

**University of Michigan's Survey of Consumers:
Measuring and Interpreting Economic Expectations¹**

Richard Curtin
University of Michigan

Pierrick Dechaux
University Paris 1

Introduction

This presentation is composed of three distinct topics driven by changes in the social, demographic, and economic structure of the U.S. population. These changes have significant implications for measuring and interpreting economic expectations, and have prompted significant revisions in the consumer sentiment surveys conducted by the University of Michigan. The three basic changes involve: first, changes in the sampling method and mode of interviewing, second, the inclusion of a wider range of economic expectations as well as different measurement scales, and third, a renewed recognition of the importance of expectations of longer term economic outcomes on the current spending and saving behavior of consumers.

Specifying these changes in detail could easily extend far beyond the time that is allotted for this presentation. Moreover, it is a research agenda that naturally entails uncertainties about what will eventually be the most productive solutions to the current challenges. The “new normal” in survey research will not simply be a new set of techniques that gives birth to a new and stable era similar to what we have enjoyed in the past. The new normal will be to devise the best means to constantly adapt to a changing environment (Link, 2015). The hope for an elegant and simple solution will no doubt continue, and, perhaps, some day that dream may be fulfilled. In the interim, we must face today’s challenges in a manner that maintains a consistent, unbiased, and representative time-series of leading economic indicators. Successful cyclical indicators require robustly measured economic expectations that primarily reflect underlying economic factors rather than being shaped by fads in measurement methodology or subject to discontinuous shifts due to changes in social, demographic, or economic environments. Yesterday’s challenge to tame rising costs pales in comparison with today’s challenges to maintain representative samples, minimum measurement errors, and robust time-series analyses.

Richard Curtin, Research Professor and Director, Consumer Sentiment Survey, University of Michigan, curtin@umich.edu. Pierrick Dechaux, Ph.D student, University Paris 1, pierrick.dechaux@gmail.com.

Survey Methodology

The goal of any survey is to produce the highest quality data at the lowest cost possible. It is undeniably true that costs get most of the attention since monetary expenses are more easily measured. Quality is much more difficult to assess. The traditional concerns of sampling are known probabilities of selection, population coverage, response rates, and mode and techniques of interviewing. Each area will be briefly discussed, starting with population coverage.

Population Coverage. The goal of the Surveys of Consumers is to estimate economic expectations on a representative sample of all households located in the United States. While that goal has not changed since 1946, the sampling frame from which the interview samples are drawn has changed. The surveys were based on area probability samples from 1946 to 1976, with the interviews conducted in person. Coverage of the entire population was nearly complete, although the homeless and those living in institutions such as nursing homes were excluded. By the mid-1970's telephone ownership became sufficiently widespread so that a new sampling frame was adopted based on telephone numbers. The shift in the late 1970's from face-to-face to telephone interviews involves a trade-off between costs and population coverage: about 90% of household owned telephones by the mid-1970's, which rose to 95% by 1990, and to 98% in 2000. While the initial shift to RDD landline telephone samples was propelled by rising costs of face-to-face surveys, the shift to cellular interviews was primarily due to coverage issues.

The ownership of landline telephones has declined at a more rapid pace than had been predicted, and the ownership of cellular phones has increased at a more rapid pace than anticipated. In 2014, household ownership of landline phones fell to just 51%, while ownership of cellular phones became the most quickly adopted consumer technology ever recorded. In 2014, 88% of all households and 90% of all adults owned a cell phone. Both types of phones were owned by 43%, while 45% of all households only owned wireless phones and just 8% only owned landline phones (Blumberg & Luke, 2015). Moreover, even among households that have both a landline and cell phone, people have increasingly displayed a propensity to only answer the cell phone since regulations reduce nuisance calls on cell phones but not for landline calls. Based on this shift in phone ownership, the University of Michigan survey began to base its samples on a dual landline and cellular frame in July 2012, with the shift to an all cellular sample by the end of 2014.

Prior to the shift by the Michigan surveys toward a cellular sampling frame, experiments were conducted on the differences between estimates based on cellular compared with landline phone samples. The same random digit dialing (RDD) procedures were applied to both frames. When the focus was on measurement errors, both samples yielded essentially comparable results for the key survey measures (Jaing et al. 2015). The big difference was in coverage, with rapidly falling coverage rates of the landline sample rising rapidly in the past few years, and increasing coverage rates of cellular samples. The main demographic correlate of coverage differences was in the age of the respondent, with landline interviews increasingly limited to older respondents. The current cellular samples are no longer younger than average but yield a close approximation to Census data across all age groups.

Once a household was selected by probability methods, the Michigan surveys used a probability method to select a respondent from among the eligible adults living in the household. That process changed with the move to an all-cellular sample. Landline phones, with a few exceptions, were mainly considered household phones. Cell phones, in contrast, are mainly considered personal phones. The cell phone sampling procedures confined respondent selection to whoever answered the phone, with the only qualification that they be 18 years or older. Questions are included about other cell and landline phones owned by household members so that household selection probabilities could be correctly computed and incorporated into survey weights. In addition, all of the questions asked in the survey focus on the household's financial situation. As a result, this shift in respondent selection did not affect the representativeness of the sample.

The other option to improve population coverage is to use a listing of all residential addresses maintained by the U.S. postal system. This technique is widely known as "address-based sampling" or ABS (Link et al. 2009). The population coverage exceeds both landline and cellular samples since it includes the small share of households that do not own any phone. Given that a postal address but no phone number is included, ABS samples are ideally suited to mail surveys. The Michigan surveys have conducted extensive experiments with ABS samples.

