LFP fell a stunning 7.5 percentage points, compared with a decline in the overall rate of only 0.6 percentage points. Currently, the LFP for teenage boys is the lowest since at least 1948 and for teenage girls is the lowest since the early 1970s.

Figure 1 also shows that the decline since 2000 in the LFP rate for those 20 and older is considerably less dramatic than the fall in the overall rate, which includes those aged 16 to 19. Although those between the ages of 16 and 19 represent only 4.2 percent of employment (and 8.2 percent of population aged 16 to 69), they account for over half of the fall in aggregate LFP since 2000. Strikingly, 16 year olds to 17 year olds, who account for only 1.6 percent of workers and 4.3 percent of the population aged 16 to 69, explain over one-third of the fall in aggregate participation since 2000. Thus, a better understanding of the forces shaping the labor force participation of teens may shed significant light on recent trends in overall participation.

Another reason to look more closely at teen labor force participation is to understand what this major shift in the allocation of young people’s time may mean for future productivity. The answer to this question likely depends on what teens are doing instead of working and whether those activities contribute to human capital development. On the one hand, if the reduction in time spent working in the market has been accompanied by a concomitant increase in the time spent in school or doing homework, one might reasonably expect an eventual increase in productivity consistent with the well-documented returns to education. The impact of the increase in schooling investments on the overall economy might also include the positive externalities associated with education, including spillover productivity effects on peers and other workers, lower crime, and greater civil involvement in the public policy process.

On the other hand, a shift in teens’ time allocation from market work to leisure or other activities that do not increase their human capital may negatively affect their future productivity. In general, labor market experience tends to raise subsequent earnings. Moreover, it is easy to imagine that moderate amounts of time devoted to a part-time job during the summer or while in school might inculcate good work habits and allow young people to make more informed educational and career choices.

Daniel Aaronson is a senior economist and an economic advisor. Kyung-Hong Park is an associate economist, and Daniel Sullivan is a senior economist and vice president in the Economic Research Department of the Federal Reserve Bank of Chicago. The authors thank Merritt Lyon for valuable research assistance and Craig Furfine and Leslie McGranahan for helpful comments.
In this article, we examine the facts about teen labor force participation in more detail. We show that, although there is some variation in the magnitude, the decline in teens’ labor force participation is extremely widespread. Virtually all groups of teens have seen a decline in LFP. We then discuss a number of possible explanations for this decline in teen labor force participation over the past quarter century as well as the sharper drop of the early 2000s. The possible explanations that we consider can be grouped into two categories: demand and supply. Those that would tend to lower the wage associated with current work can be thought of as reducing teen labor demand. Those that increase the value of human capital investments or tilt teens’ choices toward more leisure can be thought of as reducing the per capita supply of teen labor.

In the end, it seems likely that the most important factor behind the long-term decline in teen LFP over the past 25 years is a supply-side development. The significant increase in the rewards from formal education (in the form of higher future earnings) began to take hold shortly before teen participation peaked. The fact that the average hourly wage rate of teens relative to adult workers has changed relatively little as teen labor supply has shifted in over the last quarter century suggests either that the relative demand for teen labor is relatively elastic or that it also has been shifting in over time. The former possibility is consistent with evidence we present on the impact of increases in the number of competing workers on teen participation. The latter possibility would be consistent with the existence of skill-biased technical change, the tendency for recent technological innovations to raise the productivity of highly educated workers relative to those who are less educated, including teens.\(^4\) Both possibilities may be true.

It is less clear what caused the more recent acceleration in the decline of teen LFP. Wage trends suggest that a softening in teen labor demand may have played some role. Other evidence, however, suggests that the recent drop is unlikely to represent a significant margin of additional labor market slack.

**Trends in teen labor force activity**

We begin our analysis by reviewing the history of LFP among 16 year olds to 19 year olds since 1948, the earliest year for which we have data derived from the *Current Population Survey* (CPS). The CPS interviews a nationally representative sample, which is currently approximately 60,000 households per month. It collects information about the labor market activities of all those at least 16 years of age. The LFP rate shown in figure 1 is the share of civilian noninstitutionalized 16 year olds to 19 year olds who are either working or unemployed (available to work and actively looking for work) in a given month.\(^5\)

As the figure shows, there have been long periods of expansions and contractions in teen participation rates. Coming out of World War II, just over half of teenagers were in the labor force. But, soon thereafter, LFP began to fall, reaching a low of just under 45 percent in the early 1960s. Over the next two decades, teenagers slowly rejoined the labor market, with their LFP rates peaking at 59 percent in the late 1970s. Since then, teen participation has pulled back again, with LFP rates falling steadily, punctuated by a particularly large decline starting around 2000. Currently (as of December 2005), teen LFP stands at 43.3 percent, over 15 percentage points below its peak 25 years earlier, and at the lowest rate in our 50-plus-year sample.\(^6\)

The broad swings in teen LFP may be partially obscured by shorter-run fluctuations associated with the business cycle. As one way to more clearly isolate the longer-term movements from the business cycle, figure 1 identifies periods, like the third quarter of 2005, in which the aggregate unemployment rate was approximately equal to the Congressional Budget Office’s (CBO) estimate of the non-accelerating inflation rate
of unemployment (NAIRU) after having been above it for some time. Changes in teen LFP between such quarters should be little affected by changes in business cycle conditions.

As the figure displays, the rate of decline in teen LFP over the latest full business cycle was much more rapid than over the previous two cycles. The average drop of about 1 percentage point per year between the first quarter of 1997 and the third quarter of 2005 was about three times faster than the pace of decline going back to the third quarter of 1987. If the slower rate of decline in place between 1987 and 1997 had been maintained, the current teen LFP rate would be about 5.5 percentage points higher than it is currently.

Teen LFP patterns differ by gender. Historically, male teens were more likely to work than females. However, teenage female LFP grew dramatically during the late 1960s and 1970s, likely reflecting the same economic and cultural forces underlying the increase in adult female LFP. As a result, by the early 1980s, there was virtually no gender difference among 16 year olds to 17 year olds. For 18 year olds to 19 year olds, the gender gap, while narrowing, did not disappear entirely until the mid-1990s. This likely reflects the especially significant increase in female college attendance that took place over this period.

As one way to isolate the trend in teen LFP separately from developments related to gender, figure 2 shows the labor market activity of teenagers relative to the gender-specific LFP rates of prime age adults (25 years to 54 years of age). Specifically, we display the percentage difference between the teen LFP rate and the same gender’s adult rate. The relative LFP of female 18 year olds to 19 year olds has fallen the most steadily. In the late 1940s, 18-year-old to 19-year-old females were as much as 60 percent more likely to work than adult women, but now are about 25 percent less likely to work than adult females. The steady drop in the relative LFP of 18-year-old to 19-year-old females likely reflects their equally impressive increases in college attendance. For the other three age–gender groups, the relative teen LFP rate fell from the late 1940s until the mid-1960s, when it began to rise. Between 1979 and 2000, these rates have fallen steadily, accelerating again beginning around 2000. For all four age–gender groups, the ratio of teen LFP to the LFP of adults of the same gender reached an all-time low during the current cycle.

Generally, LFP is procyclical, rising during expansions and falling during recessions. Figure 3 presents teenage LFP rates since 1979 adjusted for normal business cycle fluctuations in two alternative ways. The first version (the black dashed line), which we label the “time-series adjustment,” takes advantage of the time-series relationships between LFP and aggregate labor market conditions. In particular, we run the regression \( L_i = \alpha + \beta_i (U_r - \bar{U}_r) + \beta_t + \beta_t' + \beta_t' + \epsilon_i \), where \( L_i \) is the LFP rate of group \( i \) at time \( t \), \( U_r \) is the overall unemployment rate at time \( t \), \( \bar{U}_r \) is the CBO’s estimate of NAIRU, \( t \) is a time trend (1979 = 1, 1980 = 2, and so on), and \( \epsilon_i \) is a white noise term. We define the cyclically adjusted LFP at time \( t \) as \( \bar{L}_i = L_i - \beta_i (U_r - \bar{U}_r) \). This assumes the business cycle effect is proportional to the gap between the actual unemployment rate and CBO’s NAIRU.

The second version (the green dashed line), which we label the “cross-sectional adjustment,” also subtracts a constant multiple of the unemployment gap, but uses differences in state experiences to estimate the parameter relating LFP to unemployment. Specifically, we regress state-level teen LFP on state-level aggregate unemployment. To control for long-term differences in LFP across states, we also add state fixed effects. Thus, the identification of \( \beta \) is based on within-state changes in teen LFP and unemployment. As figure 3 shows, there are

![FIGURE 2](attachment:figure2.png)

**FIGURE 2**

Labor force participation of teens relative to 25–54 year olds of the same gender

Note: The shaded areas are recessions as identified by the National Bureau of Economic Research.

Source: Authors’ calculations based on data from Haver Analytics.
three periods since 1979 when the cyclical adjustment is important, although the degree depends somewhat on which technique is used. In the early 1980s and early 1990s, the economy slowed, unemployment rates rose, and the teen labor market activity declined. Had the unemployment rate remained at the natural rate, the teenager labor market activity would have risen by roughly 1 percentage point to 3.5 percentage points in the early 1980s and 1 percentage point to 2 percentage points in the early 1990s. Given the former adjustment, it might be the case that the underlying trend in teen labor market activity peaked in the early 1980s rather than the late 1970s. Likewise, the booming economy of the late 1990s pushed up teenage labor force participation by roughly 0.5 percentage points to 1.2 percentage points, thus exaggerating the decline since then.

Table 1 shows that the unadjusted series falls by 8.4 percentage points between 2000 and 2005. In rows 2 and 3, we report how much of this decline is due to previous secular trends and the cycle, as computed using our two techniques. We compute trend as the slope of the line between 1987:Q3 and 1997:Q1, two quarters when the unemployment rate and the CBO’s natural rate were roughly the same. Between those two periods, teen LFP fell 0.3 percentage points per year. This trend suggests that teen LFP would have fallen by about 1.8 percentage points between 2000 and 2005. The cycle adds another 1 percentage point to the decline. So just over 5.5 percentage points of the 8.4 percentage point fall over this period remains unexplained.

These cyclically adjusted figures are derived from micro (that is, individual-level) data from the CPS. This has the advantage of allowing us to explore heterogeneity in labor market activity across the teenage population. For example, we can ask whether the labor market activity of teens from high-income families looks different than that of teens from low-income families. For the rest of this section, all figures and tables use cyclically adjusted (with the time-series adjustment) rates in order to get a cleaner picture of secular trends.

Table 2 shows the change in teenage LFP from 1979 to 2005, as well as between 1987 and 1997 and since 1997, by gender, race, and region. We also compute changes by family income and school enrollment but begin these calculations in 1984, when the variables become consistently available. Note that each group’s series is cyclically adjusted separately, resulting in some groups, such as enrolled students, having much of the LFP decline explained by the business cycle.

The most striking aspect of table 2 is how widespread the decline is. Although it is clearly not uniform, the rate for every subgroup reported in the table has fallen since the early 1980s, typically 2 percentage points to 20 percentage points for 16 year olds to 17 year olds and 1 percentage point to 17 percentage points for 18 year olds to 19 year olds. For nearly all groups, the majority of the cyclically adjusted decline in LFP has occurred just in the past five years: LFP has fallen 5 percentage points to 9 percentage points among younger teen groups and 2 percentage points to 7 percentage points among older teens. While there is substantial variation by age and school enrollment
status, the patterns are fairly similar within race and family income groups.

Of course, many of these measures are correlated. To isolate which of these groups experienced economically and statistically significant drops, conditional on other characteristics, we ran multivariate regressions of a teen's decision to be in the labor force (a dichotomous 0–1 variable for whether they are in the labor force) on their background characteristics, two linear time trends—one that begins in 1984 and the other in 1997—and each of their characteristics interacted with the time trends. Level shifts across background characteristics are picked up by the covariates themselves (for example, the female indicator measures the average gender gap for a person of the same race, age, family income, and region). The interaction terms measure differences in average growth rates across groups, after conditioning on other characteristics of the teen and her family. The results are reported in table 3 separately by age (16 year olds to 17 year olds versus 18 year olds to 19 year olds). For exposition purposes, we only report the coefficients of the time trends and their interactions with the background characteristics. However, all regressions include level shifters for income, race, gender, and region. The regression model is parameterized so that the time trend coefficients show the average time trend over all individuals in the sample and the interaction term coefficients show how the time trend for a given group differs from the average trend.

On average, between 1984 and 1997, LFP fell by 0.22 percentage points and 0.28 percentage points per year among 16
year olds to 17 year olds and 18 year olds to 19 year olds, respectively. Since 1997, the decline has significantly accelerated: to almost 1 percentage point per year among 16 year olds to 17 year olds and 0.7 percentage points per year among 18 year olds to 19 year olds. The decline varies somewhat across groups, especially post-1997. Since then, teen LFP has fallen fastest among 16-year-olds to 17-year-old boys and 16 year olds to 19 year olds in the middle part of the family income distribution (between the 25th and 75th percentiles). Racial gaps are negligible once income is controlled.16

All calculations discussed thus far have been limited to the “extensive” margin of teens’ labor supply—whether they are in or out of the labor force. Similar developments have occurred on the “intensive” margin—the time spent working conditional on participation. For example, among those that work at all, the average work week length has declined almost 3.5 hours, or 12 percent, since 1979. This is somewhat offset by an increase in the number of weeks worked per year.17 Combining the two figures gives us an estimate of annual hours worked, conditional on working at all. Between 1979 and 2004, teens that work reduced their market work activity by 70 hours per year or 9 percent, and as with LFP, much of this decline has transpired recently. Thus, a substantial decline in teen work activity has occurred at both the extensive and intensive margins over the past two and a half decades.

Has demand for teen labor been weak recently?

As we noted earlier, a drop in LFP could, under some circumstances, be a sign of some additional labor market slack. At least in the case of teenagers, we think that such an interpretation of current developments is hard to square with several facts.

First, the CPS asks whether those out of the labor force want a job, and in recent years there has not been a notable increase in the number of such teens. As can be seen in figure 4, the fraction of the teen population that is out of the labor force but wants a job increased in the wake of the 1980–82 and 1990 recessions. But the most recent downturn saw much less of an increase. The long-term trend, moreover, is toward a lower fraction of teens being classified as wanting a job, but not employed.

A second difficulty with the weak demand explanation is apparent in the relative employment growth of the industries most likely to hire teens. If the sharp absolute and relative decline in their participation was primarily due to weak demand, we would expect to see that the industries that have traditionally hired teenagers had fallen on hard times, disproportionately impacting teenage work activity. However, we know of no evidence that traditional employers of young people have performed poorly recently. If anything, the top five industry employers of teenagers (in order:

---

**TABLE 3**

<table>
<thead>
<tr>
<th></th>
<th>16-17 year olds</th>
<th>18-19 year olds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time trend 1</td>
<td>Time trend 2</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-0.22** (0.045)</td>
<td>-0.97** (0.059)</td>
</tr>
<tr>
<td>Male</td>
<td>-0.02 (0.042)</td>
<td>-0.16** (0.053)</td>
</tr>
<tr>
<td>Female</td>
<td>0.02 (0.045)</td>
<td>0.17** (0.058)</td>
</tr>
<tr>
<td>1st quartile income</td>
<td>0.03 (0.057)</td>
<td>0.17* (0.170)</td>
</tr>
<tr>
<td>2nd quartile income</td>
<td>-0.06 (0.108)</td>
<td>-0.32* (-0.320)</td>
</tr>
<tr>
<td>3rd quartile income</td>
<td>0.08 (0.085)</td>
<td>-0.20 (0.108)</td>
</tr>
<tr>
<td>4th quartile income</td>
<td>-0.13 (0.084)</td>
<td>-0.05 (0.107)</td>
</tr>
<tr>
<td>White</td>
<td>-0.05 (0.012)</td>
<td>-0.03* (-0.030)</td>
</tr>
<tr>
<td>Black</td>
<td>0.13 (0.121)</td>
<td>-0.03 (0.159)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.19 (0.148)</td>
<td>0.01 (0.159)</td>
</tr>
<tr>
<td>Other race</td>
<td>-0.02 (0.205)</td>
<td>0.55* (0.550)</td>
</tr>
</tbody>
</table>

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

Note: Standard errors are in parentheses.

Source: Authors’ calculations based on data from the Current Population Survey.
eating and drinking places, grocery stores, miscellaneous entertainment and services, construction, and department stores), accounting for almost half of all 16 year olds to 19 year olds employed in 1999, have together experienced employment growth well above the national average. Since 2000, payroll employment in these five industries combined rose 3.6 percent, while employment in the remaining industries fell by 2.0 percent.

Trends in teens’ wage rates provide another piece of evidence on the reasons for the decline in their LFP. If the decline in teen LFP was primarily due to weak demand, one would expect their relative wages to have fallen. Over the ten-year period prior to 2002, that was clearly not the case, as can be seen in figures 5 and 6. Figure 5 plots teenage real wages (in 2000 dollars), as computed in the CPS and deflated by the Personal Consumption Expenditures (PCE) Price Index, along with the real value of the federal minimum wage (green line) and the real value of the minimum wage after accounting for state laws (dashed line). Actual real wages of teens were flat during the latter half of the 1980s and early 1990s but grew by 21 percent, or roughly 2 percent per year, between 1993 and 2002, which more than kept pace with the wages of less-educated adults. The latter point can be seen in figure 6, which plots the ratio of teen wages to adult wages and teen wages to less-educated (high school diploma or less) adult wages. In the 20 years prior to 2002, the average hourly wage rate of teens rose roughly 5 percentage points relative to prime-age workers without any college education, although it fell 2 percentage points relative to all prime-age workers.

However, since 2002, the real wage rates of teen workers, though still well above their levels in the late 1980s and 1990s, have fallen modestly. This is undoubtedly partly the consequence of a declining real minimum wage. Although a number of states have increased their minimum wages recently, the average real minimum wage remains roughly 8 percent below 1998 levels. Declining real wages could also be consistent with some softening in the demand for teen labor in the last few years. However, given the lack of an increase in the rates at which teens report they want a job, it is unlikely to be the major factor in the decline in teen LFP.

**Crowding out by adult low-skilled workers**

One possible demand-side explanation for lower teen work activity is that teens are facing stiffer competition for jobs from other workers. Card (1990) provided a classic analysis of a similar question—the effect of increased numbers of immigrants on native workers’ labor market outcomes—by studying the case of the large and likely exogenous increase in the number of workers in the Miami labor market after the Mariel boatlift of 1980, when a mass exodus of Cuban refugees landed on Florida’s (particularly Miami’s) shore. He finds that this influx of roughly 7,000 low-skilled Cubans had a positive impact on the employment of native Miamians, particularly relative to the employment of similar workers in four comparable cities. Lewis (2004) shows that the boatlift caused industries in Miami to adapt to less skill-intensive technologies, allowing the economy to painlessly absorb new workers.

When we extend Card’s analysis to teenagers, comparing how the teenage labor force participation rate in Miami looked pre- and post-Mariel and relative to Card’s four comparable cities (Atlanta, Houston, Los Angeles, and Tampa–St. Petersburg), our results are quite similar to his. Table 4 shows that teenage labor force participation rates rose absolutely (by 4.1 percent) and relative to the comparison cities (by 8.2 percent) in the year after Mariel. Likewise, teenage unemployment rates fell by over 6 percent in Miami and almost 8 percent in the comparison cities. Furthermore, when we extend the analysis past 1981,
it is apparent that the 1981 (non)effect remains sturdy years after the boatlift, suggesting that there is no evidence of a delayed reaction to the influx of workers. The boatlift is a valuable experiment because the influx of workers into Miami likely had little to do with the area’s pre-boatlift labor market conditions but more to do with its geographic proximity to Cuba and the decisions of the Cuban government. But, of course, it is possible that Miami’s experience in the wake of Mariel is not representative of other cases in which the number of low-skilled workers increased. Therefore, we explore two alternative analyses.

The first is the sizable influx of low-education single mothers with children after the 1996 Welfare Reform Act. Since 1995, the LFP of such women rose 30 percent, while it increased only 5 percent for low-education single women with no children and fell for the population at large. We concentrate on single mothers with two or more children and a high school diploma or less, given Meyer and Sullivan’s (2004) evidence that the law primarily impacted such women. We break the data into individual states and regress state teenage LFP on year and state fixed effects, the share of low-education single moms with two or more children, and that group interacted with an indicator of whether the year is 1996 or later. This interaction tells us whether the influx of such women post-reform had an impact on teenage work activity. In fact, we find no evidence that increases in low-education single women with children crowd teens out of the work force. Consistent with the Mariel evidence, an F-test fails to reject the hypothesis that the post-1996 year dummies interacted with share of such women in the state differ from zero.

Our second analysis is not tied to specific exogenous events. We created a state panel from 1979 to 2004 of teenage labor force participation rates, along with the share of the state population that 1) has less formal education (high school diploma or less) or 2) has less formal education and is unemployed or out of the labor force. We then regressed teen LFP on each of these, including state and year fixed effects in order to identify the association between within-state changes in teen work activity and within-state changes in the share of potentially substitutable workers. We also allowed each state to have its own time trend. Here, we do find results consistent with crowding out. However, the size of the effect is often economically small and statistically insignificant. More importantly, the size of the unskilled adult work force has been shrinking over time. In 1979, about 49 percent of the 25-year-old to
65-year-old population had a high school diploma or less. In 2004, only 35 percent did. Thus, in the aggregate, this cannot explain the large secular decline in teen participation over this period.

**Crowding out by peers**

The LFP of teens could also be affected by the sheer size of their peer group. Like the crowding out story described previously, increases in the size of teen cohorts could cause their wages and LFP to decline. However, the share of the working-age population accounted for by teens fell substantially from roughly 12.5 percent in the 1970s to 8.5 percent in the mid-1990s and has been relatively flat since then. Thus, if anything, the trend in teen cohort sizes should have pushed their wages and LFP up through the mid-1990s and been neutral since then.

**Supply explanations**

We suspect that teen LFP declines, particularly over the long run, are driven primarily by labor supply choices. This section describes three possibilities: the increased time devoted to school, the increased time spent helping out at home as mothers return to the labor force, and increases in wealth.

### Increased time devoted to school

A massive literature has documented that the financial return to obtaining more education has increased significantly in recent decades. This can be seen in figure 7, which is based on a standard methodology to value the effects of increasing educational levels on hourly wage rates. As the figure shows, the return to having a college education began to rise substantially in the late 1970s, shortly before teen LFP began to decline.

