Seiler, Christian

Working Paper

On the Robustness of the Balance Statistics with respect to Nonresponse

ifo Working Paper, No. 126

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

ifo Working Papers

On the Robustness of the Balance Statistics with respect to Nonresponse

Christian Seiler

Ifo Working Paper No. 126

March 2012

Ifo Working Paper No.126

On the Robustness of the Balance Statistics with respect to Nonresponse

Abstract

Business cycle indicators based on the balance statistics are a widely used method to monitor the actual economic situation. In contrast to official data, indicators from business surveys are early available and typically not revised after their first publication. But as surveys can be in general affected by distortions through the response behaviour, these indicators can also be biased. In addition, time-dependent nonresponse patterns can produce even more complex forms of biased results. This paper examines a framework which kind of nonresponse patterns lead to biases and decreases in performance. We perform an extensive Monte Carlo study to analyse their effects on the indicators. Our analyses show that these indicators are extremely stable towards selection biases.

JEL Code: C81, C83.
Keywords: Business survey, Monte Carlo study, nonresponse.

Christian Seiler
Ifo Institute – Leibniz Institute
for Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
Phone: +49(0)89/9224-1248
seiler@ifo.de

1 Introduction

Monitoring and forecasting of economic activity nowadays is one of the main topics for public as well as private institutions. As official data are commonly released with high delay, timely business cycle indicators are needed. Balance statistics indicators based on survey data play a central role in this context. Its major advantages are that they are available in real-time and are almost not subject to any revisions. Indeed, aggregated business survey data has proved to be one of most competitive indicators for analysing macroeconomic variables, e.g. the Consumer Confidence Index in the United States (Ang et al., 2007), the Economic sentiment indicator for the European Union (Gayer, 2005) or the Ifo Business Climate Index for Germany (Drechsel and Scheufele, 2010, Kholodilin and Siliverstovs, 2006 and Robinzonov and Wohlrabe, 2010)

Most papers using survey based business cycle indicators (usually not explicitly) assume that the indicator is measured without any selection bias, i.e. the survey results are not affected by any bias through nonresponse. While there exists a large literature concerning nonresponse in household or individual surveys, less is known about the processes and reasons for participation and responding behaviour in business surveys, see Janik and Kohaut (2012), especially in cases of surveys with the purpose to evaluate the business cycle. Therefore, it can not generally be stated that the indicators are unbiased. For example, Seiler (2010) found evidence for a dependence of the responding behaviour in the Ifo Business Survey (IBS) from the business cycle which confirmed the results of Harris-Kojetin and Tucker

(1999) for population surveys. So, the questions arises how a potential selection bias affects the results of these type of business surveys. How do indicators look when they are biased by systematic nonresponse patterns? How strongly is the performance reduced?

To give answers on these questions, the paper is organised as follows: In Section 2 we define the methodologic framework for the calculation of the balance statistics and show how selection biases affect the indicators. Afterwards, we introduce measurements to explore the biases. In Section 3, we perform an extensive Monte Carlo study for a wide variety of nonresponse patterns and different types of business cycles. We show that the bias is minimal even for very strong bias structures in the data. Finally, Section 4 sums up the results.

2 Methodological framework

In market economies, one can typically observe that the general long-term growth underlies temporary fluctuations. In general, the classical business cycle is divided into four phases: upswing, boom, downswing and recession. In ideal form, the cycle would have a sinus-like shape. The business cycle itself is a theroetical construct which can not be measured directly, but can be visualised by observing economic indicators, such as GDP growth rates or production in the manufacturing sector. We can think of a business cycle $g(t)$ as a stationary function in time, e.g. $g(t) = \sin(t)$. Due to the fact that official data is published with high delay (and, in addition, commonly revised after the first publication), business cycle tendency surveys can monitor the actual economic situation considerably quicker. The Ifo Institute for Economic Research was one of the first which conducted such surveys and this method has been accepted widely in the OECD countries, for an overview see OECD (2003). In line with the Joint Harmonised EU Programme of Business and Consumer Surveys (see European Union, 2006), the indicators base on two variables (business situation and business expectations) which are measured on a 3-level-Likert scale representing a good, equal or bad state, i.e. $y \in S = \{+, =, -\}$. Due to the construction of the questions in the questionnaire, the indicators obtained in fact measure the business cycle without trend, see OECD (2003). For more theoretical aspects of the balance statistics see Anderson (1951,1952) and Theil (1952).

For the current study, we abstract from any formal definition of the business cycle. We define a macroeconomic time series for which a qualitative survey with a 3-level Likert scale is constructed or referenced to. Usually, a variable on a 3-level Likert scale is influenced by an unobserved process (e.g. the business cycle in our case). However, also continuous variables may be asked in the survey on a 3-level Likert scale which realisations are only available with high delay, e.g. the World Economic Survey of the Ifo Institute asks for inflation in the appropriate country (see Stangl, 2007). In this section, we leave the exact shape of $g(t)$ undefined, but will show the effects for different types of $g(t)$ in Section 3.

2.1 Construction of the balance statistics

All surveys asking on a 3-level Likert scale mentioned above calculate a so-called *balance statistics* after data collection. The balance statistics is defined as the fraction of positive replies minus the fraction of negative replies. We assume that respondent i is affected in his opinion formation by the cycle function $g(t)$, for which the survey is conducted, and an individual error term ϵ_i:

$$y_i^* = g(t) + \epsilon_i, \qquad \epsilon_i \sim N(0, \sigma^2), \tag{1}$$

$i = 1, \ldots n$, with

$$\mathrm{E}_t(y_i^*) = g(t), \qquad \forall i. \tag{2}$$

As each survey participant i is restricted to give answers on a 3-level Likert-scale, we observe

$$y_i = \begin{cases} + & \text{if} \quad y_i^* > \tau^+ \\ = & \text{if} \quad \tau^+ > y_i^* > \tau^- \\ - & \text{if} \quad \tau^- > y_i^* \end{cases}$$

where $\tau^s, s \in S = \{+, =, -\}$, defines time-independent thresholds. This is the standard model for latent variable models, see Bartholomew and Knott (1999). The assumption of time- and individual-independent thresholds as well as time- and individual-independent normality of ϵ_i is required for the calculation of the balance statistics. In addition, τ^s are assumed to be symmetric around zero, i.e. $-\tau^- = \tau^+$ (see Wollmershäuser and Henzel, 2005). Therefore, the balance statistics is a special case according to the Carlson-Parkin framework (Carlson and Parkin, 1975) which allows for differences for thresholds and error terms in time and across individuals. These very strong assumptions have been criticised by several papers, see Nardo (2003) for an overview. However, as these thresholds are 'contained' in the cumulative distribution function of ϵ, we do not go more into detail as this is less an issue regarding nonresponse. For a discussion on the shifts of thresholds τ^s, see Stangl (2009). The same applies to the distribution of ϵ_i. Of course, the distribution of ϵ can be generalised to a wider class. Due to the fact that these extensions would affect the biased indicator as well as the unbiased indicator in the same way and this paper is focused on a relative comparison between unbiased and biased indicators, we consider τ^s and $\epsilon_i = \epsilon$ as independent from time and individuals.

To enable the calculation of means, we define '+' $\equiv$ 1, '=' $\equiv$ 0 and '-' $\equiv -1$. Then, the mean of y for given t is defined as

$$\begin{aligned}
\mathrm{E}_t(y) = \mathrm{E}_t(y_i) &= \int_y y_i \cdot f(y_i)\, dy \qquad (3)\\
&= \sum_s y_i \cdot p^s \\
&= 1 \cdot P(y_i^* > \tau^+) + (-1) \cdot P(y_i^* < \tau^-) \\
&= P(g(t) + \epsilon_i > \tau^+) - P(g(t) + \epsilon_i < \tau^-) \\
&= P(\epsilon_i > \tau^+ - g(t)) - P(\epsilon_i < \tau^- - g(t)) \\
&= [1 - P(\epsilon_i < \tau^+ - g(t))] - P(\epsilon_i < \tau^- - g(t)) \\
&= \left[1 - \Phi\left(\frac{\tau^+ - g(t)}{\sigma}\right)\right] - \Phi\left(\frac{\tau^- - g(t)}{\sigma}\right)
\end{aligned}$$

with $p^s = P(y = s)$ and $\Phi(\cdot)$ as the cumulative distribution function of the standard normal distribution.[1] $\mathrm{E}_t(y) = \mathrm{E}_t(y_i)$ in equation (3) is given since $\left[1 - \Phi\left(\frac{\tau^+ - g(t)}{\sigma}\right)\right] - \Phi\left(\frac{\tau^- - g(t)}{\sigma}\right)$ includes no individual information i. In the following, we call $\mathrm{E}_t(y)$ the unbiased mean function or the unbiased indicator. In appendix A we show that (3) is the same as the balance statistic for a 3-level Likert-scale.

