Measuring Weekly Consumer Confidence

Rene Segers* Philip Hans Franses
Econometric Institute and Tinbergen Institute
Erasmus University Rotterdam

Econometric Institute Report EI 2008-01

February 6, 2008

Abstract

This paper puts forward a data collection method to measure weekly consumer confidence at the individual level. The data thus obtained allow to statistically analyze the dynamic correlation of such a consumer confidence indicator and to draw inference on transition rates, which is not possible for currently available monthly data collected by statistical agencies on the basis of repeated cross-sections. An application of the method to various waves of data for the Netherlands shows its merits. Upon temporal aggregation we also show the resemblance of our data with those collected by Statistics Netherlands.

Keywords: Consumer confidence, randomized sampling
JEL Classification Codes: C33, C42, C81, E20

*Corresponding author. Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands. Tel.: +31-10-4082524, fax: +31-10-4089162. E-mail addresses: `rsegers@few.eur.nl` (R. Segers), `franses@few.eur.nl` (P. H. B. F. Franses). All results presented in the paper are available at the question level upon request. Any remaining errors are ours.

1 Introduction

Consumer confidence indicators (CCIs) are often regarded as useful variables to measure the current state of the economy as well as to forecast its future states at reasonably short horizons, see Ludvigson (2004) for a recent assessment. Most industrialized countries report such indicators at a monthly level. Statistical agencies typically survey one thousand or more individuals each month. They ask whether these individuals believe that their situation has improved in the previous period or will improve in the next period, concerning their financial situation, employment, and for example their purchases of durable and more expensive products. The answer categories are (very) positive, neutral, and (very) negative, and their origin goes back to Katona (1951). The final indicator is constructed by subtracting the percentage of negative answers from the percentage of positive answers. Many countries also report more specific indicators, which are confined to just the financial position or just employment. Publicly available data are published in original format as well as after seasonal adjustment.

Despite their widespread use and interpretation, it can be of interest to investigate if the way consumer confidence is measured can be improved. One research angle can concern the very questions asked and the way indicators are constructed from these questions. One may for example consider replacing the traditional qualitative questions by probabilistic questions inquiring about more well-defined events, as suggested in Dominitz and Manski (2004). Also the fact that consumer confidence data show signs of seasonality can be viewed as inconvenient, and perhaps a rephrasing of the questions can overcome this potential drawback.

A second angle for potential improvement of consumer confidence indicators would be to better understand how consumer confidence varies across individuals with different socioeconomic and demographic characteristics. These insights could be exploited to reduce sampling error due to the use of small and possibly unrepresentative samples, which improves the reliability of the indicators. We believe that improvement in these two directions can be relevant, but the two research angles to be discussed below seem more promising.

A third angle concerns the fact that the data are only available at the monthly level. Indeed, businesses tend to operate in terms of weeks, and also many other economic indicators, like stock market returns, interest rates, and industry-specific figures like the number of temporary employees, are available at the weekly level. In fact, it seems that a weekly figure of consumer confidence, reported at the beginning of a new week, would be a helpful indicator for many people in business and industry. In this paper we therefore aim to propose a method to collect such data.

A fourth angle, which is also addressed in the present paper, is that consumer confidence data are usually so-called repeated cross-sectional data. That is, each month, say, one thousand individuals are interviewed, but each month this concerns one thousand *different* individuals. A major consequence of this way of collecting data is that developments over time are difficult to interpret. Basically, when an indicator is -18 in December 2005, while it was -21 in November 2005, we must conclude that the average fraction of more negative answers in December was smaller than in November. We could even say that in December 2005 consumer confidence has increased with 3 points, but we must be aware that this does not concern the same individuals. Hence, an interpretation of a sequence of monthly confidence indicator values is prone to the so-called ecological fallacy. This fallacy concerns the situation where we seek to derive micro behavior from aggregated data. In the literature there are various suggestions to circumvent or solve this problem, see King (1997), Moffitt (1993), Sigelman (1991), and the collection of papers in King *et al.* (2004), among many others. In the present paper we also seek to do that, but now by proposing an alternative method of data collection.

In sum, in this paper we put forward a method to collect weekly consumer confidence data at the individual level. We keep the Katona-type questions intact, but we merely focus on the collection of the data, which should be comparable from week to week, that is, we try to prevent facing ecological fallacy problems. To that end, we need to collect data such that we have the same (though not all) individuals being interviewed from one week to another, without them being annoyed or becoming uninterested.

The outline of our paper is as follows. In Section 2 we present our method of data collection, and we argue that it has various convenient properties for the purpose at stake. Next, we introduce the model that will be used to describe longitudinal developments in consumer confidence at the individual level. In Section 3 we illustrate the usefulness of our method by surveying individuals during periods of at most three months. We show that we obtain weekly confidence data that can be compared across the weeks. We also show how to compute confidence bounds around these numbers, so that one can infer whether this week's figure is significantly different from the previous week's figure. When we compare our figures with the actually published figures by Statistics Netherlands (SN), we observe a remarkable resemblance. Hence, when we temporally aggregate our weekly figures, we get the actually published monthly data. In Section 4 we conclude with an outline of various areas for further research.

2 Methodology

In this section we present our new method to collect consumer confidence data at the individual level. Additionally, we introduce the model that will be used in our empirical work.

2.1 Data Collection Method

To measure developments in consumer confidence over time it is desirable to conduct a longitudinal or panel study where the same individuals are surveyed at multiple points in time. This allows the researcher to study developments in confidence at the individual level and to capture the dynamic relationships between confidence and events. However, surveying the very same individuals frequently, such as weekly, likely deteriorates the quality of the survey. People get irritated and they disconnect from the panel, thereby making the panel less efficient. Or perhaps worse, respondents' (reported) confidence levels may change due to being a member of a panel, which is called panel conditioning.

For this reason, most statistical agencies decide to collect repeated cross-sections instead of panel data. This amounts to surveying a new group of individuals at each survey occasion, which implies that individuals are surveyed only once. The design is illustrated in Figure 1, panel (a). Here we index time by t, where $t = 1, \ldots, T$, individuals by i, where $i = 1, \ldots, N$ and groups of individuals by g, where $g = 1, \ldots, G$. Although clearly this design reduces respondent burden and eliminates potential panel conditioning biases, obviously one does not have the advances of a panel data set. Therefore it seems promising to collect longitudinal data nevertheless, but to calibrate the design of the panel carefully, such that the above adverse effects are negligible or, at least, manageable.

In panel design, three key decisions have to be made. Firstly, as individuals obviously cannot be surveyed continuously, one has to decide on the total time-span a panel member is requested to join the panel, to be denoted by T^*. As panel members leave the panel, one may decide to invite new individuals to join the panel in such a way that the total number of panel members remains constant. Such a strategy is referred to as rotation, see Patterson (1950) and Kish and Hess (1959). Naturally, the next step would be to decide upon the number of survey requests within this period, to be labelled n. Note that T^* and n together constitute the sampling frequency $f = n/T^*$ of the survey, which is equal to the reciprocal of the time between subsequent survey occasions, or waves. Thirdly and finally, one needs to decide when to conduct the n surveys within the time-span T^*. We will refer to this aspect as date selection. A natural way is to divide the time-span T^* into n equally long time periods, and to survey around the beginning of each subperiod. Typically, in this case the implied sampling frequency f is lower than the desired data frequency. Again one may therefore apply rotation, such that at each point in time t a new group of panel members is surveyed and data are collected continuously. The above strategy is mostly referred to as time sampling.