Figure 8 shows the substantial rise in the fraction of 16 year olds to 19 year olds enrolled in school, particularly in the 1980s. For each age group, it displays two measures of the fraction of the population enrolled in school. The lines labeled “October” are estimates of the enrollment rate for the month of October that are derived from a special supplement to the CPS that has been done every October since the late 1960s. The lines labeled “all months” are estimates of the average enrollment rate over the entire year. They are derived from a question on enrollment status that was added to the basic CPS in 1985. The all months lines are substantially lower than the October lines because they include the summer months of June through August when most students have traditionally been on vacation from school. Both the October lines and the all months lines show increases in enrollment over time, but the slope of the all months lines have been steeper recently. This is because enrollment increases have been especially great in the summer months. For example, summer enrollments were only 20.5 percent in 1992, when the increases began, but 44.3 percent in 2005.

Table 5 reports a simple decomposition of the change in teen LFP into components due to 1) the increase in enrollments given constant within-enrollment-status-group LFP rates, and 2) the fall in LFP within-enrollment-status group given a constant enrollment rate. The calculations are based on the more comprehensive all months measure of enrollment mentioned earlier.

Panel A of table 5 shows the decomposition for the drop in LFP between 1987 and 1997. As we noted above, aggregate labor market conditions were similar in the two years, so the changes reported should be largely free of business cycle effects. As the table shows, the enrollment rate increased by 0.65 percentage points per year over this period.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Teenage LFP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>39.6</td>
<td>43.7</td>
<td>4.1</td>
</tr>
<tr>
<td>(4.2)</td>
<td></td>
<td></td>
<td>(6.2)</td>
</tr>
<tr>
<td>Card (1990) comparison cities</td>
<td>56.8</td>
<td>52.7</td>
<td>-4.1</td>
</tr>
<tr>
<td>(1.3)</td>
<td></td>
<td></td>
<td>(1.8)</td>
</tr>
<tr>
<td>Miami-comparison difference</td>
<td>-17.2</td>
<td>-9.0</td>
<td>8.2</td>
</tr>
<tr>
<td>(4.4)</td>
<td></td>
<td></td>
<td>(6.5)</td>
</tr>
<tr>
<td><strong>B. Teenage unemployment rates</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>27.3</td>
<td>21.2</td>
<td>-6.2</td>
</tr>
<tr>
<td>(5.5)</td>
<td></td>
<td></td>
<td>(7.8)</td>
</tr>
<tr>
<td>Card (1990) comparison cities</td>
<td>17.7</td>
<td>19.5</td>
<td>1.8</td>
</tr>
<tr>
<td>(1.4)</td>
<td></td>
<td></td>
<td>(2.0)</td>
</tr>
<tr>
<td>Miami-comparison difference</td>
<td>9.6</td>
<td>1.7</td>
<td>-7.9</td>
</tr>
<tr>
<td>(5.6)</td>
<td></td>
<td></td>
<td>(8.1)</td>
</tr>
</tbody>
</table>

Notes: The comparison cities are Atlanta, Houston, Los Angeles, and Tampa/St. Petersburg. Standard errors are in parentheses. Source: Authors’ calculations based on data from the Current Population Survey.
The contribution of this enrollment change to the overall decline in teen LFP was 0.18 percentage points per year. This is the change in overall teen LFP that would have occurred if LFP had remained constant for both enrollees and non-enrollees. The increase in enrollments at constant within-enrollment-status-group LFP accounts for 60 percent of the decline in teen LFP over the period.

The table also shows the contributions of within-enrollment-status-group LFP to the overall decline.

Among teens who were enrolled in school, LFP declined by 0.12 percentage points per year. This was a little over one third of the 0.31 points per year rate at which LFP declined overall. Given that 68 percent of teens were enrolled in school, the contribution of their LFP decline to the overall LFP decline was 0.08 percentage points per year, or 26 percent of the total. The rate of LFP decline for non-enrollees was slightly faster at 0.14 points per year. But because they are a smaller fraction of the teen population than enrollees, the non-enrollees’ decline in LFP only accounted for 14 percent of the total drop in LFP.

Panel B of table 5 shows the same decomposition for the change in teen LFP between 1997 and 2005. Again, these were two years in which, by standard measures, aggregate labor market conditions were similar. As we discussed earlier, the rate of decline in teen LFP increased over this period to about 1 percentage point per year. Table 5 shows that most of this acceleration was due to faster declines in LFP within-enrollment-status groups. The rate at which enrollments rose did increase somewhat relative to the earlier period, resulting in about a 10 percent increase in the annual contribution of enrollment increase to teen LFP decline. But, the biggest factor in the acceleration was the significant increase in the rate at which LFP declined for those enrolled in school. The contribution of that factor to the decline in teen LFP increased by over 0.5 percentage points per year and its share of the entire decline increased to 62 percent. A faster rate of decline in LFP for those not enrolled also contributed to the faster rate of overall teen LFP decline.

The calculations just described only capture the effects of increased schooling at the extensive margin. However, similar effects may be at work on the intensive margin—conditional on being in school, students may be devoting more time to their studies and less to part-time or full-time jobs. However, the evidence on this point is quite a bit sketchier. A U.S. Department of Education (2005) publication reports time spent in school increased 30 hours to 40 hours per year (or about one hour per week) between 1987

---

FIGURE 7

Returns to education

![Chart showing wage rate changes for college versus high school and high school versus dropout.]

Source: Authors’ calculations based on data from the March Current Population Survey.

FIGURE 8

School enrollment rates

![Chart showing enrollment rates for 16-17 year olds and 18-19 year olds.]

Source: Authors’ calculations based on data from the March and October Current Population Surveys.
and 1999. Another U.S. Department of Education (2001) report found that for ages 13 to 17, the amount of homework time increased between 1984 and 1999. Juster, Ono, and Stafford (2004) find large increases in schooling and studying time between 1981 and 2002, although as we discuss later, there are reasons to think that the time-use data on which that study is based may be subject to substantial measurement error.

It is possible, albeit a bit speculative, that the increasing recognition of the value of more education in recent years has played a role in the sharp recent decline in teen LFP. For example, after falling fairly steadily by 3.3 percentage points between the 1983–84 and 1999–2000 school years, the high school graduation rate, defined as the number of diplomas issued as a fraction of the population of 17 year olds, rose 5.1 percentage points to 74.9 percent in the 2003–04 school year. Perhaps recognition that schooling is increasingly valuable is causing teens who are enrolled in school to study harder and graduate more frequently. As a side effect, it may be lowering their rate of labor force participation.

**Substituting house work for market work**

Among the biggest developments in labor markets over the past several decades has been the increased participation of women, particularly those with children. There is substantial evidence that technological innovations, such as the washing machine, dishwasher, and the like have aided in this transition. Furthermore, there has likely been an important reallocation of home production from wives to husbands. But how has the increase in female labor force participation affected teenage children? Specifically, has it led teenagers to substitute house work for market work?

As part of a pilot study on 322 children aged six to 17 in the early 1980s, the Institute for Social Research (ISR) at the University of Michigan conducted a time-use survey, where parents filled out time diaries in five minute increments. Juster, Ono, and Stafford (2004) compared this survey to a similar one conducted in 2001–02 using families from the Panel Study of Income Dynamics (PSID). For 15 year olds to 17 year olds, they show that market work fell by over one hour per week over the two decades, while home production work increased by two hours per week. However, there are some serious problems with this survey, particularly in the earlier years. The authors warn that the definition of home and market work may have been altered between surveys. Furthermore, many hours in the early 1980s survey are simply unclassified. But if we assume that these unaccounted hours are not work hours, and even if we combine the two work activities, we can infer that teen home production must have increased given the sizable fall in teenager market work hours documented in the CPS.

Nevertheless, because these results are based on small samples with highly imperfect data, we turn to

### TABLE 5

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage enrolled</th>
<th>LFP of enrolled</th>
<th>LFP of not enrolled</th>
<th>Overall LFP</th>
<th>1987</th>
<th>1997</th>
<th>Annual change</th>
<th>Contribution to LFP decline</th>
<th>Percent of total LFP decline</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 1987–97</td>
<td>61.07</td>
<td>67.60</td>
<td>0.652</td>
<td>-0.184</td>
<td>60.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage enrolled</td>
<td>43.71</td>
<td>42.55</td>
<td>-0.116</td>
<td>-0.079</td>
<td>25.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LFP of enrolled</td>
<td>71.69</td>
<td>70.56</td>
<td>-0.136</td>
<td>-0.044</td>
<td>14.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LFP of not enrolled</td>
<td>54.69</td>
<td>51.62</td>
<td>-0.307</td>
<td>-0.307</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall LFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage enrolled</th>
<th>LFP of enrolled</th>
<th>LFP of not enrolled</th>
<th>Overall LFP</th>
<th>1997</th>
<th>2005</th>
<th>Annual change</th>
<th>Contribution to LFP decline</th>
<th>Percent of total LFP decline</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. 1997–2005</td>
<td>67.60</td>
<td>73.16</td>
<td>0.696</td>
<td>-0.195</td>
<td>19.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage enrolled</td>
<td>42.55</td>
<td>35.80</td>
<td>-0.843</td>
<td>-0.617</td>
<td>62.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LFP of enrolled</td>
<td>70.56</td>
<td>65.17</td>
<td>-0.674</td>
<td>-0.181</td>
<td>18.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LFP of not enrolled</td>
<td>51.62</td>
<td>43.68</td>
<td>-0.993</td>
<td>-0.993</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall LFP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
more recent time-use data from the U.S. Bureau of Labor Statistics to uncover within-household home-market work distinctions.26 Because this survey only began in 2003, we must rely on cross-sectional evidence, in this case, differences in the number of earners in the family. We compare the work activity of teenagers with two parents working to the work activity of teenagers with one parent working and one parent at home. Our informal test asks whether teenagers with a mother working out of the home spend more of their day doing housework and, consequently, less of their day in market activity. In fact, we find no evidence of this trade-off. Among teenagers living in a home with both parents working, 11.3 percent of their typical nonsummer day is spent working away from home and 8.9 percent working within the home. Among teenagers living in a two-parent home with only their father working, market activity is lower (9.8 percent) and home production work is higher (9.7 percent), a result inconsistent with mothers shifting more home production work to their older children. Moreover, we find that number of siblings has little impact on market activity and, if anything, leads to more home production work, not less.

Finally, we find similar patterns when looking directly at labor force participation by number of parental earners in the family. If both parents work, teenagers are more likely to work as well. This is true even if we stratify the sample into family income quartiles and look at teen work activity within income quartiles or, in a regression context, if we control for family income (and number of siblings). Consequently, we cannot conclude from these data that the rise in adult female labor force participation has led to the decrease in teenager labor force participation.

Wealth effects

A final explanation of the decline in teen work activity that we explore is the role of increases in wealth among families with teenage children. Basic economic theory predicts that when wealth increases, and the wage available to a worker in the marketplace, as well as preferences for leisure, remain the same, people will want to work less and consume more leisure. In practice, pure wealth effects are hard to uncover because they require situations where these assumptions (especially a constant market wage) hold. Nevertheless, researchers have exploited a number of clever examples where increased sources of wealth are likely exogenous to the person supplying the labor, including bequests, war reparations, and lottery winnings.

At first glance, we find little support for such a possibility. In particular, parental income and teen LFP are, if anything, positively related in 2004 (as well as all other years). However, this could be because parental income is correlated with many other factors that might influence teen work. Moreover, inflation-adjusted median net worth has barely budged for families with heads aged 35 to 54, the families where the vast majority of teens reside. However, for such families real mean net worth increased 1.5 percent to 2 percent per year between 1983 and 2001 (the latest year of publicly available data) and aggregate real median and mean net worth increased 2.2 percent and 3.7 percent, respectively, per year over the same period. The vast majority of the aggregate increase is due to older households.

Here, we provide several pieces of evidence to quantify the role wealth may play in explaining the recent acceleration in the decline in teenage LFP. All three revolve around college pricing. A fall in the price of college can have two implications for work activity. First, cutting prices causes demand for that product to rise. Since time is constrained, an increase in enrollment pushes people out of work and other activities. Second, as the cost of college falls, students, particularly those at the margin of the enrollment decision, need to work less to afford it. Keane and Wolpin (2001) offer an example of this result within a dynamic model of the school–work decision for young men. Among the exercises they present is a simulation of a $3,000 per semester tuition subsidy. Their results suggest that the average full-time student earns over $450 less (and consumes over $1,200 more) per school year than a baseline group that does not receive this subsidy. Using the outgoing rotation files of the CPS to compute hourly wages allows us to infer that someone in their sample (white male full-time student) will work 89 fewer hours per school year after such a subsidy. In other words, a transfer of wealth to students and their families can significantly reduce work time.

We attempted two simple exercises to test the predictions of their simulations. First, we compared the work participation rates of teenagers in states that have introduced state-wide merit scholarships, often called Hope Scholarships, with rates for states that have not. The Hope Scholarship program, initiated in Georgia in 1993 and adopted in some form by 15 other states since, offers students a free or highly reduced tuition to in-state universities so long as they meet minimum entrance requirements, minimum college performance criterion, and attend an in-state
college. In Georgia, for example, qualified in-state students receive up to $4,500 ($3,000 for private school) per academic year for tuition, fees, and book expenses, regardless of family income.

Cornwell, Mustard, and Sridhar (2005) find that the program is working as intended—in-state college enrollment has increased. But their research describes several other important results as well. First, Cornwell, Lee, and Mustard (2005) document a number of “grade-enhancing” strategies—including enrolling in fewer classes and withdrawing from those where performance is subpar—used by students to ensure qualification for the scholarship. Second, roughly two-thirds of the increase in in-state enrollment is due to students switching from out-of-state colleges to in-state colleges, particularly four-year institutions. Finally, in line with the Keane and Wolpin (2001) results on consumption, Cornwell and Mustard (2005) show a positive association between county-level car purchases and Hope Scholarship grantees. Together, these results are consistent with the notion that these programs are transferring wealth to college-attending children and their families with relatively little direct impact on skill accumulation and current market wage rates.

Consequently, the Hope program can be thought of as a useful experiment to analyze labor supply wealth effects—what happens if we increase wealth leaving all else unchanged, including a worker’s potential market wage. As shown in table 6, among 16 year olds to 17 year olds in Hope states, LFP fell by 10.4 percentage points between 2000 and 2004. By comparison, in states without a Hope program, the decline was 8.7 percentage points. That is, young teen LFP fell 1.7 percentage points more in Hope states after 2000. Since 24 percent of all teens in the country reside in states with merit scholarship programs like Hope, we can estimate the impact these scholarships had on aggregate teen LFP. We find that roughly 5 percent (0.24 times 1.7 divided by 8.8) of the decline in young (16 to 17) teen LFP could be traced to differences in Hope and non-Hope states. The actual impact is likely bigger once we account for timing and generosity differences across states, which we plan to do in follow-up research.

Columns 2 and 3 of table 6 repeat this exercise for 18 year olds to 19 year olds by school enrollment. We find a small negative impact among those in school (about 2 percent of the total decline among 18 year olds to 19 year olds between 2000 and 2004) but no effect, at least on the extensive margin, among those not currently enrolled. While these effects are relatively small, they also represent just one of many financial aid programs offered in the U.S. (see Wirtz, 2005). A natural way to corroborate and generalize these findings is to see how changes in tuition, more broadly defined, influence work decisions. Typically, such studies examine the impact of tuition on college enrollment decisions. Instead, we analyze the teenager labor force participation rates of tuition using real annual tuition and fees data from the College Board (2005). The data are available back to 1975 for four-year private, four-year public, and two-year public institutions.

In general, the tuition results seem consistent with those for Hope Scholarships. Overall, we find that tuition changes are positively correlated with teen work activity. From a statistical perspective, the strongest results are those for two-year college tuition rates. This is what we would expect since these are the rates that likely affect students whose enrollment decision are most price sensitive. Furthermore, we find that families from the upper middle of the income distribution are more likely to respond to college price changes. This strikes us as plausibly the part of the income distribution for which enrollment decisions are particularly sensitive to tuition. Finally, while tuition at four year colleges has risen in recent years, the cost of attending community college is now substantially lower that during the second half of the 1990s. For instance, the College Board reports that community college tuition, net of grants and education tax benefits, fell from $1,000 for the 1997–98 school year to $200 for 2001–02. Our results suggest this decline could have lowered LFP for some teens.

Overall, we view the evidence as consistent with the hypothesis that increased wealth, via lower education prices, can reduce teen labor supply. The importance of this effect for recent trends depends critically on the real net price of schooling over time, which we believe has fallen at the margin. While the Keane

<table>
<thead>
<tr>
<th>TABLE 6</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Percentage point change in teen LFP, by Hope Scholarship status, 2000–04</strong></td>
</tr>
<tr>
<td>16–17</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Hope states (24% of total pop.)</td>
</tr>
<tr>
<td>Other states</td>
</tr>
<tr>
<td>Difference</td>
</tr>
</tbody>
</table>

Source: Authors’ calculations based on data from the Current Population Survey.

14 1Q/2006, Economic Perspectives
and Wolpin (2001) results are based on a larger tuition reduction program than recent experience, the flavor of their structural model matches our empirical findings.

Conclusion

Teens can be thought of as allocating their time between current market work, current leisure, and human capital investment. Since the late 1970s and especially since 2000, they have devoted less of their time to current market work. To a significant extent, they have also been increasing the time they devote to human capital investment. The increased value of education for their future earnings has apparently caused teens to increase their school enrollments and likely also the intensity with which they pursue their studies when enrolled. We know less about any possible changes in their leisure time. However, we have found some preliminary evidence that wealth effects from increased financial aid may have reduced their work effort as well.

It is possible that a sudden drop in demand for teen labor has played a role in the recent, sharp decline in teen participation rates. The modest decline in relative teen wages would be consistent with some role for weakened labor demand. We doubt, however, that this is the main explanation. The latest recession ended more than four years ago. In an unusual development, teens who are out of the labor force are not likely to report that they want a job, and the industries that typically employ them have been reporting stronger than usual overall employment growth. Of course, only time will tell whether the recent drop in teen participation is a manifestation of a weak labor market or a new equilibrium. The increases that we have noted in teen’s human capital investments, however, do suggest some reason for optimism for future levels of productivity.

NOTES

1See, for example, Aaronson and Sullivan (2001) for a discussion of the impact of greater educational attainment on aggregate productivity.

2See Moretti (2004) for a review of this evidence.

3See Ruhm (1997) and Stinebrickner and Stinebrickner (2003) for interesting discussions of these issues.

4For example, see Katz and Autor (1999). See Card and DiNardo (2002) for a skeptical view of the skill-biased hypothesis.

5This definition ignores several interesting groups. First, the data does not include those who are under 16. Second, by concentrating on the noninstitutionalized population, we are ignoring the sizable increase in incarceration over the last three decades. The adult prison population has grown from 0.2 percent of the adult population in the early to mid 1970s to almost 1 percent by the late 1990s. See Katz and Krueger (1999). Their study assumes that 35 percent of the incarcerated would be employed if not in jail. A similar assumption for incarcerated teenagers would lead to an even stronger trend down in the teenager LFP over time. Finally, the civilian population ignores the military. This might be of particular concern during the 1960s.

6Recent declines in teenage work participation have occurred throughout much, but not all, of the developed world, according to data from the Organization for Economic Cooperation and Development.

7For example, a “20” reveals that that teenage group’s LFP is 20 percent lower than the same gender’s adult population.

8For example, in the early to mid-1970s, female school enrollment was 3 percentage points to 6 percentage points lower than males among 18 year olds to 19 year olds (calculated from the October files of the CPS). By the early 1990s, this gap disappears. Several years later, school enrollment among the same aged females was 1 percentage points to 4 percentage points higher than their male peers.

9We focus on the period since 1979 because the CPS outgoing rotation files begin in that year. As a third alternative, we have also used the Hodrick-Prescott Trend, a standard statistical tool to isolate a long-term trend from short-term fluctuations in time-series. Those results provide a similar story to the two cyclically adjusted series presented in figure 3.

10We have also estimated this equation with a lag in U, to allow for delays in responding to aggregate conditions. This has no appreciable difference on the results. This equation seems adequate for picking up the time trend since 1979 but would not work as well over a longer period since there appears to be trend breaks in this series in late 1970s and early 1980s and perhaps in the early 1960s as well. In that case, we would simply estimate the time trends separately for different periods. We chose to focus our analysis on the post-1979 period.

11However, Staiger, Stock, and Watson (1997) show how imprecisely estimated the natural rate is.

12These regressions also use time dummies rather than linear time trends. Time dummies are unidentified in the time-series version.

13We use the outgoing rotation files of the CPS. Participating households are surveyed for four months, left out of the sample for eight months, and finally surveyed again for four additional months. Those households in the fourth and eighth months of their participation are known as the outgoing rotation groups.
Our technique for matching teenagers with their parents exploits the family relationship variable in the outgoing rotation files. This variable begins in 1984. We subtract the teenager’s own income from the family income measure.

We specified the time trends so that there is a kink, rather than a discontinuous jump, at 1997.

We have also run these regressions with controls for school enrollment status and its interaction with the time trends. Adding these additional regressors does not impact the gender, race, income, or (unreported) regional time trends in a significant way. The impact of enrollment on LFP is discussed later.

The weeks worked calculation is based on the March CPS. We are able to compute family income back to 1979 because the March files contain an explicit measure of family income (that is, there is no reason to have to match teenagers with other family members). Again, we use family income less the teenager’s own income.

Over the period shown, between 66 percent and 80 percent of teenagers (and 80 percent to 90 percent of 16 year olds to 17 year olds) had wage rates within 50 percent of the minimum wage.

Of course, one possibility is that teens report not wanting a job because they know wages are not above their reservation price (that is, the lowest wage at which they are willing to work). This story has particular resonance if we believe that teens look at the minimum wage, which has declined steadily since last raised at the federal level in 1997, rather than actual market wages when deciding whether to work.

Card (1990) selected these cities because of their similarity to Miami in terms of racial composition and economic growth during the late 1970s and early 1980s.

For example, the LFP rate of single mothers with two or more children has grown by 30 percent since 1996, while the rate for single women with no children has been relatively flat.

These results, as well as others referenced in the text without tables and figures, are available upon request from the authors.

See, for example, Katz and Autor (1999).