2.2 Inclusion of nonresponse

The mean function $\mathrm{E}_t(y)$ in equation (3) is the expected function in time for $n \to \infty$ when no selection bias is present. To include the decision of the respondent to participate in the survey, we define a binary variable m

[1] We abstract from different types of weighting that may occur in real data sets.

which indicates this decision, so that $m = 1$ if the observation is missing and $m = 0$ if not. According to Little and Rubin (2002), not missing at random (NMAR), i.e. a selection bias, occurs if the probability for missing $P(m = 1)$ depends on the outcome of the variable of interest $y = s$, i.e. $P(m = 1|s, \xi)$, with ξ as additional parameters which are unrelated to the variable of interest. Therefore, the missing process would depend on the state s in our case. To evaluate the effects of the bias patterns on the balance statistics, we assume that every respondent has the same probability not to respond to the survey for given s and t, i.e. $P(m = 1|s, t, \xi) = P(m = 1|s, t), \forall i$. Then, $\pi^s(t) = P(m = 0|y = s, t)$ is the (not necessarily time-dependent) probability to be observed at time t being in state s. Therefore, the mean function of the *observed units* is given by

$$
\begin{aligned}
\mathrm{E}_t(y_{obs}) &= \int_y y_{obs} \cdot f(y_{obs})\, dy \qquad (4) \\
&= \int_y y_{obs} \cdot f(y|m = 0)\, dy \\
&= \int_y y_{obs} \cdot \frac{f(m = 0|y) \cdot f(y)}{f(m = 0)}\, dy \\
&= \sum_s y_{obs} \cdot \frac{P(m = 0|y = s) \cdot P(y = s)}{P(m = 0)} \\
&= \sum_s y_{obs} \cdot \frac{P(m = 0|y = s) \cdot P(y = s)}{\sum_s P(m = 0|y = s) \cdot P(y = s)} \\
&= \frac{\pi^+(t) \cdot [1 - \Phi((\tau^+ - g(t))/\sigma)]}{\sum_s \pi^s(t) \cdot \Phi^s(t)} - \frac{\pi^-(t) \cdot \Phi((\tau^- - g(t))/\sigma)}{\sum_s \pi^s(t) \cdot \Phi^s(t)} \\
&= \frac{\pi^+(t)}{\bar{\pi}(t)} \cdot \Phi^+(t) - \frac{\pi^-(t)}{\bar{\pi}(t)} \cdot \Phi^-(t)
\end{aligned}
$$

with $\bar{\pi}(t) := \sum_s \pi^s(t) \cdot \Phi^s(t)$, $\Phi^+(t) := 1 - \Phi((\tau^+ - g(t))/\sigma)$, $\Phi^-(t) := \Phi((\tau^- - g(t))/\sigma)$, $\Phi^=(t) := 1 - [\Phi^+(t) + \Phi^-(t)]$ and $0 \leq \pi^s(t) \leq 1$. Note that $\pi^s(t)$ is the acceptance rate being in state s at time t and $\Phi^+(t)$ and $\Phi^-(t)$ directly depend on $g(t)$ and beyond only on time-invariant variables. Therefore, NMAR occurs in cases of $\exists t \in \{1, \ldots, T\} \cap \exists (r,s) : \pi^r(t) \neq \pi^s(t)$ with $r, s \in S$. Note that the probability $\pi^=(t)$ for the center category enters the mean function $\mathrm{E}_t(y_{obs})$ in $\bar{\pi}(t) = \sum_s \pi^s(t) \cdot \Phi^s(t)$. Thus, $\mathrm{E}_t(y_{obs})$ is the mean function in t of our (potentially biased) indicator for $n \to \infty$. Therefore, $\mathrm{E}_t(y_{obs})$ is designated as the observed or biased mean function or the observed or biased indicator. In addition, $\mathrm{E}_t(y)$ denotes the unbiased case of $\mathrm{E}_t(y_{obs})$.

2.3 Theoretical considerations

When inspecting selection biases and their effects on the observed indicator $\mathrm{E}_t(y_{obs})$, we have to define which type of bias we want to analyse. The first 'natural' comparison would be to evaluate the difference between the observed indicator $\mathrm{E}_t(y_{obs})$ and the unbiased mean function $\mathrm{E}_t(y)$, e.g. by (absolute) differences. However, we have two notes on this issue: First, survey indicators which are constructed by a balance statistics are, at least, artifical. In constrast to e.g. official data, their level does not reflect a certain quantity. In particular, this holds for latent variables such as the 'business situation'. At best, a positive/negative value indicates an increase/decrease. Second, these indicators are constructed to display the cyclical development $g(t)$ for different economic parameters over time. A shift of the whole time series $\mathrm{E}_t(y)$ or a stretch or a compression by a constant factor would leave the relationship between $\mathrm{E}_t(y)$ and the cycle function $g(t) = \mathrm{E}_t(y^*)$

unaffected. Therefore, the main subject of investigation is (in contrast to most bias analyses) not the level but the correlation between a (possibly) biased indicator and the cycle function. We consider two different correlations to inspect: The correlation $\rho(\mathrm{E}_t(y_{obs}), g(t))$ between the observed (biased) indicator $\mathrm{E}_t(y_{obs})$ and the cycle function $g(t)$ and the correlation $\rho(\mathrm{E}_t(y_{obs}), \mathrm{E}_t(y))$ between the observed (biased) indicator $\mathrm{E}_t(y_{obs})$ and the (unbiased) mean function $\mathrm{E}_t(y)$.

2.3.1 Correlation with the cycle function

Before we start to evaluate the effects on the correlation with the cycle function $g(t)$, we have to think about which correlation can be obtained in the unbiased case. We define

$$\rho(\mathrm{E}_t(y), g(t)) \underset{n,t \to \infty}{=} \rho_g^{unbiased}, \tag{5}$$

where $\rho_g^{unbiased}$ denotes the maximum correlation in cases of an unbiased mean function $\mathrm{E}_t(y)$. As Pearsons correlation coefficient measures the linear relationship between two variables, it is only invariant towards linear transformations $h(\cdot)$. Since $\mathrm{E}_t(y)$ only depends on the cycle function $g(t)$ and $\Phi(\cdot)$ is continuous but not linear in equation (3), $\rho^P_{unbiased} < 1$. This result is not surprising since a 3-level trait always includes less information. For this reason, we rescale ρ_g^{biased} to

$$\tilde{\rho}_g^{biased} = \frac{\rho_g^{biased}}{\rho_g^{unbiased}} \tag{6}$$

with $\rho^P_{biased} = \rho^P_{biased}(\mathrm{E}^{bias}_t(y_{obs}), g(t))$ and $\rho^P_{unbiased} = \rho^P_{unbiased}(\mathrm{E}_t(y), g(t)) < 1$. We notice that $\tilde{\rho}^{biased}_g$ is not a real correlation, but the adjustment in equation (6) is necessary to display the differences in correlation. However, in nearly all of the cases $\rho^{unbiased}_g$ is close to 1. The nonlinearity arises from the cumulative density function of ϵ. Linearity can only be obtained in cases of a rectangular distribution $\epsilon \sim U(c_l, c_u)$. But this might be unrealistic in most settings. Stangl (2009) evaluated both variables of the Ifo index on a visual analog scale and found evidence that the distribution of y^* is similar to a bell curve. A consequence of $\rho^{unbiased}_g < 1$ is that an observed biased indicator $\mathrm{E}_t(y_{obs})$ may have a higher correlation with the cycle function $g(t)$ than the unbiased indicator $\mathrm{E}_t(y)$. We can not analytically derive in which cases this effect would appear but we will show later that a perfect correlation can never be obtained.