To the best of our knowledge, the only consumer confidence indicator that is not measured through repeated cross-sections is the Index of Consumer Sentiment of the University of Michigan. Michigan adopts a rotating panel design in which the respondents are requested to be re-interviewed six months after the first interview, see Curtin (1982) for details. This design is illustrated in Figure 1, panel (b). In our terminology, we would characterize the Michigan panel as a rotating panel where $T^* = 12$ months, $n = 2$ survey occasions per individual and time sampling is applied to obtain monthly data.

In a recent study Segers and Franses (2007) propose a new date selection approach, which seems very useful here. The authors choose the n survey occasions at random, independently for each panel member. They show that in this case of randomized sampling, data is collected to measure every possible autocorrelation up to $T^* - 1$ lags, where the lower lag orders are sampled most frequently. This facilitates the identification of any type of individual dynamics in the data and it allows for efficient estimation. Additionally, Segers and Franses (2007) show that randomized sampling may have a positive effect on response rates and response quality in case one does not inform individuals in advance about the fact that they are members of a panel. Their results provide support for the conjecture that randomized sampling increases response because it takes longer for panel members to learn the (average) sampling frequency of the panel. Response quality tends to be higher due to less panel conditioning bias. This finding seems to be explained by the effect that in a randomized panel individuals are less likely to develop expectations as to when they will be surveyed again. Having this in mind, it is demonstrated that response rates and response quality can be significantly improved if the design of the panel is carefully calibrated. A straightforward way to do so is to measure the effects of panel design characteristics, that is, of T^*, n, and the date selection strategy, on response rates and response quality by means of a pilot study. The results of this study can be used to calibrate the actual panel design to be used, as we will illustrate in the next section. An example of a randomized rotating panel, where two new individuals are invited to join the panel in each time period is shown in Figure 1, panel (c). In this example, we set the maximum time-span a panel member is requested to join the panel, T^*, equal to 8 and the number of survey requests, n, equal to 4. As a consequence, the sampling frequency f is 0.5. Each dotted area encloses all survey requests assigned to one particular cohort of individuals.

- Insert Figure 1 about here -

2.2 Modelling Consumer Confidence

To model respondents' answers to the five questions which together summarize their consumer confidence level, we employ a dynamic panel version of the Ordered Probit model, as originally developed by McKelvey and Zavoina (1975). The model is relevant in applications such as surveys, in which respondents express their preferences on an ordinal scale. In our case, they are requested to indicate whether their economic situation has been or will be better (-1), the same (0) or worse (1), see Appendix A for details. We index consumers by i, where $i = 1, \ldots, N$, questions by j, where $j = 1, \ldots, J$ and time by t, where $t = 1, \ldots, T$. A respondent's unobserved assessment of the change in his or her economic situation is denoted by $y^*_{i,j,t}$ while we only observe the multinomial variable $y_{i,j,t} \in \{-1, 0, 1\}$. We assume that the latent variable $y^*_{i,j,t}$ can be explained by a set of explanatory variables $\boldsymbol{z}_{i,j,t}$ and the previous assessment through

$$y^*_{i,j,t} = \boldsymbol{z}_{i,j,t}\boldsymbol{\beta} + \rho_j^{d_{i,t}} y_{i,j,t-d_{i,t}} + a_{i,j} + e_{i,j,t} \tag{1}$$

where $\{\boldsymbol{\beta}, \rho_1, \ldots, \rho_J\}$ are unknown parameters. The variable $y_{i,j,t-d_{i,t}}$ denotes the previous observation, which is measured $d_{i,t}$ time periods before $y_{i,j,t}$. The question-specific factor $\rho_j^{d_{i,t}}$ is a finite duration adjustment of the geometric lag or Koyck model, see Ansari *et al.* (2008) for a recent application. We assume that $\rho_j < 1$ for $j = 1, \ldots, J$. This implies that the effect of an individual's previous opinion on his or her present opinion decreases, as the time between the present and the previous survey gets longer. The above representation allows us to analyze any incomplete panel directly, even if the observations are unequally spaced. Finally, $a_{i,j}$ denotes an individual- and question-specific random effect, and $e_{i,j,t}$ denotes an idiosyncratic error term. It is assumed that $a_{i,j}$ and $e_{i,j,t}$ are mutually independent and independent of the regressors.

The latent variable $y^*_{i,j,t}$ gets mapped onto $y_{i,j,t}$ by the rule

$$\begin{aligned} y_{i,j,t} &= -1 && \text{if } \gamma_{-2} < y^*_{i,j,t} \leq \gamma_{-1} \\ y_{i,j,t} &= 0 && \text{if } \gamma_{-1} < y^*_{i,j,t} \leq \gamma_0 \\ y_{i,j,t} &= 1 && \text{if } \gamma_0 < y^*_{i,j,t} \leq \gamma_1, \end{aligned} \tag{2}$$

where the parameters γ_{-2} to γ_1 are unobserved thresholds which must satisfy $\gamma_{c-1} < \gamma_c$ for $c = -1, 0, 1$. Because the boundary values of our latent variable $y^*_{i,j,t}$ are unknown, we set γ_{-2} and γ_1 equal to $-\infty$ and $+\infty$, respectively. We normalize γ_{-1} to 0 in order to be able to include an intercept in the model. The threshold γ_0 will be estimated from the data.

In dynamic nonlinear panel data models with unobserved heterogeneity, such as the model specified above, the treatment of the initial observations is an important issue. An incorrect treatment of the initial observations may lead to a bias in the parameter estimates which only gets reduced when T is large, see Heckman (1981) for details. To deal with this problem, we apply the Wooldridge (2005) approach. This amounts to approximating the distribution of the random effects $a_{i,j}$ conditional on the initial conditions $y_{i,j,1}$ rather than the distribution of $y_{i,j,1}$ conditional on $a_{i,j}$, as suggested by Heckman (1981). Specifically, we assume that

$$a_{i,j}|y_{i,j,1}, \boldsymbol{z}_{i,j} \sim N(\alpha_0 + \alpha_1 y_{i,j,1} + \boldsymbol{z}_i \boldsymbol{\alpha}_2, \sigma_a^2), \tag{3}$$

where $\boldsymbol{z}_{i,j}$ is the subset of all nonredundant explanatory variables in $\boldsymbol{z}_{i,j,t}$ in all time periods. The idiosyncratic error is assumed to satisfy

$$e_{i,j,t}|y_{i,j,t-1}, \ldots, y_{i,j,1}, \boldsymbol{z}_{i,j}, a_{i,j} \sim N(0, 1). \tag{4}$$

The variance of $e_{i,j,t}$ is set equal to 1 as no scaling of the underlying utility model can be deduced from the observed data. We assume that the dynamics are correctly specified, which means that at most one lag of $y_{i,j,t}$ appears in the distribution given outcomes back to the initial time period. Secondly, the variables $\boldsymbol{z}_{i,j}$ are assumed to be strictly exogenous conditional on $a_{i,j}$, see Wooldridge (2005) for details.

