Analysis of Indexes of Consumer Sentiment

Jirka Slacalek
DIW Berlin
jslacalek@diw.de

***Abstract*— I examine the relations between individual questions on the Michigan Survey of Consumer Sentiment and their impact on consumption growth. I find (1) the Survey information is captured well by three common factors, one highly correlated with the Michigan Index of Consumer Sentiment, (2) future consumption growth is sensitive to the individual questions and common factors, (3) many of the individual questions and common factors are stable predictors consumption growth over all periods of time considered, and (4), while the common factors typically dominate the individual questions in predictive power, there are questions, such as that on expected unemployment, that surpass even the factors. The econometric contribution is an application of the bootstrap method to estimate uncertainty about the common factors. This uncertainty is substantially higher than is implied by existing asymptotic confidence intervals.**

***Index Terms*— consumer sentiment, factor analysis, consumption.**

I. Introduction

Businessmen, forecasters and policy-makers keenly await the release of new figures on consumer sentiment because these data contain valuable information about the future evolution of the economy and the stock market. Recently, academic researchers have joined the crowd and begun to investigate the relationships between sentiment and other variables (see Ludvigson, 2004 for a review). Typically, however, the literature analyzes the properties of aggregate indexes of consumer sentiment, not the original disaggregated survey questions, that respondents are asked. This paper attempts to fill this gap and examine the role of individual components of the survey. I apply the methods of dynamic factor analysis to extract common factors from the Michigan Survey of Consumer Sentiment and study the relationships between the common factors, individual questions, aggregate sentiment indexes and consumption growth.

The paper combines insights from two strands of recent economic research. First, I build on work that analyzes indexes of consumer sentiment and their relationship to future consumption growth. Starting with Acemoglu and Scott (1992) and Carroll et al. (1994), economists realized that sentiment provides yet another example of economic variable (in addition to stock returns and disposable income) to which consumption growth is "excessively" sensitive. Sommer (2002) provides a possible explanation for this finding, which relies on a combination of habit formation and measurement error in consumption. In that case, Sommer argues, the lagged sentiment is just a proxy for the "true" lagged consumption growth (that appears in the Euler equation due to habits). The literature on out-of-sample predictive power of sentiment was pioneered by Bram and Ludvigson (1998) and Slacalek (2004), who document that consumer sentiment can be used in real time to improve forecasts of consumption growth. Ludvigson (2004) provides a recent review of the current state of this literature.

Second, there is a rapidly growing econometric literature on extracting information from datasets with a large number of series. The literature assumes that the information in the datasets can be summarized by a small number of underlying common factors. These factors can then be used for forecasting (see for example Forni and Reichlin, 1998; Stock and Watson, 2002 and Boivin and Ng, 2005), structural modelling (see Bernanke and Boivin, 2003 for an application to forward-looking Taylor rules and Forni et al., 2004 to larger-scale structural models) or other purposes. This work now provides well-developed estimation and inference methods of the dynamic factor analysis. See Stock and Watson (2005) for a summary of this literature.

This paper uses the techniques of the dynamic factor analysis to examine the properties of the individual questions in the Michigan Survey of Consumer Sentiment. I believe this is the first paper that investigates the relationships between the *disaggregated* sentiment ques-

First version: December 2004; this version: February 4, 2006.
Jirka Slacalek, German Institute for Economic Research, DIW Berlin, Department of Macro Analysis and Forecasting, Königin-Luise-Strasse 5, 14195 Berlin, Germany.I thank Christopher Carroll, Ulrich Fritsche and Jonathan Wright for helpful suggestions and comments. I also thank seminar audiences at Freie Universität, Berlin and Verein für Socialpolitik 2005 Annual Meeting. Replication files are available at http://www.slacalek.com/.

tions, common factors and consumption growth. I report several findings. First, the common factors summarize very well the information in the Survey; the median R^2 of 94 regressions of the disaggregated survey answers on the first three factors is 0.75. Second, future consumption growth is excessively sensitive to the individual questions and the common factors. Third, many individual questions and the common factors are stable predictors consumption growth over all periods of time considered. Fourth, while the common factors typically dominate the individual questions in the predictive power, there are questions, in particular the question on expected unemployment (question 12), that beat even the factors. Fifth, the factors the estimates of the common factors are not sensitive to various specifications and estimation techniques. In addition, I propose a bootstrap method to evaluate uncertainty about the estimated common factors. I document that in this application uncertainty implied by the bootstrap is much higher than is implied by the most common alternative method of Bai (2003).

The plan of the paper is as follows. Section II provides a brief description of the construction of the survey questions, a preliminary correlation analysis of the questions and an assessment of their suitability for factor analysis. Section III summarizes the main econometric results and methods of the dynamic factor analysis. Section IV brings the main findings on the relationship between the individual questions and the common factors for the baseline model. Section V analyzes the link between the questions and future consumption and its stability. Section VI reports alternative estimates of the common factors. Finally, Section VII concludes. Appendixes provide detailed descriptions of the questions in the Michigan Survey and the bootstrap procedure for evaluating uncertainty about the common factors.

II. Preliminary Analysis

A. Construction of Michigan Indexes of Consumer Sentiment

This section discusses the construction of the Michigan Index of Consumer Sentiment, which is published by the Survey Research Center at the University of Michigan.[1] Every month the Center asks approximately 500 respondents to answer 21 questions. The questions are concerned with current and expected personal financial situation and current and expected overall economic conditions.[2]

Most questions ask the respondent to pick from three broad choices. For example, question 1 asks: "…Would you say that you …are better off or worse off financially than you were a year ago?" The respondent then chooses the answer from: Better off, Same, Worse off (or declines to answer). Given the proportion of respondents with each answer, the response is summarized as the *relative score* (balance statistic), or the proportion giving favorable responses minus the proportion giving unfavorable responses plus 100. The Survey Research Center constructs three indexes of consumer sentiment by averaging the relative scores of the following five questions:

- Q1 "…Would you say that you …are *better off* or *worse off* financially than you were a year ago?"
- Q3 "…[D]o you think that *a year from now* you …will be *better off* financially, or *worse off*, or just about the same as now?"
- Q10 "Now turning to business conditions in the country as a whole—do you think that during the *next twelve months* we'll have *good* times financially, or *bad* times, or what?"
- Q11 "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?"
- Q16 "…Generally speaking, do you think now is a *good* or *bad* time for people to buy major household items?"

The Center calculates three aggregate indexes of consumer sentiment:

Index of Consumer Sentiment (ICS)
Scaled simple average of relative scores of Q1, Q3, Q10, Q11 and Q16.

Index of Consumer Expectations (ICE)
scaled simple average of relative scores of Q3, Q10 and Q11.

Index of Current Economic Conditions (ICC)
scaled simple average of relative scores of Q1 and Q16.

Thus, only five questions of 21 asked are actually used in the construction of indexes of consumer sentiment.

[1] An alternative measure of consumer sentiment for the United States is released by the Conference Board. The aggregate Michigan and Conference Board indexes are strongly correlated and probably all findings on sentiment hold for both measures. The reason this paper analyzes the Michigan Survey is that the Michigan disaggregated data are more easily available.

[2] The exact wording of all questions is recorded in Survey of Consumers and reproduced in Appendix I.

This paper, in contrast, investigates the relationships among *all* questions in the survey, common factors and consumption growth.

B. Individual Questions

Responses to most of the 21 questions asked in the survey can be summarized with relative scores.[3] As a first step in my analysis of individual sentiment questions I compare correlations between relative scores. The correlation matrix is shown in Table I. The *absolute values* of correlations are displayed in Figures **??** and 1.

While one might a priori suspect that many of the questions will be strongly correlated, it turns out that this correlation is not extremely pervasive. This is documented in Figure 1, where many of cells are quite light, corresponding to relatively low correlations. Specifically, of 136 absolute correlations reported in Table I, the absolute values of 108 are greater than 0.3 and 71 are greater than 0.5. On the other hand the individual questions are designed to capture the respondents' overall feelings about the current and expected states of the economy as a whole and their personal financial situations in particular. This suggests that (i) one might summarize much of the information contained in the individual questions by a small number of factors and (ii) the individual questions are independent enough that they are not going to coincide with the extracted factors.

The next section presents a summary of the principal component analysis, a statistical method that serves to extract common factors from a large dataset.

III. Factor Analysis

A. Dynamic Factor Model

The dynamic factor model (DFM) assumes that the relationship among a large number, N, series $x_{i,\cdot}$, $i = 1, \ldots, N$ can be captured by a few, $\bar{r}$, typically unobservable, underlying factors f:[4]

$$x_{i,t} = \lambda_i(L) f_t + u_{i,t} \qquad i = 1, \ldots, N, \tag{1}$$

where f_t is the $\bar{r} \times 1$ vector of factors, $\lambda_i(L)$ is an $\bar{r} \times 1$ vector lag polynomial, called the "dynamic factor loadings," and $u_{i,t}$ is an idiosyncratic, possibly serially correlated, disturbance. The factors and disturbances are assumed to be uncorrelated at all leads and lags, $\mathbf{E}\big(f_t u_{i,s}\big) = 0$ for all i, s and t.