The correlation $\rho(\mathrm{E}_t(y_{obs}), g(t)) = 1$ if

$$\begin{aligned} \mathrm{E}_t(y_{obs}) &= a \cdot g(t) + b \qquad (7) \\ \frac{1}{\bar{\pi}(t)} \cdot \big(\pi^+(t) \cdot \Phi^+(t) - \pi^-(t) \cdot \Phi^-(t)\big) &= a \cdot g(t) + b. \end{aligned}$$

Assuming that $a \neq 0$, equation (7) can be written as

$$\left(\frac{\pi^+(t)}{\bar{\pi}(t)} \cdot \Phi^+(t) = a^+ \cdot g(t) + b^+\right) \cap \left(\frac{\pi^-(t)}{\bar{\pi}(t)} \cdot \Phi^-(t) = a^- \cdot g(t) + b^-\right) \qquad (8)$$

with $a^+, a^- \neq 0$. It is important to notice that a, a^+, a^-, b, b^+ and b^- have to be constant and therefore independent from t. Unfortunately, a perfect correlation of 1 can not be received for any NMAR process: As $\Phi^+(\cdot)$ and $\Phi^-(\cdot)$

are nonlinear if $\epsilon \nsim U(c_l, c_u)$, linearity in $\Phi^s(\cdot)$ can only be obtained directly via the inverse of $\Phi^s(\cdot)$. As the $\pi^s(t)$ have a *multiplicative effect* on $\Phi^s(\cdot)$ in equation (4), this way for correction is not possible. Another approach to obtain a linear relationship would be to constrain $\pi^+(t) = a \cdot g(t) + b$ or $\pi^-(t) = a \cdot g(t) + b$ and to oblige that $\Phi^+(t)$ *and* $\Phi^-(t)$ are reduced to time-invariant constants via division by $\bar{\pi}(t)$. But because we have stated above that $\pi^+(t)$ or $\pi^-(t)$ have to be a linear transformation of $g(t)$ and they are also included in $\bar{\pi}(t)$, $\Phi^+(t)$ and $\Phi^-(t)$ can not be changed simultaneously into time-invariant effects. Therefore, it is not possible to obtain a perfect correlation regardless of the type of $\pi^s(t)$. However, this fact does not exclude the possibility of receiving a higher correlation for a biased indicator $\mathrm{E}_t(y_{obs})$. In these cases, the functions $\frac{\pi^+(t)}{\bar{\pi}(t)}$ and $\frac{\pi^-(t)}{\bar{\pi}(t)}$ eliminate the bias of $\Phi(\cdot)$.

2.3.2 Correlation with the unbiased indicator

Similarly, we can derive conditions for a perfect correlation between the observed mean function $\mathrm{E}_t(y_{obs})$ and the 'ideal', unbiased mean function $\mathrm{E}_t(y)$. The correlation $\rho(\mathrm{E}_t(y_{obs}), \mathrm{E}_t(y)) = 1$ holds if

$$\begin{aligned}
\mathrm{E}_t(y_{obs}) &= a \cdot \mathrm{E}_t(y) + b && (9)\\
\frac{1}{\bar{\pi}(t)} \cdot \left(\pi^+(t) \cdot \Phi^+(t) - \pi^-(t) \cdot \Phi^-(t)\right) &= a \cdot [\Phi^+(t) - \Phi^-(t)] + b \\
\frac{\pi^+(t)}{\bar{\pi}(t)} \cdot \Phi^+(t) - \frac{\pi^-(t)}{\bar{\pi}(t)} \cdot \Phi^-(t) &= a \cdot \Phi^+(t) - a \cdot \Phi^-(t) + b.
\end{aligned}$$

As before, a and b have to be time-independent. This is only the case, if $\pi^+(t) = \pi^-(t) \cap \pi^+(t)/\bar{\pi}(t) \perp t \cap \pi^-(t)/\bar{\pi}(t) \perp t$. Since $\bar{\pi}(t) = \sum_s \pi^s(t) \cdot \Phi^s(t)$ and $\sum_s \Phi^s(t) = 1, \forall t$, we can conclude that $\rho^P(\mathrm{E}_t(y_{obs}), \mathrm{E}_t(y)) = 1$ holds, if $\pi^+(t) = \pi^=(t) = \pi^-(t)$, i.e. nonresponse is MAR. This result is essential: It shows that any type of NMAR would decrease this type of correlation and therefore the linear correlation between $\mathrm{E}_t(y_{obs})$ and $\mathrm{E}_t(y)$.

In the statements made above, we focused on the relationship between our observed (and potentially biased) indicator $\mathrm{E}_t(y_{obs})$ and the cycle function $g(t)$ which is the main interesting part. However, we also display differences in correlation between a biased indicator $\mathrm{E}_t(y_{obs})$ and the 'ideal', unbiased indicator $\mathrm{E}_t(y)$. In contrast to the correlation between $\mathrm{E}_t(y)$ and $g(t)$, we do not need any adjustment for this correlation. Therefore, we analyse the following correlations when inspecting possible bias patterns for observed data: $\tilde{\rho}_g = \rho(\mathrm{E}_t(y_{obs}), g(t))/\rho_g^{unbiased}$ to inspect the (adjusted) linear relationship between the observed (potentially biased) indicator and the cycle function and $\rho_E = \rho(\mathrm{E}_t(y_{obs}), \mathrm{E}_t(y))$ to inspect the linear relationship between the biased and the unbiased indicator.

3 Simulation study

3.1 Definition of cycle functions

Before we run our Monte Carlo study, we have to specify our cycle functions $g(t)$. As noted in Section 3.1, the ideal case of a business cycle in a sinus function in time, i.e. $\sin(t)$. In addition, historical business cycle theory identifies four types of operlapping cycles: Kitchin (3-5 years), Juglar (7-11), Kuznets (15-25) and Kondratiev (45-60).[2] We first want to inspect the effects of our acceptance functions $\pi^s(t)$ on a simple sinus-function for the cycle $g(t)$, so $g(t) \approx \sin(t)$. We refer t to represent months and define our simple sinus-function as a Juglar cycle with 10 years, i.e. 120 months. So, we have to rescale $\sin(t)$ by a constant k to $\sin(t/k)$. The first full cycle is reached at 2π, so $\sin(120/k) \approx \sin(2\pi)$.[3] With this, $k \approx 120/2\pi \Rightarrow k = k_{Jug} \approx 20$. Then, our first cycle $g(t)$ to inspect is defined as

$$g_1(t) := \sin(t/k_{Jug}) = \sin(t/20).$$

Second, we define a cycle function $g(t)$ including all four cycles mentioned above. One full Kitchin 4-years cycle covers 48 months, one Kuznets 20-years cycle 240 months and one Kondratiev 50-years cycle 600 months. As above, the scaling parameters k turn to: $k_{Kit} \approx 8$, $k_{Kuz} \approx 38$ and $k_{Kon} \approx 95$.

[2]For an actual study to the 2008-2009 economic crisis with respect to the different cycles see Korotayev and Tsirel (2010).

[3]We notice that this procedure is inaccurate but the cycles, as defined in the literature, do also not have an exact length. The cycles in this section are only selected for representation of the selection bias effects.

Therefore, the cycle function is defined as

$$\begin{aligned} g_2(t) &:= \sin(t/k_{Kit}) + \sin(t/k_{Jug}) + \sin(t/k_{Kuz}) + \sin(t/k_{Kon}) \\ &= \sin(t/8) + \sin(t/20) + \sin(t/38) + \sin(t/95) \end{aligned}$$

Of course, $g_1(t)$ and $g_2(t)$ display a very ideal case of a business cycle which is usually not observed in reality. To come closer to a more relatistic process, we define a third cycle function $g_3(t)$ by an AR(1)-process

$$g_3(t) := 0.9 \cdot g_3(t-1) + u_t$$

with $g_3(0) = 0$ and $u_t \sim N(0,1)$. The series $g_3(t)$ can be regarded as a monthly growth rate. One example for such an latent process is the inflation question of Ifo's World Economic Survey, which is also asked on a 3-level Likert scale. Our fourth cycle function is based on a real data example, the U.S. industrial production. We define $g_4(t)$ by extracting the output gap with the Hodrick-Prescott band pass filter for time series (see Hodrick and Prescott, 1997).

All cycle functions $g_1(t), g_2(t), g_3(t)$ and $g_4(t)$ are displayed in Figure 1 for $T = 500$. For an easier comparison, we standardise all cycle functions to $g_j(t) \in [-1, 1]$. The thresholds τ^s are defined as $\tau^+ = 1/3$ and $\tau^- = -1/3$ so that the range of every function is divided into three parts of equal size which fit the assumptions for the calculation of the balance statistics in Section 2.1.

3.2 Definition of response functions

After defining cycle functions $g_j(t)$ we introduce different nonresponse patterns defined by the acceptance functions $\pi^s(t)$. Of course, the number of different functions $\pi^s(t)$ is infinite (and thus the number of combinations), as well as the number of cycle functions $g_j(t)$. Nevertheless, we examine the effects of some main types of $\pi^s(t)$ on the correlations and perform a Monte Carlo study on this.