Comparisons of the results from the ABS mail and RDD landline surveys were done using the following three criteria (Peytchev et al. 2010). The first was coverage bias, since the ABS frame included all households regardless of telephone ownership. The second was differential measurement bias, since the different modes of interviewing may lead to differential measurement errors. The third was the differential overall bias between the two survey modes. Small errors on these measures would indicate that the shift toward ABS samples would yield estimates similar to the telephone frame, avoiding a break in the times series due to methodological changes. The results were confounded by differences in how the responses to a few questions were coded. One question on the economic outlook is treated as an open-ended question in telephone interviewing but fixed response categories were substituted in the mail survey; another question on personal finances omitted the "same" category when the question was asked in the phone interview but the mail interview included that response. In addition, mode differences could be expected since the questions were verbally presented in one mode and visually presented in the other mode. Aside from the two questions which were handled differently in the two modes, only one other question out of 13 measures recorded significant measurement and overall bias. Aside from differential measurement errors, differences in coverage did not affect the measures of economic expectations (Elkasabi et al. 2014).

Mode Effects on Time-series Measures. The use of mixed-mode designs must directly confront the issue of how the mode of data collection impacts measures of change in expectations over time. After all, the most important result from the Michigan surveys is not the point estimate but how it has changed from the prior measurement. Most of the research on the effect of alternative modes on survey responses has been based on cross-sectional designs (e.g., Dillman and Christian, 2005; Jäckle, Roberts, and Lynn, 2010), with most of the focus on response rate differences or differences related to the demographic characteristics of the respondents (e.g., Link and Mokdad, 2006; Millar and Dillman, 2011; Olson, Smyth, and Wood, 2012). Different survey modes,

however, may yield different response distributions to the same question. A common explanation for a mode effect on responses is that they are due to the difference between a verbal and a visual presentation—between a telephone interview and a mail interview, for example. While there is research on differences in estimates across modes in estimates from cross-sectional surveys (Voogt and Saris, 2005; Kreuter, et al., 2008; Dillman, et al., 2009; Heerwegh and Loosveldt, 2011), little research has been done on the effect of different survey modes on the measurement of wave-to-wave change.

The standard hypothesis is that mode differences cancel out in estimates of change over time: since mode effects are typically represented by a constant factor, these effects are assumed to essentially disappear in measures of month-to-month change. Of course, this hypothesis only applies when the same mode is used, and not for the increasingly common practice of using multiple modes in each survey wave. Moreover, mode effects are typically assumed to be constant across demographic groups as well as across the range of the estimated change in expectations. The Michigan survey is in the process of conducting experiments to understand the effects of alternative modes and mixed-mode designs on estimates of change. Measurements of mode effects on the estimates of month-to-month changes in expectations are critical, at least in terms of minimizing the size of the effect. A proposal to the National Science Foundation has been submitted to further this important research.

Response Rates. While population coverage improved with the recent shift to a cell phone frame, response rates continued on the same downward trend. This longstanding problem was the subject of intensive research some year ago (Curtin, Singer, and Presser, 2000, 2005). Time-series data from the Michigan surveys was utilized to test the sensitivity of the various measures of economic expectations to increasing unit non-response. The findings were encouraging in that rather large simulated declines in response rates did not bias the survey findings. A critical question that has emerged over the past few years is how low can unit response rates fall before a probability sample can no longer claim representativeness. If response rates fell below 10%, can we claim the survey results still retain the desirable statistical attributes of a probability based sample?

Change is sometimes prompted by the availability of superior methods, and sometimes change is forced by deteriorating circumstances. The only alternative is non-probability samples. Such samples are now commonplace on the internet as well as in commercial applications conducted by traditional survey research organizations. Moreover, new statistical methods are being developed to enable statistical inferences based on non-probability samples (see Baker et al., 2013 for an overview of non-probability samples). It is hard to envision the next several decades of the Michigan surveys without the incorporation of non-probability samples. Nonetheless, we have moved from an unquestioned aversion to non-probability samples to question whether probability sampling is necessary at all. Of course, big data samples drawn from various sources are the primary candidates, which will be discussed shortly.

In the interim, the Michigan surveys have experimented with mixed mode surveys in an attempt to improve response rates. Specifically, we have experimented with adding Internet and mail surveys to augment the cell phone samples. Importantly, the internet and mail surveys were

based on probability samples of U.S. households using ABS methodology. Letters were sent to the selected random sample with invitations to do the interview on the Web or by mail. Two findings emerged. First, rather than mail and internet surveys appealing to different demographic subgroups, no survey mode significantly enhanced coverage rates. Second, response rates from mail/internet surveys were comparable to landline phone surveys, and significantly higher than those obtained using cellular phone samples. Along with much lower costs, mail surveys appear so promising that the Michigan survey is scheduled to begin adding a mail/internet component in 2016. As previously noted, there are still unresolved issues about the mode effect on the measurement of change, although it is now planned to only estimate change, not level, and to only measure change within each survey mode.

The mail survey experiments will also include a component that will allow respondents to complete the interview on the internet. This option is even more cost effective since the data entry is completed by the respondent. On most metrics of quality including response rates, internet surveys were comparable to mail surveys. Looking ahead, however, the rapid pace of change in computing devices available to households may ultimately result in a vast decline in screen size. There has already been falling sales of desktops and laptops, with rising sales of tablets and smart phones. The small screen size limits the size of question and answer categories, and more page turns increases the effective length of the interview. The same issues are also present in mail interviews, although not to the same extent. Experiments are now planned that would divide the overall interview into a few overlapping sections as well as conduct multiple mailings to the same person each month.

Finally, the potential use of “big data” as a substitute for probability sampling deserves mention. The Michigan surveys are just starting a project to investigate the usefulness of big data to estimate trends in consumer confidence. First, some definitions. The term “big data” is used here to signify data resources that are primarily collected to serve other interests, ranging from digital records of social media posts to transactions automatically recorded by companies. The data are generally based on self-selected samples, with any single individual present in proportion to the frequency they, for example, post tweets or engage in certain transactions. The digital data may be textual, visual, or verbal. The data is usually scanned to determine the frequency of some targeted phrase, amount, or other indicator of interest. Some researchers have used these techniques to estimate an index of consumer confidence, and have demonstrated its usefulness by its correlation with established indexes.