Following Chamberlain and Rothschild (1983), many authors distinguish between exact and approximate DFMs. The exact DFM postulates that the disturbances $u_{i,t}$ are mutually uncorrelated,

$$\mathbf{E}\big(u_{i,t} u_{j,t}\big) = 0 \qquad \text{for } i \neq j.$$

The approximate DFM allows for a limited amount of correlation among the disturbances. Stock and Watson (2002) require that

$$\lim_{N \to \infty} 1/N \sum_{i=1}^{N} \sum_{j=1}^{N} \big|\mathbf{E}(u_{i,t} u_{j,t})\big| < \infty.$$

Recent econometric research extended the classical methods of principal components to the approximate DFMs and made it possible to apply the factor analysis to large economic datasets, in which the exact DFM structure hardly ever holds.

B. Estimation

There are two ways to estimate the dynamic factor model (1). If N, the number of series in x, is small the standard maximum likelihood and the Kalman filter algorithm can be used. However, in many recent applications (Forni and Reichlin, 1998; Stock and Watson, 2002; Bernanke and Boivin, 2003; Boivin and Ng, 2005 and others) the number of series in x ranges from scores to hundreds. In such cases the exact maximum likelihood is not computationally feasible due to a large number of parameters to be estimated (curse of dimensionality).

An alternative to the exact maximum likelihood is the principal component analysis. Stacking the lags of factors in $F_t = [f_t' \, f_{t-1}', \ \ldots, \ f_{t-p+1}']'$, where p is the (maximum) order of the lag polynomials $\lambda_i(L) = \sum_{j=0}^{p} \lambda_{i,j} L^j$, the DFM (1) can be written as

$$x_t = \Lambda F_t + u_t, \tag{2}$$

where the i^{th} row of Λ is $(\lambda_{i,0}, \ldots, \lambda_{i,p})$. This is the static representation of the dynamic factor model (1), for which the classical statistical method of principal components can be used to estimate. The method solves the nonlinear least squares problem

$$\min_{F_1,\ldots,F_T,\Lambda} 1/T \sum_{t=1}^{T} \big(x_t - \Lambda F_t\big)'\big(x_t - \Lambda F_t\big) \tag{3}$$

subject to $\Lambda'\Lambda = I_r$, where $r \leq (p+1)\bar{r}$. Denoting Σ_{xx} the $N \times N$ variance–covariance matrix x, the solution of this problem is to set $\hat{\Lambda}$ equal to the first r eigenvectors of $\hat{\Sigma}_{xx} = 1/T \sum_{t=1}^{T} x_t x_t'$ and $\hat{F}_t = \hat{\Lambda}' x_t$. Thus F_t is

[3] The exceptions are questions 2, 17, 19 and 21. These questions are about the reasons for respondent's opinion, i.e. "Why do you say that …" and the answers cannot be summarized as "Better off," "Same" or "Worse off."

[4] This material is standard; the exposition in this section mostly follows Stock and Watson (2005) and Stock and Watson (2002).

the vector consisting of the first r principal components of $\hat{\Sigma}_{xx}$. These estimators of factors and factor loadings are consistent and asymptotically normal under suitable conditions (see Bai, 2003).

IV. Individual Questions and Common Factors

Preliminary results of section II-B suggest that the panel of individual questions of the Survey of Consumer Sentiment is a suitable object for more rigorous factor analysis. Before I apply this analysis I want to discuss the specification of my baseline model. Alternative specifications are investigated in section VI below.

Since the methods of factor analysis are especially suited for datasets with many time series (large cross section dimension), I further disaggregate the answers in the survey. Instead of using just the relative scores, i.e. the difference between the fractions "better off" and "worse off" we will include all the possible answers separately. This makes it possible to extract information from:

- All questions, including those that cannot be summarized with simple relative scores statistics (e.g. What favorable news about the changes in business conditions have you heard?).
- The fraction of consumers with the "no change" and "don't know" answers. The relative scores only change when the relative number of "better off" and "worse off" answers changes. These scores will not react to other changes in the answers, such as an increase in the number of "no change" answers that leaves the relative number of "better off" and "worse off" answers unchanged. Thus, in a sense the relative scores only capture a shift in the mean, but not in the variance of the answers.
- Finally, relative scores effectively postulate that the "better off" and "worse off" answers have symmetric effects on the aggregate index of consumer sentiment. Entering the answers separately enables us to allow for possible asymmetric effects on the extracted factors (since the loadings of the "better off" and "worse off" questions are not restricted to have the same absolute values).

Once we include all the disaggregated answers, the total number of series in our dataset increases to 94.[5]

A. *Econometric Issues I*

Before estimating the dynamic factor model we need to tackle several econometric issues. First, the baseline dynamic factor model requires that the data as well as the common factors are stationary. Clearly, aggregate indexes of consumer sentiment and the individual questions are persistent variables. Since the individual questions are summarized as relative scores (which are by construction bounded), it is reasonable to model them as stationary. This is confirmed by unit root tests. Using the conventional unit root tests and critical values, majority of questions and aggregate sentiment indexes are either stationary or borderline stationary. In particular, the Elliott et al. (1996) test (ERS) reveals that 9 of 19 series investigated (3 aggregate indexes and 17 questions—relative scores) were found stationary (rejection of the unit root null at less than 1% confidence level), 7 borderline (between 1% and 10% confidence levels) and 4 I(1) (at more than 10% confidence level).[6] Therefore, I will proceed by estimating the factors from levels, rather than differences of sentiment questions.[7]

We also need to choose the number of factors, $\bar{r}$ and the number of lags, p. In this section I report my baseline results with $\bar{r} = 3$ and $p = 0$. The number of factors $\bar{r}$ was chosen by the Bai and Ng (2002) selection criteria.[8] In section VI below I explore the implications of alternative values of p.

It turns out that the static representation (2) identifies the individual factors only up to multiplication by a non-singular matrix.[9] Thus, without imposing further identifying restrictions, one should not interpret the *individual* factors obtained from (3) and relate them to entities these are supposed to represent (i.e. real activity, nominal factor, financial factor, ...). However, it is still possible to analyze how all the factors jointly explain other variables or how close the linear space spanned by the factors is to observable series, e.g. if a given series is actually one of the factors (see Bai, 2004). This is of course also useful for forecasting and testing the implications of economic models.

[5] For an extension of the dynamic factor model that treats missing data (following Stock and Watson, 2002) see section VI.

[6] Using the ADF test (with 1 lag) 7 series were stationary, 3 borderline and 10 nonstationary (I(1)). The Im et al. (2003) panel unit root test overwhelmingly rejects the null of nonstationarity.

[7] It should, however, be noted that even if the sentiment questions and the underlying factors were cointegrated non-stationary (I(1)) variables, the estimation technique I use and much of the inference still remains valid, as recently shown by Bai (2004). (One major difference is that, similarly to cointegrating regressions, the estimates converge faster to their true values when the series and factors are cointegrated I(1) processes than when they are stationary.)

[8] Specifically, I investigated the PC_{p1}, PC_{p2} and PC_{p3} criteria (Bai and Ng, 2002, p. 201). When fixing the maximum number of lags to 10, PC_{p1} and PC_{p2} picked $\bar{r} = 3$; PC_{p3} selected $\bar{r} = 4$.

[9] If $\tilde{F}$ solves (3), then $\tilde{F}H$ solves (3) for any $r \times r$ non-singular matrix H.

B. Results

After addressing some of the econometric issues let us have a look at the results. Figure 2 shows the Michigan Index of Consumer Sentiment and the first three estimated factors together with 95% asymptotic confidence intervals obtained using the approximations of Bai (2003).[10] It is obvious that the first factor captures extremely well the index; the correlation is 0.93. The correlation between the sentiment index and the other two factors is much lower, -0.32 and -0.06, respectively. This finding extends to the other two sentiment indexes (ICC and ICE; see also Table IV).

The first panel of Figure 2 is enlarged in Figure 3. Figure 3 tests how far the ICS is from (the closest linear combination of) the factors. Since consumer sentiment correlates very strongly with the first factor, this linear combination has a very high weight on the first factor (but is not exactly equal to the first factor since the other factors have non-zero weights).[11]

Formally, one has to reject the hypothesis that sentiment is in the linear space spanned by the factors, since about 40% of time sentiment lies outside the gray band (95% confidence interval). However, one limitation of the Bai's results is that for the approximations to be accurate one has to make several potentially quite restrictive assumptions. In particular, the Bai's approximations may not hold in small samples when the disturbances are cross-sectionally dependent (i.e. $\mathbf{E}u_{i,t}u_{j,t} \neq 0$), non-normally distributed or serially correlated (i.e. $\mathbf{E}u_{i,t}u_{i,t-k} \neq 0$). To investigate the implications of these limitations in my empirical application I carry out a bootstrap procedure which is an alternative to the method proposed by Bai (2003).[12] The bootstrap procedure typically provides a more precise small sample approximation of the distribution of the statistic of interest under the above data irregularities.