To derive the likelihood function of the model, it is convenient to replace $a_{i,j}$ by

$$a_{i,j} = \alpha_0 + \alpha_1 y_{i,j,1} + \boldsymbol{z}_{i,j} \boldsymbol{\alpha}_2 + a_{i,j}^*, \tag{5}$$

so that

$$a_{i,j}^*|y_{i,j,1}, \boldsymbol{z}_{i,j} \sim N(0, \sigma_a^2). \tag{6}$$

The model specified in (1) now reads as

$$\begin{aligned} y_{i,j,t}^* &= \boldsymbol{z}_{i,j,t}\boldsymbol{\beta} + \rho_j^{d_{i,t}} y_{i,j,t-d_{i,t}} + \alpha_0 + \alpha_1 y_{i,j,1} + \boldsymbol{z}_{i,j}\boldsymbol{\alpha}_2 + a_{i,j}^* + e_{i,j,t} \\ &= f(\boldsymbol{x}_{i,j,t}, \boldsymbol{\theta}) + a_{i,j}^* + e_{i,j,t}, \end{aligned} \tag{7}$$

where $\boldsymbol{x}_{i,j,t}$ summarizes the explanatory variables $\{\boldsymbol{z}_{i,j,t}, y_{i,j,t-d_{i,t}}, y_{i,j,1}, \boldsymbol{z}_{i,j}\}$ and $\boldsymbol{\theta}$ summarizes the parameters $\{\boldsymbol{\beta}, \rho_1, \ldots, \rho_J, \alpha_0, \alpha_1, \boldsymbol{\alpha}_2\}$. For notational convenience we write $\boldsymbol{z}_{i,j} = \boldsymbol{z}$ and $a_{i,j}^* = a^*$ in the remainder. The model's conditional density of $\{y_{i,j,2}, \ldots, y_{i,j,T}\}$ given $\{y_{i,j,1}, \boldsymbol{z}, a^*, \boldsymbol{\theta}\}$ is given by

$$\begin{aligned} g(y_{i,j,2}, \ldots, y_{i,j,T}|y_{i,j,1}, \boldsymbol{z}, a^*, \boldsymbol{\theta}) &= \prod_{i,j,t,c} \Pr[y_{i,j,t} = c|\boldsymbol{x}_{i,j,t}]^{\mathbb{I}[y_{i,j,t}=c]} \\ = \prod_{i,j,t,c} \Big(\Phi\big(\gamma_c - f(\boldsymbol{x}_{i,j,t}, \boldsymbol{\theta}) - a^*\big) - \Phi\big(\gamma_{c-1} &- f(\boldsymbol{x}_{i,j,t}, \boldsymbol{\theta}) - a^*\big)\Big)^{\mathbb{I}[y_{i,j,t}=c]}, \end{aligned} \tag{8}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the Normal distribution, and $\mathbb{I}[\cdot]$ the indicator function, which takes the value 1 if the condition in brackets is true and zero otherwise. Integrating out a^*, we obtain the likelihood of $\{y_{i,j,2}, \dots, y_{i,j,T}\}$ conditional on $\{y_{i,j,1}, \boldsymbol{z}, \boldsymbol{\theta}\}$

$$\mathcal{L}(y_{i,j,2}, \dots, y_{i,j,T} | y_{i,j,1}, \boldsymbol{z}, \boldsymbol{\theta}) = \int_{\mathbb{R}} g(y_{i,j,2}, \dots, y_{i,j,T} | y_{i,j,1}, \boldsymbol{z}, a^*, \boldsymbol{\theta}) \times (1/\sigma_a)\phi(a^*/\sigma_a)\mathrm{d}a^*, \tag{9}$$

where $\phi(\cdot)$ denotes the probability density function of the Normal distribution. The model parameters can be estimated by maximizing the log conditional likelihood, $\log \mathcal{L}(y_{i,j,2}, \dots, y_{i,j,T} | y_{i,j,1}, \boldsymbol{z}, \boldsymbol{\theta})$.

3 Empirical Illustration

In this section we apply our data collection method to measure consumer confidence as defined by SN among students at Erasmus University Rotterdam over two periods of three months. Using the model presented in Section 2 we illustrate the advances of collecting weekly consumer confidence panel data rather than monthly repeated cross-sections.

3.1 Panel Calibration and Data Collection

As the target population of our survey coincides with the target population in Segers and Franses (2007), we use the results as presented therein to calibrate the design of our panel. Particularly, we inspect the response curves as shown in Figure 4 of their paper to determine the maximum sampling frequency f to be used for the present study, provided that we want to survey our respondents over a period as long as three months and apply a randomized sampling strategy. The curves indicate that in order to have at least 50% expected response in the twelfth week, we need to set the sampling frequency f equal to 0.47 or lower. In order to collect as much data as possible, we decide to set f equal to this maximum value, which implies that we will survey each student close to biweekly on average, or 5.6 times within three months. Once the data are collected we verify whether indeed the response rates are as high as expected, and, perhaps more importantly, we use our model to check whether no signs of panel conditioning are apparent from the data.

We applied the design to two different cohorts of students. We surveyed the first cohort from October 2nd, 2005, to January 7th, 2006, and the second cohort from

October 1st, 2006, to January 6th, 2007. These periods both span 14 weeks instead of 12 weeks as we also allowed students to join the panel in the second and third week of data collection. Note that we did not apply a rotation strategy. In total, 78 students agreed to participate in 2005 and 52 students in 2006. We measured the respondents' confidence levels using the questionnaire as developed by SN, see Appendix A for details. The survey was conducted online through an interactive website. All correspondence, including the participation requests, was generated automatically and sent by e-mail. No survey follow-up was performed to increase response. The response rates of the survey are shown in Figure 2, separately for each cohort. During the first 5 weeks, typically the response rate decreased by about 8% per week, whereas in the later weeks it decreased by about 4%. The response rate of the 12th week was still just above 50% for both cohorts, as desired.

- Insert Figure 2 about here -

3.2 Comparison with Officially Published Statistics Netherlands Data

In Table 1 we compare our *weekly* consumer confidence indicator to *monthly* consumer confidence as measured by SN. SN report their indicator for the Netherlands around the 22nd of each month, and the surveys are conducted during the first 10 working days of that particular month. Therefore we compare the average of our weekly indicator over the first two weeks, which is in fact a biweekly indicator, to the monthly indicator of SN. As students, especially those close to graduation, are generally more positive about their future financial situation as compared to the overall Dutch population, we anticipate our levels of CCI to be higher as compared to the levels as recorded by SN, who use a representative sample of the population. The longitudinal developments in our indicator, however, should be roughly similar. Especially in 2005 this is indeed the case. SN measured an increase of 4% in November, where we measured an increase of 5%. Also the changes in December (3% versus 2%) and January (6% versus 4%) compare well. There is slightly more variation in the 2006 estimates, possible due to the smaller sample of only 52 respondents. Overall, however, we believe our longitudinal changes resemble those of SN quite closely.

For ease of comparison, we also plotted the weekly CCI levels as collected using our randomized panel against the monthly CCI levels of SN in Figure 3. It is interesting to see that the changes in CCI from week to week can be substantial. This supports our conjecture that it is useful to measure CCI at a higher frequency. Also the figure nicely illustrates that the developments in weekly and monthly CCI are equal. Both indicators show a steady upward trend in 2005, a slight decrease in confidence from October to November 2006, and another increase in confidence towards the end of 2006.