Most analysts are well aware of the drawbacks of this approach, and equally as important, most analysts are convinced that these new approaches will dominate the future. Perhaps the most important drawback of the use of big data is that it rules out the use of deductive logic and standard statistical tests of hypotheses. Moreover, the data does not usually include the socio-economic characteristics of sample members that allow a more detailed analysis of trends, distributions of benefits and negative outcomes. While information on the socio-economic characteristics may improve over time, big data is passive as it can never actively investigate any topic before it becomes a stable of big data. Despite these drawbacks, interesting information may be garnered from big data sources.

The experiments planned by the Michigan surveys involve two inter-related topics. Which big data source has the best coverage to measure consumer sentiment, and how should consumer sentiment be estimated using big data. Recent attempts have used tweets to measure the occurrence of some key words, such as jobs, incomes, and prices, along with whether they were used in a positive or negative context. One recent study claimed success by simply scanning for jobs and incomes based on tweets (O'Connor et al., 2010). An attempt that also claimed success dates back several decades, and is based on scans of a broad sample of newspapers and used a much more complete set of key words (Fan, 1988, 2010). These attempts shared the common approach of measuring the relevant elements that constitute consumer optimism or pessimism. Presumably, as economic conditions and events change, they would simply adjust the set of keywords.

Another approach would be to estimate not the elements that determine confidence, but the outcomes in terms of expressed optimism or pessimism. This approach has the advantage of not having to forecast the appropriate keywords for new economic events, and it could also be based on an indirect measure of confidence. Initial work on coding differences in sentiment from images on the social web has already been undertaken (Siersdorfer, 2010). Perhaps the most creative application of this principle was when the most popular songs in the U.S. were coded for expressions of optimism and pessimism, and the resulting index exhibited the same trends as the Michigan Sentiment Index (Zullo, 1991). At the time this research was conducted, it had the additional behavioral property that it reflected an actual market transaction. For this same reason, economists may favor big data that reflects actual transactions rather than the more transient hot media items.

Overall, the methodology used to measure trends in the economic expectations of consumers is in a state of flux. Rapid changes in technology and social media are unlikely to slow anytime soon. It is hard to imagine that in 10 or 20 years the measurement methodology will not have been transformed in some as yet undermined ways. The fast pace of technological and social change must be confronted with a robust research methodology that preserves a consistent series of high quality measurements on economic expectations. Even the American Association for Public Opinion Research sees its future as moving beyond the reliance on traditional survey methods as the primary means for collecting valid data (Link, 2015).

Questionnaire Content

When George Katona first developed the consumer confidence measure in the late 1940's, he did so as a means to directly incorporate empirical measures of expectations into models of spending and saving behavior. Katona is famously known for his dictum that spending depends on both the "ability and willingness" of consumers to buy. By spending, he meant discretionary purchases; by ability, he meant the current income of consumers; and by willingness, he meant consumers' assessments of their future job and income prospects, including the certainty or uncertainty consumers attached to those expectations. Katona hypothesized that spending would increase when people became optimistic, and precautionary saving would rise when they became pessimistic. This straightforward formulation proved to be a useful predictor of the future course

of the macroeconomy and has been adopted by more than six dozen countries.

In the past half century the social, demographic, and economic environment has changed considerably. Whereas in the 1940's and 1950's consumers viewed all aspects of the economy through the single dimension of how it affected income prospects; in the advanced economies of the world, the focus of consumers has broadened considerably to include many other economic factors that affect consumer's wealth and entitlements, including the details of fiscal and monetary policies. Importantly, it has been the interaction of social, demographic, and economic changes that has accelerated the overall pace of change. For example, the baby boom that followed World War II synchronized movements through the economic life cycle for an unusually large proportion of the population, a cohort that acquired more formal education and economic sophistication than any prior generation, and are now leaving the labor force and entering retirement. It should be no surprise that the economic concerns of this generation will shift from labor market uncertainty to risks originating in financial markets—and to the more volatile worldwide financial markets. In addition, it has been theorized that once sufficient wealth is accumulated during preretirement years, there is no need for additional precautionary savings. Precautionary motives, however, will not disappear in retirement. Uncertainty about the future course of the economy will prompt older consumers to engage in precautionary decreases in spending. Finally, expectations about potential changes in the provisions of private and public health and retirement programs will increase in importance.

The Michigan surveys have expanded its questionnaire to cover a much broader list of topics (see Table 1 for details). Detailed questions on household wealth, its composition, assessments of its sufficiency as well as the adequacy of public and private pensions have been added in the past decade. In addition, several questions have focused on longer term horizons, expected changes over the next five to ten years instead of only asking about expected changes during the year ahead. Initial experiments are now underway to test if a new confidence index can be constructed whose predictions are more accurate than the current confidence index, as well as whether different sub-indexes can be constructed that aim to predict subsets of total consumption or for demographic subgroups rather than the entire population. Needless to say, a sufficiently long time series of measurements is needed to support a robust analysis.

The Michigan surveys have always used a range of response scales, including the verbal likelihood scale, percentage, and dollar amounts, and more recently, probability response scales. The choice is determined by response error considerations as well as respondent burden. As these are familiar issues, no further comments will be made, with the sole exception of the probability response scale. It should be no surprise that probability response scales that range from 0 to 100 allow more precise answers than a three or five point verbal likelihood scale. The increased response variations provide the potential for better fitting models, but this advantage only accrues to a static cross-section analysis and not to time-series models. Indeed, it has been shown that the time-series trends are virtually identical in questions that only differ in whether the response scales were presented in terms of probabilities or verbal likelihood scales (Curtin, 2007).