The estimates are based on a regression of the natural logarithm of wage rates on standard variables such as potential experience, gender, and race and indicator variables for different levels of schooling. The data are from the March CPS. See Aaronson and Sullivan (2001).

An alternative way that time in school may have increased is through changes in legally mandated years in school. Aaeomoglu and Angrist (2001) find that the number of years required in school has not changed much since the middle of the 1900s. See also Lochner and Moretti (2004).


In this article, the term “home production work” includes all work performed within a household for which no compensation is received from outside parties.

The data are available at www.bls.gov/tus/home.htm. We use both the 2003 and 2004 surveys. Hamermesh, Frazis, and Stewart (2005) provide background.

In some states (like Georgia), there are two components to the Hope program—a merit scholarship that requires minimum grades and is applied to degree programs and a grant that can be applied to two-year and less than two-year programs but has no grade requirements.

However, Cornwell, Mustard, and Sridhar (2005) and Cornwell, Lee, and Mustard (2005) find that a sizable fraction of the college enrollment effect happens among freshmen who delayed college enrollment by more than 12 months past their high school graduation.

See Mazumder (2003) for a nice review.

To allow for information delays, we include two years of lag on tuition. LFP is computed from September to August to correspond with school year tuition data.

When we estimate teen LFP regressions separately by family income quartile, data limitations only allow the series to start in 1985. We find that the only income quartile where there is a statistically significant response to price changes is the second highest (income between the median and 75th percentile), although all quartiles have a positive, albeit imprecisely estimated, point estimate.

Despite the small sample sizes, we found that none of these results are sensitive to outliers. We also tried using a separate dataset on two-year college tuition rates provided by the Washington State Education Group. The advantage of their data is that it is disaggregated by state. When we aggregate their data to the national level, we find correlations that are very similar to the College Board data. However, our attempts to use panel methods to take advantage of state differences in tuition and teen LFP growth are unreliable. We suspect measurement error is severe at the state level, which attenuates estimates of the betas. Mazumder (2003) finds little correlation between this tuition measure and enrollment, which is generally contrary to the literature. Obviously, there are many refinements that could be made to each of these analyses that would limit the damage from measurement error, including looking at teenagers who are at the margin of deciding whether to go to college and improving our understanding of what the relevant tuition measure is. The latter, for example, would entail better information on financial aid.

See College Board (2005). Since 2002, two-year college costs have begun to rise and aid has stagnated, but net costs remain historically low.
REFERENCES


Cornwell, Christopher, and David Mustard, 2005, “Merit-based college scholarships and car sales,” University of Georgia, working paper.


Variations in consumer sentiment across demographic groups

Maude Toussaint-Comeau and Leslie McGranahan

Introduction and summary

Consumer sentiment is one of the many macroeconomic indicators tracked by policymakers. Consumer sentiment—as measured by indexes such as the Index of Consumer Sentiment (ICS) and the Consumer Confidence Index (CCI)—is seen as a barometer of economic activity, one that is a reliable indicator of the way people plan to spend their money. Consumer sentiment is important because it affects household spending. Nationally, household spending on final goods and services (retail sales) represents about 65 percent of all expenditures for final goods and services, the nation’s gross domestic product (GDP). Since private consumption expenditure accounts for such a large proportion of GDP, consumer sentiment can signal changes in the direction of the economy. Numerous studies have assessed the extent to which consumer sentiment is related to fluctuations in GDP, the stock market, and other outcomes.

While the overall index scores, so closely watched by the public, are important, these aggregate numbers conceal a wealth of demographic-specific information contained in the survey data. Analyzing the survey data at disaggregated levels enhances the indexes’ informative power (Dominitz and Manski, 2004). Consumers’ expectations about specific sectors of the economy, such as expectations of inflation, income, employment, and home values, usually differ by demographic group and often move in opposite directions by group. These disparities in expectations translate into distinct spending patterns by different groups. Additionally, personal spending patterns vary across demographic groups. For example, older consumers tend to spend more on health care; also, poor consumers spend a higher proportion of their income on food and shelter. Because of these and other differences, examining disaggregated consumer sentiment survey data can provide us a more detailed picture of future expenditure.

Beyond predicting expenditure, household-level sentiment data tell us something about the current welfare of vulnerable populations. There is increasing evidence that consumer expectations vary systematically across demographic and socioeconomic groups. As policymakers seek to better understand the economic experiences of various societal groups over the business cycle, disaggregated consumer sentiment data can be a useful tool. For example, if a certain subpopulation expresses pessimism about general business conditions during an economic recovery or growth period, there is good reason to think that the benefits of economic expansion may not be reaching that group. These insights can inform policy initiatives aimed at assisting these populations.

In this article, we use household micro-level data to investigate the determinants of consumer sentiment. We use data from the University of Michigan’s Surveys of Consumers, grouping respondents by characteristics such as race, ethnicity, gender, and income, among others. We examine responses to the questions that go into calculating the University of Michigan’s Index of Consumer Sentiment, as well as responses to other questions in the survey. One important finding is that sentiment differences across groups persist regardless of whether the question asks about personal situations or general situations—that is, groups have different views not only of their own outlook, but of the outlook for the country as a whole. We look into consumers’ explanations of their sentiment to investigate why this is, considering group-level subjective

Maude Toussaint-Comeau and Leslie McGranahan are economists in the Consumer and Community Affairs Division of the Federal Reserve Bank of Chicago. The authors thank Dorothy Kronick and Lori Timmins for providing valuable research assistance.
experiences and differences in information sets across individuals as possible explanations for this gap in sentiment.

We proceed with a brief literature review that outlines the basic theoretical framework for understanding the relationship between consumer sentiment and consumer behavior. Then we provide a description of how consumer sentiment is measured. After this, we continue with an analysis of the variations in sentiment across groups, while exploring explanations for the differences. Finally, we discuss the implications of our findings and comment on areas for future research.

**Literature review**

There is a large amount of literature that deals with the role of consumer sentiment in explaining consumption. The point of departure for much of this literature is the permanent income hypothesis (PIH) (Friedman, 1957). The PIH maintains that consumption is determined solely by individuals' incomes over their lifetimes—that is, expenditure depends only on permanent income (wealth). Consistently, Hall (1978) concluded that under conditions of perfect capital markets, the PIH can be approximated by a random walk, meaning that no past information (aside from that needed to measure lifetime income) is required to predict current consumption.

Research has found that consumption is partly determined by current income, a notion that is referred to as "excess sensitivity" of consumption relative to income. For example, Campbell and Mankiw (1990) find that only half of consumers tend to be "life-cyclers," following the PIH assumptions, while the others tend to be "rule-of-thumbers," or those who consume from their current income rather than just from their lifetime income. Studies have attributed excess sensitivity of consumption relative to current income to liquidity constraints and precautionary savings motives (Shea, 1995; Flavin, 1991; Alessie and Lusardi, 1997). Liquidity constraints mean that individuals may not be able to borrow as they desire. That is, even if consumers anticipate more income (and consumers' confidence increases), with binding liquidity constraints, they will not be able to immediately act on the improvement in permanent income; the consumers will increase consumption only when the rise in income materializes. Some consumers accumulate precautionary savings when there is uncertainty relative to future income, which will cause them to have higher expected utility, since they reduce current consumption in case of a drop in income. In other words, even if consumers' financial positions remain unchanged, greater uncertainty about their future positions (hence a decrease in confidence) might cause consumers to engage in precautionary savings, which would affect their marginal propensity to consume. If lower consumer confidence reflects higher uncertainty about the future and enhances the precautionary motive for savings, then lower consumer confidence today causes consumption to decrease today relative to tomorrow. In contrast, higher consumer confidence is associated with lower savings and more consumption in the present, as well as lower consumption growth in the future. In the PIH framework, the ability of consumer sentiment to explain consumption arises from the fact that consumer sentiment serves as a proxy for liquidity constraints and precautionary savings motives (Acemoglu and Scott, 1994).

Empirical research using micro-level sentiment data has focused on inflation expectations of households. For example, Bryan and Venkatu (2001b) find that predictions of inflation significantly differ by socioeconomic and demographic characteristics of consumers. Palmqvist and Strömberg (2004), Lombardelli and Saleheen (2003), and Ranchhod (2003) find similar results in Sweden, the United Kingdom, and New Zealand. Souleles (2004) provides some explanations for people's differences in sentiment as he suggests that differences in people's expectations may be due to time-varying group-specific shocks (for example, during a recession the less educated may be disproportionately adversely affected). Another set of research examines the nature of the information to which consumers have access. This includes information that might help form their expectations, such as private local information, information they have gathered from their industries, or news media reports. For example, Dunn and Mirzaie (2004) calculate manufacturing employment concentration as a proxy to measure agents' private information to explain regional variations in consumer confidence. Sims (2003) presents a theoretical framework for evaluating the way people process information, accounting for the fact that people might have capacity constraints in processing information and extracting signals from the information that is transmitted to them. (That is, two people may be exposed to the same information, but they may not assimilate or use the information the same way. Therefore, their expectations of the same event may be different.) Dom and Morin (2004) analyze the role of the news media. They suggest that even if media coverage affects consumer sentiment, the effects are very short-lived. These findings underscore the difficulty in assessing the role of information in consumer sentiment and, ultimately, consumer behavior.
Friedman (1957) and Hall (1978), under the PIH framework, assume specific types of preferences—exogenous and stochastic income, no borrowing constraints, and rational expectations. However, the field of behavioral economics has extended our understanding of preferences to account for “psychological factors,” such as addiction and lack of self-control (Gul and Pesendorfer, 2002) and discrimination (Becker, 1976). There is in fact a tradition of portraying the connection between sentiment and behavior in psychological terms: John Maynard Keynes (1936) wrote that household consumption is influenced by “spontaneous optimism” and “animal spirit.” Similarly, George Katona (1975)—the founder of the Survey Research Center (SRC) at the University of Michigan, which generates the ICS—explained that in addition to factors that affect a consumer’s ability to pay, consumption is based on a consumer’s “willingness to pay.” These suggest that households form their expectations about the future based on preferences, technology, market frictions or borrowing constraints, and subjective experiences; indexes like the ICS and CCI are summaries of their views. This article builds upon this literature with its exploration of the possible links between consumer sentiment, personal characteristics of individuals, their subjective experiences, and exposure to news information.

**Measuring consumer sentiment**

While sentiment surveys are well known, their methods of construction are more obscure. Here, we describe how the Index of Consumer Sentiment from the University of Michigan’s Survey Research Center is designed, and examine business cycle components of its trends. This aggregate index, the ICS, is constructed using a formula based on responses to the following five survey questions. (The names of the variables, as identified by the survey, are in parentheses after the questions.)

1) We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago? *(PAGO)*

2) Now looking ahead—do you think that a year from now you (and your family living there) will be better off financially, or worse off, or just about the same as now? *(PEXP)*

3) Now turning to business conditions in the country as a whole—do you think that during the next 12 months we’ll have good times financially, or bad times, or what? *(BUS12)*

4) Looking ahead, which would you say is more likely—that in the country as a whole we’ll have continuous good times during the *next five years* or so, or that we will have periods of widespread unemployment or depression, or what? *(BUS5)*

5) About the big things people buy for their homes—such as furniture, a refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or bad time for people to buy major household items? *(DUR)*

To compute the ICS, first an index for each of the five questions is constructed as the “net balance,” where the proportion of negative responses is subtracted from the proportion of positive responses. The overall ICS is then calculated as an average of the net balance for these questions. There are two other indexes derived from these questions. The Index of Current Economic Conditions (ICC) is based on the two questions that ask about present personal and economic situations, *PAGO* and *DUR*. The Index of Consumer Expectations (ICE) is based on the three questions that ask about consumer-expected changes in business conditions and respondents’ income, *PEXP*, *BUS12*, and *BUS5*.

Figure 1 plots the three indexes—the ICS, the ICE, and the ICC—using quarterly data for the period 1978 to 2004. (By design, the ICS lies between the ICE and the ICC. The correlation between the ICE and the ICC is 0.82.) Looking at the figure, one can note the relationship between the indexes and business cycles. From 1978 to 2004, there were four recessions. These are shown as shaded regions in the figure. The most recent was from March 2001 to November 2001 (2001Q1–2001Q4). The ICS always takes a dip during a recession, although there are some brief intervals outside of the recession periods when this index also takes a dip. The three indexes begin decreasing one to four quarters ahead of three of the four recessions. (The one exception is the 1980 recession, at which time the indexes fell as the recession began.) The indexes rise prior to all upturns. These observations suggest the potential predictive power of the indexes. The indexes climbed to historically high levels throughout the expansionary years of the 1990s, before a reversal of the trend prior to the 2001 recession. Various researchers have found that the Index of Consumer Expectations has some predictive power for GDP, consumption, and the stock market, among other outcomes.

**Sentiment and demographic characteristics**

Next, we examine the ICS by group on a quarterly basis (due to space constraints, we do not report
we populations males, populations 22 1Q/2006, the ing it the in the characteristics.
We (age, age, for example, the poor relative to the nonpoor. The set of plots in figure 2 shows that the populations that we classify as vulnerable populations have, on average, lower confidence than their counterparts. Figure 2 reveals a number of patterns: For the poor and nonpoor, we observe a large gap in the ICS in the 1980s, a sharp contrast to what later occurred with the expansion of the 1990s. In both periods, the trend in the ICS for the poor is more variable than that of the nonpoor. The confidence pattern for the elderly (age 65 or older) and non-elderly is similar to that which we observe for the poor and nonpoor—that is, it is lower and more variable for the elderly. Comparing the ICS of blacks and whites, we find that during the 1980 and the 1981–82 recessions there was an increase in the gap between the two groups—consistent with findings in other research that blacks may have been disproportionately affected by these recessions (Wall, 2003). Starting after the 1990–91 recession, there was a tendency for the gap in the ICS between blacks and whites to narrow. The gap appears to have wide-

Because the ICS is made up of disparate questions, the changes over time in the ICS by group as indicated by figure 2 are not easy to interpret. The ICS is made up of five component questions concerning personal financial situations as well as general business conditions. These responses can move in opposite directions, even within a demographic group. If more blacks since 1991 are responding that their personal financial situations have improved relative to the proportion of blacks who are projecting that the economy will do well, then the improvement in their ICS may be interpreted as reflecting an improvement in their personal financial situations. The opposite would indicate that blacks simply think that the economy will do well, although they do not think that their own situations have changed. We take a closer look at the components of the index to ascertain the factors that might explain changes over time in the ICS by group.

Figures 3 through 6 show the results of disaggregating the ICS into its component questions for selected groups. (The DUR variable is not reported in our figures because the differences in responses by group are minor by comparison to those of the other variables.) First, consider the results for blacks and whites (figure 3). Before 1991, the fact that consumer sentiment among blacks was lower can be attributed to their lower confidence in both the overall economy and their personal financial situation. This is indicated by the gap in the BUS5, BUS12, and PAGO series over that period. In contrast, after 1991 the relatively higher consumer confidence of blacks can be attributed to the fact that they had relatively more confidence regarding their financial situation, as indicated by the small gap in the PAGO series between blacks and whites and as evidenced by higher PEXP among blacks in the post-1991 period. In short, the convergence of black and white sentiment measures since 1991 can be largely attributed to improvement in black consumers’ reports of their own financial situation.
FIGURE 2
Index of Consumer Sentiment by demographic group

A. Poverty status

B. Age

C. Gender

D. Race

E. Education attainment

Source: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers.
Turning to figure 4, we note that the elderly have a similar level of confidence in the overall economy compared with the non-elderly. This is indicated by the fact that the BUS12 and BUS5 series of the elderly and non-elderly virtually coincide in both level and pattern. The triggering factor for the lower overall consumer sentiment shown by the elderly is their lower assessment of their financial situation, at present and as predicted for the future. This can be seen from the gap between the PAGO and the PEXP series of the elderly and the non-elderly. Similarly, for those without a high school diploma (figure 5) and the poor (figure 6), a lower confidence in their personal financial situation (PAGO) seems to be a contributing source of lower consumer sentiment overall. (The pattern for the PEXP series of the poor and nonpoor is less clear than that for PAGO.) In addition, those without a high school diploma and the poor are also less confident about the economy as a whole (BUS12 and BUS5).

It is not surprising to find that respondents have different expectations concerning their personal experiences. However, it is puzzling that they should also have different expectations of the same economic events (business conditions). One possibility is that they form their expectations of the economy based on their own subjective experiences. We investigate this possibility by looking at respondents’ expectations of unemployment and their actual (group-level) experiences of unemployment. Besides the five questions mentioned previously, respondents are asked in
FIGURE 4
Components of Index of Consumer Sentiment by age

A. PAGO index

<table>
<thead>
<tr>
<th>Year</th>
<th>1978</th>
<th>'83</th>
<th>'88</th>
<th>'93</th>
<th>'98</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>190</td>
<td>170</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Non-elderly</td>
<td>150</td>
<td>130</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
</tbody>
</table>

B. BUS5 index

<table>
<thead>
<tr>
<th>Year</th>
<th>1978</th>
<th>'83</th>
<th>'88</th>
<th>'93</th>
<th>'98</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>190</td>
<td>170</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Non-elderly</td>
<td>150</td>
<td>130</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
</tbody>
</table>

C. PEXP index

<table>
<thead>
<tr>
<th>Year</th>
<th>1978</th>
<th>'83</th>
<th>'88</th>
<th>'93</th>
<th>'98</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>190</td>
<td>170</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Non-elderly</td>
<td>150</td>
<td>130</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
</tbody>
</table>

D. BUS12 index

<table>
<thead>
<tr>
<th>Year</th>
<th>1978</th>
<th>'83</th>
<th>'88</th>
<th>'93</th>
<th>'98</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly</td>
<td>190</td>
<td>170</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>Non-elderly</td>
<td>150</td>
<td>130</td>
<td>110</td>
<td>90</td>
<td>70</td>
<td>50</td>
</tr>
</tbody>
</table>

Note: See the text for definitions of the variables.
Source: Authors' calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers.

the SRC survey: “How about people out of work during the coming 12 months—do you think that there will be more unemployment than now, about the same, or less?” We calculate the coefficient of correlation between the response to this question and actual unemployment to ascertain whether a relationship exists between the two series. The correlation coefficient is a measure of the degree of linear association between two variables, with −1 indicating perfect negative association, +1 indicating perfect positive association, and 0 indicating no association. The results, which are summarized in table 1 (p. 28), indicate that a respondent’s expectation of unemployment corresponds to the experience of her own group (in the second column) more closely than it corresponds to the experience of the population as a whole (in the first column), even though the question asks about the general situation. This suggests that group-based aggregate experiences tend to inform individuals’ expectations of the economy.

We also consider people’s expectations of price changes and the actual changes in the Consumer Price Index. We find evidence consistent with previous studies that expectations of inflation vary systematically by demographic and socioeconomic characteristics. In particular, female unmarried heads of households, the poor, the less educated, and blacks have higher expectations of inflation. Several studies have offered potential explanations of the sources of differences in inflation forecasts. These include differences in information sets across agents and substantial variation in
FIGURE 5
Components of Index of Consumer Sentiment by education attainment

A. PAGO
index

B. BUS5
index

C. PEXP
Index

D. BUS12
Index

1978  '83 '88 '93 '98 '03
Without high school diploma College graduate
Without high school diploma College graduate
Without high school diploma College graduate
Without high school diploma College graduate

Note: See the text for definitions of the variables.
Source: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers.

the cost of consumption baskets across individual households (Carlson and Valev, 1999; Michael, 1979). However, McGranahan and Paulson (2005) derive inflation rates for specific population groups by re-weighting price index components (price indexes for individual items) based on the market basket consumed by members of the population group of interest. They find that from 1983 to 2004, the series are similar for all the groups. Given this, variation among people’s perceptions of inflation is difficult to explain in the context of people’s own subjective experiences with inflation (since inflation does not seem to vary by group).

Next, we review a two-part question about news in the SRC survey to explore the role of information in consumer sentiment. This survey question asks whether the respondent has “heard of any favorable or unfavorable changes in business conditions.” If the respondent answers “yes,” it further asks, “What did you hear?” Respondents can provide up to two responses to this two-part question. From the responses, we can examine whether different groups have different levels of exposure to the news. Table 2 presents the percentage of people that have heard any news—those who responded “yes” to the first part of the question above—by demographic group. In table 2, we see that only 58 percent of the sample reports hearing any news concerning the economy. The differences across groups are quite substantial. While 72 percent of college graduates report
having heard about business conditions, only 38 percent of individuals without a high school diploma report having heard about any such news. Similarly, 60 percent of the nonpoor report having heard news about the economy as opposed to 41 percent of the poor.

To further investigate the role of news, we divided the news into five areas—namely, employment, prices, government programs or decisions, output/GDP, and conditions in a respondent’s own industry. Tabulations of news sources by group are presented in table 3. Among those who have heard news, the most common type of news pertains to the employment situation. A large number of people also have heard news about GDP, news about their own industry, and (a disproportionate amount of bad) news about prices. The prevalence of news about the respondent’s own line of work suggests that news is not necessarily objective in nature, but is filtered through a subjective lens. Furthermore, considering demographic characteristics, we found that most groups get a consistent fraction of their news on the same topics. However, there is one exception to this pattern: Information on GDP is more common among more educated and wealthier households than among those that are less educated and have lower incomes. Less than half of the individuals in the vulnerable groups we investigate report having heard any news about business conditions. This fraction becomes even smaller if we exclude individuals who have
only heard news about conditions concerning their own industries. If individuals have no exposure to news, they must be forming their assessments of the macroeconomy based on other information. This other information would likely come from their personal experiences, such as noticing prices in local stores and conversing with peers. If this is the case, the differences in expectations among the different groups are easier to explain. Their disparate experiences, as evidenced in the PMGO responses, translate into different expectations for the economy, especially given the relative absence of objective information from news sources within certain groups.