[10]The factors have been normalized so that they are orthogonal, $\hat{F}'\hat{F} = I_r$.

[11]The testing procedure follows Bai (2003) and can be summarized as follows. I first rotate the factors towards the index of consumer sentiment using the regression

$$S_t = \delta'\hat{F}_t + e_t$$

and collect the fitted values as $\hat{S}_t = \hat{\delta}'\hat{F}_t$. The 95% confidence band for the closest linear combination of factors to the consumer sentiment is then $\left(\hat{S}_t - 1.96\left(\hat{\delta}'\hat{\Pi}_t\hat{\delta}/N\right)^{1/2}, \hat{S}_t + 1.96\left(\hat{\delta}'\hat{\Pi}_t\hat{\delta}/N\right)^{1/2}\right)$, where $\hat{\Pi}_t$ is (a consistent estimate of) the variance matrix of the factors, given in Bai (2003), formula (7). If the sentiment series, S_t lies (mostly) within this band, the null of sentiment being one of the factors is not rejected.

[12]I thank Jonathan Wright for suggesting that I apply the block bootstrap procedure. The procedure is described in detail in Appendix II.

The bootstrap procedure consists of drawing a large number of artificial samples (pseudo-data) from the available dataset, calculating "artificial" factors from these artificial samples, and inferring the confidence intervals from the empirical distribution of the artificial factors. To preserve the time persistence of data (and factors), I apply the block bootstrap procedure—I sample blocks of series, rather than individual time observations.

The confidence bands for the Michigan Index of Consumer Sentiment obtained from the block bootstrap procedure are displayed in Figure 4. It turns out that these bands suggest that there is substantially more uncertainty about the estimated factors than implied by the Bai method. Two principal reasons for this difference are perhaps the non-normality and positive autocorrelation of residuals. The Shapiro–Francia test for normality reveals that 38 of 94 (40%) are not normally distributed on the 5% probability level and 49 (52%) are non-normal on 10% level. This is probably caused by the discreteness of the data; the individual answers are mostly expressed as integers between 0 and 100. Often however, the number of responses is quite small, so that the observations lie between 0 and 10. The residuals are also quite persistent; the mean first order autocorrelation over 94 series of residuals is 0.63.

Since the bootstrap confidence bands are wider than the Bai's bands, sentiment lies in the 95% band more often. Consumer sentiment lies within the 95% confidence band about 95–96% of the time (depending on the block size). This implies that the null hypothesis that sentiment lies in the space spanned by the factors cannot be rejected.

Table II lists the correlations between factors and individual questions and factors.[13] The findings in Table II can be summarized as follows:

- The first three common factors explain more than 78% of the variation in 11 questions and 46–60% of the variation in the remaining questions.[14] These are very high numbers; for example, in Stock and Watson's (2002a) panel consisting of 215 series the first *six factors* explain about 39% of the variance (measured by the trace R^2). This is probably for two reasons. First, Stock and Watson's panel is much more heterogenous, consisting of very diverse series. Second, their panel consists of monthly data, which are inherently noisier and more volatile. As a result, the factors in the Michi-

[13]This follows a similar exercise of Stock and Watson (2002), section 4.2.

[14]Question 4 (How much do you expect your income to increase during the next 12 months?) is the exception and is explained relatively badly by the factors.

gan dataset are very effective in summarizing the information from the relative scores of individual questions.[15]
- The factors explain well the questions that are highly correlated with the first factor. This is due to the fact that the first factor typically captures most of the variation in the questions, even though it explains some questions much better than others. While the correlations with the second and third factors tend to be higher for questions that are not strongly correlated with the first factor, this has only a "second-order" effect on the $\bar{R}^2$.

Leading Indicator of Consumer Sentiment: Stock and Watson (1999), Evans et al. (2002) and others use factor analysis to extract leading and coincident of the business cycle from large datasets. These measures of economic activity attempt to signal downturns in real time, well before a business cycle dating committee, such as NBER, officially declares a recession. Since many practitioners, policy-makers and academics believe indexes of consumers sentiment provide useful information about the economy, it is interesting to do some further investigation about how to extract a "leading indicator of consumer sentiment" from the Michigan dataset.

Table II and Figure 4 document extremely strong contemporaneous correlation between the Michigan Index of Consumer Sentiment and the first common factor. In contrast, I will now briefly consider how to produce good *forecasts* of consumer sentiment. Some guidance is given by Table III, which reports adjusted $\bar{R}^2$s of regressions of consumers sentiment on lagged individual questions,

$$S_t = \beta_0 + \beta_1 Qi_{t-k} + \varepsilon_t, \qquad k = 1, \ldots, 4,$$

where S is the Michigan Index of Consumer Sentiment and Qi_t is the individual question (or sentiment index or the first common factor). The individual questions vary substantially in their ability to predict the overall index. While some questions capture the index badly, others—questions 3, 10, 11, 13 and 20—explain more than 40% of its variation. I use a panel of the relative scores of these five questions to extract the first common factor. The last line of the Table, denoted "Lead F1," reports that this "leading" common factor predicts consumer sentiment somewhat better than any individual questions, overall and expected consumer sentiment or the first common factor from the baseline model. The leading common factor performs relatively better at longer forecasting horizons (i.e. larger k). At shorter forecasting horizons there does not seem to be much advantage in using the leading common factor (over simple univariate AR forecast, reported in line "lagged").

V. Consumer Sentiment and Consumption

Hall (1978) showed that the standard certainty-equivalent model of consumption dynamics implies that aggregate consumption follows a random walk. Subsequently, large literature arose that investigated the empirical relevance of the Hall's model. This literature found that consumption growth is "excessively sensitive" in that it reacts to past variables—including past income growth, past consumption growth, past sentiment—to which a rational optimizing consumer with time separable utility would not respond. The excess sensitivity of consumption growth to sentiment indexes was first noted by Carroll et al. (1994) and Acemoglu and Scott (1992) and further investigated by Sommer (2002). This section provides an extension of findings of these authors and detailed analysis of the sensitivity of consumption growth to individual sentiment questions and the common factors.

The findings are summarized in Table IV. The Table considers the regression of consumption growth on (one lag of) past sentiment questions, Qi,

$$\Delta \log C_t = \beta_0 + \beta_1 Qi_{t-1} + \varepsilon_t. \qquad i = 1, \ldots, 21.$$

The $\bar{R}^2$s and t statistics on β_1 are displayed and the stability of this relationship is investigated for three periods: 1960:Q1–2004:Q1, 1960:Q1–1981:Q4 and 1982:Q1–2004:Q1.

Several findings emerge:
- 16 of 17 individual sentiment questions are statistically significant (in-sample) predictors (on the 5% confidence level) of consumption growth over the whole sample (1960:Q1–2004:Q1).
- Sentiment indexes tend to have higher t statistics (in absolute values) in the first subsample than in the second.
- Still, 7 of 17 individual are statistically significant predictors in all three subsamples considered.
- Some questions are overwhelmingly significant, with p values of 0.001 or less in all subsamples. They are: Q6—news about the change in business conditions in the *past* few months, Q9—expected overall business conditions a year from now and especially Q12—expected unemployment a year from now. These questions beat even the Index of Consumer Expectations.
- The common factors, especially the first factor, are significantly related to future consumption growth

[15] The factors are similarly successful at explaining the disaggregated answers, with the median R^2 of 0.75 for the 94 answers.

in all subsamples.[16] Judging by the $\bar{R}^2$s, the factors are as good predictors as the "best" individual question in the whole sample and the first subsample. This is in line with findings of many other researchers (including Forni and Reichlin, 1998; Stock and Watson, 1999 and Boivin and Ng, 2005) that in various empirical applications forecasts generated by the common factors are superior to those from the individual series. However, I also find that the factors get beaten by several questions after 1981 (questions 6, 9 and 12 among others).

- The Table confirms the findings of previous authors that the aggregate overall index of consumer sentiment and especially the ICE predict consumption growth.
- Finally, comparing Tables II and IV, the factors explain relatively badly the questions that do a really good job in predicting consumption growth (questions 6, 9 and 12).

These results deserve some discussion. First, my findings confirm that sentiment indexes are statistically significant predictors of consumption growth. This holds for most individual sentiment questions and subsamples, as well as the common factors.[17] Second, while the aggregate indexes of consumer sentiment, in particular the Index of Consumer Expectations, are significant, these are dominated both by some individual questions and by the common factors. For example, the unemployment expectations question (Q12) is among the strongest single predictors of consumption growth; it does a particularly good job after 1981.[18]

VI. Alternative Models

A. Econometric Issues II

In the previous section I assumed away some econometric issues related to the principal component analysis. I will now investigate the robustness of the above baseline estimation method to some modifications proposed in the literature. First, I briefly describe these alternative algorithms and then report the results.