Table 1: University of Michigan Survey Measures

Question	Response Scale
Summary Indices	
1 Index of Consumer Sentiment	Index
2 Index of Current Economic Conditions	Index
3 Index of Consumer Expectations	Index
Personal Finances	
4 Current Financial Situation Compared with a Year Ago	Verbal Likelihood
5 Selected Reasons for Opinions About Household Financial Situation	Free Response
6 Expected Change in Financial Situation in a Year	Verbal Likelihood
7 Current Financial Situation Compared with 5 Years Ago	Verbal Likelihood
8 Expected Change in Financial Situation in 5 Years	Verbal Likelihood
9 Expected Change in Household Income During the Next Year	Verbal Likelihood & Percent
10 Expected Change in Real Household Income During the Next Year	Verbal Likelihood
11 Probability that Personal Income will Increase During the Next Year	Probability
12 Probability of Real Income Gains During the Next 5 Years	Probability
13 Probability of Losing a Job During the Next 5 Years	Probability
14 Total Household Income	Dollar Amount
Savings and Retirement	
15 Probability that Social Security/Pensions Provide Adequate Retirement Income	Probability
16 Change in Likelihood of a Comfortable Retirement Compared with 5 Years Ago	Verbal Likelihood
17 Probability of Increase in the Stock Market Prices in the Next Year	Probability
18 Current Value of Stock Market Investments	Dollar Amount
19 Current Market Value of Primary Residence	Dollar Amount
Economic Conditions	
20 News Heard of Recent Changes in Business Conditions	Verbal Likelihood
21 Selected Items of News Heard of Recent Changes in Business Conditions	Free Response
22 Current Business Conditions Compared with a Year Ago	Verbal Likelihood
23 Expected Change in Business Conditions in a Year	Verbal Likelihood
24 Business Conditions Expected During the Next Year	Verbal Likelihood
25 Business Conditions Expected During the Next 5 Years	Verbal Likelihood
Unemployment, Interest Rates, Prices, Government Expectations	
26 Expected Change in Unemployment During the Next Year	Verbal Likelihood
27 Expected Change in Interest Rates During the Next Year	Verbal Likelihood
28 Expected Change in Prices During the Next Year	Verbal Likelihood & Percent
29 Expected Change in Prices During the Next 5 Years	Verbal Likelihood & Percent
30 Opinions About the Government's Economic Policy	Verbal Likelihood
Household Durables Buying Conditions	
31 Buying Conditions for Large Household Durables	Verbal Likelihood
32 Selected Reasons for Buying Conditions for Large Household Durables	Free Response
Vehicle Buying Conditions	
33 Buying Conditions for Vehicles	Verbal Likelihood
34 Selected Reasons for Opinions About Buying Conditions for Vehicles	Free Response
35 Expected Change in Gasoline Prices During the Next Year	Percentage
36 Expected Change in Gasoline Prices During the Next 5 Years	Percentage
Home Buying and Selling Conditions	
37 Buying Conditions for Houses	Verbal Likelihood
38 Selected Reasons for Opinions About Buying Conditions for Houses	Free Response
39 Selling Conditions for Houses	Verbal Likelihood
40 Selected Reasons for Opinions About Selling Conditions for Houses	Free Response
41 Change in Home Values During the Past Year	Verbal Likelihood
42 Expected Change in Home Values During the Next Year	Verbal Likelihood & Percent
43 Expected Change in Home Values During the Next 5 Years	Verbal Likelihood & Percent

Consumer Demand in a New Economic Era

The methodological issues about sampling, modes of interviewing, or questionnaire content are not independent of the primary purpose of the surveys, which is to understand how changes in expectations influence macroeconomic cycles. One of the most interesting challenges that lie ahead is how changing social, demographic, and economic conditions have influenced and will continue to influence the analysis of the relationship between subjective expectations and objective economic conditions. While short term economic expectations have always been the workhorse of cyclical economic analysis, diminished long term expectations may have contributed to concerns about secular stagnation. The term secular stagnation was first used by Alvin Hansen in 1939 to describe the negative impact on the macroeconomy from what he incorrectly anticipated would be lower population growth in the decades ahead. More recently, the same concept has been used to describe the slower pace of growth in the economy following the financial crisis in 2008. The theory behind secular stagnation has typically been conceptualized as a supply side issue: it is due to a reduced rate of growth in productive capital, labor, and technology. The causes and cures of secular stagnation are typically tied to the cost of capital and the expected return on productive investments. The best policy response based on this thesis has been to lower nominal interest rates to zero, and even below zero, to spark renewed growth in productive investments.

A slowdown in the pace of demand growth is rarely considered a critical determinant of secular stagnation, although it has been mentioned in passing by Keynesian economists. The dismissal of demand has been primarily associated with the strongly held belief that consumers always want more, with incomes being the only limitation on their spending. That view fails to take account of opportunity costs of spending, especially the substantial risks to their future financial security from too much spending. Prior to the Great Recession, debt financing of material wants was as commonplace as increased household labor force participation to support the higher incurred expenses. The economy could be relied on to continuously advance living standards and wages, with the government as a backstop to correct cyclical imbalances. The lasting lesson of the Great Recession, however, was that when economic disaster struck, households paid a heavy price for the loans they could no longer afford with weakened wage and jobs prospects. The government focused on repairing the balance sheets of the very biggest, while individual households were too small, and too many, to directly provide effective relief. As a result, consumers became more sensitive to the potential risks that accompanied their economic decisions. It was this heightened risk aversion that doomed the power of traditional counter-cyclical policies aimed at stimulating demand. Consumers wanted to reduce their indebtedness and rebalance their finances, not to step up their spending and incur additional debt. The resulting slow recovery could be more accurately described as insufficient demand by consumers rather than insufficient investments by businesses.

This description of the subpar performance of the economy is hardly new. What may be more surprising is how it has influenced consumers' judgments about the economy. The conventional hypothesis is that it would lower consumer optimism. It surely did lower consumer sentiment during several years following the start of the Great Recession, but in the past several years consumer sentiment has regained levels that have been traditionally associated with high levels

of consumer expenditures and economic growth. To be sure, consumer spending has been the primary factor behind the improved performance of the U.S. economy. Nonetheless, the average level of economic growth has remained about one percentage point below its historical average. Just as important, the labor force participation has been reduced to the lowest levels in decades, which can only partially be attributed to the increased pace of retiring baby-boomers.