- Insert Table 1 about here -

- Insert Figure 3 about here -

3.3 The Dynamics of Consumer Confidence

The second angle for potential improvement of consumer confidence indicators which we pursue with our data collection method is the fact that we collect observations of the same individuals at multiple moments in time. This allows us to study developments in consumer confidence at the individual level, correcting for heterogeneity among individuals. In particular, we may assess to what extent developments in consumer confidence are driven on the one hand by the observed and unobserved characteristics of respondents and on the other hand by state dependence. For this purpose, we analyze our consumer confidence panel data set using the dynamic Ordered Probit model with random effects, as put forward in Section 2.

We include the following explanatory variables. Firstly, to explicitly control for some observed heterogeneity, we include the demographics age and gender as $\boldsymbol{z}_{i,j}$ variables. Secondly, individual specific dummies κ_i are included to account for unobserved heterogeneity among respondents. Similarly, we aim to capture possible heterogeneity among specific questions and weeks by the variables λ_j and μ_t, respectively. Finally, we want to verify whether there are no signs of panel conditioning bias apparent from the data. This requires a careful comparison of the responses given in the first wave of data collection, which is free of panel conditioning bias by definition, and in the next waves. We extend the notion of Hansen (1980) who argues that there should not be a difference in the response distribution of different subgroups of panel members who have been exposed to different methods of data collection. This implies that responses should not depend on the particular panel design chosen, nor on the wave of data collection. This can be verified by including the panel design parameters as explanatory variables, and testing for their (joint) significance. We include n directly through a variable that indicates the number of times respondents have been requested to be surveyed previously, which is $n-1$. As the sampling frequency f is set equal to 0.47 for all panel members, we cannot include this variable directly. However, we can include a variable that indicates the number of weeks since the previous participation request, which is $1/f$ in expectation. Note that by construction the above two variables change over time. For this reason, we include the values of the two variables in all time periods as $\boldsymbol{z}_{i,j}$ variables, and their values in week t as $\boldsymbol{z}_{i,j,t}$ variables.

Estimation Results

We estimate the parameters of the above model on the 2005 data, and use the 2006 data for an out-of-sample forecasting evaluation. The estimation results are shown in Table 2. In order to delete possibly redundant explanatory variables, we consider in each column a different subset of these variables. First of all, we note that there are no signs of panel conditioning in the collected data, as none of the variables that are based on the panel design parameters have an effect on $y_{i,j,t}$ or on unobserved heterogeneity as measured by $a_{i,j}$ at any reasonable level of significance, irrespective of the model specification chosen. This indicates that the design is calibrated properly. A second general observation is that we do not find evidence for differences in consumer confidence based on age or gender. The absence of an age effect, however, is most likely simply due to the fact there is not enough variation in age among the students interviewed.

- Insert Table 2 about here -

For the other variables the results differ. The first column displays the most general version of the model, where we included the question-, time-, as well as individual-specific dummy variables. For this specification, there seems to be state dependence only in the answers given to Question 4 and to a lesser extent in the answers given to Question 5, as only the coefficients $\hat{\rho}_4$ and $\hat{\rho}_5$ are statistically different from zero. The bottom panel displays the in- and out-of-sample hit rates, as defined in Franses and Paap (2001, Section 5.3). The in-sample hit rate indicates the fraction of correctly explained answers to the five CCI questions as given by the 2005 cohort, whereas the out-of-sample hit rate indicates the fraction of correctly predicted answers as given by the 2006 cohort. Model 1 explains 70% of the answers correctly, which is promising. As the data for 2006 are collected using a new cohort of students, we substitute the average estimate of the individual- and time-specific coefficients for every individual in every week of the out-of-sample period. Even though this greatly simplifies the model, it still predicts 55% of the answers as given by this cohort correctly.

The second column presents the results for the model without individual specific dummy variables. Perhaps not surprising, this specification allows more room for state dependence, as unobserved heterogeneity is only captured by the model's random effects, see Keane (1997) and the discussion in Erdem and Sun (2001) in the context of the Wooldridge (2005) approach. This also holds for Model 3, where also the time specifics are excluded, and Model 4, where even the question specifics are excluded. Looking across the results for Models 2 to 4, we conclude that there is more state dependence in the answers given to Questions 3, 4 and 5, which together constitute the Willingness To Buy indicator, than in the answers given to Questions 1 and 2, which together constitute the Economic Climate indicator, see Appendix A. Nevertheless, in the rightmost column we present the result of a specification in which all questions share the same dynamic parameter ρ. This simplification, however, leads to a substantial decrease in likelihood. The initial values of $y_{i,j,t}$ is highly significant across all but the first specification, which indicates that in general there is substantial correlation between the initial condition and unobserved heterogeneity.

Finally, inspecting the hit rates across the different specifications, we observe that the in-sample hit rate differs only very marginally. This indicates that there is not much difference in terms of in-sample fit. In terms of out-of-sample forecasting performance, however, Model 3 appears to be the clear winner, with an out-of-sample hit rate of 61%. This suggests that the time- and individual-specific variables have no added value for out-of-sample classification. For this reason we use Model 3 in the remainder of our analysis.

The prediction-realization table for both the in-sample estimates and the out-of-sample predictions obtained using Model 3 is shown in Table 3. This table allows us to check whether the model tends to under- or overpredict certain answer categories. As a consequence, it would owe its hit rate solely to good predictions in a subset of the answer categories. In our case, however, we conclude that the model predicts well in all answer categories. Particularly in the two largest categories, which are neutral and positive. The model tends to have most difficulties predicting negative answers correctly. In fact, the majority of the predictions in this category are wrong. However, as the share of negative answers is very small, this does not affect the model's overall performance substantially.

- Insert Table 3 about here -

3.4 Correcting for Changes in the Sample Composition

Consumer confidence indicators may change due to a change in the average level of consumer confidence among the population, as desired, but also simply due to fact that at each survey occasion different individuals are surveyed. Obviously, data collection agencies seek to ensure that each time a representative sample is drawn from the population. Nevertheless, this causes additional variation in the indicator, which is undesirable. Our data collection method reduces the uncertainty due to changes in the sample composition as respondents are not replaced at each wave of data collection. Instead, they join the panel for a prolonged period of time. Still, because at each time period we only request a subsample of our panel members to be surveyed rather than all members, and some do not respond, our index is prone to this type of uncertainty. However, by imputing the missing values in our panel by simulated model predictions, we can correct for this.

We apply a multiple imputation approach, as in Schafer (1997) and Little and Rubin (2002). This implies that we replace all missing values by simulated model predictions not just once, but S times, where S is the number of simulation runs. The S complete panels are then used to compute an average consumer confidence index, which accounts for imputation uncertainty and sampling uncertainty. It proves to be advantageous to impute forward in time, so that a realization of the past observations, $\hat{y}_{i,j,t-1}$, is always available by the time $y_{i,j,t}$ has to be imputed. This allows us to substitute the dynamic component of the model, $\rho_j^{d_{i,t}} y_{i,j,t-d_{i,t}}$, by $\hat{\rho}_j \hat{y}_{i,j,t-1}$.