Missing Data: One of the advantages of principal components method is that it is able to handle various data irregularities. Stock and Watson (2002) proposed a solution based on the EM algorithm that makes it possible to estimate panels with observations of different frequencies, (occasionally) missing observations and unbalanced panels.[19] This method is useful for the sentiment panel as an alternative to the above baseline estimates since some of the data are not available for the whole time range 1960–2004 but start later in the sample. In the above calculations I discarded the series that start later in the sample. Alternatively, one can include the series in the calculation and apply the EM algorithm instead of simple dynamic principal components. This is of course a preferable if these series contain valuable independent information about the common factors.

Weighted Principal Components: Boivin and Ng (2005) argue that cross-correlation between disturbances $u_{i,t}$ and $u_{j,t}$ can give rise to suboptimal small sample properties of the simple principal components estimator (3). They propose several *weighted* principal components estimators, some of which perform better in the empirical applications they consider (forecasting various economic activity and inflation series). Below I investigate two alternative weighting schemes that worked well in the empirical forecasting applications of Boivin and Ng (2005). These modifications, counterparts of the generalized least squares in the DFM world, are obtained by minimizing

$$\min_{F_1,\ldots,F_T,\Lambda} 1/T \sum_{t=1}^{T} \left(x_t - \Lambda F_t\right)' \Omega^{-1} \left(x_t - \Lambda F_t\right)$$

subject to $\Lambda'\Lambda = I_r$ instead of (3) for a given weighting matrix Ω. Below I investigate two of the weighting schemes proposed by Boivin and Ng:

1) Ω is diagonal with zeros and ones on the main diagonal. The exact DFM assumes that the idiosyncratic terms $u_{i,t}$ are not cross-correlated. The weighting proposes to discard the series whose residuals obtained by regressing on factors are most strongly correlated with some other residuals (and gives them zero weight in Ω).

[16]Since the common factors are not observed but rather have to be estimated, any regression that includes factors among the explanatory variables potentially suffers from the generated regressor bias. Bai (2003) shows that estimates of factors converge to their true values at rate $\min\{N, T\}$ and for large N and T such that $\sqrt{T}/N \to 0$ the factors can be treated as known. Based on the Monte Carlo simulations, Bai finds that "[f]or $N = 1000$, the confidence intervals collapse to the true values." However, figures similar to Figure 4 imply that in my empirical application there remains sampling uncertainty about the factors. If present, this uncertainty is likely to bias the t statistics in Table IV towards 0 (similarly to the regression with a measurement error). Thus if the sampling error is non-negligible, the t statistics on the "true" factors are actually likely to be higher than those reported in the Table.

[17]These in-sample results complement similar out-of-sample results of Slacalek (2004).

[18]This finding confirms a result first noted by Carroll and Dunn (1997).

[19]The EM (expectation–maximization) algorithm starts with picking a subset of series that are available for the whole sample. It then iterates until convergence between two steps: (i) given the series, calculate an estimate of common factors and factor loadings and (ii) given the factor and factor loadings, estimate the missing observations as fitted values. For a detailed description of the algorithm see the Appendix of Stock and Watson (2002).

2) Ω is diagonal with with the diagonal terms inversely related to variances of residual from the regression of individual series on factors.

As an alternative to the baseline model ($p = 0$), I consider including increasing the number of lags in equation (2) to $p = 1$. This is equivalent to creating a stacked panel by augmenting the dataset with its lagged copy,

$$\tilde{X} = [X_{2:T} \;\; X_{1:T-1}],$$

where $X_{2:T}$ is a $T-1 \times N$ matrix of all N series in periods $2, \ldots, T$. The principal components are then estimated using the augmented dataset $\tilde{X}$, instead of the original dataset.

B. Alternative Results

Figures 5–7 and Table V compare the estimation results for these 4 alternative specifications to the baseline "balanced" panel procedure reported in section IV. The alternative procedures are:

- The unbalanced panel that explicitly treats missing observations (denoted "Unbal (EM)" in graphs). Including the series for which some of observations are missing increases the number of series from 94 to 117.
- The weighed principal components with weighting scheme (1), which discards the series with residuals strongly cross-correlated with others (denoted "W: Discard"). This in turn reduces the number of series to 40.
- The weighed principal components with weighting scheme (2) with weights inversely proportional to the variance of residuals (denoted "W: Est Diag").
- Stacked balanced panel that estimates the baseline model for $p = 1$ (denoted "Stacked").

The first feature that one immediately notices looking at Figures 5–7 is how similar the alternative estimates of the factors are. The alternative methods produce basically the same factors. The correlations between the five factor estimates are greater than 0.87 (for each factor).

Table V documents that the findings of Table IV are replicated for the alternative factor estimates. In general, the factors, irrespective of the estimation method, robustly forecast consumption growth. The $\bar{R}^2$s of regressions in Table V vary only negligibly across rows.

VII. Summary and Conclusions

To the best of my knowledge, this paper is the first to provide an analysis of disaggregated questions from the Michigan Survey of Consumer Sentiment. I apply the methods of factor analysis to investigate the relationships between the individual questions, their underlying common factors and future consumption growth. I find that the common factors summarize very well the information in the Survey. Second, I confirm and extend the findings of previous authors on the excess sensitivity of future consumption growth with respect to the individual questions and common factors. Third, I document that many individual questions and the common factors are stable predictors consumption growth over all periods of time considered. While the common factors typically dominate the individual questions in the predictive power, there are questions, in particular the question on expected unemployment, that beat even the factors. Fourth, the estimates of the common factors are not sensitive to various specifications and estimation techniques. In addition, I propose a bootstrap method to evaluate uncertainty about the estimated common factors. I document that uncertainty implied by the bootstrap confidence intervals is much higher than is suggested by the most common alternative method.

References

Acemoglu, Daron, and Andrew Scott (1992), "Consumer Confidence and Rational Expectations: Are Agents' Beliefs Consistent with the Theory?" *The Economic Journal*, 104, 1–19.

Bai, Jushan (2003), "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71, 135–171.

Bai, Jushan (2004), "Estimating Cross-section Common Stochastic Trends in Nonstationary Panel Data," *Journal of Econometrics*, 122, 137–183.

Bai, Jushan, and Serena Ng (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221.

Berkowitz, Jeremy, and Lutz Kilian (2000), "Recent Developments in Bootstrapping Time Series," *Econometric Reviews*, 19, 1–48.

Bernanke, Ben, and Jean Boivin (2003), "Monetary Policy in a Data-Rich Environment," *Journal of Monetary Economics*, 50, 525–546.

Boivin, Jean, and Serena Ng (2005), "Are More Data Always Better for Factor Analysis?" *Journal of Econometrics*, ...

Bram, Jason, and Sydney Ludvigson (1998), "Does Consumer Confidence Forecast Household Expenditure? A Sentiment Index Horse Race," *FRBNY Economic Policy Review*, 4(2), 59–77.

Carroll, Christopher D., and Wendy E. Dunn (1997), "Unemployment Expectations, Jumping (S,s) Triggers, and Household Balance Sheets," *NBER Macroeconomics Annual*, ...

Carroll, Christopher D., Jeffrey C. Fuhrer, and David W. Wilcox (1994), "Does Consumer Sentiment Forecast Household Spending? If So, Why?" *American Economic Review*, 84(5), 1397–1408.

Chamberlain, Gary, and Michael Rothschild (1983), "Arbitrage Factor Stucture, and Mean–Variance Analysis of Large Asset Markets," *Econometrica*, 1281–1304.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock (1996), "Efficient Tests for an Autoregressive Root," *Econometrica*, 64, 813–836.

Evans, Charles L., Chin Te Liu, and Genevieve Pham-Kanter (2002), "The 2001 Recession and the Chicago Fed National Activity Index: Identifying Business Cycle Turning Points," Economic perspectives, Federal Reserve Bank of Chicago.

Forni, Mario, Domenico Giannone, Marco Lippi, and Lucrezia Reichlin (2004), "Opening the Black Box: Structural Factor Models versus Structural VARs," mimeo, Université Libre de Bruxelles.
Forni, Mario, and Lucrezia Reichlin (1998), "Let's Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycle," *Review of Economic Studies*, 65, 453–474.
Hall, Robert E. (1978), "Stochastic Implications of the Life Cycle–Permanent Income Hypothesis," *Journal of Political Economy*, 86(6), 971–987.
Howrey, Philip E. (2001), "The Predictive Power of the Index of Consumer Sentiment," *Brookings Papers on Economic Activity*, (1), 175–216.
Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin (2003), "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115, 53–74.
Ludvigson, Sydney (2004), "Consumer Confidence and Consumer Spending," *Journal of Economic Perspectives*, 18(2), 29–50.
Slacalek, Jirka (2004), "Forecasting Consumption," mimeo, Johns Hopkins University.
Sommer, Martin (2002), "Habits, Sentiment and Predictable Income in the Dynamics of Aggregate Consumption," working paper, Johns Hopkins University.
Stock, James H., and Mark W. Watson (1999), "Forecasting Inflation," *Journal of Monetary Economics*, 44, 293–335.
Stock, James H., and Mark W. Watson (2002), "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20(2), 147–162.
Stock, James H., and Mark W. Watson (2005), "Forecasting with Many Predictors," in Graham Elliott, Clive Granger, and Allan Timmermann, editors, *The Handbook of Economic Forecasting*, Elsevier/North Holland.
Survey of Consumers (1999), "Data File Documentation," *Survey Research Center*.
http://www.sca.isr.umich.edu/

APPENDIX I.: MICHIGAN INDEXES OF CONSUMER SENTIMENT

This Appendix describes in detail the construction of Michigan Indexes of Consumer Sentiment and other series used in the paper. For more documentation refer to Survey of Consumers and other documents available on the web page of the Survey Research Center.