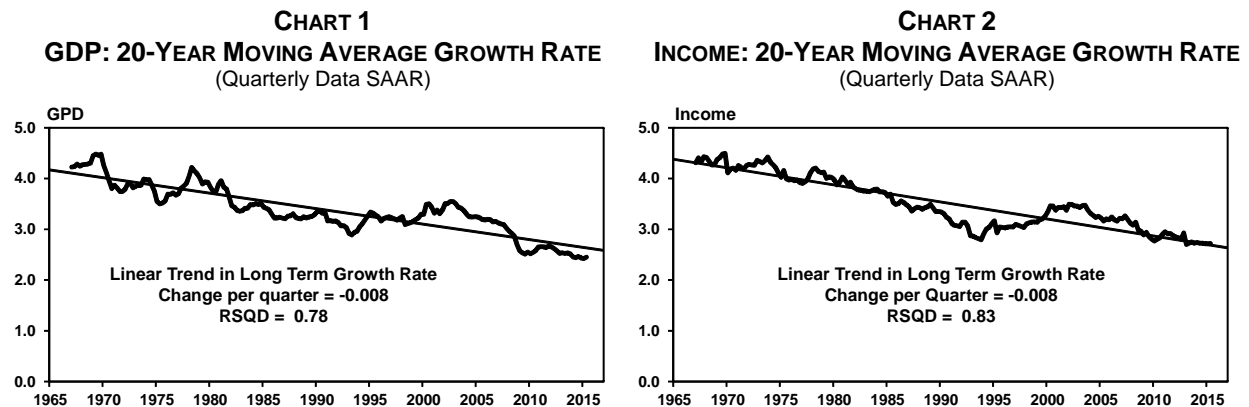
The basic premise that will be advanced is centered on two developments. The first is that a slowdown in the pace of economic growth has persisted for decades. A slower pace of growth in incomes had been offset by rising levels of household debt, in part supported by higher home and stock values as well as other increases in household wealth. The second development, the financial crises, caused the sudden bursting of the debt bubble. The lasting lesson for consumers was an elevated appreciation of economic risk. Economic losses not only had severe financial consequences, they also had a devastating psychological impact. A well-known research finding in behavioral economics is the asymmetry between losses and gains (Kahneman and Tversky, 1979). Loss of what one once had can do much more psychological damage than never having had it in the first place. The appreciation of higher economic risks meant that people now exercise new limitations on their economic decisions. To be sure, the change has not resulted in dramatic turnaround in the consumer culture that defines the U.S. Rather, the shift toward greater risk averse behavior operates at the margin for individual households. It is at the aggregate level that the slightly greater precautionary behavior cumulates to represent a slowdown in the pace of growth in the macro economy.

From countless examples over time and across the world, people quickly adapt to the best and worst economic circumstances. It would be unreasonable to believe that people's reactions to secular stagnation would be to permanently adopt unfavorable economic evaluations. People use a relative standard to judge economic prospects. To steadfastly refuse to adjust to new economic conditions would be irrational. The final element of this thesis is that the renewed appreciation of economic risks meant that people could no longer ignore the longstanding declines in the pace of economic growth. They were forced to adopt a new and lower standard to judge current economic prospects. With this transition, consumers voiced more optimism about slower economic growth than they would have in the past. Economists should not be surprised by such a shift since they have used the term "secular growth" to describe the same concept, which they have lowered to take account of the "new normal."

Performance Benchmarks. Accurate reference standards or economic benchmarks are needed for people to make robust assessments of the current performance of the economy. The rationality hypothesis suggests that the most appropriate benchmarks should be equal to the actual long term growth prospects for the economy. How long a term should be considered? The term would have to include at least one complete economic cycle, but ideally the term should extend over multiple cycles. That is the approach that is favored by economists. Ordinary people might favor a time period that corresponds with life events, such as their early or late career when the pace of economic growth could have a substantial impact on their economic lives. From either perspective, the time period could be defined anywhere from 10 to 15, or 20 years. Since the theory holds that expectations for long term growth should mirror actual economic developments, actual long term

averages can be used as proxies.

The data indicate a clear downward trend in the pace of GDP growth over the past half century. The quarterly inflation and seasonally adjusted annual rate of change in U.S. GDP was used to calculate 10, 15, and 20 year moving averages. The data ranged from 1947 to mid 2015, although the moving averages began in 1957 to 1967. The estimated downward slope in the growth rate was nearly identical, falling by 0.008 percentage points per quarter for the 10, 15, and 20 year moving averages. Over ten years, this would amount to a decline of a third of a percentage point in the average GDP growth rate. From 1967 to 2015, the decline in average trend in GDP totaled 1.5 percentage points. Of course, variations around the trendline were smaller the longer the moving average period. The data for the 20 year moving average is shown in Chart 1. While the chart suggests a rather continuous slow decline in the performance of the economy, when tested, the trend since 2007 was significantly lower. Also note that the debt boom is clearly reflected by higher levels of average GDP growth in the 2000's. Overall, the major finding is that the long term trend in the real growth rate of GDP has steadily declined over the past half century.