**Empirical analysis**

Previously, we have shown descriptively that the Index of Consumer Sentiment differs across demographic groups. In particular, we have shown that respondents’ perceptions are less positive for those groups that we label as vulnerable based on relative income. Here, we take a closer look at the responses of the individuals who make up the groups. Specifically, we look at the microdata to gain a better measure of the contributions of each demographic attribute to each index. We base the analysis on measures of the ICS, ICC, and ICE provided in the data for each individual in the sample.14

We can ask how the different demographic attributes of individuals contribute to their measures in these indexes. We do this via regression analysis. We ask what the contribution of each demographic characteristic is to the different indexes, while holding the other characteristics constant. The results from a series of regressions are presented in table 4 for the ICS. The table contains three separate regressions. In the regression presented in the first column, we predict the index based only on demographic attributes and region of residence. In the second column, we add in four measures of the conditions of the macroeconomy during the month of the survey—the unemployment rate, the percent change in real personal income from one year ago, the year-over-year inflation rate, and the percent change in the real value of the Dow Jones Industrial Average from one year ago. Higher income is likely to trigger higher consumption, with accompanying stronger consumer confidence. Therefore, we expect a positive relationship between past income and confidence. An increase in the unemployment rate is likely to generate an increase in uncertainty among consumers, even though they may not themselves be unemployed. This is likely to increase precautionary

**TABLE 1**

<table>
<thead>
<tr>
<th>Group</th>
<th>Correlation with overall unemployment</th>
<th>Correlation with own group’s unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>0.814</td>
<td></td>
</tr>
<tr>
<td>Without high school diploma</td>
<td>0.490</td>
<td>0.541</td>
</tr>
<tr>
<td>College graduate</td>
<td>0.804</td>
<td>0.826</td>
</tr>
<tr>
<td>White</td>
<td>0.680</td>
<td>0.790</td>
</tr>
<tr>
<td>Black</td>
<td>0.676</td>
<td>0.808</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.360</td>
<td>0.607</td>
</tr>
</tbody>
</table>

Sources: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers and U.S. Bureau of Labor Statistics.

**TABLE 2**

<table>
<thead>
<tr>
<th>Group</th>
<th>Heard any news</th>
<th>Heard news within own industry</th>
<th>Heard only news within own industry and no other news</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>57.63</td>
<td>21.17</td>
<td>10.11</td>
</tr>
<tr>
<td>Elderly</td>
<td>47.80</td>
<td>20.62</td>
<td>10.04</td>
</tr>
<tr>
<td>Non-elderly</td>
<td>59.55</td>
<td>21.30</td>
<td>10.13</td>
</tr>
<tr>
<td>Poor</td>
<td>40.77</td>
<td>18.13</td>
<td>9.47</td>
</tr>
<tr>
<td>Nonpoor</td>
<td>59.78</td>
<td>21.51</td>
<td>10.18</td>
</tr>
<tr>
<td>Top income quartile</td>
<td>68.00</td>
<td>20.67</td>
<td>8.62</td>
</tr>
<tr>
<td>Bottom income quartile</td>
<td>43.58</td>
<td>19.26</td>
<td>11.22</td>
</tr>
<tr>
<td>Without high school diploma</td>
<td>38.46</td>
<td>20.23</td>
<td>10.81</td>
</tr>
<tr>
<td>College graduate</td>
<td>72.19</td>
<td>19.22</td>
<td>7.80</td>
</tr>
<tr>
<td>White</td>
<td>59.23</td>
<td>21.83</td>
<td>10.35</td>
</tr>
<tr>
<td>Black</td>
<td>47.55</td>
<td>15.98</td>
<td>7.80</td>
</tr>
<tr>
<td>Hispanic</td>
<td>49.63</td>
<td>18.25</td>
<td>10.20</td>
</tr>
</tbody>
</table>

Note: See the text for further details.

Source: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers.
saving and lower consumption and confidence. We therefore expect a negative relationship between consumer confidence and unemployment. Increased inflation decreases the purchasing power of the consumer. Rising inflation can create an erosion of purchasing power that could lower consumer confidence. Greater price volatility or inflation would create more uncertainty surrounding real wage changes. Because of this, changes in inflation are expected to be negatively related to consumer sentiment. Stock market prices may affect consumer confidence in two ways: An increase in stock market prices may increase wealth and directly boost confidence, or rising stock markets may act as an indicator of higher expected labor income, which would also increase confidence.

In the third column, we replace the macroeconomic variables with a series of month-year dummies. These dummies control for any changes in the economy or overall national situation that affect all respondents in a given month.

We see a number of patterns in table 4. First, each of our attributes indicating vulnerability, in terms of relative income, has an independent, statistically significant negative effect on the index measure. The poor, females, the less educated, the elderly, blacks, and Hispanics are less optimistic about the economy.15 Second, the condition of the macroeconomy has a strong effect on consumer sentiment. We can see this in two ways—first, through the statistically significant effect of the macroeconomic variables on the index measure and, second, through the increase in the explanatory power of the regression as a whole (as measured by the adjusted R-squared presented in the final row of the table) once these independent variables are added. At the same time, the coefficients on the attributes change only slightly with the addition of the macroeconomic measures or the time dummies. The one exception to this is the Hispanic indicator, which goes from being positive to negative once the macroeconomic measures are included. Further investigation suggests that this is the result of the larger Hispanic population in the later years of the sample when the economy is also doing well. As a result, the Hispanic measure in the initial regressions is partly picking up the positive association between the condition of the economy and the size of the Hispanic population. We also find that most of the contribution from the time dummies is captured by the four measures of the macroeconomic situation. Although the adjusted R-squared increases when the dummies are introduced, the jump is not dramatic. If we look at the individual macroeconomic measures, we find that a respondent’s sentiment is positively correlated with the increase in stock market prices and changes in disposable income and negatively correlated with the unemployment and inflation rates. All of these signs are in the direction we would anticipate because increasing income and stock market prices are indicators of economic strength and a rising unemployment rate is a sign of economic weakness. While high inflation can be a sign of rapid economic activity, it negatively affects consumer well-being. In the remainder of this article, we include our macroeconomic measures rather than the series of time dummies because the macroeconomic variables lead to more straightforward interpretations.

We ran a similar regression analysis for the ICE and the ICC.16 The results are broadly similar to those for the ICS. Groups with relatively lower income are significantly less optimistic and have lower assessments

<table>
<thead>
<tr>
<th>Group</th>
<th>Employment</th>
<th>Prices</th>
<th>Government</th>
<th>GDP</th>
<th>Own industry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(………………)</td>
<td>(………………)</td>
<td>(………………)</td>
<td>(………………)</td>
<td>(………………)</td>
</tr>
<tr>
<td>All</td>
<td>38.01</td>
<td>22.89</td>
<td>12.14</td>
<td>25.56</td>
<td>22.07</td>
</tr>
<tr>
<td>Elderly</td>
<td>44.42</td>
<td>15.64</td>
<td>12.35</td>
<td>20.47</td>
<td>21.52</td>
</tr>
<tr>
<td>Non-elderly</td>
<td>36.81</td>
<td>24.23</td>
<td>12.11</td>
<td>26.75</td>
<td>22.20</td>
</tr>
<tr>
<td>Poor</td>
<td>45.80</td>
<td>16.94</td>
<td>14.10</td>
<td>15.61</td>
<td>18.81</td>
</tr>
<tr>
<td>Nonpoor</td>
<td>37.50</td>
<td>23.42</td>
<td>12.03</td>
<td>26.78</td>
<td>22.43</td>
</tr>
<tr>
<td>Top income quartile</td>
<td>36.60</td>
<td>21.59</td>
<td>13.69</td>
<td>25.55</td>
<td>21.82</td>
</tr>
<tr>
<td>Bottom income quartile</td>
<td>47.47</td>
<td>10.80</td>
<td>14.06</td>
<td>17.10</td>
<td>20.16</td>
</tr>
<tr>
<td>Without high school diploma</td>
<td>42.41</td>
<td>20.47</td>
<td>11.93</td>
<td>14.73</td>
<td>20.72</td>
</tr>
<tr>
<td>College graduate</td>
<td>35.13</td>
<td>26.92</td>
<td>13.29</td>
<td>36.21</td>
<td>20.17</td>
</tr>
<tr>
<td>White</td>
<td>37.18</td>
<td>23.34</td>
<td>11.78</td>
<td>26.63</td>
<td>22.77</td>
</tr>
<tr>
<td>Black</td>
<td>46.90</td>
<td>20.09</td>
<td>14.78</td>
<td>17.23</td>
<td>16.56</td>
</tr>
<tr>
<td>Hispanic</td>
<td>40.61</td>
<td>18.38</td>
<td>13.38</td>
<td>21.76</td>
<td>18.95</td>
</tr>
</tbody>
</table>

Note: See the text for further details.
Source: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers.
### TABLE 4

Determinants of the Index of Consumer Sentiment

<table>
<thead>
<tr>
<th></th>
<th>Demographic characteristics</th>
<th>Add macroeconomic measures</th>
<th>Add month-year dummies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor</td>
<td>-6.447***</td>
<td>-7.340***</td>
<td>-7.986***</td>
</tr>
<tr>
<td></td>
<td>(0.346)</td>
<td>(0.337)</td>
<td>(0.336)</td>
</tr>
<tr>
<td>Resides in Northeast</td>
<td>-1.859***</td>
<td>-1.307***</td>
<td>-1.304***</td>
</tr>
<tr>
<td></td>
<td>(0.263)</td>
<td>(0.248)</td>
<td>(0.245)</td>
</tr>
<tr>
<td>Resides in South</td>
<td>2.024***</td>
<td>1.954***</td>
<td>1.956***</td>
</tr>
<tr>
<td></td>
<td>(0.230)</td>
<td>(0.218)</td>
<td>(0.216)</td>
</tr>
<tr>
<td>Resides in West</td>
<td>-0.172</td>
<td>0.286</td>
<td>0.314</td>
</tr>
<tr>
<td></td>
<td>(0.264)</td>
<td>(0.251)</td>
<td>(0.248)</td>
</tr>
<tr>
<td>Female</td>
<td>-8.694***</td>
<td>-8.652***</td>
<td>-8.737***</td>
</tr>
<tr>
<td></td>
<td>(0.180)</td>
<td>(0.171)</td>
<td>(0.169)</td>
</tr>
<tr>
<td>Without high school diploma</td>
<td>-17.516***</td>
<td>-12.918***</td>
<td>-12.868***</td>
</tr>
<tr>
<td></td>
<td>(0.307)</td>
<td>(0.298)</td>
<td>(0.297)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>-8.380***</td>
<td>-6.341***</td>
<td>-6.363***</td>
</tr>
<tr>
<td></td>
<td>(0.225)</td>
<td>(0.214)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>Some college</td>
<td>-3.296***</td>
<td>-1.959***</td>
<td>-1.983***</td>
</tr>
<tr>
<td></td>
<td>(0.247)</td>
<td>(0.233)</td>
<td>(0.230)</td>
</tr>
<tr>
<td>Black</td>
<td>-5.457***</td>
<td>-6.141***</td>
<td>-6.121***</td>
</tr>
<tr>
<td></td>
<td>(0.326)</td>
<td>(0.307)</td>
<td>(0.305)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1.822***</td>
<td>-1.898***</td>
<td>-1.818***</td>
</tr>
<tr>
<td></td>
<td>(0.452)</td>
<td>(0.436)</td>
<td>(0.434)</td>
</tr>
<tr>
<td>Other race (nonwhite)</td>
<td>-0.608</td>
<td>-2.786***</td>
<td>-2.773***</td>
</tr>
<tr>
<td></td>
<td>(0.957)</td>
<td>(0.574)</td>
<td>(0.563)</td>
</tr>
<tr>
<td>Elderly</td>
<td>-7.731***</td>
<td>-9.072***</td>
<td>-9.043***</td>
</tr>
<tr>
<td></td>
<td>(0.263)</td>
<td>(0.254)</td>
<td>(0.251)</td>
</tr>
<tr>
<td>Family size</td>
<td>0.192***</td>
<td>0.434***</td>
<td>0.469***</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.068)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Married</td>
<td>-1.080***</td>
<td>-1.152***</td>
<td>-1.134***</td>
</tr>
<tr>
<td></td>
<td>(0.204)</td>
<td>(0.193)</td>
<td>(0.191)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td></td>
<td></td>
<td>-3.439***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.060)</td>
</tr>
<tr>
<td>Percent change in real disposable income (year over year)</td>
<td>2.594***</td>
<td>(0.046)</td>
<td></td>
</tr>
<tr>
<td>Inflation rate (year over year)</td>
<td>-1.280***</td>
<td>(0.034)</td>
<td></td>
</tr>
<tr>
<td>Percent change in real Dow Jones Industrial Average (year over year)</td>
<td>0.205***</td>
<td>(0.006)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>100.782***</td>
<td>119.720***</td>
<td>105.505***</td>
</tr>
<tr>
<td></td>
<td>(0.305)</td>
<td>(0.493)</td>
<td>(1.869)</td>
</tr>
<tr>
<td>Observations</td>
<td>167.507</td>
<td>167.507</td>
<td>167.507</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.06</td>
<td>0.16</td>
<td>0.18</td>
</tr>
</tbody>
</table>

*Significant at the 10 percent level.

**Significant at the 5 percent level.

***Significant at the 1 percent level.

Notes: Robust standard errors are in parentheses. See the text for further details.

Sources: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers; U.S. Bureau of Economic Analysis; U.S. Bureau of Labor Statistics; and Yahoo! Finance.
of the current state of the economy than their comple-
ments. For these other measures we also observe the
same pattern for Hispanic respondents—positive co-
efficients that reverse sign once the macroeconomic
measures or dummies are added in.

If we compare the results across the three indexes,
we find that being poor lowers the ICC by 12,
ICE by 5, and ICS by 7 index points relative to being
nonpoor. It is not surprising that the poor differ most
from the nonpoor in their assessments of current eco-
nomic conditions because being poor is partly the re-
sult of current financial distress. For individuals with
a high school diploma, we find that their ICC, ICE,
and ICS are all lower by a similar amount—between
12 and 13 index points—relative to those of college
graduates (the omitted category). Because low edu-
cation affects both current and future employment
prospects, this similarity across results is also not sur-
prising. For females, we find that relative to males
their ICC is lower by 6, their ICE by 10, and their ICS
by 9 index points—a pattern that suggests greater
pessimism on the part of women, despite their compara-
able assessment of the current economic conditions.
This is consistent with other research that has found
women to have lower consumer confidence (Bryan
and Venkatu, 2001a).

The effects of macroeconomic variables remain
similar to the effects found for the ICS as discussed
previously. We find that the coefficient on the unem-
ployment rate is larger in absolute value in the ICC
regression than in the ICE regression. In contrast, the
coefficients on the three other macroeconomic vari-
ables—disposable income, inflation, and stock mar-
kets—are larger in absolute value in the ICE regression
than in the ICC regression. By design, the ICS coef-
ficients lie between those of the other two indexes.

To gain a better understanding of the pattern of these
responses and the rationale for them, we turn to an in-
vestigation of the five questions from which the three
indexes are calculated. As argued previously, in addi-
tion to giving us further insight into the rationale be-
hind the differences across the indexes, investigating
the responses themselves allows us to move away from
the arbitrary nature of the index calculations. In the
regressions presented in table 4 and the results for ICC
and ICE, we are explaining continuous index calcula-
tions, but those calculations are based on discrete answers to
a series of five questions. We now look at the discrete
answers to the questions with respect to the ICS.

Table 5 presents the results from a series of ordered
logit regressions where the dependent variable indi-
cates whether the respondent is positive, neutral, or
negative about the question being asked. For instance,
for the PAGO question, the respondents are separated
into individuals who are better off than a year ago,
the same as a year ago, and worse off than a year
ago. For BUS12 and BUS5, respondents can say that
they expect good times, good times with qualification,
mixed experiences, bad times with qualifications,
and bad times. For these regressions, we group the
two positive and two negative responses together, in
order to be more consistent with the other questions.
These responses are also grouped in the calculation
of the ICS. Using the ordered logit framework, we
are able to include individuals who are neutral or the
same. These individuals are omitted from the calcu-
lation of the published indexes.17

For ease of interpretation, we present odds ratios
and z-statistics rather than coefficients and standard
errors in table 5. The odds ratios indicate how being
in the underlying group contributes to the probability
of responding more positively to the question relative
to the probability of responding more negatively. A co-
efficient less than one indicates that belonging to the
group leads to more negative responses relative to being
in the omitted category. The asterisks indicate whether
the odds ratio is significantly different from one,
which is equivalent to asking whether the estimated
coefficient is different from zero.

If we look at the first column of table 5 (the pre-
dictors of the responses to the PAGO question), we
see that the poor, females, the less educated, blacks,
and the elderly are all more likely to be negative than
positive about their previous year relative to individu-
als in the omitted categories (nonpoor, males, college
graduates, whites, and the non-elderly). The coeffi-
cients on the education categories are monotonically
increasing with education level. The smallest coeffi-
cients (those furthest from one) are on the groups we
know from other sources to have poor earnings poten-
tial—the poor, those without a high school diploma,
and the elderly. The effects of the macroeconomic
variables remain very similar to the effects found in
the continuous regressions presented earlier. The PAGO
question is subjective—that is, it asks individuals
about their own economic experiences. Given that
members of groups have different economic experi-
ences, it is not surprising that we find differences
across groups.

The other subjective question is PEXP, which
asks individuals to anticipate whether they will be
financially better off in a year. These results are quite
different from those found with PAGO. The poor do
not have responses significantly different from the
nonpoor. Hispanics and blacks are both more optimis-
tic than whites. However, the less educated, females,
### TABLE 5

Determinants of component questions of the Index of Consumer Sentiment

<table>
<thead>
<tr>
<th></th>
<th>1 PAGO</th>
<th>2 DUR</th>
<th>3 PEXP</th>
<th>4 BUS12</th>
<th>5 BUS5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor</td>
<td>0.649***</td>
<td>0.742***</td>
<td>0.998</td>
<td>0.805***</td>
<td>0.756***</td>
</tr>
<tr>
<td></td>
<td>(23.56)</td>
<td>(13.72)</td>
<td>(9.13)</td>
<td>(9.96)</td>
<td>(13.25)</td>
</tr>
<tr>
<td>Resides in Northeast</td>
<td>0.904***</td>
<td>0.961**</td>
<td>0.98</td>
<td>0.937***</td>
<td>1.002</td>
</tr>
<tr>
<td></td>
<td>(7.46)</td>
<td>(2.32)</td>
<td>(1.39)</td>
<td>(4.16)</td>
<td>(0.12)</td>
</tr>
<tr>
<td>Resides in South</td>
<td>1.079***</td>
<td>0.982</td>
<td>1.163***</td>
<td>1.088***</td>
<td>1.049***</td>
</tr>
<tr>
<td></td>
<td>(6.36)</td>
<td>(1.21)</td>
<td>(12.14)</td>
<td>(6.09)</td>
<td>(3.68)</td>
</tr>
<tr>
<td>Resides in West</td>
<td>0.983</td>
<td>0.970*</td>
<td>1.177***</td>
<td>0.994</td>
<td>0.978</td>
</tr>
<tr>
<td></td>
<td>(1.24)</td>
<td>(1.76)</td>
<td>(11.35)</td>
<td>(0.37)</td>
<td>(1.50)</td>
</tr>
<tr>
<td>Female</td>
<td>0.848***</td>
<td>0.771***</td>
<td>0.835***</td>
<td>0.680***</td>
<td>0.620***</td>
</tr>
<tr>
<td></td>
<td>(17.68)</td>
<td>(22.20)</td>
<td>(18.56)</td>
<td>(35.67)</td>
<td>(47.24)</td>
</tr>
<tr>
<td>Without high school diploma</td>
<td>0.587***</td>
<td>0.739***</td>
<td>0.571***</td>
<td>0.766***</td>
<td>0.557***</td>
</tr>
<tr>
<td></td>
<td>(33.42)</td>
<td>(15.58)</td>
<td>(32.83)</td>
<td>(13.95)</td>
<td>(31.96)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>0.722***</td>
<td>0.98</td>
<td>0.742***</td>
<td>0.907***</td>
<td>0.719***</td>
</tr>
<tr>
<td></td>
<td>(27.66)</td>
<td>(1.33)</td>
<td>(24.51)</td>
<td>(7.29)</td>
<td>(26.18)</td>
</tr>
<tr>
<td>Some college</td>
<td>0.825***</td>
<td>1.039***</td>
<td>0.946***</td>
<td>1.015</td>
<td>0.908***</td>
</tr>
<tr>
<td></td>
<td>(14.80)</td>
<td>(2.37)</td>
<td>(4.18)</td>
<td>(0.98)</td>
<td>(7.05)</td>
</tr>
<tr>
<td>Black</td>
<td>0.851***</td>
<td>0.841***</td>
<td>1.221***</td>
<td>0.709***</td>
<td>0.588***</td>
</tr>
<tr>
<td></td>
<td>(9.28)</td>
<td>(8.45)</td>
<td>(11.16)</td>
<td>(17.60)</td>
<td>(27.57)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.978</td>
<td>0.823***</td>
<td>1.170***</td>
<td>0.958</td>
<td>0.846***</td>
</tr>
<tr>
<td></td>
<td>(0.93)</td>
<td>(8.62)</td>
<td>(6.25)</td>
<td>(1.55)</td>
<td>(6.39)</td>
</tr>
<tr>
<td>Other race (nonwhite)</td>
<td>0.908***</td>
<td>0.819***</td>
<td>0.978</td>
<td>0.905***</td>
<td>0.945*</td>
</tr>
<tr>
<td></td>
<td>(3.05)</td>
<td>(6.16)</td>
<td>(6.68)</td>
<td>(2.75)</td>
<td>(1.66)</td>
</tr>
<tr>
<td>Elderly</td>
<td>0.600***</td>
<td>0.814***</td>
<td>0.365***</td>
<td>0.953***</td>
<td>1.075***</td>
</tr>
<tr>
<td></td>
<td>(38.29)</td>
<td>(11.93)</td>
<td>(67.86)</td>
<td>(2.83)</td>
<td>(4.59)</td>
</tr>
<tr>
<td>Family size</td>
<td>1.031***</td>
<td>1.005</td>
<td>1.029***</td>
<td>1.015***</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td>(7.96)</td>
<td>(1.10)</td>
<td>(7.42)</td>
<td>(3.50)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>Married</td>
<td>0.975**</td>
<td>0.962***</td>
<td>0.812***</td>
<td>0.993</td>
<td>1.014</td>
</tr>
<tr>
<td></td>
<td>(2.35)</td>
<td>(2.94)</td>
<td>(18.78)</td>
<td>(0.57)</td>
<td>(1.21)</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.904***</td>
<td>0.804***</td>
<td>0.976***</td>
<td>0.874***</td>
<td>0.906***</td>
</tr>
<tr>
<td></td>
<td>(30.66)</td>
<td>(54.67)</td>
<td>(6.94)</td>
<td>(35.40)</td>
<td>(27.63)</td>
</tr>
<tr>
<td>Percent change in real disposable income (year over year)</td>
<td>1.067***</td>
<td>1.080***</td>
<td>1.028***</td>
<td>1.195***</td>
<td>1.082***</td>
</tr>
<tr>
<td></td>
<td>(25.61)</td>
<td>(24.69)</td>
<td>(10.69)</td>
<td>(61.15)</td>
<td>(28.85)</td>
</tr>
<tr>
<td>Inflation rate (year over year)</td>
<td>0.981***</td>
<td>0.963***</td>
<td>0.941***</td>
<td>0.943***</td>
<td>0.940***</td>
</tr>
<tr>
<td></td>
<td>(10.20)</td>
<td>(17.31)</td>
<td>(31.25)</td>
<td>(27.15)</td>
<td>(30.04)</td>
</tr>
<tr>
<td>Percent change in real Dow Jones Industrial Average (year over year)</td>
<td>1.004***</td>
<td>1.009***</td>
<td>1.001***</td>
<td>1.016***</td>
<td>1.004***</td>
</tr>
<tr>
<td></td>
<td>(14.11)</td>
<td>(23.56)</td>
<td>(4.46)</td>
<td>(43.57)</td>
<td>(11.48)</td>
</tr>
<tr>
<td>Observations</td>
<td>167,012</td>
<td>158,277</td>
<td>163,085</td>
<td>152,227</td>
<td>155,809</td>
</tr>
</tbody>
</table>

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

Note: The absolute value of z-statistics are in parentheses. See the text for definitions of the variables and for further details.