Individual Questions

The list of all individual questions and answers from the Michigan Survey of Consumer Sentiment follows.

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?"

   1 Better Off
   2 Same
   3 Worse Off
   4 DK; NA

2) "Why do you say so?"

   5 Higher income
   6 Lower income
   7 Higher prices

3) "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?"

   8 Better Off
   9 Same
   10 Worse Off
   11 DK; NA

4) "During the next 12 months, do you expect your (family) income to be higher or lower than during the past year?" and "By about what percent do you expect your (family) income to increase during the next 12 months?"

   19 Expect Increase: 1–4%
   20 Expect Increase: 5%
   21 Expect Increase: 6–9%
   22 Expect Increase: 10–24%
   23 Expect Increase: 25% or more
   24 Expect Increase: DK how much up
   25 Expect Same
   26 Expect Down
   27 DK; NA

5) "How about the next year or two—do you expect that your (family) income will go up more than prices will go up, about the same, or less than prices will go up?"

   28 Income will go up more than prices
   29 Income will go up same as prices
   30 Prices will go up more than income
   31 DK; NA

6) "During the last few months, have you heard of any favorable or unfavorable changes in business conditions?" and "What did you hear?"

   32 Heard Favorable News
   33 Heard Unfavorable News
   34 No mentions

7) "What did you hear?"

   35 Favorable News: Government; Elections
   36 Favorable News: Employment
   37 Favorable News: Higher Consumer Demand
   38 Favorable News: Lower Prices
   39 Favorable News: Easier Credit
   40 Favorable News: Stock Market
   41 Favorable News: Trade Deficit
   42 Unfavorable News: Government; Elections
   43 Unfavorable News: Unemployment
   44 Unfavorable News: Lower Consumer Demand
   45 Unfavorable News: Higher Prices
   46 Unfavorable News: Tighter Credit
   47 Unfavorable News: Energy Crisis
   48 Unfavorable News: Stock Market
   49 Unfavorable News: Trade Deficit

8) "Would you say that at the present time business conditions are better or worse than they were a year ago?"

   50 Better Now
   51 Same
   52 Worse Now
   53 DK; NA

9) "And how about a year from now, do you expect that in the country as a whole business conditions will be better, or worse than they are at present, or just about the same?"

54 Better
55 Same
56 Worse
57 DK; NA

10) "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?"

65 Good Times
66 Uncertain; Good & Bad
67 Bad Times
68 Don't Know
69 Not Ascertained

11) "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 5 years or so, or that we will have periods of widespread unemployment or depression, or what?"

70 Good Times
71 Uncertain; Good & Bad
72 Bad Times
73 Not Ascertained

12) "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?"

74 Less Unemployment
75 Same Unemployment
76 More Unemployment
77 DK; NA

13) "No one can say for sure, but what do you think will happen to interest rates for borrowing money during the next 12 months—will they go up, stay the same, or go down?"

78 Go Up
79 Stay the Same
80 Go Down
81 DK; NA

14) "During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?" and "By about what percent do you expect prices to go up, on the average, during the next 12 months?"

82 Prices will stay the same or go down
83 Prices will go up by: 1–2%
84 Prices will go up by: 3–4%
85 Prices will go up by: 5%
86 Prices will go up by: 6–9%
87 Prices will go up by: 10–14%
88 Prices will go up by: 15% or more
89 Prices will go up by: DK how much up
90 DK; NA

15) "As to the economic policy of the government—I mean steps taken to fight inflation or unemployment—would you say the government is doing a good job, only fair, or a poor job?"

91 Good Job
92 Only Fair
93 Poor Job
94 DK; NA

16) "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 a bad time for people to buy major household items?"

95 Good Time to Buy
96 Uncertain; Depends
97 Bad Time to Buy

17) "Why do you say so?"

Good Time to Buy:
98 Prices are low; good buys available
99 Prices won't come down; are going higher
100 Interest rates are low; credit is easy
101 Borrow-in-advance of rising interest rates
102 Times are good; prosperity

Bad Time to Buy:
103 Prices are high
104 Interest rates are high; credit is tight
105 Times are bad; can't afford to buy
106 Bad times ahead; uncertain future

18) "Generally speaking, do you think now is a good time or a bad time to buy a house?"

107 Good Time to Buy
108 Uncertain; Depends
109 Bad Time to Buy

19) "Why do you say so?"

Good Time to Buy:
110 Prices are low; good buys available
111 Prices won't come down; are going higher
112 Interest rates are low; credit is easy
113 Borrow-in-advance of rising interest rates
114 Good investment
115 Times are good; prosperity

Bad Time to Buy:
116 Prices are high
117 Interest rates are high; credit is tight
118 Times are bad; can't afford to buy
119 Bad times ahead; uncertain future

20) "Speaking now of the automobile market—do you think the next 12 months or so will be a good time or a bad time to buy a car?"

120 Good Time to Buy
121 Uncertain
122 Bad Time to Buy

21) "Why do you say so?"

Good Time to Buy:
123 Prices are low; good buys available
124 Prices won't come down; are going higher
125 Interest rates are low; credit is easy
126 Borrow-in-advance of rising interest rates
127 Times are good; prosperity
128 New fuel efficient models

Bad Time to Buy:
129 Prices are high

130 Interest rates are high; credit is tight
131 Times are bad; can't afford to buy
132 Bad times ahead; uncertain future
133 Price of gas; shortages
134 Poor selection; poor quality

The relative scores, reported examined in Tables I–II were constructed as the the proportion of respondents giving favorable responses minus the proportion giving unfavorable responses plus 100. For questions 2, 7, 17 and 19 this was not possible. The relative score for question 7 was constructed as the proportion of respondents giving favorable responses about *employment* minus the proportion giving unfavorable responses about *(un)employment* plus 100. The reason for this is that the "employment" responses make up vast majority of the news heard in question 7. The construction of the relative score for question 7 or any other relative score of course does not have any effect on the estimates of the underlying factors F_t since these are estimated using the *disaggregated* answers, not the relative scores.

Other Data

I use quarterly data, 1960:Q1–2004:Q1. All consumer sentiment data were obtained from the web site of the Survey Research Center. The consumption data come from the FRED II database of the Federal Reserve Bank of St. Louis (available at http://research.stlouisfed.org/fred2/). The series is seasonally adjusted (total) per capita real personal consumption expenditures in chained 2000 dollars, originally produced by the Bureau of Economic Analysis.

Appendix II.: Bootstrap Procedure for Factor Confidence Intervals

This Appendix describes the bootstrap procedure used in section IV and Figure 4 to estimate confidence bands for the common factors (and their linear combinations). The procedure is an application in the factor model setup of the block bootstrap procedure. (For a general description of the block bootstrap procedure see Berkowitz and Kilian, 2000 or Howrey, 2001.)

Suppose we have an $T \times N$ matrix of data X with N series and T observations with a typical row x_t'. Given a suitable length of block, k, the data are first partitioned into $b = T - k + 1$ overlapping blocks $\tilde{x}_t = [x_t, \ldots, x_{t+k-1}]'$, $t = 1, \ldots, b$. Draw a large number (e.g. 2000) of random samples $x^*_{(i)}$, $i = 1, \ldots, 2000$ of $l = T/k$ blocks with replacement from b blocks $\tilde{x}_t$. Given these samples, calculate the statistics of interest (estimates of the common factors). Sort these factors and select the empirical 2.5% and 97.5% quantiles in each time period as the 95% confidence band for $\hat{F}_t$.

The size of the block, k, is a crucial parameter in implementing the block bootstrap. Unfortunately, there is not much guidance in the literature for our particular situation. The general guidelines are that the block size k should increase with the sample size T and persistence of the data. Berkowitz and Kilian (2000) note that "…choosing a block size involves a tradeoff. As the block size becomes too small, the moving blocks bootstrap destroys the time dependency of the data and its average accuracy will decline. As the block size becomes too large, there are few blocks and pseudo-data will tend to look alike. As a result, the average accuracy of the moving blocks also will decline." In the univariate framework Berkowitz and Kilian propose a data-based procedure that searches over alternative block sizes to maximize accuracy of the bootstrap. Unfortunately, this procedure is not computationally feasible in our setup (because of the large cross-section dimension of the panel).

For that reason I investigate the sensitivity of confidence intervals to the choice of block size in Figure 8. Fortunately, the shape (and size) of confidence intervals do not depend much on the block size. The overwhelming common feature of all of block bootstrap confidence intervals is that they are much wider than the ones implied by the Bai (2003)'s normal approximations. For the results shown in the paper I chose the block size $k = 40$.[20]

[20] Some guidance for this was provided by an example of Berkowitz and Kilian (2000). They estimate impulse response on a series S&P stock earnings–price ratio, that is somewhat more persistent that the factors I investigate and find that the optimal choice of block size in their application is about 48 (quarters). This is a considerably larger number than sometimes chosen in the toy examples of the block bootstrap where the block size ranges around 5.