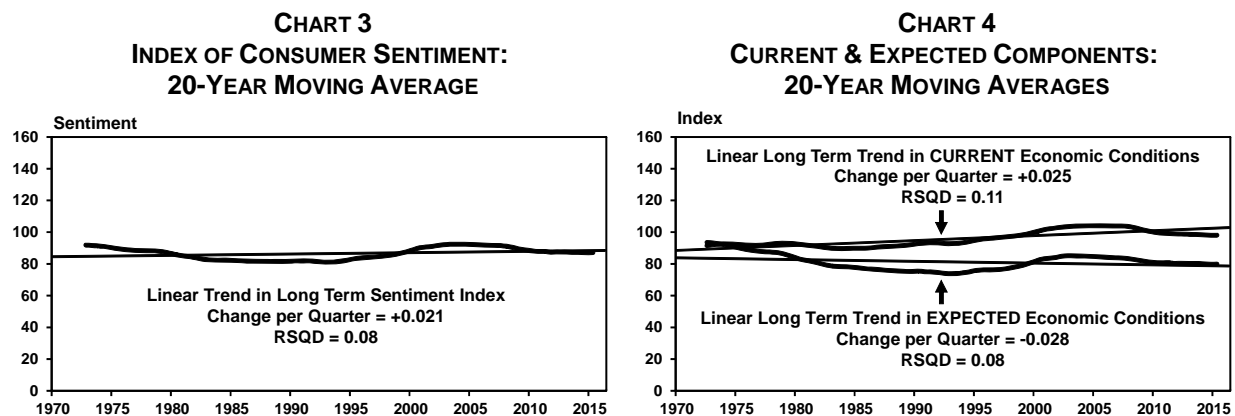


Personal income is another major indicator of the health of the economy that attracts widespread attention. The long term growth rate in personal disposable income corrected for inflation and seasonal variations is shown in Chart 2. Not surprisingly, the same rate of decline per quarter was estimated for long term trends in personal incomes as for GDP, despite the fact that the overall level of real income growth was slightly higher on average than for GDP. When total income was replaced by real disposable income per capita, the same results were obtained. The same finding was also true for real personal consumption expenditures, not surprisingly since consumption accounts for about 70% of GDP. The only notable difference was that the quarterly decline in the 20-year average growth rate in consumption expenditures was slightly lower, at 0.006. This reflected a long term decline in the savings rate. The means of the 20-year moving averages were 3.4% for GDP, 3.5% for income, and 3.6% for consumption.

Relative reference standards implies that people would have rationally incorporated the downward sloping long term trends in GDP and personal income, for example, into their reference standards, leaving their average assessments essentially trendless. If absolute rather than relative

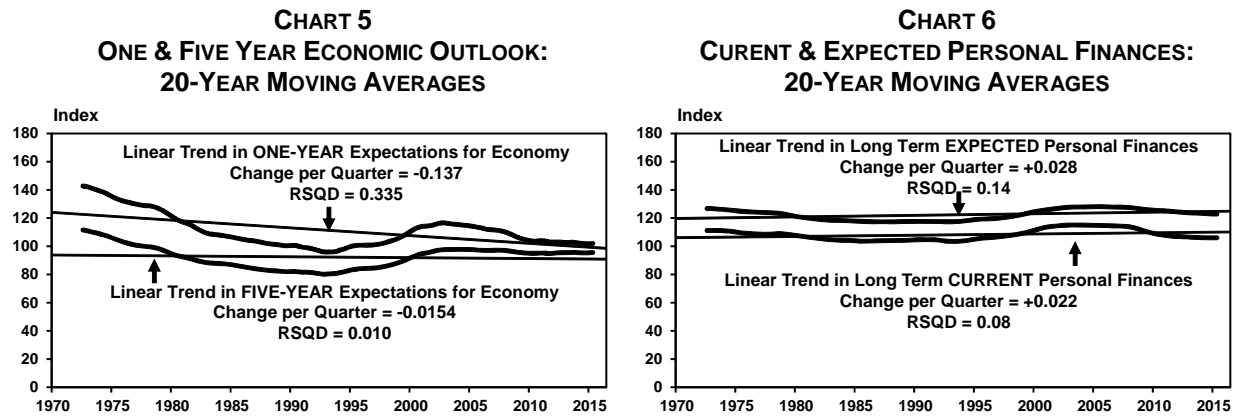
standards were used, it would suggest that declines in people's economic assessments should match the secular declines in the growth rates of GDP and personal incomes. Katona believed that reference standards changed relative to performance of the economy even over much shorter periods of time (Katona, 1959).

The Index of Consumer Sentiment is the most widely known and used measure from the University of Michigan's survey. The Sentiment Index was first estimated in 1952, with the trend in the 20 year moving average shown in Chart 3. While there is a small degree of variation about the trendline, the slope of that line is virtually zero. Over ten years, the average Index reading changed by just 0.8 Index-points, and over the entire period from 1972 to 2015, the decline in the trend amounted to only 2.5 Index-points, or less than 3% of the series mean. This was a tiny increase compared with the decline in the trend growth rate in GDP, which fell by 1.3 percentage points or by nearly 40% of the mean over the same time period. The two major components of the Sentiment Index, the Current Conditions and Expected Conditions, show divergent long term trends (Chart 4). The Current Conditions component shows a slight upward trend (+0.025) and the Expectation Index shows a slight downward trend (-0.028). The R-squared statistic provides a convenient way to compare the importance of the time trend in the variables since it indicates the proportion of variance that is "explained" by the time trend. The R-squared for GDP and personal income were relatively high, with the time trend explaining 78% and 83% of the change in the long term moving averages. In comparison, the trend variable explained just 8% for the Sentiment Index and the Expectations Index, rising to only 11% for the Current Economic Conditions Index.



Since the Indices include several questions on personal finances, overall economic conditions, and purchase conditions, the second set of charts focuses on trends in two questions on the outlook for the economy in the year ahead and next five years (Chart 5) and on current and expected personal finances (Chart 6). The estimated trends in personal financial assessments are quite flat, especially when compared with the pronounced trend in personal income. The same is true for assessments of the economic outlook over the next five years, with the trend variable accounting for just 1% of the variance. Assessments of the outlook for the year ahead were the clear exception, with an estimated negative time trend (-0.137) than accounted for 34% of the variance. Interestingly, the decline was consistent with the overall decline in GDP, although the trend rate of

decline in GDP explained a much higher proportion of variance.