We focus on fixed (time-independent), cycle-dependent, cycle-shifted, monotone and random types of nonresponse patterns in t. All time-independent structures describe general differences in the responding behaviour when being in the appropriate state. For example, one can imagine that the respond rate is higher when the firms' situation is bad and the company wants to complain about that. Also the other effect can occur, when the firms' state is good and they want to tell about that. Trends, nonlinear as well as linear, may be possible as the nonresponse rates in general raised throughout the last decades (see, for example, Brehm, 1994). So, it is possible that these trends can also be different according to the state of the company. Maybe, the probability to respond has been decreased for those firms which are in a good situation. A dependence of the overall respond rate from the business cycle in general was found by Seiler (2010) for the Ifo Business Survey and Harris-Kojetin and Tucker (1999) for the U.S. Current Population Survey. Although nothing is known directly about the state of the company, but it would be as conceivable that the company has a higher probability for nonresponse when the firm remarks that their situation is bad in relation to the market. Also a dependence from the cycle, as well as cycle shifts may occur as the decision to respond may be affected by the latent variable behind. For robustness checks, we also include random probabilities for $\pi^s(t)$. Regardless of the type of $\pi^s(t)$, we have to define $0 \leq \pi^s(t) \leq 1, \forall t$, to enable a correct specified probability function.

3.3 Monte Carlo setup

As we are able to evaluate the mean function for the observed units $\mathrm{E}_t(y_{obs})$ directly with equation (4), we focus our Monte Carlo study on analysing different situations of nonresponse patterns. In Section 3.2, we defined five general types of acceptance rates: random, fixed, cycle-dependent, cycle-shifted and monotone. For every cycle function $g_j(t)$ and each of these bias patterns, we draw $Z = 1000$ different situations for the triple $\pi^+(t), \pi^=(t)$ and $\pi^-(t)$. For example, for a time-independent bias pattern we draw a probability for each $\pi^s(t)$ and fix these over t. In contrast, a random bias pattern would lead to a draw of each $\pi^s(t)$ for every t. For cycle-dependent and cycle-shifted bias patterns, $\pi^s(t)$ are functions of $g_j(t)$ whereas monotone and linear bias patterns are functions in t. In every iteration z, different scaling parameters (which are defined in the subsequent paragraphs) are drawn. Therefore, we are able to reflect very different situations of nonresponse patterns. For each of these patterns, we calculate the correlations defined in Section 2.3 and the dispersion measure v from equation (12).

Random and fixed

We start our analysis with random, unstructured response functions $\pi^s(t)$.

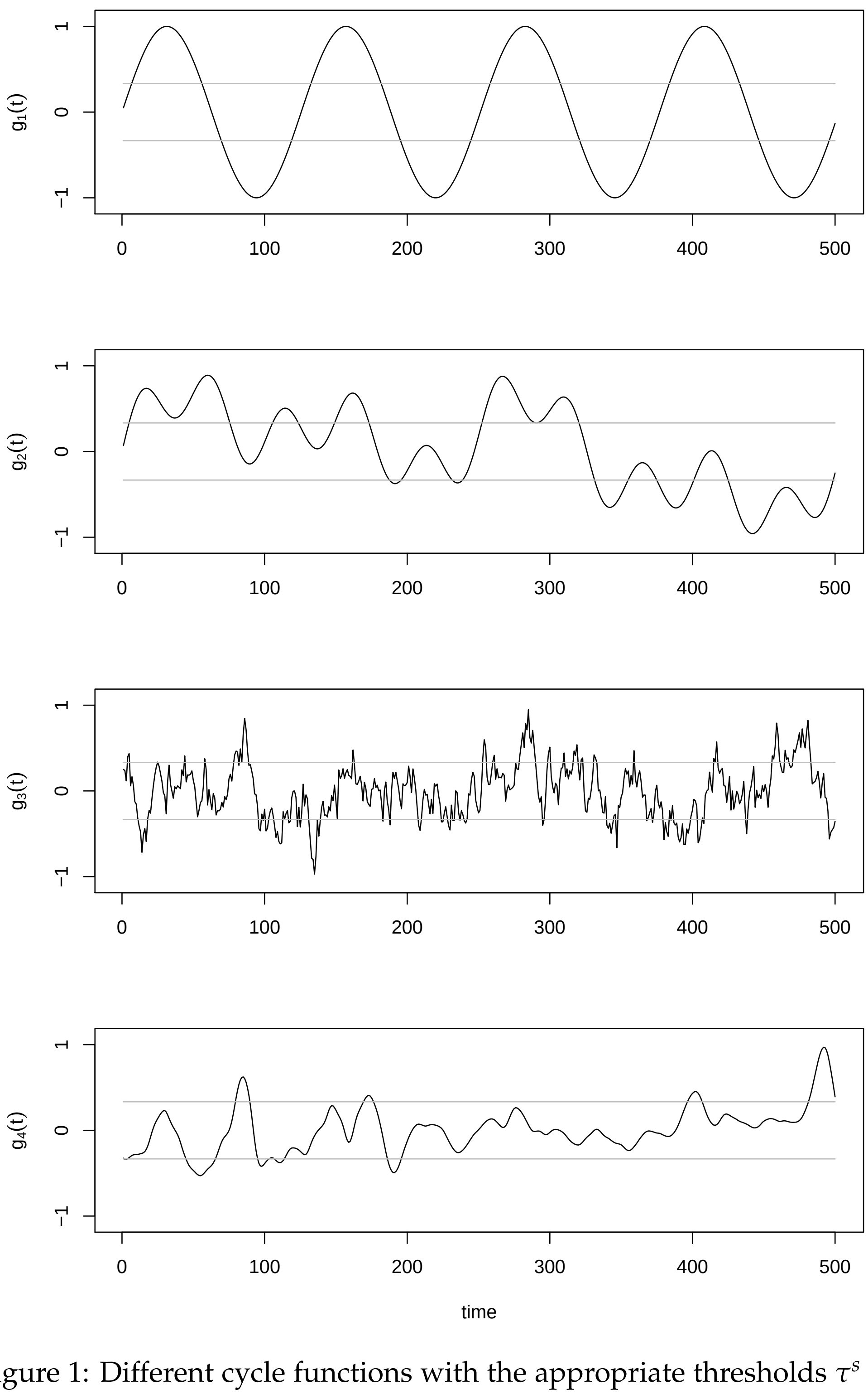

Figure 1: Different cycle functions with the appropriate thresholds τ^s (grey lines).

For every iteration z, response function $\pi^s(t)$ and t, we draw from the Uniform distribution $U(0,1)$ to obtain randomness, i.e. $\pi_z^s(t) \sim U(0,1), \forall s,t,z$. For the time-independent case, we draw only once for every iteration z and state s and fix this value over t, i.e. these functions may be different across s but do not fluctuate over t which leads to $\pi_z^s(t) = \pi_z^s \sim U(0,1)$.

Cycle-dependent

In contrast to random and fixed bias patterns, all cycle-related as well as monotone patterns are functions in $g_j(t)$ or t. As we have to ensure that $0 \leq \pi^s(t) \leq 1, \forall t$, we can not use $g_j(t)$ directly and have to standardise it. In Section 3.1 we defined $g_j(t) \in [-1,1]$. So, our cycle-dependent acceptance rate function has the form

$$\pi_{j,k_C}^C(t) := \frac{g_j(t)+1}{2k_C}, \tag{10}$$

for $j = 1,\ldots,4$. For our Monte Carlo study, we draw the scaling parameters $k_C \in [1,10]$. As (10) defines a positive relationship between $\pi^s(t)$ and $g_j(t)$, we also allow negative relationships $1 - \pi_{j,k_C}^C(t)$. This is done by drawing a 0/1-dummy d_I if to use $\pi_{j,k_C}^C(t)$ or $1 - \pi_{j,k_C}^C(t)$. d_I will also be used in the cycle-shifted and montone response functions.

Cycle-shifted

As survey business cycle indicators including expectation questions are commonly used for now- and forecasting, i.e. should contain leading information, we now assume that the selection process also includes leading or lagging information. It can be assumed that the respondents might anticipate coming development[4] or might react on the most recent development. To reflect this, we define the acceptance rate functions $\pi^s(t)$

$$\pi^S_{j,k_C}(t + k_T) := \pi^C_{j,k_C}(t + k_T) \tag{11}$$

for $j = 1, \ldots, 4$. Equation (11) has the same form as the cycle-dependent acceptance rates in (10), but is shifted in time by k_T units. Therefore, we first draw $k_T \in [-12, 12]$, i.e. we allow a maximum of 12 months for leads or lags. Second, we draw k_C and d_I as in the cycle-dependent case.