In order to be able to compare the model imputations with the observed data, we first plot the consumer confidence indicator solely based on the imputations in Figure 4, panel (a). The imputed indicator is based on $S = 10,000$ imputed panels. A comparison between the imputed indicator and the observed indicator as displayed in Figure 3 shows that the imputed indicator resembles the observed indicator rather closely, also during the out-of-sample period. This reinforces our confidence in the model. In Figure 4, panel (b) we show the indicator based on the complete panel, that is, on the observed panel where all missing values are substituted by model imputations. Note that missing values are either due to nonresponse or due to the design of the panel. As nonresponse increases over time, typically, during the first weeks a smaller portion of the data is missing as compared to the final weeks. This explains why the confidence bounds around the observed indicator are tighter during the first weeks and less tight during the final weeks, whereas for the imputed indicator the opposite holds. As a consequence, the combined indicator relies more heavily on model imputations during the final weeks. This potential adverse effect can be avoided by applying a rotation strategy. Obviously, for the complete index to be correct, we do have to assume that our model is correctly specified. Whether it is to be preferred to use the complete index instead of the observed index, depends on the particular application of the indicator and one's belief in the imputation model.

- Insert Figure 4 about here -

3.5 Weekly Changes in Consumer Confidence

We conclude our analysis with an assessment of changes in consumer confidence. Recall that one of the advances of collecting panel data rather than repeated cross-section is that panel data allow us to exactly track changes in respondents' opinions over time at the individual level. This allows us, for example, to identify whether an increase in the index is due to respondents changing their opinions from negative to neutral, or from neutral to positive. On the other hand, suppose that the index does not change significantly, we may assess whether this is because respondents' opinions did not change, or whether they did change, but in such a way that the share of respondents who became more negative is equal to the share of respondents who became more positive. Such an assessment would not have been possible on the basis of repeated cross-sections. This type of polarization in opinions may however have important implications to policy makers.

In Table 4, we tabulate the transition rates from any previous state $y_{i,j,t-d_{i,t}}$ to the next state $y_{i,j,t}$, separately for our two cohorts. This table reveals that respondents changed their answers frequently. In fact, in 2005, 28% $(1 - 0.718)$ of the answers given to the five questions which together comprise a respondent's confidence level changed as compared to the previous answers given by the same respondent. Strikingly, in 2006, even 44% $(1-0.564)$ of the answers changed. Note that, based on the overall index in Figure 3, it would have been tempting to conclude that there were no significant changes in opinions during this period, which clearly is totally wrong.

The changes in opinions as tabulated in Table 4 occurred between time $t - d_{i,t}$ and time t, that is, between the previous and the present survey occasion. As in our panel design the time between subsequent waves is random, this complicates the interpretation of these changes. It would be more insightful to study weekly changes instead. For this purpose, we may again repeatedly impute the missing values in our panel data set, and calculate the transition rates separately for all imputed panels. The variation in these transition rates across the different imputed panels allows us to derive confidence bounds around the point estimates of the transition rates. Both the point estimates as well as the 95% confidence bounds are shown in Table 5. Clearly, the percentage of unchanged opinions in this figure is lower as compared to Table 4. Note that, effectively, in the latter table we showed the transition rates based on an average time between the past and present state of $1/f = 2.13$ weeks. As the percentages of unchanged answers is significantly lower at the weekly level, we conclude that individuals are more likely to change their opinions in the longer run.

Finally, we look at the developments over time in the transition rates. To visualize these it is convenient to classify the transitions into three categories, which are transitions towards a more positive state, transitions to the same state, and transitions towards a more negative state. In Figure 5, we indicate the shares of the transitions within each of these three categories by the white, light and dark grey area's, respectively. The dotted lines indicate the confidence bounds around the shares of positive and negative transitions. This figure illustrates perhaps most convincingly that, while the CCI was stable over time in 2006, there was a lot of variability in individuals' answers over this period. Especially in October and December

2006 we observe large up- and downswings. This could not have been noticed on the basis of repeated cross-sections.

- Insert Table 4 about here -

- Insert Table 5 about here -

- Insert Figure 5 about here -

4 Conclusions

In this paper we considered two angles for potential improvement of consumer confidence indicators. Firstly, we considered measuring consumer confidence at the weekly instead of the monthly level. Secondly, we considered collecting panel data rather than repeated cross-sections. This allowed us to measure and statistically test longitudinal changes in weekly consumer confidence. We verified that upon temporal aggregation these changes matched with the officially published ones, and obtained evidence that reliable and more insightful indicators may be constructed on the basis of relatively small panels rather than on larger repeated cross-sections.

There are various directions for further research. The first is that we can now correlate significant weekly changes with weekly observed macroeconomic variables, in order to study whether consumer confidence has predictive value. Indeed, currently most such studies concern monthly observed cross-sectional data, and it may well be that substantial information is lost. Secondly, we can use our techniques in other application areas such as customer monitoring in marketing. An application in finance would be to monitor the perceived conditions of financial markets with the aim to construct a financial barometer. Finally, it may be important to study whether the parameters of our model vary over time. This is particularly relevant if the model is used for imputation. Possibly it is necessary to regularly update the parameter estimates or to use a time-varying parameter version of the model. We may also consider enlarging our model with various other explanatory variables, as it is well known that factors as mood, temperature and specific events have an impact on consumer confidence.

A The Consumer Confidence Survey of Statistics Netherlands

As opposed to the consumer confidence indicator measured by the European Commission, the indicator measured by Statistics Netherlands not only concerns consumers' opinions on their financial situation, the economy in general, willingness to save and unemployment in the next twelve months, but also consumers' present situations and their opinions on the previous twelve months. The two indicators show roughly the same developments over time[1].

Consumer confidence is based on five questions from a more elaborate consumer survey. These questions are subdivided into a section on the economic climate and a section on the respondent's willingness to buy. The questions are formulated as follows:

Economic Climate

1. How do you think the general economic situation in this country has changed over the last twelve months?

   Possible answers: At present, it is better (1) / the same (0) / worse (−1)

2. How do you think the general economic situation in this country will develop over the next twelve months?

   Possible answers: It will be better (1) / the same (0) / worse (−1)

Willingness To Buy

3. How does the financial situation of your household now compare to what it was twelve months ago?

   Possible answers: At present, it is better (1) / the same (0) / worse (−1)

[1] See `http://www.cbs.nl/en-GB` for details.

4. How do you think the financial situation of your household will change over the next twelve months?

   Possible answers: It will be better (1) / the same (0) / worse (-1)

5. Do you think that at present there is an advantage for people to make major purchases, such as furniture, washing machines, TV sets, or other durable goods?

   Possible answers: Yes, now it is the right time (1) / It is neither the right nor the wrong time (0) / No, it is the wrong time (-1)

The economic climate indicator is computed as the percentage of positive answers minus that of negative answers, averaged over Questions 1 and 2. Similarly, the willingness to buy indicator is computed as the percentage of positive minus negative answers, averaged over Questions 3 to 5. Finally, the consumer confidence indicator is defined as the average of the economic climate indicator and the willingness to buy indicator.