Sources: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers; U.S. Bureau of Economic Analysis; U.S. Bureau of Labor Statistics; and Yahoo! Finance.
and the elderly remain more negative than the relevant omitted categories. The odds ratio for the elderly is especially low, 0.365. This indicates that the elderly are nearly three times as likely to anticipate being worse off as being better off in the coming year. This probably results from the limited scope for financial improvement among the elderly, many of whom are on fixed incomes and out of the labor force. The results for blacks and Hispanics are more difficult to explain. Looking at the raw data, we observe that the relative optimism of both groups arises from a higher likelihood of being positive and a lower likelihood of being neutral, not from a difference in pessimism. Given that blacks and Hispanics, on average, have lower incomes than whites, this finding may demonstrate that they anticipate a forthcoming improvement to their current relative income status. In other words, they expect their future financial experience to conform more closely to the overall population average. However, the response to the PAGO question indicates that this anticipation is misplaced. One year later a sample representing the same population is more likely than whites to report being worse off financially.

While the PAGO and PEXP questions are subjective and directed toward individual experiences, the remainder of the questions are more objective and ask respondents about their perceptions of the overall macroeconomic climate. The effect of demographic attributes on these perceptions is very consistent across the three responses. For all three outcomes, the poor, females, the less educated, and the nonwhite are all more pessimistic or negative. The magnitude of the odds ratios on these attributes is also consistent across the three outcomes. The results for the elderly are less consistent. Elderly respondents are more pessimistic about purchasing durable goods, slightly more pessimistic about the coming year, and slightly more optimistic about the next five years.

It is challenging to explain why different groups would have such different impressions about the prospects for the macroeconomy. For these objective questions all the different individuals are being asked about the same phenomenon—namely, their perception of the prospects for, or condition of, the general economy. In fact, the coefficients of the DUR, BUS12, and BUS5 regressions are not all that different from the coefficients of the PAGO regressions.

This led us to ask from where individuals are getting their expectations for the macroeconomy. If respondents are basing their expectations on their own experiences, then we would expect their different economic realities to translate into different expectations, as we find. On the other hand, if respondents are getting their information from a common source, such as the national news media, it is more difficult to explain this pattern in which demographic attributes affect expectations of the macroeconomy. Here, we look at this issue econometrically and ask how PAGO responses and news exposure translate into BUS12. In doing so, we are inquiring about the source of the BUS12 response. The results for the determinants of BUS12 are presented in table 6.

As before, we estimate an ordered logit model. We group the two positive and two negative responses together, and present odds ratios in the tables. The first column of table 6 investigates how good and bad changes in personal economic experiences over the past year translate into expectations for the national economy in the coming year, controlling for the macroeconomic climate. We find that individuals who are better off than a year ago are more optimistic about the coming year than those who are the same (the omitted category), and we find that those who are worse off are more pessimistic. This tells us that national expectations are partly driven by recent individual experiences. The macroeconomic variables have the expected magnitudes. In the second column, we add indicators of whether individuals have heard any good or bad news about business conditions. Individuals who have heard good news are more than twice as likely to report being optimistic as being pessimistic about economic prospects relative to those who have heard no news (the omitted category), while individuals who have heard bad news are only half as likely to be optimistic as pessimistic. The addition of these variables only changes the odds ratios on other variables by a small amount. In the third column of the table, we add interactions between hearing no news and recent past experience. We are inquiring whether individuals with no exposure to news place more weight on their own recent experience. We find that hearing no news and having a good past year render respondents more optimistic, while hearing no news and having a bad year render respondents more pessimistic. In other words, in the absence of external sources of information, individuals place more importance on their own experiences.

In the fourth column of the table, we add in the demographic characteristics. For the sake of comparison, we include the coefficients from the regression estimating BUS12 from table 5 in the final column of table 6. We find that even when controlling for past experiences and news exposure, the demographic characteristics matter. In fact, comparing the fourth and fifth columns of table 6, we find that the
### TABLE 6

**Determinants of BUS12**

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PAGO response</strong></td>
<td><strong>Add news</strong></td>
<td><strong>Add news x PAGO</strong></td>
<td><strong>Add demographic characteristics</strong></td>
<td><strong>Original demographic characteristics</strong></td>
</tr>
</tbody>
</table>

- Better off than a year ago: 1.486*** (32.38) | 1.430*** (28.00) | 1.371*** (19.62) | 1.338*** (16.81) |
- Worse off than a year ago: 0.575*** (41.28) | 0.585*** (38.49) | 0.635*** (25.51) | 0.651*** (22.71) |
- Unemployment rate: 0.889*** (32.53) | 0.869*** (36.80) | 0.869*** (36.77) | 0.873*** (33.57) |
- Percent change in real disposable income (year over year): 1.179*** (58.97) | 1.144*** (45.77) | 1.144*** (45.88) | 1.150*** (44.98) |
- Inflation rate (year over year): 0.946*** (27.12) | 0.938*** (29.70) | 0.938*** (29.56) | 0.937*** (27.89) |
- Percent change in real Dow Jones Industrial Average (year over year): 1.015*** (41.97) | 1.011*** (29.72) | 1.011*** (29.56) | 1.012*** (29.57) |
- Heard good economic news: 2.201*** (60.00) | 2.187*** (48.90) | 2.070*** (42.65) | 2.070*** (42.65) |
- Heard bad economic news: 0.454*** (72.09) | 0.450*** (51.75) | 0.433*** (50.91) | 0.433*** (50.91) |
- Heard no news x better off than a year ago: 1.108*** (4.58) | 1.124*** (4.93) | 1.108*** (4.58) | 1.124*** (4.93) |
- Heard no news x worse off than a year ago: 0.829*** (7.52) | 0.855*** (5.92) | 0.829*** (7.52) | 0.855*** (5.92) |
- Poor: 0.862*** (6.46) | 0.805*** (9.96) | 0.862*** (6.46) | 0.805*** (9.96) |
- Female: 0.713*** (29.66) | 0.680*** (35.67) | 0.713*** (29.66) | 0.680*** (35.67) |
- Without high school diploma: 0.791*** (11.46) | 0.766*** (13.95) | 0.791*** (11.46) | 0.766*** (13.95) |
- High school graduate: 0.917*** (5.98) | 0.907*** (7.25) | 0.917*** (5.98) | 0.907*** (7.25) |
- Some college: 1.029* (1.82) | 1.015 (0.98) | 1.029* (1.82) | 1.015 (0.98) |
- Black: 0.715*** (16.12) | 0.709*** (17.60) | 0.715*** (16.12) | 0.709*** (17.60) |
- Hispanic: 0.932** (2.39) | 0.958 (1.55) | 0.932** (2.39) | 0.958 (1.55) |
- Elderly: 1.065*** (3.48) | 0.953*** (2.83) | 1.065*** (3.48) | 0.953*** (2.83) |
- Family size: 0.999 | 1.015*** (3.50) | 0.999 | 1.015*** (3.50) |

Observations: 167,622 | 164,501 | 164,501 | 149,142 | 152,227

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

Notes: The absolute value of z-statistics are in parentheses. Region of residence, marital status, and other races are controlled for, but not reported. See the text for definitions of the variables and for further details.
Sources: Authors’ calculations based on data from the University of Michigan, Survey Research Center, Surveys of Consumers; U.S. Bureau of Economic Analysis; U.S. Bureau of Labor Statistics; and Yahoo! Finance.
The magnitudes of the odds ratios are little changed from the earlier regressions. Most of the odds ratios are closer to one (indicating that the underlying coefficients are closer to zero), but the differences are not large. This indicates that the PAGO and news variables can explain a small part of the contribution of demographics to expectations. Individual experiences may play a larger role in influencing expectations, but these experiences are not fully captured in the PAGO variable.

**Conclusion**

Policy decisions that are made using aggregate data are often ultimately aimed at particular income and demographic groups. Therefore, it might be useful to have an alternative measure of macroeconomic situations from the perspective of lower-income populations. We investigate this possibility with an analysis of group differences in consumer sentiment. Our findings suggest that index disaggregation by group matters because sentiment varies systematically by group attributes. In addition, demographic characteristics are found to influence responses to all five of the component questions that contribute to the index measure of the ICS. That is, the importance of demographic characteristics holds for both subjective and objective questions. Individuals’ attributes not only influence perceptions of their own experiences, but also their expectations of the economy more generally. Further investigation into this result shows that individuals form their expectations based on both their individual experiences and their exposure to news. However, many individuals in the sample report that they have heard no news, leaving them dependent on their idiosyncratic experiences and perceptions.

Future research might test whether consumer sentiment forecasts the behavior of households actually surveyed (as opposed to merely capturing broad aggregate economic trends). It might also be interesting to determine whether differences in consumer sentiment might explain or predict groups’ differences not only in consumption but also in savings and investment behavior. Future research using household microdata might test whether accounting for the distribution of sentiment across different groups might provide additional information to forecasts of macro-level models. For instance, the current aggregate consumer sentiment index, the ICS, is an equally weighted average of the sentiment of the survey respondents, which ignores the scale of the differences in consumption across respondents. Future research might construct a new sentiment series by taking into account the distribution in sentiment across groups. For instance, a new index could use group-level consumption-to-weight sentiment. Such a series could potentially assist in the forecasting of consumption.
NOTES

1We are grateful to Richard Curtin, director of the University of Michigan’s Surveys of Consumers, for providing us with the data.

2A short review of papers that use micro-level data is provided in Souleles (2004). The papers include Leeper (1992); Matsusaka and Sbordone (1995); Berg and Bergstrom (1996); Batchelor and Dua (1998); Bram and Ludvigson (1998); Acemoglu and Scott (1994); and Carroll, Fuhrer, and Wilcox (1994).

3Carroll, Fuhrer, and Wilcox (1994) and Bram and Ludvigson (1998) find no correlation between sentiment and future spending, which is inconsistent with the precautionary savings motive assumption. However, Souleles (2004) analyzes micro-level data of consumer sentiment and finds that higher confidence is correlated with lower consumption growth or more savings, which are consistent with precautionary motives. Aggregate indexes sum up responses of individuals that have different sentiments. These authors attribute the discrepancy in the results with previous studies to potential “aggregation bias” in the macro-level analysis approach of other studies. The aggregation bias stems from the fact that it is possible that the differences may not aggregate up.

4The idea is that the kinds of information that come from the manufacturing sector may be better known to the population that is geographically closer to its source. For example, layoffs may be more visible and may have a bigger impact on the population’s region. These consumers may have an earlier signal of change on which to base their assessments of future economic trends.

5The University of Michigan, Survey Research Center’s ICS was first introduced in 1952. In 1976 the index’s baseline was set at the 1966 level of 100, which is a level generally considered to represent a high level of optimism.

The survey population now consists of 500 nationally representative individuals in the coterminous United States (prior to 1976, in the earlier years of the survey, two to three times as many individuals were interviewed). This cross-sectional sample is constructed using a stratified system that assures proportional representation of different states, geographic regions, and metropolitan areas of varying sizes. (The survey also employs a rotating-panel design in which respondents are reinterviewed six months after the initial questioning, resulting in a monthly sample that is, on average, 40 percent first-time respondents and 60 percent second-time respondents.)

6More specifically, to generate the index number, a “score” for each question is created. The score is equal to the difference between the percent of respondents giving unfavorable (pessimistic) responses to each question and the percent of respondents giving favorable (optimistic) responses to each question, plus 100. For example, if 55 percent of interviewees expect to be better off next year, 30 percent expect to be worse off, and 15 percent expect no change, the score for question two is 125 (= 55 - 30 + 100). The SRC adds the scores of the five questions together and divides that sum by 6.7558, a constant which makes the index relative to the 1966 base score of 100. Finally, the number 2.0 is added to the index in order to “correct for sample design changes from the 1950s” (prior to December 1981, n = 2.7). This process is represented in the formula:

\[ \text{ICS} = \frac{\text{PAGO} + \text{PEXP} + \text{BUS12} + \text{BUS5} + \text{DUR} + n}{6.7558} \]

The Index of Consumer Expectations (ICE) and the Index of Current Economic Conditions (ICC) are calculated as follows:

\[ \text{ICC} = (\text{PAGO} + \text{DUR} + n) \times 2.6424 \quad \text{and} \quad \text{ICE} = (\text{PEXP} + \text{BUS12} + \text{BUS5} + n) \times 4.1134. \]

7The sampling error is an important consideration in correctly interpreting the ICS. With a monthly sample size of 500 and a quarterly sample size of 1,500, small shifts in the index may not be significant. Specifically, the 95 percent confidence interval for the monthly ICS is +/- 3.29 points (Curtin, 2002). The 95 percent confidence interval for the quarterly data is +/- 1.91 points.

8See, for example, Carroll, Fuhrer, and Wilcox (1994) and Bram and Ludvigson (1998).

9The Survey Research Center at the University of Michigan weights responses in order to generate a representative sample of all U.S. households (or all individual adults, depending on which set of weights is used). The weights correct for undersampling of certain populations, such as the poor. After weights are applied, most subpopulations seem to be well represented in the 2000 Surveys of Consumers, compared with population data from the 2000 Census. However, it does appear that undersampling of the Hispanic population is not corrected when weights are applied. Additionally, the population without a high school diploma is underrepresented in the year 2000, although over all years of the SRC survey, 15.96 percent of respondents have less than a high school education (after weighting).

10In particular, the number of negative responses given to the underlying questions is subtracted from the positive responses. This number is divided by the number of questions asked. Then 100 is added to this number which is then multiplied by two and divided by a scaling factor that depends on which index is being calculated. Two is then added to this number after 1982 and 2.7 before 1982.

11We calculated the index of individuals based on their regions of residence. There is no noticeably strong difference in the consumer sentiment across respondents living in different regions. We therefore do not report these results, but they are available from the authors upon request.

12We use the annual poverty thresholds calculated by the U.S. Census Bureau. The thresholds differ based on family composition and the ages of household members. A household is considered poor if household income falls below the threshold.

13Respondents’ expectations of the unemployment rate over the 12 months following the survey are based in part on what has happened to unemployment over the previous six to 12 months. They also predict, to some degree, unemployment over the subsequent six to 12 months. We measure change in actual unemployment as a four-quarter moving average of the one-year change in the unemployment rate. For example, the change in actual unemployment in 1990-Q1 is recorded as the average of the difference between the unemployment rates between: 1) 1990-Q1 and 1989-Q1; 2) 1989-Q4 and 1988-Q4; 3) 1989-Q3 and 1988-Q3; and 4) 1989-Q2 and 1988-Q2. In other words, “change in actual unemployment” actually reflects the way the unemployment rate changed over the previous year relative to the year before it. Therefore we can interpret the two-quarter lagged correlation between expectation and actual unemployment as the relationship between expectations of unemployment over the coming year and actual changes in unemployment during the six-month periods immediately preceding and following the survey.

14Please refer to note 10 for more details.
Some studies have found that, at the local level, demographic characteristics explain variation in consumer confidence. We run these regressions separately by region and find that the effects of demographic characteristics are still significant in explaining differences in the indexes.

Due to space constraints, the results for the ICE and the ICC are not given in tables. They are available upon request.

We ran these regressions using different specifications of the dependent variables, including omitting neutral individuals, separating the two positive and two negative BUS3 and BUS12 responses, and measuring the net number of positive reasons given for the PIAGO question. These other specifications provided substantively similar results. The only difference was for the elderly—when we dropped individuals who were the same, the odds ratio moved farther from one. Because the individuals who are the same are better off than those who are worse off, and because more elderly report that they are worse off than that they are better off, including individuals who are the same, this leads to an increase in the odds ratio among the elderly.

We perform a parallel set of analyses with the BUS5 variable. The results are substantively similar to the BUS12 results except for the odds ratio for the interaction between a good past year and hearing no news. It is less than one and insignificant for BUS5, while this odds ratio was greater than one and statistically significant for BUS12.

REFERENCES


Earnings announcements, private information, and liquidity

Craig H. Furfine

Introduction and summary

Efficient financial markets facilitate the smooth transfer of money from those who save to those with profitable investment opportunities. Such markets generally exhibit high levels of trading volume and widespread market participation. Investors are willing to participate because they are convinced that the prices at which securities can be bought and sold are reasonably efficient. For example, a market participant should be able to buy or sell a share of stock in XYZ company at a price very close to the present discounted value of the market’s best estimate of XYZ’s future dividend payments.¹

So where do these market estimates come from? There are two main types of information underlying these estimates—one, information that is common to all market participants (I call this public information), and two, information that is specific to individual investors (I call this private information). An example of public information in this context would be a news release about company XYZ that might be expected to move the company’s share price. A company making a surprisingly good earnings announcement typically sees an immediate rise in its stock price. Bad news generally has the reverse effect. Clearly, this type of information impacts market prices.

What is less clear is whether and to what extent private information impacts stock prices. On most trading days, there is no obvious “news” (that is, public information) regarding the value of a particular stock, yet stocks still trade and often show noticeable price changes. Some widely held stocks trade every few seconds on all trading days. Although some of this activity can be linked to public news, much of the trading and related price changes occur when there is no easily observable event or publicly conveyed information believed to be relevant to a given company’s stock price. Suppose that an investor places a large order to purchase shares of XYZ on a day when no public news about XYZ is released. Depending on the characteristics of this trade, the price of XYZ may change. For example, if market participants believe this trade was made by an investor who believes the stock is undervalued, others may revise their own expectations and afford shares of XYZ a higher price. Alternatively, participants observing a large purchase order may attribute the purchase to the fact that the given purchaser of XYZ is a manager of an index fund that is well known to have been receiving large inflows of investment capital. Thus, the purchase of XYZ contains no information about the value of XYZ shares. In this case, one might expect the purchase to have a more limited impact on the stock’s price. In reality, the underlying purpose of individual trades is not generally known, and therefore one can characterize trades as containing some degree of private information.

Understanding how both public and private types of information influence security prices is one of the main goals of financial market microstructure analysis. Earnings announcements are perhaps the most visible form of public information. At the most extreme, insider trading by an executive knowing the contents of a forthcoming news release is an example of private information. However, private information can simply be thought of as all information about a given security price that is not known by all who trade it. For example, a mutual fund manager’s decision to reduce

Craig H. Furfine is a senior economist and an economic advisor in the Research Department at the Federal Reserve Bank of Chicago. The author would like to thank colleagues at the Federal Reserve Bank of Chicago for their helpful comments and Lauren Gaudino for excellent research assistance.
the holding of a given stock would be considered private information capable of affecting security prices.

Private information, however, can be even less tangible. Differing opinions as to the implications of an earnings announcement may generate important private information, since some may believe the optimal response to the news is to buy a security, whereas others may wish to sell. It is the collective trades of market participants that move prices. Market microstructure analysis presumes that trading is necessary to determine prices because it conveys private information regarding the value of the underlying asset. The intuition is relatively simple: As a sequence of sell orders arrives, prices will be adjusted lower as potential buyers incorporate a higher probability that better informed traders believe the previous price was too high.

This article attempts to shed light on the relative importance of private versus public information in moving security prices. I examine closely the intraday trading activity of ten large companies trading on the New York Stock Exchange and estimate an empirical model that relates trading activity to price changes. My focus is on how this trading–price change relationship changes on days when there is a major release of public information regarding the company in question—in particular, quarterly earnings announcements. In this way, my goal is to quantify by how much, if any, the trading–price change relationship changes with a large increase in public information.

My hypothesis is that the strength of the trading–price change relationship is a measure of the importance of private information in security price formation. An empirical implication of this is that a major release of news should be accompanied by a reduction in the strength of the relationship between trading and price changes. I conduct a series of empirical exercises that provide evidence consistent with this hypothesis. However, my results further indicate that even after an earnings announcement, private information plays a significant role in security price determination. Across the firms in my sample, I find that the strength of the trading–price change relationship, my proxy for the importance of private information, declines by no more than one-third on trading days immediately following a company’s quarterly earnings announcement. Thus, private information appears to be a significant factor in the relationship between trading and prices.

The remainder of this article is organized as follows. First, I briefly review some related work. Then, I describe the data and the empirical framework of my analysis. Finally, I present my findings and discuss their implications.

Related literature

As mentioned previously, market microstructure theory argues that order flow (that is, the sequence of buy and sell orders) affects prices because it conveys private information regarding the value of the underlying asset. In Glosten and Milgrom (1985), for example, the authors formally model why private information leads immediately to the presence of a bid–ask spread (the difference between the proposed purchase price and proposed sale price for the same security) as well as a relationship between trading and price changes. In their model, a marketmaker for a given security stands ready to buy or sell. The marketmaker believes, however, that some of the potential buyers have private information that indicates that the marketmaker’s current price of the security is too low. Alternatively, or perhaps in addition, the marketmaker believes that some potential sellers of the security have private information indicating the marketmaker’s price is too high. The result of this asymmetry, along with the marketmaker’s continued willingness to trade, is a positive bid–ask spread. That is, the price at which a marketmaker is willing to buy is lower than the price at which he is willing to sell. This spread serves as compensation for trades made with those counterparties with superior information. As a sequence of sell orders arrive, marketmakers lower bid prices, incorporating the probability that the order flow implies that better informed investors believe the previous price was too high. This adjustment of posted spreads implies an analogous change in observed transaction prices.