Monotone

Cycle-dependent and cycle-shifted response rates $\pi^s(t)$ are direct transformations of $g_j(t)$, i.e. the underlying cycle function. To allow $\pi^s(t)$ to be dependent from t but not from $g_j(t)$, we specify functions $\pi^s(t)$ which display general trends of monotone or linear form. Therefore, we define two different types of acceptance rates $\pi^s(t)$:

$$\begin{aligned} \pi^L_{k_L}(t) &:= \frac{t + 1000}{k_L} \quad \text{and} \\ \pi^M_{k_M}(t) &:= \Phi\left(\frac{t - 250}{k_M}\right), \end{aligned}$$

where $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution. $\pi^L_{k_L}(t)$ was chosen to have values between 0 and 1 for $t \in [1, T]$. k_L and k_M are scaling parameters which control for the steepness of $\pi^M_{k_M}(t)$

[4]The participants of the Ifo Business Survey are asked every month to give an assessment on the future business situation. This might also affect the decision to respond.

and $\pi^L_{k_L}(t)$. First, we draw a dummy variable d_L which decides to use $\pi^L(t)$ or $\pi^M(t)$. Conditional on d_L we draw $k_L \in [100, 500]$ or $k_M \in [2000, 4000]$. Finally, we draw d_I as in the cycle-dependent case to allow for decreasing trends e.g. by $1 - \pi^L(t)$.

Mixed

Of course, the five main bias patterns described above can also appear simultaneously. To reflect this situation, we mix these patterns by introducing a variable d_P which specifies the bias pattern to use. After this classification, the simulations are done as described above.

3.4 Dispersion measure

From Section 2.2 we know that MAR occurs in cases of $\pi^r(t) = \pi^s(t), \forall r, s \in S$, and $\forall t \in [1, T]$. We want to analyse if a higher average variation in the acceptance rates $\pi^s(t)$ between the different states for given t leads to a lower correlations. To evaluate this variation for the whole time series, we introduce a variation measure

$$v = \frac{1}{T}\sum_{t=1}^{T} \mathrm{Var}_s(\pi^s(t)|t). \tag{12}$$

The inner part $\mathrm{Var}_s(\pi^s(t)|t)$ is the variance across $\pi^s(t)$ for given t which is then averaged over t. As $\pi^s(t) < 1, \forall t$, $\mathrm{Var}_s(\pi^s(t)|t) < 1$ and therefore also $v < 1$.

3.5 Results

Figures 2 and 3 show the scatterplots for the pairs $(\tilde{\rho}_g, v)$ and (ρ_E, v) according to the five different main bias patterns plus the mixed case as defined in Section 3.3. The vertical axis are scaled to $[-1, 1]$ whereas the horizontal axis are not scaled due to the fact that the range of v differs strongly between the various patterns. In addition, boxplots for the correlations and the dispersion parameters are drawn and the underlying cycle functions $g_j(t)$ are displayed with different colors. The boxplots on the y-axis clearly show that the majority of biased indicators still receives high correlations close to one for $g_j(t)$ as well as $\mathrm{E}_y(t)$. This holds in particular for the fixed, cycle-dependent, cycle-shifted and monotone types of $\pi^s(t)$. On average, the lowest correlations appear in the random and mixed case. Furthermore, clusters according to the different cycle functions can be seen in the random case. For this bias pattern type, the ideal cycle functions $g_1(t)$ and $g_2(t)$ seem to be less affected by the selection bias. As the random bias pattern includes to more unstructered uncertainty in the data, the correlations become smaller on average. With exception of the monotone bias types, there seems to be no clear connection between the dispersion v and the correlations. Correlations remain high, even if the the dispersion increases.

To get an idea how the acceptance rates $\pi^s(t)$ transform the biased indicator $\mathrm{E}_t(y_{obs})$, we show some, partially extreme, cases for every main bias type in Figures 4 and 5. We select the cycle function $g_2(t)$ to display the effects. On the left side, the cycle function $g_2(t)$ (in grey), the unbiased mean function $\mathrm{E}_t(y)$ (in green) and the biased mean function $\mathrm{E}_t(y_{obs})$ (in red) are drawn. The right side shows the appropriate acceptance rates $\pi^{+}(t)$ (- - -), $\pi^{=}(t)$ (—) and $\pi^{-}(t)$ ($\cdots$). In Figure 4 a random, fixed and cycle-dependent case, including correlations $\tilde{\rho}_g$ and ρ_E as well as the dispersion parameter v, is displayed. The random bias pattern clearly shows that the correlation decreaes but the general underlying structure from the cycle function is still present. Smoothing approaches, such as the Hodrick-Prescott band pass filter, could still be used in this case to reduce this effect. However, we have to notice here that sampling size effects will include even more uncertainty. The second row displays the effect of time-independent acceptance rates $\pi^{s}(t)$. With $\pi^{-}(t) = 0.1$ we choosed a low acceptance for the negative replies. This led to a shift of $\mathrm{E}_t(y_{obs})$ upwards.[5] However, the correlation still remained high, as we already saw in the scatterplots in Figure 2. In the last row, the cycle-dependent case is shown. $\pi^{s}(t)$ were chosen to display a very extreme case where we receive high negative correlations. This effect appears when the acceptance rates are anti-cyclical, i.e. negative replies are seldom in recessions and positive ones in boom times. In our case, we have $\pi^{+}(t) = 1 - \pi^{C}_{2,1}(t)$ and $\pi^{-}(t) = \pi^{C}_{2,1}(t)$.[6] Figure 5 shows the other three bias types: cycle-shifted, monotone and mixed cases. Although the correlations remained high, it can clearly be seen for the cycle-shifted case that the lag/lead structure of the observed indicator has changed. The bias pattern led to a shift in the future which caused a decrease of leading information of the indicator. The second row shows the effect of monotone and

[5]A stretch of the indicator can appear in cases where $\pi^{=}(t) << \pi^{+}(t), \pi^{-}(t)$ whereas a compression of the indicators appears in cases of $\pi^{=}(t) >> \pi^{+}(t), \pi^{-}(t)$.

[6]As the negative replies have higher probability in recession times, is $\pi^{C}_{2,1}(t)$ anti-cyclical for the '-'-category.

linear acceptance rates $\pi^s(t)$ on the indicator. It can clearly be seen that this type introduced a trend movement of the indicator which is diametrical to the underlying cycle. This was caused by an increase of acceptance of the positive replies and a simultaneous decrease of the negative replies. Also in this case, the correlations are negative. The last type of bias pattern is the mixed case, where we combined a cycle-shifted (for $\pi^+(t)$), a monotone (for $\pi^=(t)$) and a fixed (for $\pi^=(t)$) bias pattern. Although the acceptance rates are of very different kind, the correlations still remain high.