TABLE I

CORRELATIONS BETWEEN INDIVIDUAL QUESTIONS

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21
Q1	1.00	–	0.80	0.39	0.77	0.54	0.51	0.78	0.23	0.79	0.79	0.42	-0.48	-0.61	0.78	0.79	–	0.60	–	0.70	–
Q2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Q3	0.80	–	1.00	0.21	0.91	0.50	0.26	0.56	0.59	0.79	0.85	0.49	-0.21	-0.85	0.72	0.63	–	0.62	–	0.77	–
Q4	0.39	–	0.21	1.00	0.15	0.08	0.32	0.29	-0.14	0.29	0.35	0.05	-0.41	0.07	0.41	0.31	–	0.05	–	0.13	–
Q5	0.77	–	0.91	0.15	1.00	0.40	0.13	0.50	0.55	0.76	0.86	0.42	-0.12	-0.85	0.77	0.65	–	0.63	–	0.79	–
Q6	0.54	–	0.50	0.08	0.40	1.00	0.81	0.85	0.60	0.78	0.59	0.83	-0.27	-0.38	0.37	0.43	–	0.27	–	0.36	–
Q7	0.51	–	0.26	0.32	0.13	0.81	1.00	0.85	0.24	0.68	0.48	0.71	-0.53	-0.08	0.31	0.39	–	0.06	–	0.15	–
Q8	0.78	–	0.56	0.29	0.50	0.85	0.85	1.00	0.32	0.84	0.70	0.69	-0.56	-0.39	0.57	0.69	–	0.43	–	0.50	–
Q9	0.23	–	0.59	-0.14	0.55	0.60	0.24	0.32	1.00	0.58	0.58	0.76	0.20	-0.59	0.36	0.16	–	0.22	–	0.40	–
Q10	0.79	–	0.79	0.29	0.76	0.78	0.68	0.84	0.58	1.00	0.92	0.77	-0.38	-0.68	0.74	0.57	–	0.41	–	0.59	–
Q11	0.79	–	0.85	0.35	0.86	0.59	0.48	0.70	0.58	0.92	1.00	0.67	-0.34	-0.73	0.85	0.59	–	0.47	–	0.67	–
Q12	0.42	–	0.49	0.05	0.42	0.83	0.71	0.69	0.76	0.77	0.67	1.00	-0.22	-0.41	0.37	0.32	–	0.20	–	0.37	–
Q13	-0.48	–	-0.21	-0.41	-0.12	-0.27	-0.53	-0.56	0.20	-0.38	-0.34	-0.22	1.00	0.02	-0.18	-0.65	–	-0.44	–	-0.39	–
Q14	-0.61	–	-0.85	0.07	-0.85	-0.38	-0.08	-0.39	-0.59	-0.68	-0.73	-0.41	0.02	1.00	-0.56	-0.46	–	-0.58	–	-0.74	–
Q15	0.78	–	0.72	0.41	0.77	0.37	0.31	0.57	0.36	0.74	0.85	0.37	-0.18	-0.56	1.00	0.64	–	0.44	–	0.61	–
Q16	0.79	–	0.63	0.31	0.65	0.43	0.39	0.69	0.16	0.57	0.59	0.32	-0.65	-0.46	0.64	1.00	–	0.85	–	0.83	–
Q17	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Q18	0.60	–	0.62	0.05	0.63	0.27	0.06	0.43	0.22	0.41	0.47	0.20	-0.44	-0.58	0.44	0.85	–	1.00	–	0.89	–
Q19	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Q20	0.70	–	0.77	0.13	0.79	0.36	0.15	0.50	0.40	0.59	0.67	0.37	-0.39	-0.74	0.61	0.83	–	0.89	–	1.00	–
Q21	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Median	0.70	–	0.62	0.21	0.63	0.50	0.32	0.57	0.36	0.74	0.67	0.42	-0.34	-0.56	0.57	0.59	–	0.43	–	0.59	–

Notes: The “Median” row shows the median of *absolute values* of correlations in the same column above it. Entries “–” are not available since answers to some questions cannot be summarized with a relative score.

TABLE II

CORRELATIONS BETWEEN INDIVIDUAL QUESTIONS AND FACTORS

	F1	F2	F3	$\bar{R}^2$
Q1	0.85	-0.30	-0.19	0.85
Q2	–	–	–	–
Q3	0.78	-0.43	0.27	0.86
Q4	0.32	0.07	-0.46	0.24
Q5	0.72	-0.76	0.19	0.78
Q6	0.74	0.03	-0.15	0.57
Q7	0.71	0.32	-0.45	0.80
Q8	0.87	-0.04	-0.38	0.90
Q9	0.46	-0.13	0.48	0.46
Q10	0.97	-0.00	0.06	0.94
Q11	0.89	-0.15	0.14	0.83
Q12	0.71	0.12	0.08	0.51
Q13	-0.50	0.18	0.65	0.55
Q14	-0.67	0.39	-0.56	0.91
Q15	0.71	-0.43	-0.16	0.52
Q16	0.64	-0.59	-0.33	0.87
Q17	–	–	–	–
Q18	0.48	-0.75	0.00	0.79
Q19	–	–	–	–
Q20	0.62	-0.68	0.11	0.85
Q21	–	–	–	–
Overall	0.93	-0.32	-0.06	0.96
Current	0.73	-0.53	-0.32	0.92
Expected	0.95	-0.15	0.10	0.93

Notes: The "$\bar{R}^2$" column shows $\bar{R}^2$s from the regressions of questions factors,

$$Qi_t = \beta_0 + \beta_1 F1_t + \beta_2 F2_t + \beta_3 F3_t + \varepsilon_t, \qquad i = 1, \dots, 21.$$

"Overall," "Current" and "Expected" rows display the results for the (overall) Index of Consumer Sentiment, Index of Current Economic Conditions and Index of Consumer Expectations, respectively.

TABLE III

HOW TO FORECAST CONSUMER SENTIMENT?

Lags k	1	2	3	4
Q1	0.66	0.55	0.43	0.34
Q2	–	–	–	–
Q3	0.69	0.63	0.55	0.49
Q4	0.05	0.02	-0.00	-0.01
Q5	0.65	0.61	0.56	0.50
Q6	0.39	0.34	0.28	0.22
Q7	0.21	0.15	0.10	0.05
Q8	0.54	0.44	0.34	0.23
Q9	0.26	0.25	0.23	0.21
Q10	0.71	0.60	0.50	0.41
Q11	0.74	0.63	0.53	0.42
Q12	0.36	0.32	0.28	0.21
Q13	0.11	0.05	0.03	0.00
Q14	0.54	0.53	0.50	0.49
Q15	0.54	0.45	0.36	0.29
Q16	0.49	0.41	0.33	0.26
Q17	–	–	–	–
Q18	0.40	0.37	0.34	0.30
Q19	–	–	–	–
Q20	0.58	0.52	0.46	0.41
Q21	–	–	–	–
Overall (Lagged)	0.82	0.70	0.57	0.46
Expected	0.79	0.67	0.56	0.46
Factor 1	0.72	0.60	0.50	0.39
Lead F1	0.71	0.60	0.50	0.38

Notes: The Table shows $\bar{R}^2$s from the regressions of the Michigan Index of Consumer Sentiment on the individual questions, sentiment indexes and first common factors,

$$S_t = \beta_0 + \beta_1 Qi_{t-k} + \varepsilon_t, \qquad k = 1, \ldots, 4.$$

The sample is 1960Q1–2004Q1.

TABLE IV

HOW WELL DO INDIVIDUAL QUESTIONS EXPLAIN FUTURE CONSUMPTION GROWTH?