B Tables and Figures

Table 1: A comparison of weekly CCI to monthly CCI

Month (Approx.)	Week	Randomized panel data: Weekly CCI: Average Level[1]		Weekly change	Biweekly CCI: Average Level[1]		Monthly change	Repeated cross-sections[2]: Monthly CCI: Average Level	Monthly change
Oct '05	41	27	(5.4)	–	23	(3.8)	–	-25	–
	42	19	(5.3)	-8					
	43	21	(5.2)	2	24	(3.6)	–		
	44	28	(4.5)	7					
Nov '05	45	26	(6.1)	-3	28	(4.4)	5	-21	4
	46	32	(6.1)	6					
	47	27	(6.6)	-6	26	(5.1)	2		
	48	25	(8.0)	-1					
Dec '05	49	32	(5.7)	6	30	(4.4)	2	-18	3
	50	28	(6.8)	-3					
	51	27	(7.4)	-1	28	(5.0)	2		
	52	29	(5.3)	2					
Jan '06	53	33	(6.3)	4	34	(5.5)	4	-12	6
	1	35	(7.3)	1					
Oct '06	40	42	(6.9)	–	42	(5.6)	–	5	–
	41	43	(7.8)	1					
	42	42	(6.5)	-1	40	(5.8)	–		
	43	37	(11.2)	-5					
Nov '06	44	38	(7.0)	1	36	(5.1)	-6	3	-2
	45	34	(7.3)	-4					
	46	37	(6.1)	3	36	(4.9)	-4		
	47	33	(8.1)	-3					
Dec '06	48	36	(4.3)	3	36	(4.1)	0	6	3
	49	36	(8.7)	0					
	50	38	(4.7)	2	37	(3.5)	1		
	51	36	(5.2)	-2					
Jan '07	52	37	(4.7)	1	38	(3.5)	4	15	9
	1	40	(5.4)	3					

[1] Standard errors are in parentheses. [2] Seasonally unadjusted data as reported by Statistics Netherlands.

Table 2: Estimation results of the dynamic Ordered Probit model with random effects

Variable		Model 1.	Model 2.	Model 3.	Model 4.	Model 5.
Panel design parameters						
No. of weeks since prev. request	β_1	0.011	−0.023	0.019	0.020	0.012
		(0.504)	(0.082)	(0.032)	(0.031)	(0.029)
No. of times requested before	β_2	0.003	−0.021	0.005	0.002	0.006
		(0.186)	(0.046)	(0.029)	(0.029)	(0.026)
Demographics						
Age	$\alpha_{2,1}$	−0.028	−0.012	−0.013	—	−0.014
		(0.124)	(0.022)	(0.014)		(0.011)
Gender	$\alpha_{2,2}$	0.175	0.240	0.248	—	0.214
		(0.850)	(0.199)	(0.164)		(0.138)
State dependence						
Question 1	ρ_1	0.814	0.857***	0.854***	0.858***	—
		(1.223)	(0.133)	(0.123)	(0.096)	
Question 2	ρ_2	0.715	0.764***	0.764***	0.762***	—
		(0.692)	(0.144)	(0.126)	(0.095)	
Question 3	ρ_3	0.847	0.875***	0.867***	0.868***	—
		(0.689)	(0.107)	(0.097)	(0.069)	
Question 4	ρ_4	0.900**	0.923***	0.917***	0.921***	—
		(0.357)	(0.106)	(0.096)	(0.077)	
Question 5	ρ_5	0.908*	0.902***	0.891***	0.898***	—
		(0.550)	(0.113)	(0.099)	(0.080)	
All questions	ρ	—	—	—	—	0.898***
						(0.043)
Additional model parameters						
Intercept	α_0	0.904	0.725	0.248	0.143	0.092
		(6.620)	(0.928)	(0.455)	(0.260)	(0.315)
Initial condition	α_1	0.281	0.282***	0.276***	0.280***	0.302***
		(0.770)	(0.069)	(0.057)	(0.056)	(0.054)
Threshold	γ_0	2.205***	2.169***	2.149***	2.149***	2.078***
		(0.288)	(0.058)	(0.053)	(0.053)	(0.042)
Variance of the random effects	σ_a^2	0.237***	0.272***	0.265***	0.284***	0.207***
		(0.532)	(0.074)	(0.054)	(0.056)	(0.053)
Question-specifics	λ_j	Included	Included	Included	—	—
Time-specifics	μ_t	Included	Included	—	—	—
Individual-specifics	κ_i	Included	—	—	—	—
Hit rate in-sample		0.712	0.716	0.727	0.722	0.713
Hit rate out-of-sample		0.546	0.556	0.607	0.538	0.548
Max. log-likelihood value		-963.1	-1007.8	-1014.1	-1016.8	-1041.4

*** Significant at the 1% level, ** at the 5% level, * at the 10% level. The estimates of the question-, time- and individual-specific parameters, as well as the $\boldsymbol{\alpha}_2$ parameters that measure the effects of the panel design parameters in all time periods on $y_{i,j,t}$, are not displayed for ease of presentation. Standard errors are in parentheses.

Table 3: Prediction-realization table

Observed		Cohort 2005: In-sample predictions Negative (−1)	Neutral (0)	Positive (1)	Total	Cohort 2006: Out-of-sample predictions Negative (−1)	Neutral (0)	Positive (1)	Total
Negative	(−1)	0.045	0.047	0.007	0.099	0.007	0.036	0.005	0.048
Neutral	(0)	0.028	0.377	0.090	0.494	0.020	0.325	0.169	0.513
Positive	(1)	0.007	0.094	0.305	0.406	0.002	0.162	0.274	0.439
Total		0.080	0.519	0.401	**0.727**	0.030	0.522	0.448	**0.607**

Note: The bold-faced numbers are the hit rates of the model, which are computed as the sum of the diagonal elements. Possible inconsistencies are due to rounding.

Table 4: Transition rates in the incomplete panel data set

From		Cohort 2005: To ($y_{i,j,t}$)				Cohort 2006: To ($y_{i,j,t}$)			
($y_{i,j,t-d_{i,t}}$)		Negative (-1)	Neutral (0)	Positive (1)	Total	Negative (-1)	Neutral (0)	Positive (1)	Total
Negative	(-1)	0.060	0.043	0.013	0.116	0.010	0.048	0.006	0.064
Neutral	(0)	0.032	0.353	0.088	0.474	0.031	0.279	0.158	0.468
Positive	(1)	0.008	0.098	0.305	0.410	0.003	0.189	0.276	0.469
Total		0.099	0.494	0.406	**0.718**	0.045	0.516	0.440	**0.564**

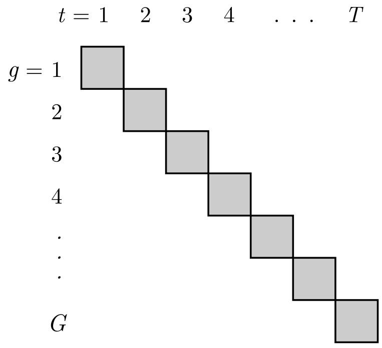
Note: The bold-faced numbers are the fractions of unchanged answers, which are computed as the sum of the diagonal elements. Possible inconsistencies are due to rounding.