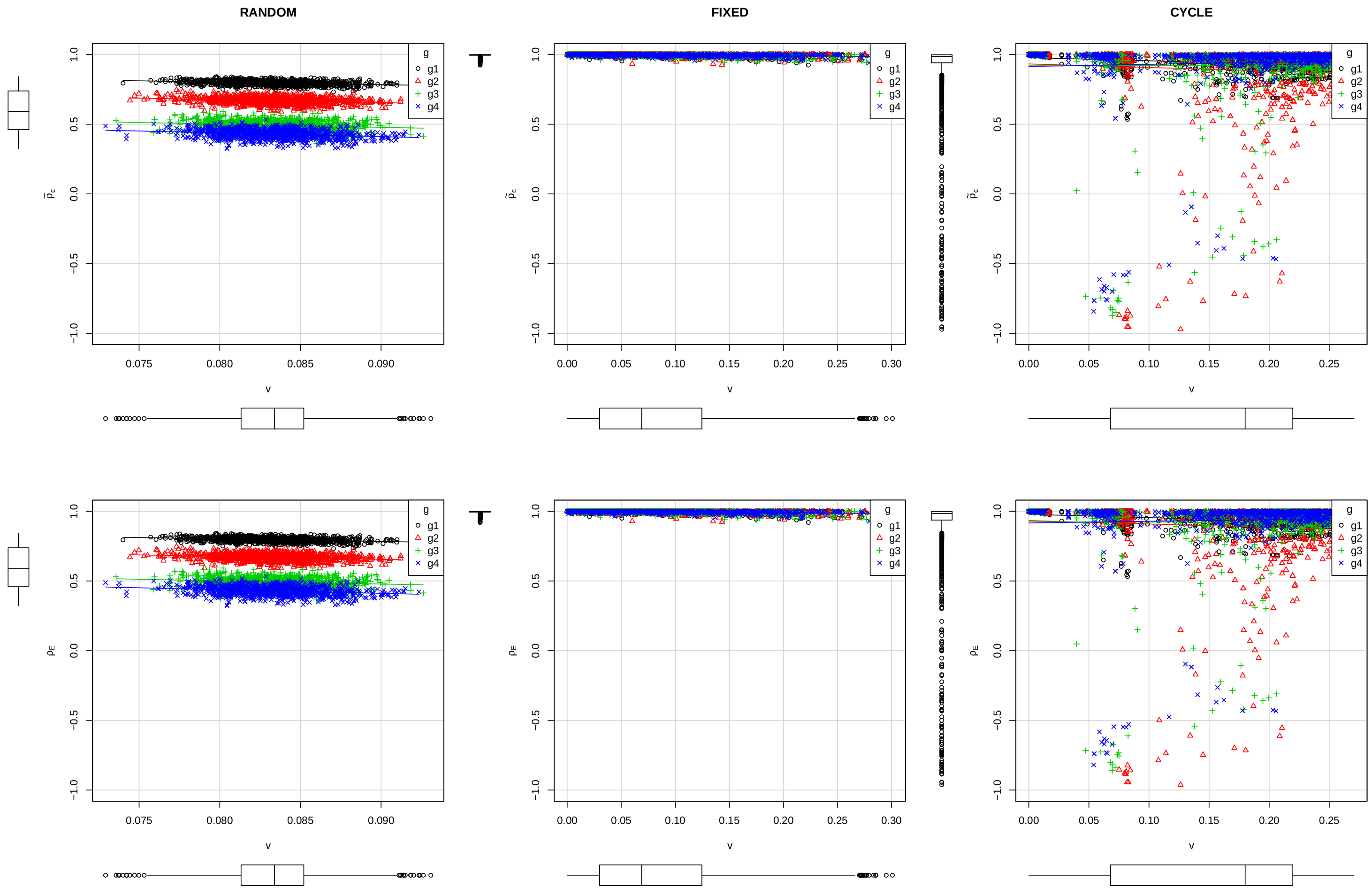

Figure 2: Scatterplots between correlations ρ and dispersions v for random, fixed and cycle-dependent acceptance rates $\pi^s(t)$. Top: Correlations $\tilde{\rho}_c$, bottom: correlations ρ_E. g1, ..., g4 denote the four different cycle functions $g_1(t), \ldots, g_4(t)$.

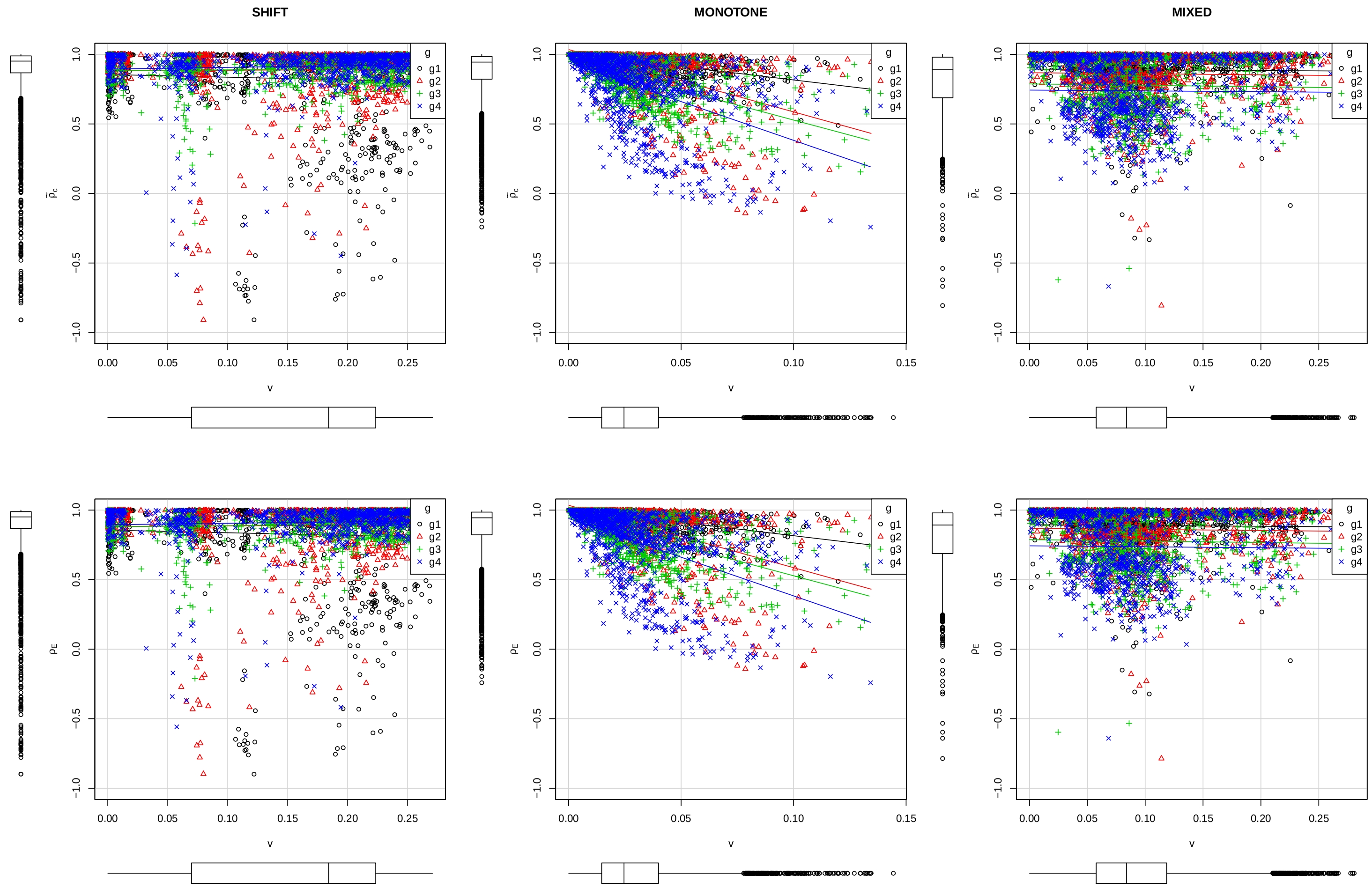

Figure 3: Scatterplots between correlations ρ and dispersions v for random, fixed and cycle-dependent acceptance rates $\pi^s(t)$. Top: Correlations $\tilde{\rho}_c$, bottom: correlations ρ_E. g1, ..., g4 denote the four different cycle functions $g_1(t), \ldots, g_4(t)$.