Diminished Material Aspirations

George Katona believed that changes in aspirations had significantly influenced Americans' willingness to incur debt to achieve their consumption aspirations as well as to make additional investments in human capital and increase their participation in the workforce. These actions created substantial expansions in markets for a wide variety of goods and services that independently added to the pace of domestic economic growth. Moreover, the strong rise in material aspirations also increased the willingness of Americans to step-up their labor force participation rates, which also acted to increase the pace of economic growth.

Katona considered aspirations as realistic goals people set for themselves, not as unrealistic hopes or dreams. The aspirations that provide the strongest behavioral motives are those that are only modestly different from recent accomplishments. Aspirations are not static, but continually change in response to accomplishment and failure. Importantly, there is an asymmetric dynamic to changes in aspirations: fulfillment quickly gives rise to new aspirations, but failure does not immediately result in diminished aspirations. Failure initially sparks renewed efforts toward attainment. No one easily nor quickly gives up their aspirations. Aspirations are finally reduced only after prolonged frustration and failure. Declines in aspirations not only indicate that people judge the probability of failure higher than the probability of success, but that an unchanged aspiration will result in net losses in utility since maintaining those aspirations would misdirect behavioral decisions.

The impact of debt on consumer purchases can be estimated by the net change in the quarterly outstanding household debt as a percentage of personal income from the Federal Reserve (see Chart 7). This figure would be an estimate of how much debt expanded the spending power of personal disposable income. High material aspirations which were accompanied by an optimistic economic outlook would be associated with larger debt increases, and lower aspirations with smaller increases or even debt reductions. Note that this measure counts the net debt addition in the quarter that the debt is incurred, while the repayment of that debt occurs over the various terms of

the loans. Over most of the past half century, net debt additions as a percent of income varied between 0.5% and 2.5%, with the exception of the recent peaks and troughs. Its recent peak was nearly 3.5% in

Chart 7
Quarterly Change in Household Debt
As a Percentage of Personal Income
 (Four-quarter moving averages)

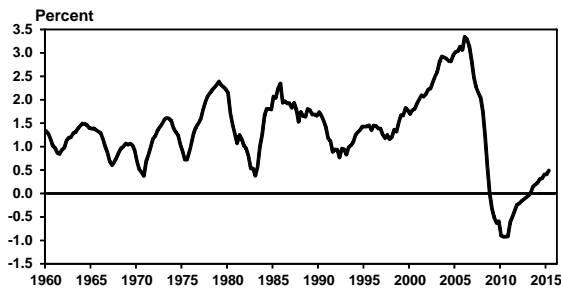
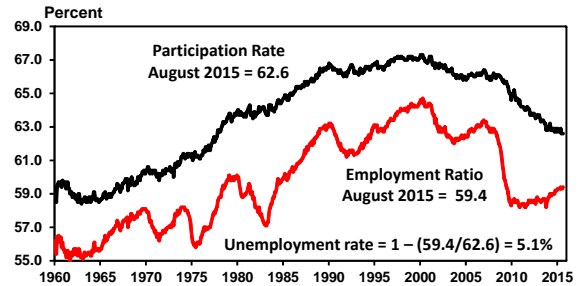


Chart 8
Labor Force Participation and
Employment Rates
 (Monthly data)



2006 and its trough was a negative 1.0% of income. Positive additions to debt decrease household savings just as negative amounts increase saving. Even though debt incurrence had begun to rebound by mid-2015, its improved level was more comparable to the lows over the prior half century. The interpretation of this decline is not as yet straightforward. Those that favor viewing secular stagnation as due to supply considerations suggest that despite the low interest rates there was a substantial tightening of loan eligibility rates as well as the need for households to pay down excessive debt. There is no doubt that these factors accounted for a share of both the decline as well as the muted rebound. Nonetheless, rational consumers adapt to their economic circumstance, as has been emphasized in this presentation. This would imply that the reduction in debt incurrence will persist for some time.

Labor force participation rates also provide similar evidence of declining aspirations. The unemployment rate can be decomposed into a ratio of two components: the percent employed and the percent in the labor force. The decline in U.S. unemployment rate has been largely due to a drop in labor force participation rather than to increases in employment (see Chart 8). The labor force participation rate, as well as the employment rate, is the lowest in a quarter-century. While the declines are partly due to retiring baby-boomers, and partly due to extended schooling for the youngest, these factors still leave a large share of the decline unexplained. The most troubling declines in participation rates have occurred among those aged 25 to 35, since these early years are critical for the development of work skills. Even with the gains in employment over the past few years, the proportion employed is still at a quarter-century low. While the atrophy of work skills among the long term unemployed is a much debated topic in economics, there has been insufficient attention to diminishing career aspirations among both the employed as well as the unemployed. These diminished aspirations are equivalent to a decline in consumers' estimates of their permanent income. This acts to diminish consumer demand and increase prospects for secular stagnation.

Katona's theory of aspirations has been largely ignored in the economic literature as irrelevant since material aspirations were always expected to increase. To be sure, there are always variations in aspirations across a population, with some people reducing their aspirations, and others increasing their aspirations. No one could imagine a coordinated reduction in material aspirations across the entire population that could have a significant impact on economic growth. The thesis of secular stagnation caused by declines in aggregate demand challenged that view. Katona would have suggested that the weakness in spending as well as labor force participation, aside from an aging population, was related to reductions in material aspirations. Aspirations that had been reduced due to reversals in income and wealth as well as lessened prospects for renewed personal financial advancement due to rising inequality.