Table 5: Transition rates in the imputed panel data set

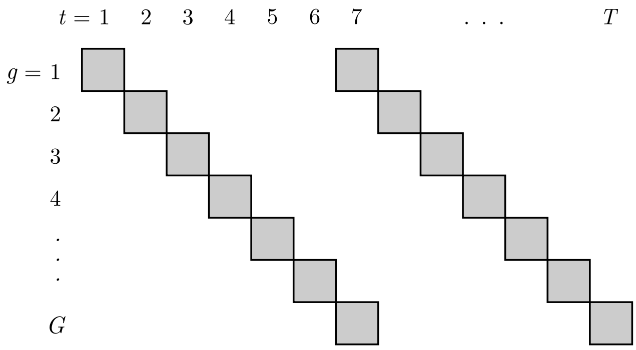
From ($y_{i,j,t-1}$)		Cohort 2005: To ($y_{i,j,t}$) Negative (-1)	Neutral (0)	Positive (1)	Total	Cohort 2006: To ($y_{i,j,t}$) Negative (-1)	Neutral (0)	Positive (1)	Total
Negative	(-1)	0.081	0.027	0.009	0.117	0.039	0.032	0.007	0.078
		(0.070, 0.093)	(0.016, 0.038)	(0.004, 0.015)	(0.090, 0.146)	(0.031, 0.048)	(0.021, 0.0440)	(0.002, 0.013)	(0.054, 0.105)
Neutral	(0)	0.023	0.430	0.059	0.512	0.023	0.367	0.107	0.498
		(0.013, 0.034)	(0.400, 0.459)	(0.043, 0.074)	(0.456, 0.568)	(0.014, 0.033)	(0.334, 0.400)	(0.089, 0.126)	(0.437, 0.559)
Positive	(1)	0.008	0.050	0.314	0.372	0.009	0.091	0.324	0.425
		(0.003, 0.014)	(0.034, 0.066)	(0.290, 0.338)	(0.326, 0.418)	(0.003, 0.017)	(0.071, 0.110)	(0.294, 0.354)	(0.369, 0.481)
Total		0.112	0.506	0.381	**0.825**	0.071	0.490	0.439	**0.731**
		(0.086, 0.141)	(0.450, 0.564)	(0.336, 0.427)	**(0.760, 0.891)**	(0.048, 0.098)	(0.426, 0.554)	(0.386, 0.493)	**(0.659, 0.802)**

Note: The lower and upper 95% confidence bounds of the transition rates are shown in parentheses. The bold-faced numbers are the fractions of unchanged answers, which are computed as the sum of the diagonal elements. Possible inconsistencies are due to rounding. The results are based on 10,000 imputed panels.

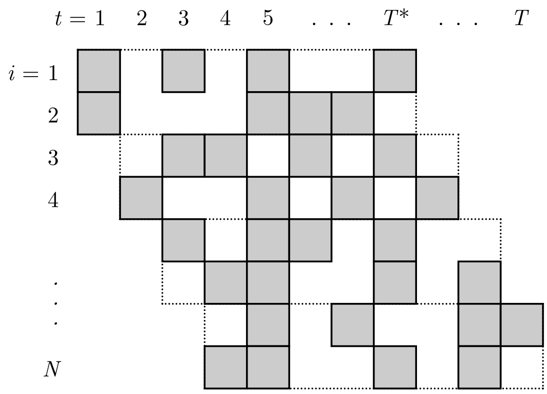
Figure 1: Panel designs that are applied in practice to measure consumer confidence



(a) Repeated cross-sections

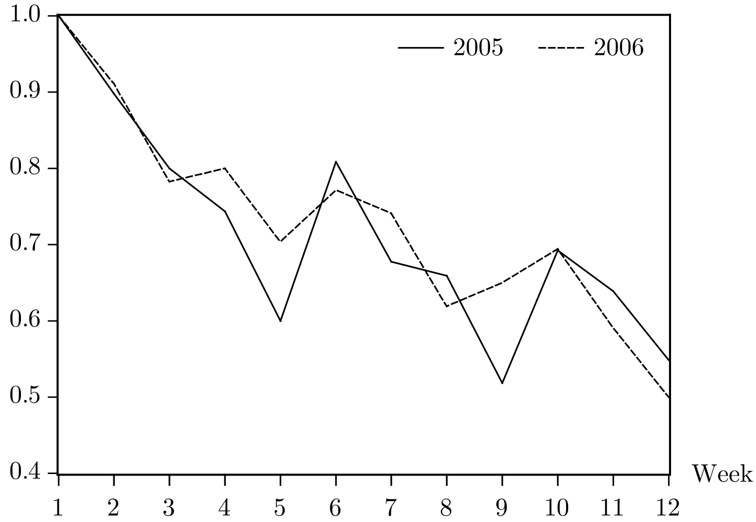


(b) The Michigan panel



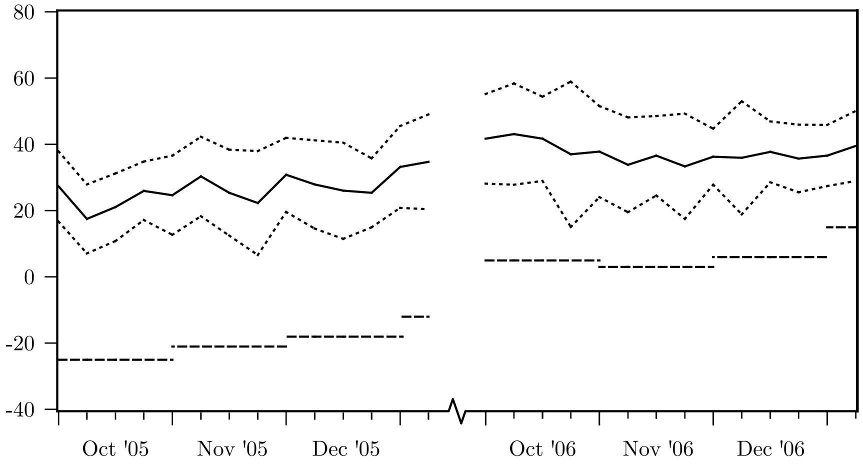
(c) An example of a randomized rotating panel

Figure 2: Weekly response rates



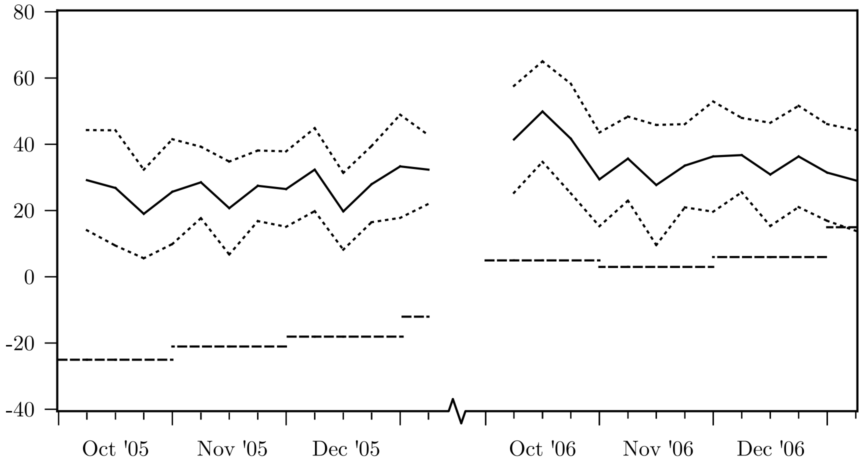
Note: This table presents the weekly fractions of panel members who completed the consumer confidence survey in Appendix A on our request.