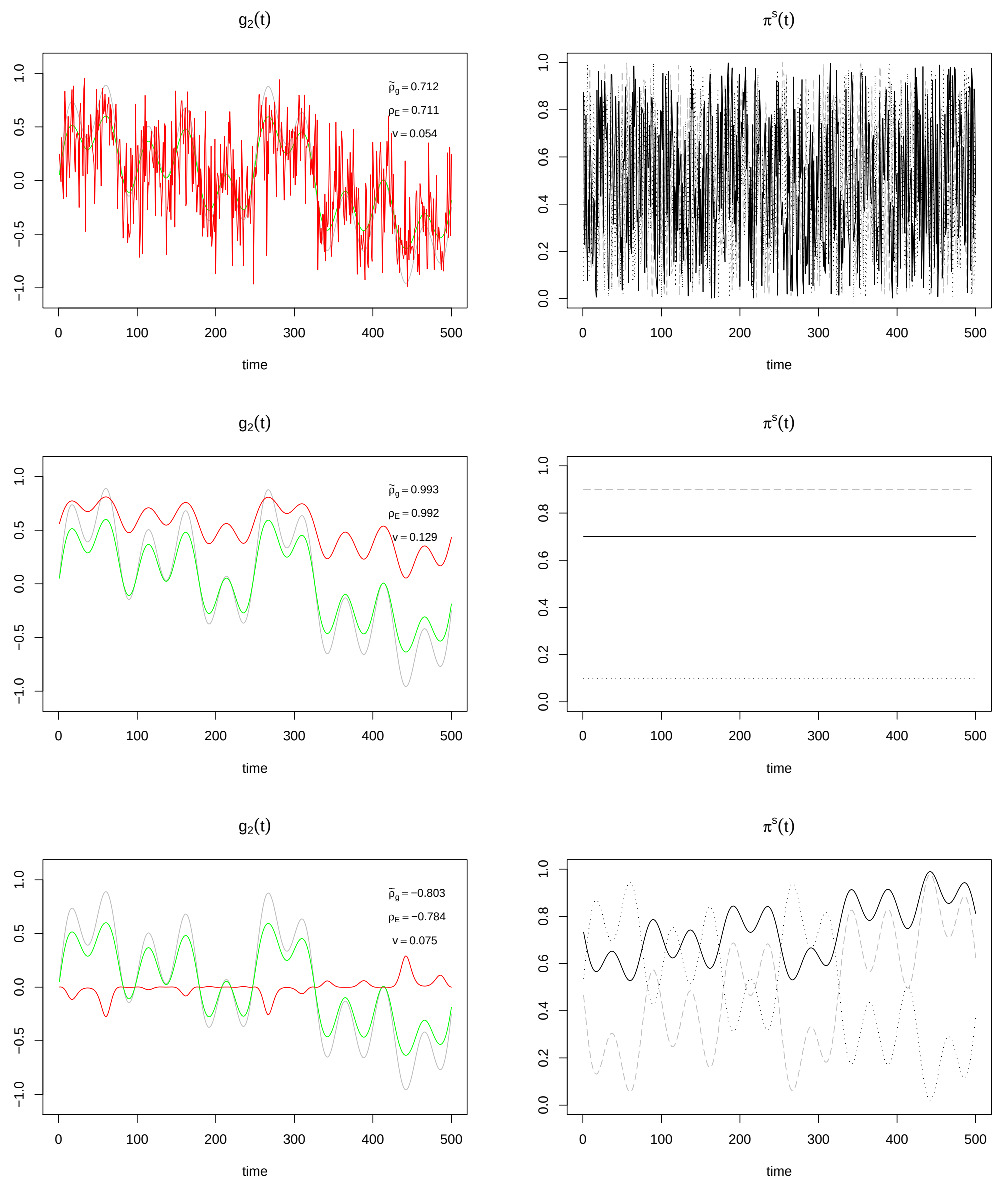

Figure 4: Effects of bias patterns I. Left: cycle function $g_2(t)$ (—), unbiased mean function $\mathrm{E}_t(y)$ (—) and biased mean function $\mathrm{E}_t(y^{obs})$ (—). Right: acceptance rates $\pi^+(t)$ (- - -), $\pi^=(t)$ (—) and $\pi^-(t)$ ($\cdots$). From top to bottom: Random, fixed (with $\pi^+(t) = 0.9$, $\pi^=(t) = 0.7$ and $\pi^-(t) = 0.1$) and cycle-dependent (with $\pi^+(t) = 1 - \pi^C_{2,1}(t)$, $\pi^=(t) = 1 - \pi^C_{2,2}(t)$ and $\pi^-(t) = \pi^C_{2,1}(t)$) patterns.

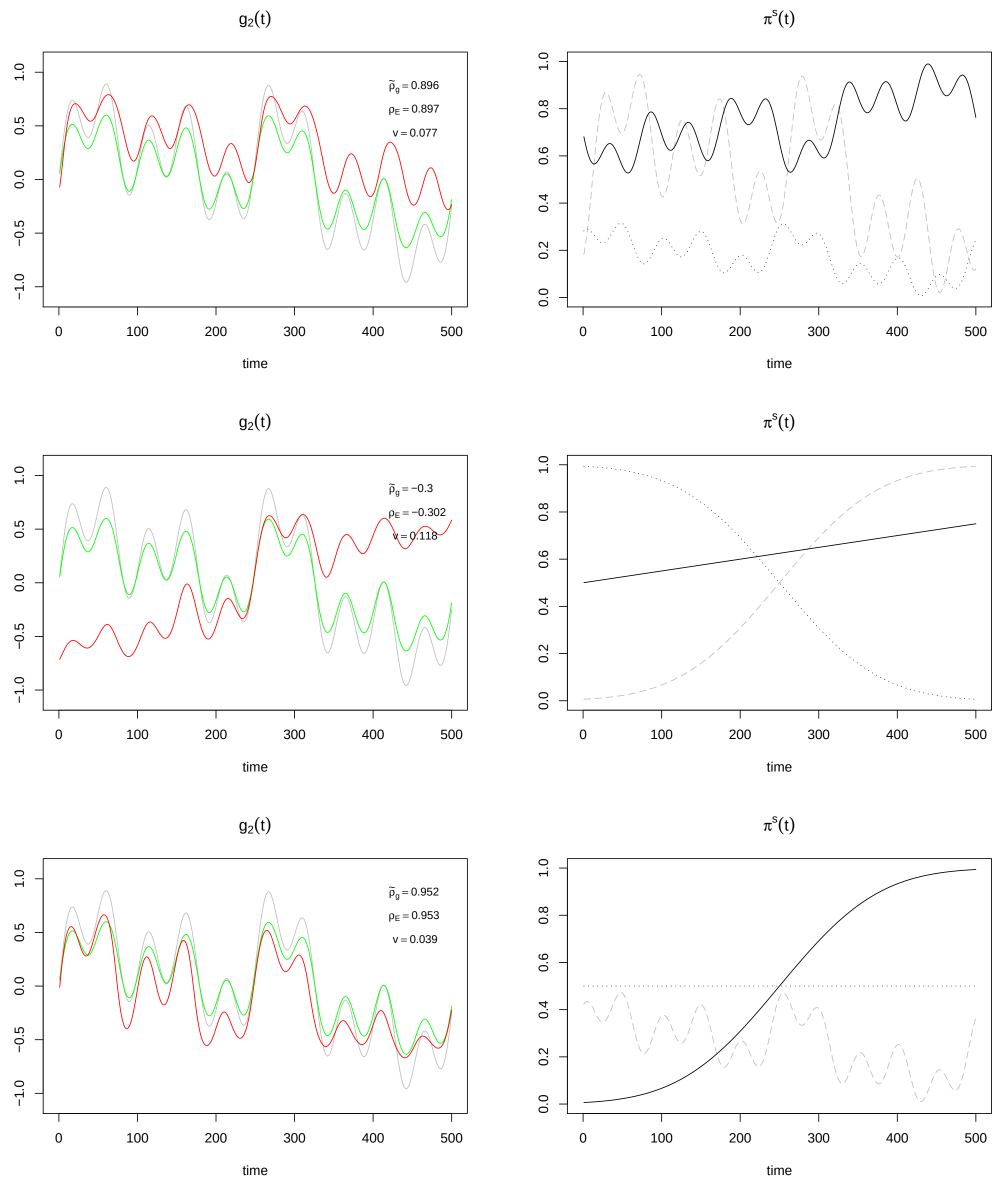

Figure 5: Effects of bias patterns II. Left: cycle function $g_2(t)$ (—), unbiased mean function $\mathrm{E}_t(y)$ (—) and biased mean function $\mathrm{E}_t(y^{obs})$ (—). Right: acceptance rates $\pi^+(t)$ (- - -), $\pi^=(t)$ (—) and $\pi^-(t)$ $(\cdots)$. From top to bottom: Cycle-shifted (with $\pi^+(t) = \pi^S_{2,1}(t-12)$, $\pi^=(t) = 1 - \pi^S_{2,2}(t+3)$ and $\pi^-(t) = \pi^S_{2,3}(t+12)$), monotone (with $\pi^+(t) = \pi^M_{100}(t)$, $\pi^=(t) = \pi^L_{2000}(t)$ and $\pi^-(t) = 1 - \pi^M_{100}(t)$) and mixed (with $\pi^+(t) = \pi^S_{2,2}(t+12)$, $\pi^=(t) = \pi^M_{100}(t)$ and $\pi^-(t) = 0.5$) patterns.

4 Conclusion

In this paper, we build a methodologic framework for the widely used balance statistics indicators for economic time series. As these indicators are based on surveys, we included nonresponse into this framework to evalute their effects. Due to the fact that the indicators are artifical, we focused our analysis on the effects of nonresponse on correlation. The analysis showed that the correlations between the observed indicators and the underlying cycle still remain high in nearly all of the cases. Of course, some bias patterns exist where the shape of the indicator is very strongly transformed but we notice that these patterns are unrealistic to appear in real-life situations. However, some patterns might cause certain problems such as a complete switch of the indicator or a shift which might affects the leading performance. The analyses also showed that there exists only a small connection between the dispersion in the acceptance rates and the correlations. If the acceptance rates differ strongly for random or monotone transformations of t, the correlation with the underlying cycle decreases on average. The reasons for this robustness are probably due to the fact that all states contain the oscillation of the underlying cycle. All of the three states are affected by the cycle and therefore include its information. When the probability for one state to be observed decreases, the other states still include enough information. Only for very extreme, in particular time- or directly cycle-dependent cases, significant decreases in correlation might appear. The robustness towards selection biases therefore supports the usage of these indicators and explains one reason for their success.