Over the past two decades, the microstructure literature has explored how, when, and how much order flow affects stock prices. Here, I focus on the work most related to my current analysis, specifically work on the relationship between trading and price changes. This relationship ultimately provides an (inverse) measure of a security’s liquidity because a stock whose price changes a lot in response to incoming trades would be deemed relatively illiquid. The literature shows that liquidity itself and the relationship between public news releases and liquidity can be measured in a number of ways.

Seppi (1992) conducts empirical tests to determine the informativeness of block trades (typically, 10,000 shares or more) and how this informativeness correlates with public news releases. In particular, he documents that the prices at which such trades are filled are positively correlated with the earnings surprises. That is, block trades occur at higher prices before positive earnings surprises and at lower prices before negative earnings surprises. This is consistent with the belief that investors making such trades, on average,
have some knowledge about the information that will be made public in a subsequent earnings announcement and are therefore anticipating the change in stock price. Lee, Mucklow, and Ready (1993) explicitly describe various alternative measures of market liquidity. They focus not only on the size of the bid–ask spread, but also on a stock’s posted depth, which measures the quantity of shares a marketmaker is willing to transact at the posted bid and ask prices. Their study documents that both spreads and depth adjust to the perceived amount of private information in the market. In particular, spreads generally widen and depth generally falls preceding earnings announcements.

Koski and Michaely (2000) extend these findings by examining the relationship between measures of liquidity (for example, spreads and depth) across key information periods, which include both earnings and dividend periods. They find that these liquidity measures do relate to the perceived information content of the trade. In particular, large trades before dividend announcement periods tend to reduce depth and increase spreads most strongly. Similar results, though smaller in magnitude, are found during announcement periods. This is consistent with the notion that private information is at its highest level just prior to a news release. However, since these authors combined information from before and after an announcement period, they could not distinguish precisely how news affects liquidity over the period immediately preceding and following announcements, nor did they formally analyze the trading–price change relationship.

Green (2004) conducts a study of the relationship between announcements and the information content of trading in U.S. Treasury bonds. For Treasury bonds, news announcements are not about corporate earnings, but rather about the latest release of economic data. Green finds that when macroeconomic news is released, the information component of trading increases. Thus, unlike Koski and Michaely (2000), Green associates public news release with an increase in private information. Perhaps macroeconomic news releases generate more information on which individual traders can disagree, generating a higher share of private information that in turn affects security prices.

Thus, the previous empirical work provides evidence that the more informative a given trade, the greater its influence on security prices. However, the evidence is somewhat mixed with regard to how the overall liquidity of a security is influenced by news. In particular, there is not yet a consensus as to whether public news arrival reduces or generates private information. Rather, it appears from previous work that public news releases have the potential either to generate or eliminate private information. Thus, it remains an empirical question to decide whether public news arrival will strengthen the trading–price change relationship (consistent with private information generation) or weaken it (consistent with private information elimination).

Data and empirical framework

My analysis relies on data from three sources. I begin with the universe of firms whose earnings history was available on Briefing.com. To select my sample, I required that Briefing.com reported the date, time, value, and market expectation of every earnings announcement that a firm reported between January 29, 2001, and December 31, 2004. For my purposes, the Briefing.com data provide an important piece of information unavailable in the more traditionally used sources of announcement histories, such as Thomson Financial’s FirstCall and the Institutional Brokers’ Estimate System: whether a firm’s earnings announcement was made prior to the stock market opening, during the trading day, or after the market close. Thus, it is possible to know precisely which day of stock market trading is associated with the reaction to the earnings announcement. This distinction will be crucial to my analysis. In what follows, I refer to the first trading day following the announcement as a company’s “announcement day.” For example, if a company announces its earnings on a Tuesday before or during trading hours, its announcement day is that Tuesday. If the announcement is made on a Tuesday after the market closes, its announcement day will be that Wednesday. Furthermore, my analysis carefully considered the role of weekends and public holidays in order to correctly pair a given announcement with the next possible trading day.

I then compared this sample of firms to the database provided by the Center for Research in Security Prices. I considered only those firms that were listed on the New York Stock Exchange (NYSE) to avoid the well-known differences in the liquidity (and by extension, the strength of the trading–price change relationship) of stocks trading on different exchanges. I then calculated each firm’s market capitalization based on stock price data as of December 31, 2001, and selected the ten largest remaining firms to be the focus of my study.

Having identified the ten stocks in my sample, I then combined the earnings announcement information from Briefing.com with high frequency data on the trading of the stocks of these ten firms from the NYSE Trade and Quote (TAQ) database. Although Briefing.com provides earnings information since
1997, I restrict my sample period to January 29, 2001, through December 31, 2004. The starting date of my sample corresponds to the first day on which all stocks listed on the NYSE began trading with a minimum price increment (tick) of one cent (that is, decimalization). This eliminates the need to consider how minimum tick size changes might influence the relationship between earnings announcements and liquidity, since a vast literature has documented the importance of minimum tick size to liquidity in general.

I then adjusted the data according to procedures common in the microstructure literature. I dropped quotes with obviously erroneous data (for example, quotes with bid or ask prices equal to zero or quotes with bid–ask spreads dramatically different from the previous or subsequent quote). Following Hasbrouck (1991), I kept only quotes originating from the NYSE and considered multiple trades on a regional exchange for the same stock at the same price and time as one trade. Then, I sorted the trade data (for each company and day) by time, with the prevailing quote at transaction t defined as the last quote that was posted at least five seconds before the transaction (Lee and Ready, 1991). I provide a complete listing of the stocks in my sample, along with summary statistics on their trading, in Table 1.

The summary statistics show many facts about stock market trading. First, these ten stocks are very heavily traded. The least actively traded stock in my sample is Coca-Cola (KO), yet shares of this stock traded over 3,800 times each day, a trading intensity of roughly once every six seconds. The most actively traded stock in my sample is General Electric (GE), whose shares traded over 11,000 times each day on average (approximately once every two seconds). Bid–ask spreads on all of the sample stocks are typically quite narrow. On average, spreads range from a low of 1.82 cents for Nortel (NT) to a high of 5.23 cents for IBM.

I am interested in changes in trading characteristics that occur on or around earnings announcement days. To present some preliminary evidence on this

<table>
<thead>
<tr>
<th>Name/Ticker</th>
<th>Average trade size, shares</th>
<th>Average number of trades</th>
<th>Average bid–ask spread, dollars</th>
<th>Average depth, round lots</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bristol-Myers Squibb Co. (BMY)</td>
<td>1,537.9490 (527.6960)</td>
<td>4,175.6380 (1,326.6590)</td>
<td>0.0334 (0.0238)</td>
<td>31.6022 (14.9068)</td>
</tr>
<tr>
<td>EMC Corp. (EMC)</td>
<td>2,343.7580 (655.9954)</td>
<td>7,070.8470 (2,606.1300)</td>
<td>0.0331 (0.0375)</td>
<td>79.0106 (36.9236)</td>
</tr>
<tr>
<td>General Electric Co. (GE)</td>
<td>2,040.1440 (703.1345)</td>
<td>11,271.3300 (3,437.3040)</td>
<td>0.0271 (0.0156)</td>
<td>68.7537 (40.4598)</td>
</tr>
<tr>
<td>Home Depot Inc. (HD)</td>
<td>1,339.4020 (339.1715)</td>
<td>6,446.9410 (2,160.3100)</td>
<td>0.0304 (0.0178)</td>
<td>36.5258 (20.9643)</td>
</tr>
<tr>
<td>International Business Machines Corp. (IBM)</td>
<td>1,100.4560 (384.0350)</td>
<td>6,776.2480 (1,680.0090)</td>
<td>0.0523 (0.0302)</td>
<td>16.3487 (8.0540)</td>
</tr>
<tr>
<td>Coca-Cola Co. (KO)</td>
<td>1,415.3070 (473.8881)</td>
<td>3,827.4650 (1,321.2770)</td>
<td>0.0290 (0.0149)</td>
<td>23.3063 (10.2120)</td>
</tr>
<tr>
<td>Merck &amp; Co. Inc. (MRK)</td>
<td>1,403.8030 (448.6887)</td>
<td>5,421.9910 (3,776.4130)</td>
<td>0.0382 (0.0258)</td>
<td>26.9202 (22.3486)</td>
</tr>
<tr>
<td>Nortel Networks Corp. (NT)</td>
<td>4,906.4550 (3099.6320)</td>
<td>5,451.2300 (3,899.5310)</td>
<td>0.0182 (0.0170)</td>
<td>720.4567 (730.3201)</td>
</tr>
<tr>
<td>Pfizer Inc. (PFE)</td>
<td>1,982.2700 (465.7550)</td>
<td>8,839.0240 (3,908.9970)</td>
<td>0.0270 (0.0152)</td>
<td>55.0667 (38.1269)</td>
</tr>
<tr>
<td>SBC Communications Inc. (SBC)</td>
<td>1,784.0950 (518.5796)</td>
<td>4,608.6720 (1,219.6850)</td>
<td>0.0282 (0.0162)</td>
<td>40.3322 (19.2419)</td>
</tr>
</tbody>
</table>

Notes: This table reports the mean (and standard deviation in parentheses) of various measures of trading activity for each stock in the sample. Averages are taken across all 959 days that are not within one day of an earnings announcement.

Sources: Author’s calculations based on data from Briefing.com and the New York Stock Exchange Trade and Quote database.

Table 1
Summary statistics
subject, I regress the daily values of various measures of trading activity on a set of dummy variables that indicate the day before, the day of, and the day after an earnings announcement. Coefficients from this regression, which represent differences relative to all other days, are presented in Table 2. This table indicates that there are noticeable changes in common proxies for stock market liquidity on earnings announcement days. Most strikingly, trading volume increases. Trade size and average depth, which represents the number of shares available at the posted spread, also tend to rise on announcement days, although these results appear to be statistically significant for only a subset of my sample firms. For KO shares, for example, average trade size increases by 587 shares, and typical depth rises by 7.4 round lots (that is, 740 shares) on announcement days. Not all statistical indicators of liquidity, however, indicate greater liquidity on announcement days. Although not statistically significant in most cases, bid–ask spreads tend to rise on announcement days. For instance, IBM’s bid–ask spread typically increases by 1.4 cents on an announcement day. Data across these ten stocks tell a similarly inconsistent story with regard to the relationship between liquidity and announcement days—namely, that announcement days witness an increase in trading volume and depth, but either little change or a widening of bid–ask spreads.

Because announcement days are correlated with heavier trading volume and higher depth but wider spreads, it would be useful to focus on a measure of stock market liquidity that may account for these changes. Here, I use the price impact of a trade as a measure of a stock’s liquidity that embeds the impact of volume, spreads, and depth. That is, I take the position that price impact is the quantity that ultimately relates to the strength of the trading–price change relationship and that volume, spreads, and depth (among other observable characteristics) are noisy indicators of such a relationship.

I adopt the general empirical framework of Hasbrouck (1991), who estimates a vector autoregression (VAR) model of two equations. The first equation models trade-to-trade stock returns as a function of past returns as well as current and past trades, explicitly considering whether the trade was to purchase or to sell shares. The second equation models the decision to buy or sell as a function of both past trading and past stock returns. In such a framework, Hasbrouck delivered some benchmark results upon which I build in my analysis. In particular, Hasbrouck documents the positive relationship between order flow and price changes using a sample of 80 NYSE and American Stock and Options Exchange (AMEX) stocks. That is, buy orders lead to price increases, and sell orders lead to price declines. Hasbrouck further extended his analysis to indicate that larger trades tend to move prices more, a finding that I incorporate into my framework.

My empirical results are based upon VAR models of increasing complexity. The first merely replicates a version of the Hasbrouck (1991) analysis. I specify this by equations 1 and 2, which I estimate separately for each of the ten firms in my sample.

\[ r_t = \sum_{i=1}^{I} a_i r_{t-i} + \sum_{i=0}^{I} y_i x_{t-i} + \epsilon_{t}, \]

\[ x_t = \sum_{i=1}^{I} a_i r_{t-i} + \sum_{i=1}^{I} y_i x_{t-i} + \epsilon_{t}. \]

The unit of observation is the trade, which is indexed by the subscript \( t \). The variable \( r \) is defined as the change in the natural logarithm of the midquote (average of the current bid and ask price) of a given stock that follows the trade at time \( t \). I use midquotes as my price variable to eliminate the well-known problems with using actual transaction prices in empirical analysis, notably the tendency of transaction prices to bounce between the bid and ask prices without indicating any true movement in the underlying security value. Also following Hasbrouck (1991), I define \( x \) as the log of the number of shares of trade \( t \), signed to indicate whether or not trade \( t \) was initiated by a buy order or a sell order. That is, a positive value of \( x \) indicates a buyer-initiated trade, and a negative value indicates a seller-initiated trade. As the TAQ data do not indicate which party initiated each trade, I follow the literature’s convention and assume that trades at a transaction price greater than the midquote were buyer-initiated and trades below the midquote were seller-initiated. For trades at the midquote, I determine the side of trade origination according to the tick rule (see Lee and Ready, 1991).

I truncate the VAR model by setting \( I \) equal to eight for all stocks and for all time periods. Though longer than the five lags adopted by Hasbrouck (1991), this reflects the higher level of trading in more recent periods. Finally, I estimate equations 1 and 2 by ordinary least squares and correct standard errors using White’s (1980) methodology.

I then expand the model in several ways to explore how the relationship between trading and price might change in ways related to earnings announcements. My first additional model can be expressed by equations 3 and 4.
<table>
<thead>
<tr>
<th>Name/Ticker</th>
<th>Average trade size, shares</th>
<th>Average number of trades</th>
<th>Average bid–ask spread, dollars</th>
<th>Average depth, round lots</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bristol-Myers Squibb Co. (BMY)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day before announcement</td>
<td>72.634</td>
<td>28.562</td>
<td>-0.005</td>
<td>-3.274</td>
</tr>
<tr>
<td></td>
<td>(105.172)</td>
<td>(257.771)</td>
<td>(0.004)</td>
<td>(2.716)</td>
</tr>
<tr>
<td>Day of announcement</td>
<td>565.982</td>
<td>1,484.962</td>
<td>0.003</td>
<td>4.404</td>
</tr>
<tr>
<td></td>
<td>(144.990)**</td>
<td>(527.086)**</td>
<td>(0.004)</td>
<td>(3.835)</td>
</tr>
<tr>
<td>Day after announcement</td>
<td>266.810</td>
<td>615.362</td>
<td>-0.000</td>
<td>-0.466</td>
</tr>
<tr>
<td></td>
<td>(155.257)</td>
<td>(296.985)*</td>
<td>(0.005)</td>
<td>(2.896)</td>
</tr>
<tr>
<td><strong>EMC Corp. (EMC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day before announcement</td>
<td>243.296</td>
<td>668.353</td>
<td>-0.002</td>
<td>9.878</td>
</tr>
<tr>
<td></td>
<td>(158.294)</td>
<td>(840.003)</td>
<td>(0.010)</td>
<td>(10.834)</td>
</tr>
<tr>
<td>Day of announcement</td>
<td>735.740</td>
<td>2,081.953</td>
<td>0.003</td>
<td>40.342</td>
</tr>
<tr>
<td></td>
<td>(176.647)**</td>
<td>(947.399)**</td>
<td>(0.009)</td>
<td>(11.384)**</td>
</tr>
<tr>
<td>Day after announcement</td>
<td>353.466</td>
<td>476.953</td>
<td>-0.003</td>
<td>15.625</td>
</tr>
<tr>
<td></td>
<td>(158.172)*</td>
<td>(614.181)</td>
<td>(0.009)</td>
<td>(8.527)</td>
</tr>
<tr>
<td><strong>General Electric Co. (GE)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day before announcement</td>
<td>97.373</td>
<td>1,056.405</td>
<td>0.000</td>
<td>-0.966</td>
</tr>
<tr>
<td></td>
<td>(133.625)</td>
<td>(1,076.020)</td>
<td>(0.003)</td>
<td>(4.692)</td>
</tr>
<tr>
<td>Day of announcement</td>
<td>129.556</td>
<td>4,079.672</td>
<td>0.001</td>
<td>12.941</td>
</tr>
<tr>
<td></td>
<td>(100.314)</td>
<td>(1,441.108)**</td>
<td>(0.003)</td>
<td>(10.381)</td>
</tr>
<tr>
<td>Day after announcement</td>
<td>-192.732</td>
<td>1,288.405</td>
<td>-0.001</td>
<td>-0.772</td>
</tr>
<tr>
<td></td>
<td>(102.162)</td>
<td>(1,190.934)</td>
<td>(0.003)</td>
<td>(7.713)</td>
</tr>
<tr>
<td><strong>Home Depot Inc. (HD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day before announcement</td>
<td>48.028</td>
<td>1,261.246</td>
<td>-0.001</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>(70.942)</td>
<td>(720.627)</td>
<td>(0.003)</td>
<td>(3.166)</td>
</tr>
<tr>
<td>Day of announcement</td>
<td>375.450</td>
<td>4,809.309</td>
<td>0.013</td>
<td>18.399</td>
</tr>
<tr>
<td></td>
<td>(86.186)**</td>
<td>(1,315.068)**</td>
<td>(0.005)**</td>
<td>(6.817)**</td>
</tr>
<tr>
<td>Day after announcement</td>
<td>219.703</td>
<td>2,343.559</td>
<td>-0.000</td>
<td>6.231</td>
</tr>
<tr>
<td></td>
<td>(96.175)*</td>
<td>(913.464)*</td>
<td>(0.003)</td>
<td>(3.282)</td>
</tr>
<tr>
<td><strong>International Business Machines Corp. (IBM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day before announcement</td>
<td>170.724</td>
<td>1,727.752</td>
<td>0.004</td>
<td>1.907</td>
</tr>
<tr>
<td></td>
<td>(104.526)</td>
<td>(380.877)**</td>
<td>(0.008)</td>
<td>(1.826)</td>
</tr>
<tr>
<td>Day of announcement</td>
<td>626.777</td>
<td>2,918.418</td>
<td>0.014</td>
<td>3.711</td>
</tr>
<tr>
<td></td>
<td>(117.181)**</td>
<td>(509.751)**</td>
<td>(0.006)*</td>
<td>(1.689)*</td>
</tr>
<tr>
<td>Day after announcement</td>
<td>213.212</td>
<td>596.018</td>
<td>-0.004</td>
<td>1.969</td>
</tr>
<tr>
<td></td>
<td>(98.206)*</td>
<td>(349.537)</td>
<td>(0.005)</td>
<td>(1.967)</td>
</tr>
<tr>
<td><strong>Coca-Cola Co. (KO)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day before announcement</td>
<td>78.408</td>
<td>-48.028</td>
<td>-0.000</td>
<td>1.610</td>
</tr>
<tr>
<td></td>
<td>(74.701)</td>
<td>(274.959)</td>
<td>(0.003)</td>
<td>(1.895)</td>
</tr>
<tr>
<td>Day of announcement</td>
<td>587.421</td>
<td>1,125.597</td>
<td>0.008</td>
<td>7.404</td>
</tr>
<tr>
<td></td>
<td>(104.822)**</td>
<td>(584.509)</td>
<td>(0.004)</td>
<td>(3.442)*</td>
</tr>
<tr>
<td>Day after announcement</td>
<td>431.017</td>
<td>202.347</td>
<td>0.003</td>
<td>1.503</td>
</tr>
<tr>
<td></td>
<td>(157.521)**</td>
<td>(365.853)</td>
<td>(0.003)</td>
<td>(1.675)</td>
</tr>
</tbody>
</table>
3) $r_i = \sum_{t=1}^{I} a_t r_{i,t} + \sum_{t=0}^{I} (r_{i,t} + \theta r_{i,t}) x_{i,t} + \nu_{it}$

4) $x_i = \sum_{t=1}^{I} a_t r_{i,t} + \sum_{t=0}^{I} (r_{i,t} + \theta r_{i,t}) x_{i,t} + \nu_{it}$

In this empirical specification, I add terms to the model that interact the trade size variable $x_i$ with a dummy variable $a_t$, which is set equal to one if trade $t$ occurs on an announcement date. This allows the relationship between trading and price changes to be different on announcement days. For example, positive estimated values for $\theta$ would indicate that the relationship between trading and stock returns becomes stronger on announcement days.

Because I have identified a positive relationship between trading volume and announcement days, it is important to confirm that any relationship I find between announcement days and price impact when I estimate equations 3 and 4 is due to the announcement and not simply an artifact of higher trading volume. To this end, I next estimate an expanded version of equations 3 and 4, where I interact the trade indicator variable with a variable $I_t$, measuring trading volume on the day on which trade $t$ occurs. This expanded specification is shown in equations 5 and 6.
5) \[ r_i = \sum_{t=1}^i a_{t} r_{t-i} + \sum_{t=0}^{i-1} (y'_t + \theta'_t a_{t-i} + \kappa'_t f_{t-i}) x_{t-i} + \nu_{it}, \]

6) \[ x_i = \sum_{t=1}^i a_{t} r_{t-i} + \sum_{t=0}^{i-1} (y'_t + \theta'_t a_{t-i} + \kappa'_t f_{t-i}) x_{t-i} + \nu_{it}. \]

My next empirical specification extends the model described by equations 5 and 6 to explore whether days immediately surrounding announcement days may be noticeably different than other days. As described in equations 7 and 8, I do this by interacting variables \( b_{t} \) defined to equal one if trade \( t \) occurs on the day before an announcement date and zero otherwise, and \( f_{t} \) defined analogously for the day after an announcement date, with the trade size variable \( x_{t} \), as follows:

7) \[ r_i = \sum_{t=1}^i a_{t} r_{t-i} + \sum_{t=0}^{i-1} (y'_t + \lambda'_t b_{t-i} + \theta'_t a_{t-i} + \phi'_t f_{t-i}) + \kappa'_t l_{t-i} x_{t-i} + \nu_{it}, \]

8) \[ x_i = \sum_{t=1}^i a_{t} r_{t-i} + \sum_{t=0}^{i-1} (y'_t + \lambda'_t b_{t-i} + \theta'_t a_{t-i} + \phi'_t f_{t-i}) + \kappa'_t l_{t-i} x_{t-i} + \nu_{it}. \]

I estimate the empirical model described by equations 7 and 8 to explore whether the private information content of a trade varies according to the proximity to a public announcement rather than only depending on whether the announcement was just made. For example, one might believe that if announcement days reduce the private information content of stock trading, then the day before such an announcement might be expected to contain a higher than average amount of private information. That is, the likelihood of a trader’s having private information regarding a future earnings announcement might be expected to be the greatest immediately before the announcement. If this were the case, one might expect that the \( \lambda \) coefficients would be greater than zero. As for the day following the announcement, allowing the relationship between trading and stock returns to differ facilitates an exploration as to whether any changes detected on the announcement day persist until the following day. To the extent that there is persistence, one might expect to estimate values for \( \theta \) very close to the values estimated for \( \phi \).