	60Q1–04Q1 t stat	$\bar{R}^2$	60Q1–81Q4 t stat	$\bar{R}^2$	82Q1–04Q1 t stat	$\bar{R}^2$
Q1	3.01	0.07	3.79	0.13	3.01	0.01
Q2	–	–	–	–	–	–
Q3	4.16	0.11	4.37	0.16	4.16	0.03
Q4	-1.26	0.01	-2.99	0.11	-1.26	-0.01
Q5	3.03	0.12	2.48	0.12	3.03	0.09
Q6	5.42	0.15	4.74	0.17	5.42	0.14
Q7	2.70	0.03	1.91	0.03	2.70	0.04
Q8	3.93	0.08	3.51	0.12	3.93	0.05
Q9	4.75	0.15	4.12	0.16	4.75	0.13
Q10	4.19	0.13	3.86	0.16	4.19	0.07
Q11	3.21	0.10	3.07	0.13	3.21	0.03
Q12	5.17	0.15	4.04	0.13	5.17	0.18
Q13	1.97	0.02	2.51	0.07	1.97	0.00
Q14	-3.51	0.11	-3.25	0.15	-3.51	0.01
Q15	2.03	0.04	1.19	0.03	2.03	0.02
Q16	1.81	0.02	2.65	0.08	1.81	-0.01
Q17	–	–	–	–	–	–
Q18	2.44	0.04	4.71	0.21	2.44	-0.01
Q19	–	–	–	–	–	–
Q20	2.59	0.06	3.74	0.16	2.59	-0.01
Q21	–	–	–	–	–	–
F1, F2, F3	4.79,-0.74,2.70	0.16	3.08,-2.59,0.90	0.23	1.71,0.46,1.22	0.03
Sentiment	3.54	0.10	3.81	0.16	3.81	0.16
Current	2.15	0.03	3.09	0.10	3.09	0.10
Expected	4.02	0.12	3.71	0.15	3.71	0.15

Notes: The "t stat" columns show the t statistics on a question from the Survey in the regression of real per capita consumption growth on a constant and the lagged question:

$$\Delta \log C_t = \beta_0 + \beta_1 Qi_{t-1} + \varepsilon_t. \qquad i = 1, \ldots, 21.$$

"$\bar{R}^2$" columns show $\bar{R}^2$s of these regressions.

For the "F1, F2, F3" row the "t stat" columns show the t statistics on the three factors in the regression of real per capita consumption growth on a constant and the lagged factors:

$$\Delta \log C_t = \beta_0 + \beta_1 F1_{t-1} + \beta_2 F2_{t-1} + \beta_3 F3_{t-1} + \varepsilon_t.$$

"$\bar{R}^2$" columns show $\bar{R}^2$s of this regression.

"Overall," "Current" and "Expected" rows display the results for the (overall) Index of Consumer Sentiment, Index of Current Economic Conditions and Index of Consumer Expectations, respectively.

TABLE V

HOW WELL DO ALTERNATIVE FACTOR ESTIMATES EXPLAIN FUTURE CONSUMPTION GROWTH?

	60Q1–04Q1 t stat	$\bar{R}^2$	60Q1–81Q4 t stat	$\bar{R}^2$	82Q1–04Q1 t stat	$\bar{R}^2$
Balanced						
F1, F2, F3	4.79, -0.74, 2.70	0.16	3.08, -2.59, 0.90	0.23	1.71, 0.46, 1.22	0.03
Unbal (EM)						
F1, F2, F3	5.00, 0.31, 2.35	0.16	4.46, -1.88, 0.72	0.22	2.53, 1.83, 2.19	0.06
W: Discard						
F1, F2, F3	4.55, -1.34, 2.02	0.15	1.15, -3.06, 2.14	0.23	1.44, 1.00, 0.81	0.03
W: Est Diag						
F1, F2, F3	3.66, -1.22, 3.08	0.14	0.80, -3.04, 1.78	0.22	1.91, 0.59, 1.87	0.04
Stacked						
F1, F2, F3	4.33, -0.50, 3.37	0.16	2.60, -1.81, 1.45	0.22	2.34, 1.39, 1.90	0.03

Notes: The "t stat" columns show the t statistics on the three factors in the regression of real per capita consumption growth on a constant and the lagged factors:

$$\Delta \log C_t = \beta_0 + \beta_1 F1_{t-1} + \beta_2 F2_{t-1} + \beta_3 F3_{t-1} + \varepsilon_t.$$

"$\bar{R}^2$" columns show $\bar{R}^2$s of this regression.

Fig. 1. Absolute Correlations between Individual Sentiment Questions

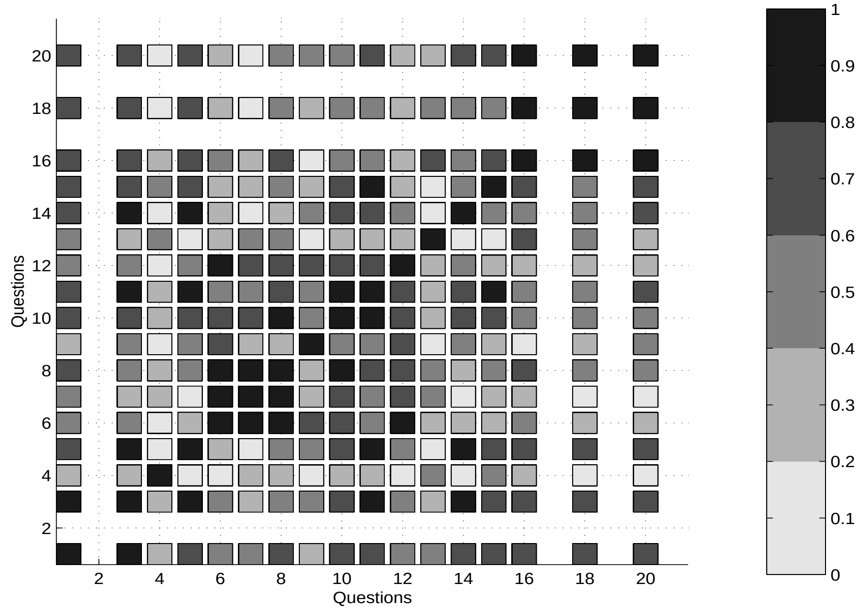
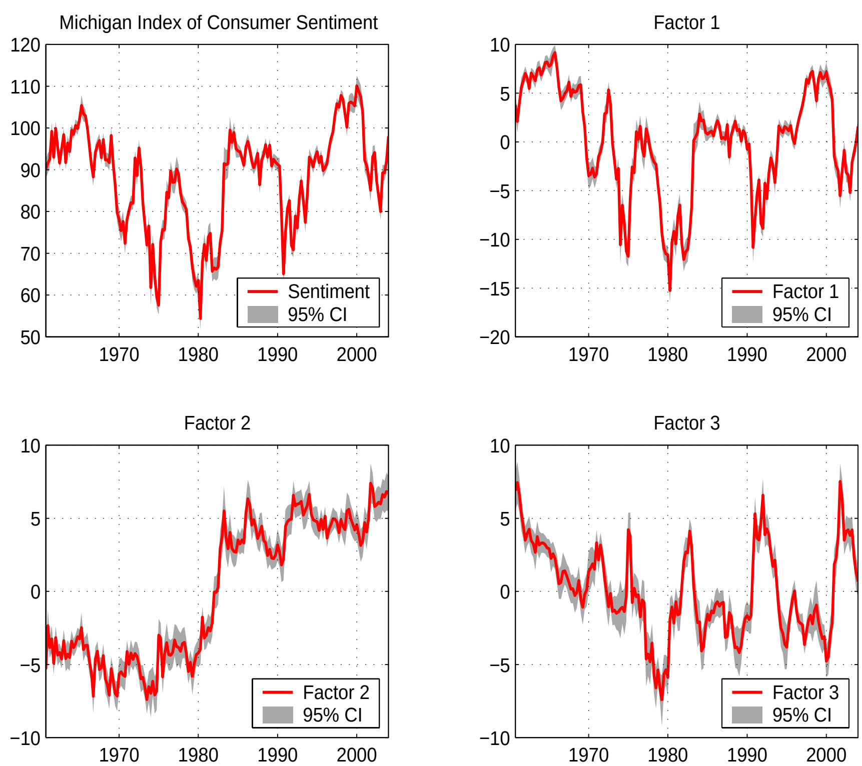
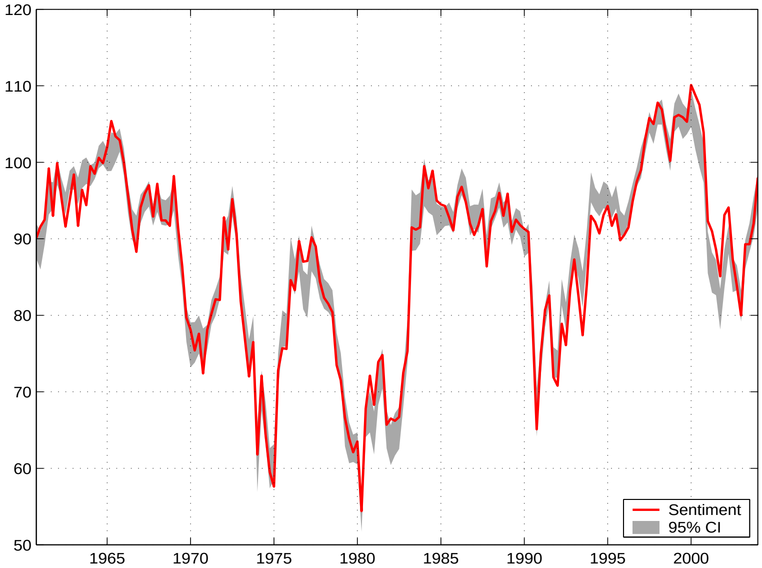