A Calculation of balance statistic

As the Ifo index is based on balances and not on the calculation of the mean, we have to ensure that we obtain the same results. We define, as in section 2.1, $+ \equiv 1$, $= \equiv 0$ and $- \equiv -1$. Then, let b_t the balance at time t, i.e.

$$b_t = \frac{\sum_{i=1}^{n_t} I(y_t = 1) - \sum_{i=1}^{n_t} I(y_t = -1)}{n_t}.$$

Therefore, the mean is defined as

$$\begin{aligned}
\mathrm{E}_t(b) &= \mathrm{E}\left(\frac{\sum_{i=1}^{n_t} I(y_t = 1) - \sum_{i=1}^{n_t} I(y_t = -1)}{n_t}\right) \\
&= \frac{1}{n_t}\mathrm{E}\left(\sum_{i=1}^{n_t} I(y_t = 1) - \sum_{i=1}^{n_t} I(y_t = -1)\right) \\
&= \frac{1}{n_t}\sum_{i=1}^{n_t} \mathrm{E}[I(y_t = 1) - I(y_t = -1)] \\
&= \frac{1}{n_t} \cdot n_t \cdot [P(y_t = 1) - P(y_t = -1)] \\
&= P(y_t = 1) - P(y_t = -1) \\
&= \left[1 - \Phi\left(\frac{\tau^{+} - g(t)}{\sigma}\right)\right] - \Phi\left(\frac{\tau^{=} - g(t)}{\sigma}\right) \\
&= \mathrm{E}_t(y).
\end{aligned}$$

References

Anderson, O. (1951). Konjunkturtest und Statistik. *Allgemeines Statistisches Archiv*, 35:209–220.

Anderson, O. (1952). The business test of the IFO-Institute for Economic Research. *Revue del'Institute International de Statistique*, 20:1–17.

Ang, A., Bekaert, G., and Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54(4):1163–1212.

Bartholomew, D. J. and Knott, M. (1999). *Latent variable models and factor analysis*. Charles Griffin.

Brehm, J. (1994). Stubbing out Toes for a Foot in the Door? Prior Contacts, Incentives and Survey Response. *International Journal of Public Opinion Research*, 6(1):45–63.

Carlson, J. A. and Parkin, M. (1975). Inflation expectations. *Economica*, 42:123–138.

Drechsel, K. and Scheufele, R. (2010). Should We Trust in Leading Indicators? Evidence from the Recent Recession. IWH-Diskussionspapiere 10/2010, Institut für Wirtschaftsforschung Halle.

European Union (2006). Joint Harmonised EU Programme of Business and Consumer Surveys. *Official Journal of the European Union*, 49(C 245):5–8.

Gayer, C. (2005). Forecast Evaluation of European Commission Survey Indicators. *Journal of Business Cycle Measurement and Analysis*, 2005(2):7.

Harris-Kojetin, B. and Tucker, C. (1999). Exploring the Relation of Economical and Political Conditions with Refusal Rates to a Government Survey. *Journal of Official Statistics*, 15(2):167–184.

Hodrick, R. and Prescott, E. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 29(1):10–16.

Janik, F. and Kohaut, S. (2012). Why don't they answer? - Unit non-response in the IAB Establishment Panel. *Quality & Quantity*. to appear.

Kholodilin, K. A. and Siliverstovs, S. (2006). On the Forecasting Properties of the Alternative Leading Indicators for the German GDP: Recent Evidence. *Journal of Economics and Statistics*, 226(3):234–259.

Korotayev, A. V. and Tsirel, S. V. (2010). A Spectral Analysis of World GDP Dynamics: Kondratieff Waves, Kuznets Swings, Juglar and Kitchin Cycles in Global Economic Development, and the 2008–2009 Economic Crisis. *Structure and Dynamics*, 4(1):3–57.

Little, R. J. A. and Rubin, D. (2002). *Statistical Analysis with Missing Data*. Wiley.

Nardo, M. (2003). The quantification of qualitative survey data: A critical assessment. *Journal of Economic Surveys*, 17(5):645–668.

OECD (2003). *Business Tendency Surveys - A Handbook*.

Robinzonov, N. and Wohlrabe, K. (2010). Freedom of Choice in Macroeconomic Forecasting. *CESifo Economic Studies*, 56(2):192–220.

Seiler, C. (2010). Dynamic Modelling of Nonresponse in Business Surveys. Ifo Working Paper 93, Ifo Institute for Economic Research.

Stangl, A. (2007). European data watch: Ifo world economic survey micro data. *Journal of Applied Social Science Studies*, 127(3):487–496.

Stangl, A. (2009). *Essays on the Measurement of Economic Expectations*. PhD thesis, Ludwig-Maximilians-Universität München.

Theil, H. (1952). On the shape of economic microvariables and the Munich business test. *Revue del'Institute International de Statistique*, 20:105–120.

Wollmershäuser, T. and Henzel, S. (2005). Quantifying Inflation Expectations with the Carlson-Parkin Method: A Survey-based Determination of the Just Noticeable Difference. *Journal of Business Cycle Measurement and Analysis*, 2(3):321–352.

Ifo Working Papers

No. 125 Arent, S., A. Eck, M: Kloss and O. Krohmer, Income Risk, Saving and Taxation: Will Precautionary Saving Survive?, February 2012.

No. 124 Kluge, J. and R. Lehmann, Marshall or Jacobs? Answers to an Unsuitable Question from an Interaction Model, February 2012.

No. 123 Strobel, T., ICT Intermediates, Growth and Productivity Spillovers: Evidence from Comparison of Growth Effects in German and US Manufacturing Sectors, February 2012.

No. 122 Lehwald, S., Has the Euro Changed Business Cycle Synchronization? Evidence from the Core and the Periphery, January 2012.

No. 121 Piopiunik, M. and M. Schlotter, Identifying the Incidence of “Grading on a Curve”: A Within-Student Across-Subject Approach, January 2012.

No. 120 Kauppinen, I. and P. Poutvaara, Preferences for Redistribution among Emigrants from a Welfare State, January 2012.

No. 119 Aichele, R. and G.J. Felbermayr, Estimating the Effects of Kyoto on Bilateral Trade Flows Using Matching Econometrics, December 2011.

No. 118 Heid, B., J. Langer and M. Larch, Income and Democracy: Evidence from System GMM Estimates, December 2011.

No. 117 Felbermayr, G.J. and J. Gröschl, Within US Trade and Long Shadow of the American Secession, December 2011.

No. 116 Felbermayr, G.J. and E. Yalcin, Export Credit Guarantees and Export Performance: An Empirical Analysis for Germany, December 2011.

No. 115 Heid, B. and M. Larch, Migration, Trade and Unemployment, November 2011.

No. 114 Hornung, E., Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia, November 2011.

No. 113 Riener, G. and S. Wiederhold, Costs of Control in Groups, November 2011.

No. 112 Schlotter, M., Age at Preschool Entrance and Noncognitive Skills before School – An Instrumental Variable Approach, November 2011.

No. 111 Grimme, C., S. Henzel and E. Wieland, Inflation Uncertainty Revisited: Do Different Measures Disagree?, October 2011.

No. 110 Friedrich, S., Policy Persistence and Rent Extraction, October 2011.

No. 109 Kipar, S., The Effect of Restrictive Bank Lending on Innovation: Evidence from a Financial Crisis, August 2011.

No. 108 Felbermayr, G.J., M. Larch and W. Lechthaler, Endogenous Labor Market Institutions in an Open Economy, August 2011.

No. 107 Piopiunik, M., Intergenerational Transmission of Education and Mediating Channels: Evidence from Compulsory Schooling Reforms in Germany, August 2011.

No. 106 Schlotter, M., The Effect of Preschool Attendance on Secondary School Track Choice in Germany, July 2011.

No. 105 Sinn, H.-W. und T. Wollmershäuser, Target-Kredite, Leistungsbilanzsalden und Kapitalverkehr: Der Rettungsschirm der EZB, Juni 2011.

No. 104 Czernich, N., Broadband Internet and Political Participation: Evidence for Germany, June 2011.

No. 103 Aichele, R. and G.J. Felbermayr, Kyoto and the Carbon Footprint of Nations, June 2011.

No. 102 Aichele, R. and G.J. Felbermayr, What a Difference Kyoto Made: Evidence from Instrumental Variables Estimation, June 2011.

No. 101 Arent, S. and W. Nagl, Unemployment Benefit and Wages: The Impact of the Labor Market Reform in Germany on (Reservation) Wages, June 2011.

No. 100 Arent, S. and W. Nagl, The Price of Security: On the Causality and Impact of Lay-off Risks on Wages, May 2011.