Concluding Remarks

This presentation has made four basic points. The first was that all economic evaluations are relative and are based on rational assessments of economic conditions. Everyone here already knew that fact. The second was that these relative evaluations can influence economic behavior in a manner largely unexpected by conventional theory. Declines in aspirations can diminish the overall rate of economic growth by reducing wants as well as by reducing the contributions of workers to the economy. As a result, variations in both demand and supply can play critical roles in creating secular stagnation in advanced economies. The third point is that the interpretation of current trends in sentiment must take into account the influence of longer term expectations. The relativity of evaluations means that identical economic conditions may not yield identical evaluations, either over time or across countries. Analysis must be sensitive to melding the absolute scales of economic data to the relative scales of sentiment data.

The final objective represents the most important challenge facing sentiment surveys: the development of new statistically-robust substitutes for current sampling and data collection methods. Conventional survey research methods are unlikely to survive the rapid decline in people's willingness to participate in traditional surveys. The Michigan surveys are attempting to delay that final reckoning by employing multimode surveys based on probability samples to attain the highest possible response rates. These adjustments, however, can only be temporary stopgaps. More research is needed on non-probability samples, on statistical inference, on new data collection technologies, and on the development of indirect measures rather than direct replies to sentiment questions. Clinging to probability samples and traditional interviewing methods is no longer an option. Our only choice is to be proactive in the development of new robust methodologies.

References

- Blumberg, S. J., and Luke, J. V. (2015). "Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July–December 2014." U.S. Department of Health and Human Services, National Center for Health Statistics.
- Baker, R., Brick, J.M., Bates, N., Battaglia, M., Couper, M. Dever, J., Gile, K., Tourangeau, R. (2013). "Summary report of the AAPOR task force on non-probability sampling." *Journal of Survey Statistics and Methodology*, 1 (2): 90-143.

- Curtin, R.T., Presser, S., and Singer, E. (2005). "Changes in telephone survey nonresponse over the past quarter century." *Public Opinion Quarterly*, 69 (1): 87-98.
- Curtin, R.T., Singer, E., and Presser, S. (2007). "Incentives in random digit dial telephone surveys: a replication and extension." *Journal of Official Statistics*, 23 (1): 91-105.
- Curtin, R. T. (2007). "Consumer Sentiment Surveys: Worldwide Review and Assessment." *Journal of Business Cycle Measurement and Analysis* 3(1): 9-45.
- Dillman, D.A., and Christian, L.M. (2005). "Survey mode as a source of instability in responses across surveys." *Field Methods*, 17 (1): 30-52.
- Dillman, D. A., G. Phelps, R. Tortora, K. Swift, J. Kohrell, J. Berck and B. L. Messer (2009). "Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet." *Social Science Research* 38(1): 1-18.
- Elkasabi, M., Suzer-Gurtekin, Z.T., Lepkowski, J.M., Kim, U., Curtin, R.T., and McBee, R. (2014). "A comparison of ABS mail and RDD surveys for measuring consumer attitudes." *International Journal of Market Research*, 56:4.
- Fan, D. P. (1988). Predictions of public opinion from the mass media: Computer content analysis and mathematical modeling. Greenwood Press, Westport, CT.
- Fan, D. P. (2010). "Predicting the index of consumer sentiment when it isn't measured." Joint Statistical Meeting Proceedings, pp. 6098–6110.
- Heerwegh, D. and G. Loosveldt (2011). "Assessing mode effects in a national crime victimization survey using structural equation models: social desirability bias and acquiescence." *Journal of Official Statistics* 27(1): 49.
- Jäckle, A., Roberts, C., and Lynn, P. (2010). "Assessing the effect of data collection mode on measurement." *International Statistical Review*, 78 (1): 3-20.
- Jiang, L., J. Lepkowski, T. Suzer-Gurtekin, M. Sadowsky, R. Curtin, R. McBee and D. Zahs (2015). "Transition from Landline-Cell to Cell Frame Design: Surveys of Consumers." AAPOR Annual Conference, Hollywood, FL.
- Kahneman, D. and Tversky, A., (1979). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2): 263-292.
- Kreuter, F., S. Presser and R. Tourangeau (2008). "Social desirability bias in CATI, IVR, and Web surveys: The effects of mode and question sensitivity." *Public Opinion Quarterly* 72(5): 847-865.
- Link, M. W., and Mokdad, A.H. (2006). "Can web and mail survey modes improve participation in an RDD-based national health surveillance?" *Journal of Official Statistics*, 22 (2): 293-312.
- Link, M. W., Battaglia, M.P., Frankel, M.R., Osborn, L., and Mokdad, A.H. (2009). "A comparison of address-based sampling (ABS) versus random-digit dialing for general population surveys." *Public Opinion Quarterly*, 72 (1): 6-27.
- Link, M. (2015) "Presidential Address: AAPOR2025 and the Opportunities in the Decade Before US." *Public Opinion Quarterly*, 79 (3): 828-836.
- Millar, M.M., and Dillman, D.A. (2011). "Improving response to web and mixed-mode surveys." *Public Opinion Quarterly*, 75 (2): 249-269.
- O'Connor, B., Balasubramanian, R., Routledge, B., and Smith, N. (2010). "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series." Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, pp. 122-129.
- Olson, K., Smyth, J.D., and Wood, H. (2012). "Does giving people their preferred survey mode actually increase survey participation rates? An experimental examination." *Public Opinion Quarterly*, 76 (4): 611-635.
- Peytchev et al. (2010). "Differences between RDD telephone and ABS mail survey design: Coverage, unit nonresponse, and measurement error." *Journal of Health Communication: International Perspectives* 15(2)
- Siersdorfer, S., Minack, E., Deng, F. and Hare, J., (2010). "Analyzing and Predicting Sentiment of Images on the Social Web," Proceedings of the international conference on Multimedia, pp. 715-718.
- Voogt, R. J. J. and W. E. Saris (2005). "Mixed mode designs: Finding the balance between nonresponse bias and mode effects." *Journal of Official Statistics* 21(3): 367-387.
- Zullo, H. M. (1991). "Pessimistic rumination in popular songs and news magazines predict economic recession via decreased consumer optimism and spending." *Journal of Economic Psychology*, 12 (3): 501-526