Fig. 2. Index of Consumer Sentiment and First Three Factors



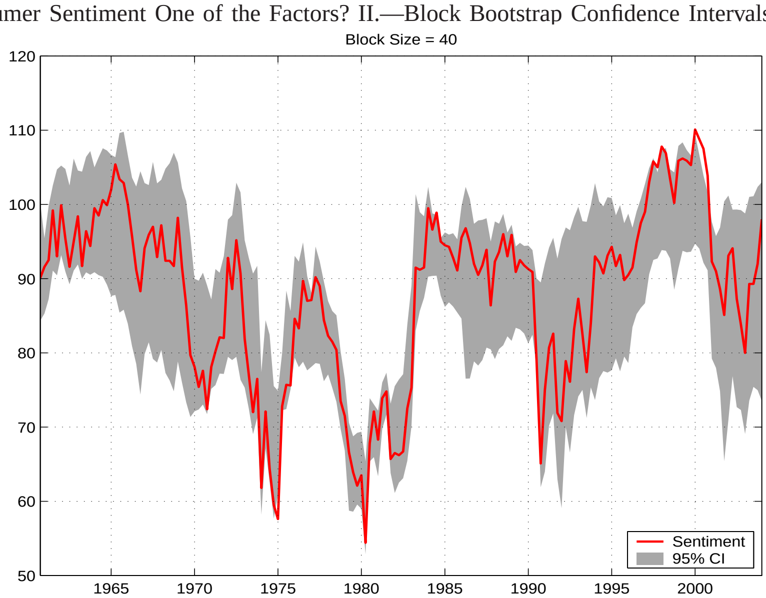
Note: The confidence intervals are 95% confidence intervals obtained using the asymptotic approximation of Bai (2003).

Fig. 3. Is the Index of Consumer Sentiment One of the Factors? I.—Bai (2003) Confidence Intervals



Note: The confidence intervals are 95% confidence intervals obtained using the asymptotic approximation of Bai (2003).

Fig. 4. Is the Index of Consumer Sentiment One of the Factors? II.—Block Bootstrap Confidence Intervals



Note: The confidence intervals are 95% confidence intervals obtained using 2000 overlapping block bootstrap samples of block size $k = 40$.

Fig. 5. Various Estimates of Factor 1

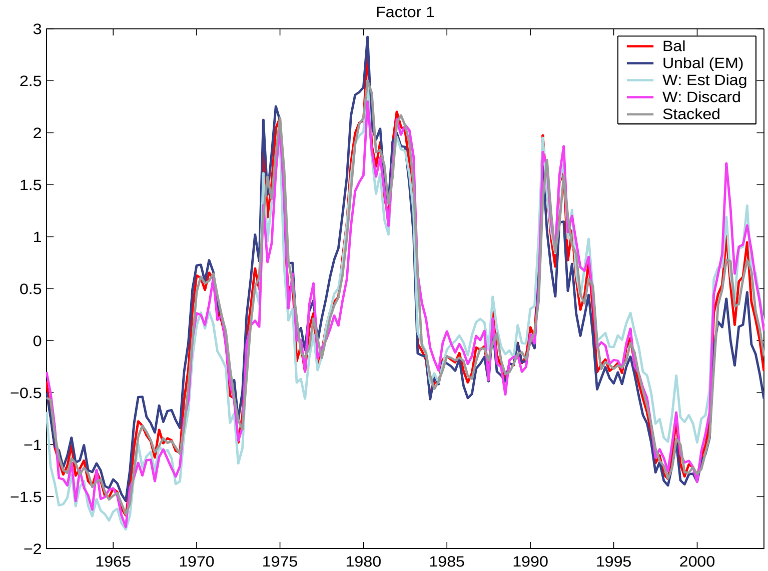


Fig. 6. Various Estimates of Factor 2

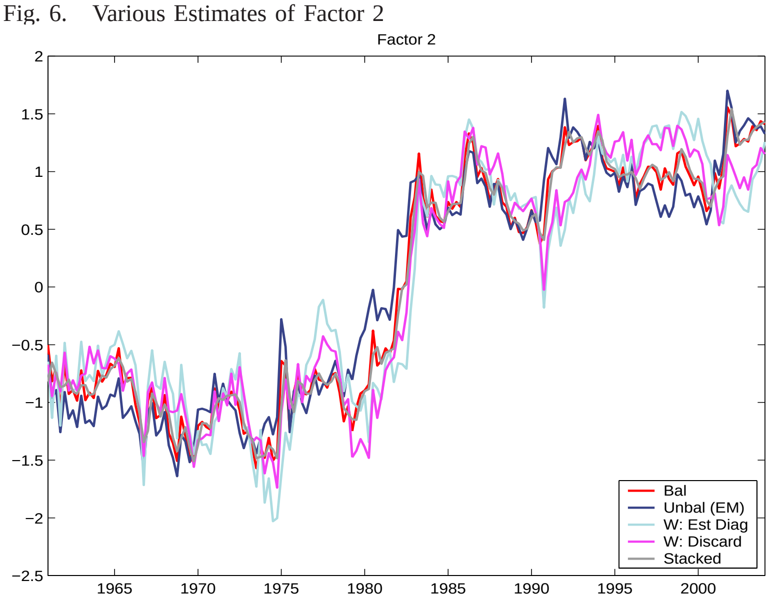


Fig. 7. Various Estimates of Factor 3

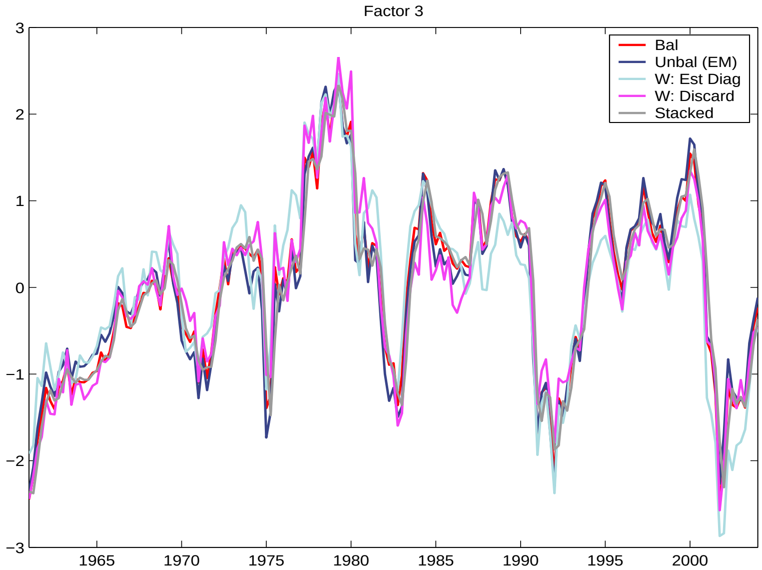
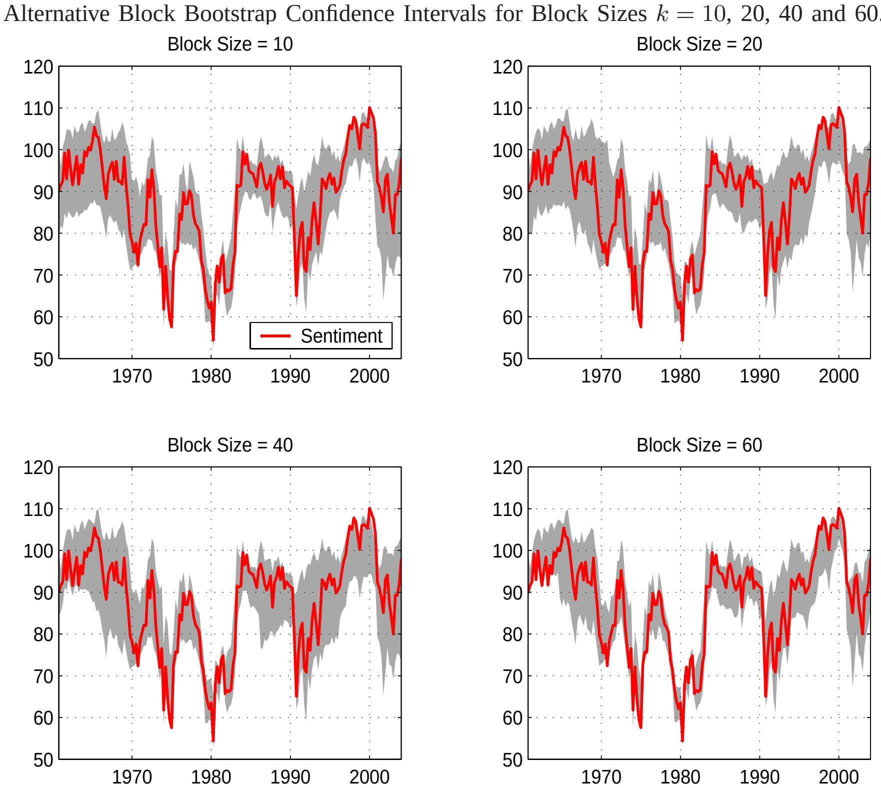


Fig. 8. Comparison of Alternative Block Bootstrap Confidence Intervals for Block Sizes $k = 10$, 20, 40 and 60.



Note: The confidence intervals are 95% confidence intervals obtained using 2000 overlapping block bootstrap samples of given block size.