Much like equations 7 and 8, my final empirical specification extends the model described by 5 and 6. However, rather than exploring whether the relationship between trading and returns varies according to the proximity in calendar time from the announcement date, I instead explore whether the importance of the announcement date varies according to the realized content of the given announcement. That is, I wish to distinguish between announcements that contain surprising information and those that do not. In particular, I estimate the model described by equations 9 and 10,

9) \[ r_i = \sum_{t=1}^i a_{t} r_{t-i} + \sum_{t=0}^{i-1} (y'_t + \theta'_t a_{t-i} + \beta'_t a_{t-i} s_{t-i}) + \kappa'_t l_{t-i} x_{t-i} + \nu_{it}, \]

10) \[ x_i = \sum_{t=1}^i a_{t} r_{t-i} + \sum_{t=0}^{i-1} (y'_t + \theta'_t a_{t-i} + \beta'_t a_{t-i} s_{t-i}) + \kappa'_t l_{t-i} x_{t-i} + \nu_{it}. \]

Here, I define \( s \) as an indicator variable that is equal to one when the actual earnings announced differed from expected earnings as reported by Briefing.com by more than \$0.01 per share. One hypothesis is that surprising earnings releases reveal more private information than those that are unsurprising. If this were true, one would expect the coefficients \( \beta \) to be negative.

**Empirical results**

I form my estimate of the price impact of a trade by calculating the cumulative impulse response of a shock to \( x \) on stock returns \( r \). As a point of departure, figure 1 graphs these responses for each firm, when the size of the shock \( x \) is set equal to each stock’s median trade size and also to the stock’s 90th percentile trade size. This allows one to judge the overall liquidity of a stock on average across the roughly four years of data and to measure by how much more a large trade moves prices than a more typical trade. For example, panel A of figure 1 depicts the impulse response functions for Bristol-Myers Squibb Co. (BMY). The graph shows that a median-sized buy order is estimated to eventually raise the price of BMY shares by approximately 1.4 basis points. A large trade that was unexpected is estimated to have a long-run impact of increasing BMY share prices by a little over 1.8 basis points. The main findings illustrated by figure 1 are that even across a sample of large firms, market liquidity varies across firms and across trades of a given firm. For instance, across these ten stocks, a median-sized trade is estimated to raise prices by between 0.7 and 1.9 basis points, depending on the
FIGURE 1

The long-run price impact of median- and large-sized trades

A. Bristol-Myers Squibb Co. (BMY)
basis points

B. EMC Corp. (EMC)
basis points

C. General Electric Co. (GE)
basis points

D. Home Depot Inc. (HD)
basis points

E. International Business Machines Corp. (IBM)
basis points

F. Coca-Cola Co. (KO)
basis points

G. Merck & Co. Inc. (MRK)
basis points

H. Nortel Networks Corp. (NT)
basis points

I. Pfizer Inc. (PFE)
basis points

J. SBC Communications Inc. (SBC)
basis points

Sources: Author’s calculations based on data from Briefing.com and the New York Stock Exchange Trade and Quote database.
stock. Furthermore, for every stock in the sample, larger trades appear to have a greater price impact.

As illustrated in figure 1, the cumulative impulse response of a trade shock on stock returns generally begins with a rapid increase immediately following the shock and then levels off at a point higher than its initial value. As I extend the initial model to account for announcement days, this shape remains. Because of this, I choose to present the remaining results by reporting values based on the “long-run” impulse response, which is derived from the cumulative impulse response function by reading off the final value calculated—in this case, the value of the cumulative shock to returns seen 16 trades after the initial shock. Note that this long run typically takes less than a minute. To illustrate, figure 2 reports the cumulative long-run impulse responses on returns of a trade shock for each of the ten stocks in my sample after 1 estimate equations 3 and 4, which extend the previous model by allowing the price impact of a trade to vary according to whether the given trade occurs on a day with an earnings announcement. Each panel of figure 2 reports two values. The first bar reports the long-run price impact of a median-sized trade on a nonannouncement day (that is, normal day). The second bar reports the long-run cumulative value of the same sized trade on an announcement date.

Qualitatively, the results are similar across all ten firms in the sample. In particular, the price impact of a median-sized trade is uniformly lower on an announcement day than on other days during the sample period. This result is consistent with the notion that price impact is partially explained by marketmakers defending themselves against asymmetric information. In other words, prices move in response to trades because marketmakers believe some traders have private information. Furthermore, this private information is reduced when a public earnings announcement is released. The magnitude of the reduction in price impact varies across the ten firms. In the case of BMY, the reduction in price impact is rather small. My model estimates that the long-run price impact of a trade declines from approximately 1.42 basis points to 1.39 basis points, a reduction of only 2.1 percent. For other companies, the reduction in price impact on announcement days is far more pronounced. The impact of a median-sized trade of Home Depot (HD) stock is roughly 1.2 basis points on nonannouncement days, but only 0.8 basis points on announcement days. This represents a reduction in price impact of 33 percent.

The results of figure 2 show that announcement days witness a decline in the price impact of trading, suggesting that the release of public information does reduce the private information embedded in a trade. I reached this conclusion by estimating a model that allowed the relationship between trading and returns to vary according to whether a given trade occurred on an announcement day. As mentioned previously, it is important to attribute the lower price impact of a trade on announcement days to the announcement and not to the typically higher trading volume witnessed on announcement days. Figure 3 reports the analogous results to those of figure 2, only with the long-run price impact measures being derived from an estimation of equations 5 and 6, which control for daily trading volume. As shown in figure 3, for nine out of ten stocks, announcement days remain correlated with a reduction in the long-run price impact of a trade. Moreover, the one stock for which this result is not found is BMY, which had a negligible decline in price impact when I did not control for trading volume. The magnitude of the decline is also largely comparable to what was reported in figure 2. The impact of a median-sized trade of Home Depot (HD) stock on a day with typical trading volume, for example, is estimated to be 0.93 basis points on nonannouncement days, but only 0.62 basis points on announcement days. This represents the same 33 percent reduction in price impact for HD that was reported in figure 2.

Next, I analyze my extensions to this basic framework. One extension is to explore whether the private information that does get released in an earnings announcement may partially “leak” to the public before the official release or, alternatively, whether the private information is at a maximum before the release. A related question is how the private information component of price impact varies after the announcement date. For example, does the relationship between trading and returns immediately revert to a more normal level or does price impact remain at a lower level for some time following the earnings release?

Figure 4 (p. 51) addresses these questions by reporting the cumulative long-run price impact for a median-sized trade, calculated from an analysis using equations 7 and 8. Recall that in this model specification, the relationship between trading and returns is allowed to vary not only on an announcement day, but also on the day before and the day after an announcement. To illustrate the information contained in figure 4, I highlight the results for shares of Nortel Networks Corp. (NT). The first bar in panel H reports that the cumulative long-run impact of a median-sized trade of NT stock is 1.6 basis points on a day not in proximity to an earnings announcement. The second through fourth bars calculate the same quantity only on the day of, the day before, and the day after an
FIGURE 2

The long-run price impact of a trade on normal days and announcement days

A. Bristol-Myers Squibb Co. (BMY) basis points

B. EMC Corp. (EMC) basis points

C. General Electric Co. (GE) basis points

D. Home Depot Inc. (HD) basis points

E. International Business Machines Corp. (IBM) basis points

F. Coca-Cola Co. (KO) basis points

G. Merck & Co. Inc. (MRK) basis points

H. Nortel Networks Corp. (NT) basis points

I. Pfizer Inc. (PFE) basis points

J. SBC Communications Inc. (SBC) basis points

Sources: Author’s calculations based on data from Briefing.com and the New York Stock Exchange Trade and Quote database.
FIGURE 3

The long-run price impact of a trade on normal days and announcement days, controlling for changes in trading volume

A. Bristol-Myers Squibb Co. (BMY)
basis points

B. EMC Corp. (EMC)
basis points

C. General Electric Co. (GE)
basis points

D. Home Depot Inc. (HD)
basis points

E. International Business Machines Corp. (IBM)
basis points

F. Coca-Cola Co. (KO)
basis points

G. Merck & Co. Inc. (MRK)
basis points

H. Nortel Networks Corp. (NT)
basis points

I. Pfizer Inc. (PFE)
basis points

J. SBC Communications Inc. (SBC)
basis points

Sources: Author’s calculations based on data from Briefing.com and the New York Stock Exchange Trade and Quote database.
The long-run price impact of a trade before, during, and after announcements, controlling for changes in trading volume

**Figure 4**

**A. Bristol-Myers Squibb Co. (BMY)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**B. EMC Corp. (EMC)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**C. General Electric Co. (GE)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**D. Home Depot Inc. (HD)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**E. International Business Machines Corp. (IBM)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**F. Coca-Cola Co. (KO)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**G. Merck & Co. Inc. (MRK)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**H. Nortel Networks Corp. (NT)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.00</td>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**I. Pfizer Inc. (PFE)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>0.60</td>
<td>0.40</td>
<td>0.20</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**J. SBC Communications Inc. (SBC)**

<table>
<thead>
<tr>
<th>Basis points</th>
<th>Normal day</th>
<th>Day of</th>
<th>Day before</th>
<th>Day after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>1.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Sources: Author’s calculations based on data from Briefing.com and the New York Stock Exchange Trade and Quote database.
earnings announcement. As reported in the second bar, the price impact of a trade of NT falls to around 1.3 basis points on an announcement day. However, the figure also indicates that the price impact remains near this lower level on the following day. On the day before the announcement, however, the price impact of an NT trade is higher than is typical, measuring approximately 1.8 basis points.

This pattern is consistent with the following story of the private information content of price impact. Suppose that every day new information about the value of NT shares is generated, but initially, this new information is private. Suppose further that none of this private information is released until the day of the announcement. In this scenario, one expects that the amount of private information is greatest just before the announcement. According to microstructure theory, this would then cause the price impact of a trade to be greatest just before an announcement and to fall after an announcement. This story is therefore consistent with the estimated price impact of NT trading around announcement dates. Four of the ten sample stocks, however, are estimated to have a greater price impact on the day before an earnings announcement, and so this story is potentially an explanation for only some firms.

Consider a different case such as the one illustrated by the results for trading of IBM stock. The long-run price impact of a trade of IBM is lower on days immediately before and immediately after an earnings announcement than it is on other days. Five of the ten stocks match this pattern. If private information was the source of the change in price impact, then these results suggest that private information is reduced before the announcement date. This would be consistent with a potential information leak or perhaps with information being intentionally released by the company prior to its formal quarter earnings release.

One final issue I explore is whether the reduction in price impact observed on announcement days is related to what news is actually released in the announcement. For example, an earnings release that is in line with market expectations may not reduce private information very much, since even in the absence of a formal announcement, market participants seemed quite knowledgeable about the announcement’s contents. A surprising announcement, however, may reveal a greater amount of private information. An alternative hypothesis is that a surprising announcement may generate more private information because there may be more differences in opinion as to the implication of an earnings surprise on the fundamental value of the stock.

I explore the relationship between price impact and announcement content by estimating equations 9 and 10, which allow the trading and return relationship to vary according to whether a trade occurs on an announcement date and whether the given announcement is surprising. I define a surprising announcement as one in which the market’s expected earnings were more than $0.01 per share away from the actual reported value. For nine of the ten firms in the sample, this identified roughly half of all announcements as surprises. The tenth firm, General Electric Co. (GE), did not have a surprising announcement over the entire sample period, with earnings never being more than a penny away from the market’s expectation. For this reason, I do not include GE in this final empirical estimation.

Figure 5 presents the long-run price impact of a median-sized trade of each of the remaining nine companies. As is illustrated in the figure, there does not appear to be a general relationship between private information content and announcement surprise content. In particular, a trade of six of the nine stocks is associated with a lower price impact when the announcement is more surprising relative to when it is not. For instance, a median-sized trade of Pfizer Inc. (PFE) stock typically moves the price by 0.66 basis points. On a day when an unsurprising announcement is made, price impact falls to 0.62 basis points. On a day when the earnings announcement is also more than a penny away from the market’s expectation, price impact falls by even more, to 0.51 basis points. The evidence from the remaining three stocks indicates the opposite relationship between announcement content and price impact reduction. For instance, a trade in the stock of SBC Communications Inc. (SBC) typically moves the share price by 1.23 basis points. This falls to 1.14 basis points on an announcement day without an earnings surprise, but falls by less than 0.01 basis points on announcement days when earnings miss expectations by more than $0.01.

**Conclusion**

In this article, I examine how the price impact of a trade varies throughout the days surrounding public earnings announcements. My results indicate that public news releases correlate with a reduction in the price impact of a trade. This finding is consistent with earnings releases generally reducing the asymmetric information component of stock trading. Moreover, this result is robust to the typical increase in trading volume generally observed on such days. Extending the sample beyond a focus on the announcement day alone, however, fails to uncover systematic relationships on either the day before or the day after earnings announcements. In particular, the reduction in price
FIGURE 5
The long-run price impact of a trade for surprising and unsurprising announcements, controlling for changes in trading volume

A. Bristol-Myers Squibb Co. (BMY)
basis points

B. EMC Corp. (EMC)
basis points

C. General Electric Co. (GE) (results not applicable)
basis points

D. Home Depot Inc. (HD)
basis points

E. International Business Machines Corp. (IBM)
basis points

F. Coca-Cola Co. (KO)
basis points

G. Merck & Co. Inc. (MRK)
basis points

H. Nortel Networks Corp. (NT)
basis points

I. Pfizer Inc. (PFE)
basis points

J. SBC Communications Inc. (SBC)
basis points

Notes: Results for General Electric Co. (GE) are not applicable. See the text for further details. Sources: Author's calculations based on data from Briefing.com and the New York Stock Exchange Trade and Quote database.
impact on announcement days does not typically persist beyond one trading day, nor do markets seem to contain higher than average levels of asymmetric information on the day prior to anticipated announcements. Perhaps most surprisingly, I do not find a predictable relationship between the change in price impact and the information content of the announcement. For some firms, surprising announcements tend to increase asymmetric information and price impact relative to unsurprising announcements, whereas for other firms the reverse is true.

NOTES

1More generally, the price of a share of stock should equal the present value of future dividends discounted at a rate commensurate with the risk of the given payment stream, where risk is measured by an asset pricing model. Thus, expectations about both future dividends and future risk are relevant in determining current market prices.

2A marketmaker is an individual or firm authorized by the stock exchange to buy and sell a particular security with an objective to provide trading liquidity for the security. Generally, a marketmaker is obliged to announce buying and selling prices for a particular security at a particular time.

3An implication of this model is that traders lacking private information face higher trading costs in that they must compensate the marketmaker for being willing to transact at posted prices in the presence of those with more information.

4Readers who are more interested may wish to begin a more in-depth review of market microstructure analysis by reading Biais, Glosten, and Spatt (2005).

5Briefing.com reports earnings expectations from Zacks Investment Research and from Reuters. As it is more complete, I choose the expectation reported by Zacks, but use Reuters data when Zacks data are missing.

6For these calculations, trading volume is set to a stock’s median (across days in the sample) daily trading volume. Regression results indicate a strong negative relationship between trading volume and price impact. That is, days with higher trading volume are associated with lower price impact of a single trade.

REFERENCES


An alternative measure of inflation

François R. Velde

Introduction and summary
Controlling inflation is a primary goal of monetary policy. In order to control inflation, central bankers need to be able to measure and forecast inflation as best they can. Forecasting is particularly important, given the fact that monetary policy operates with "long and variable lags," and therefore policymakers need to act well in advance of actual developments in inflation, on the basis of their forecasts.

Both the measurement and the forecasting of inflation have been subjects of ongoing debate and research in recent years. This article reports on my research on both aspects. Specifically, I develop a new measure of inflation, which can also be used to generate a forecast. The research is still very preliminary, but the first results are encouraging. In particular, it appears to provide some gains in forecasting compared with what remains the best and simplest forecasting model, namely, the random walk model of inflation.\(^1\)

Motivation
For the U.S. Federal Reserve (and for many other central banks), price stability is a primary goal, mandated by law. This stability is usually interpreted to mean a low level of inflation (how low will not be debated here). So what is inflation and how do we measure it?

Inflation is generally defined as the rate of change of some price index: Well-known examples are the Consumer Price Index (CPI) and the Personal Consumption Expenditures (PCE) Price Index. Price indexes, generally speaking, result from the attempt to measure with a single number a change in a collection of prices \(p_i\) (for \(i = 1, \ldots, n\)).

The simplest conceivable index is to take a straight average of the prices in each period, ignoring the quantities. But it seems more reasonable for many purposes to weight the prices. Movements in the price of an item that is of little importance relative to the others should not be given much weight. An item of little importance is one that does not represent a large share of expenditures, which naturally leads one to use expenditure shares in creating the weights (see box 1).

More generally, suppose we have observations on prices at which a given range of goods and services are bought and sold and also observations on the quantities bought and sold at those prices. We thus have a collection \(\{p_i\}\) and a collection \(\{q_i\}\). Suppose further that we have observations in two periods, 0 and 1. One period (either 0 or 1) is chosen as the reference period. The problem of constructing an index (for either prices or quantities) is that of devising a formula that takes the prices and quantities in both periods and yields a single number. The formula must be such that, if the prices and quantities are the same in the reference period and the other period, the number is 1. Note that, even if prices are unchanged between the two periods, a change in quantities will generally result in the index being different from 1. Even though prices are unchanged, the weighting of the prices, which is based in part on the quantities, will change the overall index.

From this brief overview, one can draw some general observations about price indexes. Most price indexes require information on quantities in order to weight the prices. For certain applications, the fact that quantities are measured less precisely, less easily, and less quickly than prices can be a problem. At a deeper level, there is an important connection between

François R. Velde is a senior economist at the Federal Reserve Bank of Chicago. The author thanks Michael Kouparitsas for all his advice, David Kang for outstanding research assistance, and his colleagues at the Chicago Fed, particularly Craig Furfine and Marco Bassetti, as well as Lawrence Christiano and Lars Hansen, for their comments.
Different kinds of indexes

A straightforward weighting scheme is to use the expenditure shares to weight the items. And, since the choice of the units to measure the quantities of goods, and therefore the prices per unit of goods, is arbitrary, the absolute level of an index is meaningless, and an index can only measure changes relative to a reference period. One thus arrives at the classic price indexes to measure changes between period 0 and period 1—the Laspeyres and Paasche indexes, depending on whether one chooses period 0 or period 1 as the reference period, and the Fischer index, which takes the geometric average of the two indexes and thus achieves a pleasing symmetry between the two periods.

The CPI is essentially a Paasche index, with weights based on a reference period that is changed from time to time. The current PCE index is a type of Fisher index, with no fixed reference period: The changes computed by period are chained together to form an index series. The Fisher index has another nice feature—a quantity index that can be computed in exactly the same way (weighting quantities by expenditure shares); in any period the quantity index times the price index equals the total expenditure. Thus, the price index can be seen as a deflator of the nominal expenditures that yields an index of real quantities.

The prices and quantities. In much of index theory, the two collections are intimately related, and the index makes sense with respect to a particular set of quantities. For example, the CPI is based on weights representing the typical basket of goods and services consumed by an average urban consumer. The PCE deflator is based on the quantities of goods and services consumed in the economy as a whole. These indexes yield different inflation rates, partly because they are tied to different collections of goods and services, and the computation of the index depends on the collections themselves, as well as on the quantities.

More generally, price indexes are designed for a certain purpose and have optimal properties for that purpose, but they may not be well suited for others. Monetary policy needs measures of inflation, but it may well be that indexes designed to measure the value of a basket of consumer goods or to convert nominal consumption expenditures into real consumption expenditures are not perfectly suited for the goals of monetary policy.

In fact, policymakers have come to use variants of both the CPI and PCE index, the so-called core measures. The intuition behind these measures is that monetary policy is interested in broad and persistent movements in inflation and that certain price series, being too volatile, introduce noise and confusion in the measurement of these broad movements. Therefore, the troublesome series (typically food and energy-related items) are removed altogether from the price index.

The alternative approach I propose here extends the intuition behind the core measures of inflation. My premise is that inflation is a general movement in the price level or, put differently, a movement that is common to all individual price series. Once we posit the object to be measured (inflation) as a statistical series of its own, then the measurement problem can be seen in a different light, as a signal extraction problem. Constructing (weighted) averages is a way of measuring inflation that makes particular assumptions about the movements that are specific to each series: essentially, that they are a sort of observation noise that can be removed by taking averages and counting on the law of large numbers. But these movements specific to each price series can have a more complex structure than being just noise. As it happens, statistical tools are available to measure inflation and allow for more complex structures. The result of this approach is still, in a way, a weighting scheme, but it is a dynamic weighting scheme, and it is one that weights series not by their importance in a basket, but according to the information that they contain.

Method

The Kalman filter

The Kalman filter relies on a distinction made between what is observed and what is not. This is formalized by writing two equations, known as the state equation and the observation equation. The first equation posits the evolution over time of the hidden variables, gathered in a vector called the state vector. The specification is typically dynamic, meaning that current value taken by the state depends on past values. One of these hidden variables will be our general movement in the price level. The second equation describes the relation between the state and the observables.

I call the vector of observable variables \( y_t \); That is, at every point in time \( t \), \( (y_{t1}, y_{t2}, y_{t3}, \ldots) \) represent the values of the series 1, 2, 3 at time \( t \). There is another vector, made up of variables that are not observed: It is called the state vector, \( x_t \). The state equation describes how this (unobserved) state changes over time. The general form is

\[
x_t = Ax_{t-1} + u_t,
\]
where \( A \) is a matrix and \( u_t \) is a noise or error term.