Figure 3: A comparison of weekly CCI to monthly CCI

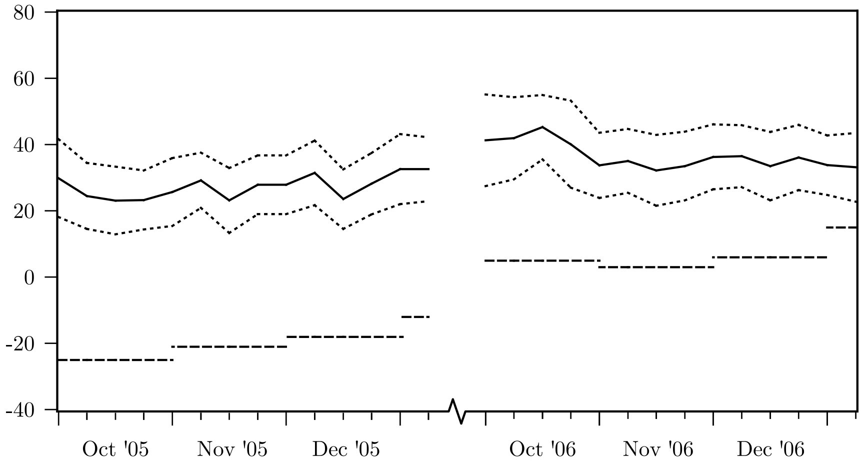


Note: This graph compares the weekly consumer confidence indicator as measured in this paper (solid line) to the monthly indicator as measured by SN (dashed line). The dotted lines represent the upper and lower 95% confidence bounds of the weekly indicator.

Figure 4: A comparison of weekly CCI using model imputations to monthly CCI



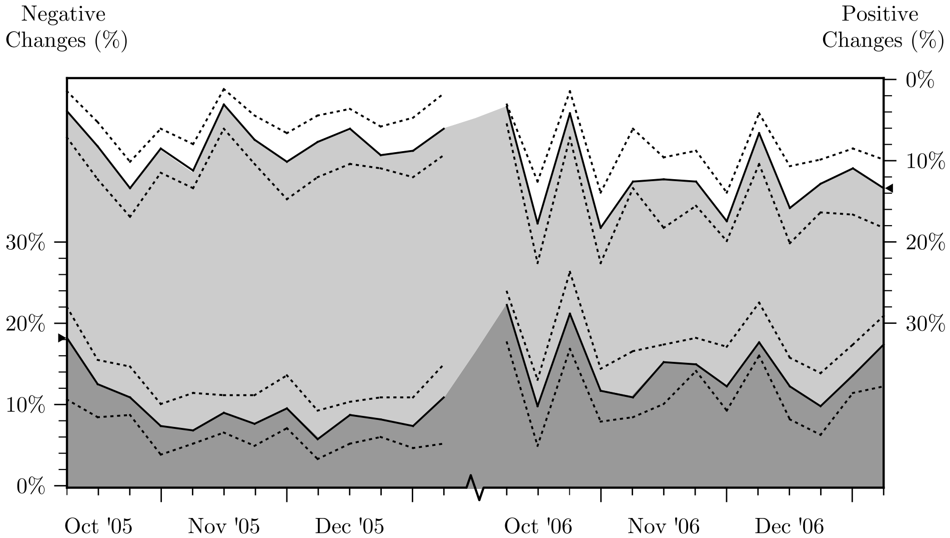
(a) Model imputations



(b) Combination of observed data and model imputations

Note: In panel (a) the model imputations of the weekly consumer confidence indicator (solid line) are compared to the monthly indicator as measured by SN (dashed line). Similarly, in panel (b) we compare the composite weekly consumer confidence indicator, which is composed of the observed data as summarized in Figure 3 and the imputed data as summarized in panel (a), to the monthly indicator as measured by SN. The dotted lines represent the upper and lower 95% confidence bounds of the weekly indicators. The results are based on 10,000 imputed panels.

Figure 5: Developments in weekly changes in consumer confidence over time



Note: This figure displays the share of respondents over time who changed their answer in the positive direction (white area, values on the right axis) and in the negative direction (dark grey area, values on the left axis), as compared to their (imputed) answer in the previous week. The share of respondents who did not change their answer is indicated by the light grey area. The dotted lines represent the upper and lower 95% confidence bounds of the shares of positive and negative answers. The results are based on 10,000 imputed panels.

References

Ansari, A., C. F. Mela, and S. A. Neslin (2008), Customer Channel Migration, *Journal of Marketing Research*, **45**, 60–76.

Curtin, R. (1982), Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers, *Public Opinion Quarterly*, **46**, 340–352.

Dominitz, J. and C. F. Manski (2004), How Should We Measure Consumer Confidence?, *Journal of Economic Perspectives*, **18**, 51–66.

Erdem, T. and B. Sun (2001), Testing for Choice Dynamics in Panel Data, *Journal of Business & Economic Statistics*, **19**, 142–152.

Franses, P. H. B. F. and R. Paap (2001), *Quantitative Models in Marketing Research*, Cambridge University Press, Cambridge, UK.

Hansen, R. A. (1980), A Self-Perception Interpretation of the Effect of Monetary and Nonmonetary Incentives on Mail Survey Respondent Behavior, *Journal of Marketing Research*, **17**, 77–83.

Heckman, J. J. (1981), The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process, in C. F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, pp. 179–195.

Katona, G. (1951), *Psychological Analysis of Economic Behavior*, McGraw Hill, New York.

Keane, M. P. (1997), Modeling Heterogeneity and State Dependence, *Journal of Business & Economic Statistics*, 310–327.

King, G., O. Rosen, and M. A. Tanner (2004), *Ecological Inference: New Methodological Strategies*, Cambridge University Press, New York.

King, G. (1997), *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton University Press, Princeton.

Kish, L. and I. Hess (1959), A "Replacement" Procedure for Reducing the Bias of Nonresponse, *The American Statistician*, **13-4**, 17–19.

Little, R. J. A. and R. B. Rubin (2002), *Statistical Analysis with Missing Data*, second edn., John Wiley & Sons.

Ludvigson, S. C. (2004), Consumer Confidence and Consumer Spending, *Journal of Economic Perspectives*, **18**, 29–50.

McKelvey, R. D. and W. Zavoina (1975), A Statistical Model for the Analysis of Ordinal Level Dependent Variables, *Journal of Mathematical Sociology*, **4**, 103–120.

Moffitt, R. (1993), Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-sections, *Journal of Econometrics*, **59**, 99–123.

Patterson, H. D. (1950), Sampling on Successive Occasions with Partial Replacement of Units, *Journal of the Royal Statistical Society Series B*, **49**, 241–255.

Schafer, J. L. (1997), *Analysis of Incomplete Multivariate Data*, Chapman and Hall, New York.

Segers, R. and P. H. B. F. Franses (2007), Panel Design Effects on Response Rates and Response Quality, *Econometric Institute Report Series*, **EI 2007-29**, Erasmus University Rotterdam.

Sigelman, L. (1991), Turning Cross Sections into a Panel: A Simple Procedure of Ecological Inference, *Social Science Research*, **20**, 150–170.

Wooldridge, J. M. (2005), Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity, *Journal of Applied Econometrics*, **20**, 39–54.