The observation equation relates the observables and the state in the following way:

\[
y_t = B x_t + v_t,
\]

with \( v_t \), another noise or error term, uncorrelated with \( u_t \). Having specified this model of the relationships between the state and observables, I need to supply initial guesses about two things: the initial value taken by the state and the uncertainty surrounding that initial value. Typically, the initial value is assumed (in the absence of any other information) to be the long-run average of the state, and the uncertainty surrounding this value can be derived from the state equation.

I am now ready to apply the Kalman filter. It may seem a little magical to estimate the value of a variable (the state) that is never observed. The way it works is as follows: The method is recursive, meaning that at any point in time it takes the most recent guess and updates it in a systematic manner based on the newly available information. Given a guess as to the value of the state yesterday, and the uncertainty around it, the mechanical application of the state equation provides a best guess as to its value today (before I introduce any of today’s information).

How do I represent today’s information? The basic rule here is learning from one’s mistakes. Since I have a guess of today’s state, I can make a guess of today’s observables, using the observation equation. Then I compare this guess with what actually happened: The difference between the two is the new information that is relevant to my model.

How do I incorporate today’s information? I project the (unknown) value of today’s state onto all of the information, which can be decomposed into the information available yesterday and the new information that became available today. A classic result of regression analysis tells us that this projection is the sum of two terms. The first is simply the best guess of today’s state using the information up to yesterday. The second corrects this guess with the new information, but weights it according to two expectations: how correlated it is likely to be with today’s state and how noisy it is. The more correlated this new information is with the state, the more weight I place on it; on the other hand, the noisier it is, the more I discount it. How these weights are determined depends on the particular values I have assigned to the coefficients of the state and observation equations.

This leads to a recursive formula: Today’s guess is yesterday’s guess updated with the (appropriately weighted) new information. Tomorrow, I will take today’s guess of today’s state, derive a guess of tomorrow’s state, and repeat the procedure.

Of course, when tomorrow rolls around, I will correct my guess of tomorrow’s state that was based on today’s information. But I could also correct my guess of today’s state based on today’s information, or even yesterday’s guess of yesterday’s state. More generally, having proceeded recursively from the beginning to the end of the available sample, it is possible to go back and correct the guesses made for the value of the state in earlier periods based on the information of the whole sample. This procedure, which is also recursive but backwards (as it updates yesterday’s guess based on today’s error), is called the Kalman smoother.

This approach to measuring inflation, like any other, has costs and benefits. Some of the benefits are apparent if we think back to the initial motivation. Modeling inflation as a hidden variable allows me to bypass a number of the issues that arise for standard indexes. For example, the basic intuition behind core measures of inflation is fully extended. Price series are not ignored or deleted when they are volatile; rather, optimal use is made of the information that they contain. The approach helps me deal with the choice of optimal weights to apply to the price series because the Kalman filter algorithm itself chooses the weights that it applies recursively. But it doesn’t choose them arbitrarily; rather, it tries to extract the information contained in the price series.\(^2\) The choice of the series themselves is not eliminated, of course, but it is of less importance. There is no conceptual problem in mixing series of different origins (say, the PCE index and CPI) or choosing a subset of either collection of series. There is no “adding up” constraint; there is no need to fully represent a given basket or bundle of goods and services. The main consideration in adding another series to the collection we use should be: Is that additional series likely to provide information about inflation that was not contained in our collection already?

Finally, one major benefit of the approach is that it yields a forecasting tool at no additional cost, so to speak. My best guess of the value of the state at time \( t+1 \) based on information available at time \( t \) is simply my best guess of the state at \( t \), projected forward one period using the state equation. The Kalman filter approach thus folds into one operation measurement and forecast.

There are disadvantages, however. One cost is of a technical nature, and another is more of a conceptual problem.

The technical difficulty becomes more apparent in the next section. Although index theory may rely on some assumptions about the economic process that
generated the prices and quantities, the considerations that lead to the choice of an index are quite general and make few or no assumptions about the prices themselves. The Kalman filter approach requires that a modeling choice be made about the statistical processes that best represent the price series and the underlying inflation as well. That is, I have to take a stand on the structure of prices, their interdependence, the correlations of a series with its past values, as well as those between the series themselves, and so on. Fortunately, some statistical tests guide the choice of that structure, as I discuss in the next section.

The conceptual difficulty is the following: One might fairly argue that I am not so much measuring inflation as inventing a concept of inflation that I can measure. The underlying inflation, or “latent inflation,” may be just a statistical artifact. My response is that, although it is indeed an invented concept, it is one that captures the intuition we have about inflation. But it would be better to think of the series I uncover as an index, perhaps not of inflation itself, but of the forces that affect inflation dynamics, at least in the short run. For lack of a simpler term, I choose to call this index latent inflation, but I need to show that, in practice, it can be closely related to more standard measures of inflation. I do this in the second part of the section that follows.

The model

General form

The general form of the model I use is relatively straightforward. Let \( Y_t \) denote the individual inflation series, with \( i = 1, \ldots, N \). Let \( P_t \) be the latent inflation. I assume that the relation between them takes the form

\[
Y_t = \lambda_t P_t + \epsilon_t \tag{1}
\]

where \( \lambda_t \) is called the “loading factor.” The term \( \lambda_t \) represents the component specific to the individual inflation series \( i \). I call it the “relative inflation” for the good or service \( i \).

One would expect the loading factors to be close to 1. (As I explain later, one of them is normalized to be 1.) Indeed, it may be hard to think of a theory in which they would not be 1, since one would expect inflation to have the same impact on all series. To the extent that I do not find them to be 1, this can be re-interpreted as capturing any immediate dependence on the relative inflation from the general inflation, for example, the product of distortions generated by inflation on the pricing decisions in one sector. Formally, the equation can be rewritten as \( P_t = P_t + \epsilon_t \) with \( P_t = (\lambda_t - 1) P_t + \epsilon_t \). An alternative explanation is that loading factors different from 1 are picking up some model misspecification, such as nonlinear time trends.

Within this general framework, I consider a variety of statistical models for the relative inflation rates \( P_t \) and the latent inflation \( P_t \).

Specific form

As I explained previously, part of the cost of the approach is that it involves many choices: not only a choice of series, but also a choice of the statistical model to apply to the series. Partly to avoid deciding, but mostly to explore the properties of the general model, I have experimented with a number of variations.

The PCE index or the CPI can be thought of as the apex of a pyramid. The general price index corresponds to the most aggregated level of observation. Immediately below, there is a first level of disaggregation. In the case of the PCE index, it contains three series: an index of durables prices, an index of nondurables prices, and an index of services prices. Further down is a second level of disaggregation, which contains 13 series, and a third level. At this stage of the research, I have experimented with a collection of three series of the PCE index (the first level of disaggregation), 13 series (the second level), and 52 series (selected from the third level).

The next decision is the choice of a statistical model for the individual series. For the sake of simplicity, I have imposed the same model on the latent and the relative inflation series, but I have varied the model. All models belong to the ARIMA (autoregressive integrated moving average) families of models, which I now explain.

The simplest statistical model one can think of is that a series is white noise; that is, it consists of realizations from uncorrelated, identically distributed statistical processes—each observation (at time \( t \)) is drawn from, say, a normal distribution with constant mean and variance. Obviously, this is not a good model for inflation, which is highly persistent, but it serves as a building block for other models. The next step is to allow for serial correlation, and imagine that inflation at time \( t \) can be decomposed into the sum of last period’s realization multiplied by a factor \( \rho \), and white noise \( \epsilon_t \):

\[
P_t = \rho N_{t-1} + \epsilon_t,
\]

This introduces some persistence in the process. More generally, one can suppose that the process depends on more than one lag. The general form is then that of an AR(\( p \)), autoregressive process with \( p \) lags:

\[
P_t = \rho_1 N_{t-1} + \rho_2 N_{t-2} + \ldots + \rho_p N_{t-p} + \epsilon_t
\]
Another step is to allow the innovation $e_i$ to have effects that extend beyond the period when it occurs, without having as much persistence as the autoregressive part. That is, the innovation $e_i$ affects not just $P_i$, but also $P_{i+1}$:

$$P_i = \rho_1 P_{i-1} + \rho_2 P_{i-2} + \ldots + \rho_p P_{i-p} + e_i + \theta_1 e_{i-1}.$$

This is a mixture of an autoregressive (AR) process with a moving average (MA) component; it is called an ARMA process. The moving average part can have $q$ terms:

$$P_i = \rho_1 P_{i-1} + \rho_2 P_{i-2} + \ldots + \rho_p P_{i-p} + \epsilon_i + \theta_1 e_{i-1} + \ldots + \theta_q e_{i-q},$$

in which case the process is called ARMA($p,q$).

Estimation is much simpler if the process is stationary; that is, its properties do not vary over time, and it tends to revert to its mean rather than drift away. This will be true if the sum of the autoregressive coefficients is less than 1 in absolute value. But this may not be a good assumption for inflation, which is so highly persistent that it can look like a random walk. One solution is to take the difference of inflation and to model that difference as an ARMA process; the original process is said to be integrated of order 1 if the first difference is stationary, and the process is called an ARIMA($p,1,q$) process, where I denotes the fact that inflation needs to be differenced once. (I do not consider higher orders of integration.)

A final variant that I consider is to allow for feedback from the relative inflation to the latent inflation. This takes the following form: The relative inflation series are modeled as ARMA($p,q$), and the latent inflation is modeled as

$$P_i = \rho_1 P_{i-1} + \rho_2 P_{i-2} + \ldots + \rho_p P_{i-p} + e_i + \theta_1 e_{i-1} + \ldots + \theta_q e_{i-q} + \psi P_i.$$

I denote the model as ARIMA($p,1,q,\psi$) if I allow for such feedback and ARIMA($p,1,q,\sim$) otherwise. The same model is imposed on all series. (Further refinement of the analysis will involve imposing different models on the different series, eliminating terms that appear to be insignificant in the estimation.) Constants

**BOX 2**

**Other methods and related literature**

In addition to presenting my results, I want to say a few words about other methods.

The approach taken here is related to other work. The model I use can be seen as a special case of what are called “dynamic factor models.” These models represent a given collection of variables $\{X_1, X_2, \ldots, X_n\}$ as being determined by a set of unobserved common factors $\{F_1, F_2, \ldots, F_n\}$ and their lags, to which observation noise is added. The general form of the equation modeling each variable would be

$$X_{it} = a_{i0} F_{1t} + a_{i1} F_{i,t-1} + \ldots + a_{i2} F_{2t} + \ldots + a_{il} F_{lt} + \ldots + u_{it},$$

where the $u_{it}$ terms are not correlated over time and with each other (Sargent and Sims, 1977).

Typically, the number of factors is kept small relative to the number of variables being modeled. More recently, researchers have found that the principal components of the collection $X$ can be used to approximate the common factors $F_i$, an approximation that becomes valid as the number of factors $X$ becomes large relative to the number of factors (Stock and Watson, 1998; Forni et al., 2000). These techniques are used by the Chicago Fed’s National Activity Index (Evans, Liu, and Pham-Kanter, 2002). Cristadoro et al. (2002) use these methods to compute a measure of core inflation for the euro area using large numbers of economic series and extracting the slow-moving component of the common factor associated with inflation.

My approach is a particular form of a dynamic factor model, where the number of factors is the number of series plus one and estimation proceeds along the more traditional (and computer-intensive) line of maximum likelihood. Bryan and Cecchetti (1983) use this method with a small-scale model to estimate the degree of bias in the CPI (the bias being the difference between actual CPI and the estimated latent variable). They do not assess the properties of their estimated variable or its forecasting ability. Jam (1992, 2001) uses the state space approach with only price series to remove seasonal fluctuations from price series, but the focus is not on estimating latent inflation. Other uses of the state space model approach to estimate or predict inflation include Bomhoff (1982), who uses a small economic model to relate inflation to money and output; Hamilton (1985) and Burmeister, Wall, and Hamilton (1986), who estimate current expectations of inflation using variables such as interest rates; and Laubach and Williams (2003).
are included in all the ARMA models, allowing for potentially different trends in relative inflation.

**Estimation method**

To estimate each model, I use the so-called estimation-maximization (EM) algorithm detailed in Watson and Engle (1983). The problem is to find the values of the parameters of the model: the loading factors; the \( \rho, \theta, \psi \) coefficients; and the variances of the innovations \( e_t \) for each relative inflation and for the latent inflation. The difficulty is that the Kalman filter and smoother formulas can compute estimates of the latent variable assuming that these parameters are known; however, they are not known, and they must themselves be estimated.

The EM algorithm uses the classic approach of assuming we know what we don’t know. Specifically, one starts with a guess for the parameters, applies the Kalman smoother, and computes estimated series (the estimation step); then, pretending that these estimated series are observed data, one finds new estimates of the parameters, essentially by regressing the observed price series on the relative and latent inflation to compute the loading factors, as well as the hidden variables on their lags to compute the parameters of the ARMA model. The main drawback of this algorithm is that it converges very slowly and makes computation time-intensive.

In box 2 (p. 59), I present a brief overview of other methods. In the following section, I discuss my results.

**Results**

**The estimated latent inflation**

As I have explained, the approach to modeling relative prices as well as latent inflation is somewhat agnostic: A variety of models have been estimated. Which does one choose? One criterion is how well the model fits the existing data (I use the sample period from 1959:Q1 to 2005:Q1). The estimation procedure tries to maximize the likelihood that the observed data were generated by the estimated model, and one can simply compare the resulting likelihood across models. Of course, models with more parameters will tend to do better, simply because they have more parameters, and ways have been devised to take this into account.

Figures 1–3 show the estimated values of the latent variable over the sample, compared with the quarterly core inflation rate, for three models: two models chosen on the basis of fit—the ARIMA(2,0,0,–) and the ARIMA(3,0,1, \( \psi \))—and a model that will turn out to have good forecasting ability in the next section, the ARIMA(2,0,2,–). The figures also show the forecasted path of the latent variable over the next 12 quarters. Note that this is not a forecast of core inflation, but only a forecast of the latent variable. I use this forecast of the latent inflation in the next section to forecast core inflation. Another important point is that neither the level nor the amplitude of the latent variable can be determined. The estimation procedure normalizes to 1 the first loading factor (in effect, I scale the latent variable so that its amplitude is comparable to that of the first price series), and in figures 1–3 I add the value of core inflation in 1959:Q2 to the level of the latent variable. Thus, the scale of the figure only applies to core inflation, and if the figures allow us to compare visually the two series, they should not be taken to mean that core inflation is higher or lower than the latent variable at any particular point in time.

Overall, the behavior of the latent variable is similar to that of core inflation. It’s worth recalling that I did not remove food or energy from the series I used to estimate the latent variable.

**Forecasting with the latent inflation measure**

As I mentioned previously, one difficulty with the latent inflation approach is that the variable I am measuring is a construct. How can I be sure that it is
measuring what I think it might be measuring? Is it of any use or, more precisely, does it capture the latent inflationary pressures that are in play, at least in the short term?

One way to evaluate the latent inflation measure is to find out if it holds any predictive power for inflation as it is commonly measured. To find out, I carry out an out-of-sample forecasting exercise similar to Fisher, Liu, and Zhou (2002) and Brave and Fisher (2004). As this research has emphasized, the naive model of inflation, which predicts that inflation in the future will be what its most recent value was, is “the man to beat.”

I proceed as follows. For each quarter $T$ between 1984:Q2 and 2002:Q2, I take the sample ranging from the beginning of the series (1959:Q1) to the chosen quarter $T$. Using only the data in this sample, I estimate a family of ARIMA models. Then, I run various regressions of core inflation over various horizons (that is, core inflation from quarter $t - H$ to quarter $t$, where $H$ ranges from 1 to 8) on the estimated measure of the latent inflation and current and lagged core inflation within the sample. Then, I construct a forecast of latent inflation over the horizon $T$ to $T + H$ and use those forecasts as well as the values for current and lagged core inflation to project core inflation over the horizon $T$ to $T + H$.

Having done this for all quarters $T$ between 1985:Q1 and 2002:Q2, I compute the root mean squared error (RMSE) of these forecasts. I compare this RMSE to the RMSE of the naive model, which simply predicts that core inflation over $T$ to $T + H$ will be what core inflation was from $T - 4$ to $T$.

Using the same notation as Brave and Fisher (2004), core inflation from $t - H$ to $t$ is

$$\pi_t^H = \ln p_t - \ln p_{t-H}.$$  

while core inflation from $t - 1$ to $t$ is simply denoted $\pi_t = \ln p_t - \ln p_{t-1}$.

Note that

$$\pi_t^H = \pi_{t-H+1} + \pi_{t-H+2} + \ldots \pi_t.$$

Latent inflation $x_t$ is calculated as the latent variable in the statistical model, and the latent variable over the $T$ to $T + H$ horizon is simply

$$x_t^H = x_{t-H+1} + x_{t-H+2} + \ldots x_t.$$
The regression I run is

1) \[ \pi_t^H = \alpha x_t^H + \beta_0 \pi_{t-H} + \beta_1 \pi_{t-H-1} + \gamma. \]

The statistical model allows me to project \( \hat{\pi}_{t+H}^H \) and then construct an estimate

\[ \hat{\pi}_{t+H}^H = \alpha \hat{x}_{t+H}^H + \beta_0 \hat{\pi}_t + \beta_1 \hat{\pi}_{t-1} + \gamma. \]

Note that, in equation 1, latent inflation is included in the regression in addition to lagged inflation. Such an inclusion usually hurts the predictive power of the forecasting equation (out of sample). By contrast, if latent inflation helps significantly, this is a success. Note also that, although a lot of work goes into coming up with the series \( x_t^H \) and the forecast \( \alpha x_{t+H}^H \), the regression itself is simple and has only three regressors.

The results in terms of relative RMSE are shown in table 1. For each model and each horizon (one to eight quarters ahead), the table shows the model’s RMSE relative to the naive model (a number lower than 1 indicates that the model performs better). The models are sorted by order of increasing likelihood.

The pattern of performance varies considerably across models. One group of models does substantially worse than the others: As it turns out, these are the models that allow feedback from lagged relative inflation to latent inflation. The models without feedback do markedly better than regressing core inflation on two quarters of inflation, the performance of which is given in the last row of table 1. In other words, the addition of the latent inflation to the regression substantially improves the forecasting performance. Which model performs best depends on the horizon: At the short horizon (one to three quarters), the ARIMA(2,1,0,~)

| TABLE 1 |
| Root mean squared error relative to naive model |

<table>
<thead>
<tr>
<th>Forecast horizon one to eight quarters ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(2,1,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,1,0) with feedback</td>
</tr>
<tr>
<td>ARIMA(3,1,0) with feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,1) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,2) without feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,1) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,0) with feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,0) with feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,1) with feedback</td>
</tr>
<tr>
<td>1 lag of inflation alone</td>
</tr>
</tbody>
</table>

| TABLE 2 |
| Performance of the moving average versions of the models’ forecasts |

<table>
<thead>
<tr>
<th>Two-quarter moving average</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(2,1,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,1) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,2) without feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,1) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,0) without feedback</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Three-quarter moving average</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA(2,1,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,1) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,2) without feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,1) without feedback</td>
</tr>
<tr>
<td>ARIMA(2,0,0) without feedback</td>
</tr>
<tr>
<td>ARIMA(3,0,0) without feedback</td>
</tr>
</tbody>
</table>
does slightly better; for all other horizons, the winner is the ARIMA(2,0,2,~), with the ARIMA(3,0,1,~) not far behind. Neither one, however, manages to do any better than the naive model, though they come reasonably close, within 10 percent of the RMSE of the naive model.

Figure 4 compares the forecasts of core inflation produced by the naive model (gray line) and the latent inflation model ARIMA(2,0,2,~) (green line) with actual core inflation (black line) at the two-year horizon. The date on the horizontal axis is the date at which the forecast is made. The gray line is the black line shifted by two years, since the naive model predicts that inflation two years hence will be the same as today. The predictions of the latent model are not substantially different from those of the naive model, and hence the latent model does not perform any better. But what is striking is how variable the green line is, relative to the gray line. The reason is as follows. The gray line averages actual inflation over the previous eight quarters and therefore smoothes out a lot of the quarter-to-quarter variability in inflation. The latent inflation model incorporates the new information that arrives in each quarter, and even though it weights it appropriately, the new information shifts the estimate of where latent inflation currently stands; this in turn shifts the whole projected path of latent inflation, and hence the forecast of core inflation. There is no smoothing mechanism here.

It is possible, of course, to add a smoothing mechanism. For example, I have tried replacing the latent model’s forecast with a two-quarter or three-quarter moving average of itself. This ad hoc procedure produces a smoother forecast. Its performance is shown in table 2, only for selected models.

The performance of both the ARIMA (2,0,2,~) and the ARIMA(3,0,1,~) models is improved markedly. Just taking a two-quarter moving average reduces the relative RMSE for the ARIMA(2,0,2,~) from 1.10 to 0.95. It becomes possible to beat the naive model, although not by a great amount. Figure 5 compares the predictions of this moving average: The green line is clearly smoother, and in some instances, it seems to do better in terms of predicting changes in inflation (for example, the downturn in the mid-1980s and the pick up in the early 1990s).
Conclusion

This article has presented recent research on measuring and forecasting inflation. The approach taken, that of state space modeling, consists of representing latent inflation as an unobserved variable affecting simultaneously a collection of individual price series, for example, the main components of an aggregate price index like the PCE deflator. The approach extends the intuition that lies behind the use of core measures of inflation in that it takes the individual price series to be noisy observations on true, underlying inflation, and filters out the noise in the individual price series. The resulting estimated latent inflation validates the use of core inflation, since the two series look very much alike. The latent inflation approach has the additional benefit of yielding a forecast of future inflation, and preliminary results indicate that some progress can be made in reducing out-of-sample forecasting error.

NOTES

2 Note that I am not fully escaping the use of weighted indexes, since the individual price series will, in practice, be indexes of their own.
3 If the model has moving average components, the $e_t$ series are treated as yet another unobserved variable.
4 Two such criteria are commonly used, the Bayesian information criterion (BIC) and the Akaike information criterion (AIC); the former tending to be stricter than the latter. In my family of models, the BIC chooses the most parsimonious model, the ARIMA(2,0,0) with no feedback, while the AIC ranks almost equally the ARIMA (3,0,0) with feedback and the ARIMA(3,0,1) with feedback.
5 I thank former Federal Reserve Board Chairman Alan Greenspan for this suggestion.

REFERENCES


