

Seit 1949 erhebt das ifo Institut monatlich Daten zur wirtschaftlichen Lage der Unternehmen in Deutschland. Das Instrument der Datenerhebung ist, wie in vielen sozio-ökonomischen Disziplinen, die direkte Befragung der Firmen. Ein bekanntes Phänomen, das im Befragungsprozess auftreten kann, sind fehlende Angaben zu einzelnen Fragen oder sogar die Nicht-Beantwortung des gesamten Fragebogens. Der vorliegende Artikel zeigt, dass die Ergebnisse des ifo Konjunkturtests sehr stabil gegenüber Antwortausfällen sind und diese insbesondere nicht zu einer Verschlechterung der Prognoseleistung des Geschäftsklimaindiktators führen.

## Mögliche Verzerrungseffekte

Ein häufig bestehendes Problem in Umfragen ist das Auftreten fehlender Werte. Hierbei unterscheidet man in der Literatur zwischen *item nonresponse*, bei dem zu einzelnen Fragen (items) keine Antworten gegeben wurden, sowie *unit nonresponse*, bei dem die befragte Einheit (unit) gar keinen Fragebogen abgibt. Die Gründe dafür können verschiedener Natur sein, beginnend mit gescheiterter Kontaktaufnahme, über zu geringe Bearbeitungszeit bis hin zu einer aktiven Antwortverweigerung.

Grundsätzlich stellen fehlende Werte zunächst nur in Bezug auf die Schätzgenauigkeit ein Problem dar, da offensichtlich weniger Information für die Analyse zur Verfügung steht. Dies kann jedoch mit entsprechenden Gewichtungsschemata, wie sie auch im ifo Institut angewendet werden, abgemildert werden. Zudem führt die große Zahl der Antworten (ca. 7 000 pro Monat, was einer durchschnittlichen Antwortquote von knapp 80% entspricht) dazu, dass der ifo Geschäftsklimaindikator wenig sensitiv bezüglich einzelner Nichtantworten ist. Ein weiteres – und erheblich schwerwiegenderes – Problem besteht allerdings dann, wenn die Entscheidung zur Beantwortung von der zur untersuchenden Variable selbst abhängig (vgl. Little und Rubin 2002 für Definitionen). Die Aussagekraft des Indiktors wäre deutlich eingeschränkt, wenn z.B. alle Unternehmen, denen es gut geht, nicht antworten würden. In einem solchen Fall spricht man von einem Ausfallprozess *not missing at random*, welcher zu verzerrten Ergebnissen führen kann.

Die Untersuchung fehlender Daten ist in der Literatur überwiegend auf Haushalts-

bzw. Individualbefragungen beschränkt. Dies liegt zum einen an der Tatsache, dass derartige Umfragen öfter durchgeführt werden als Unternehmensbefragungen, zum anderen aber auch daran, dass Verzerrungseffekte durch Nichtantworten häufiger aufgrund expliziter Fragen zu den persönlichen Umständen (z.B. Einkommen, Alkoholkonsum, Sexualverhalten) auftreten. Zur Illustration wird ein kleines Beispiel gegeben:

*In einer Umfrage wird nach dem Haushaltseinkommen gefragt. Hierbei lässt sich feststellen, dass Einkommen über einer gewissen Grenze nicht oder spärlicher in den Daten auftauchen, als man z.B. anhand von Steuerstatistiken erwarten würde. Es liegt offensichtlich der Effekt vor, dass besonders wohlhabende Haushalte eher dazu neigen, keine Angaben zu ihren Einkommensverhältnissen zu machen. Dies führt zu einer Unterschätzung des tatsächlichen durchschnittlichen Einkommens.*

Der in dem Beispiel dargestellte Verzerrungseffekt kann aufgrund von Sekundärdaten (Steuerstatistik) abgeschätzt werden. Allerdings liegt nur in seltenen Fällen eine derartige Zweiterhebung vor.

Im Gegensatz zu Individualbefragungen sind bei Unternehmensbefragungen a priori nur geringe Verzerrungseffekte zu erwarten, da die ausfüllende Person nicht über ihre persönlichen Umstände Auskunft gibt. Dies gilt insbesondere für den ifo Konjunkturtest, in dem nur wenig quantitative, sondern überwiegend Tendenzfragen gestellt werden. Dennoch können Unternehmen (zumindest zeitweise) ebenfalls eine restriktivere Informationspolitik verfolgen. Willimack und Nichols (2010) stellen fest, dass die Priorität der Firmen

immer denjenigen Aktivitäten gilt, die den Betrieb sicherstellen und zu Wachstum führen. Dementsprechend kann die Teilnahme an einer Befragung in gewissen Phasen eine niedrigere Priorität erhalten. Die Erwähnung der ökonomischen Bedingungen als Einflussgröße auf die Teilnahmebereitschaft im ersten Abschnitt spiegelt bereits die Problematik wider, dass daher auch in Unternehmensbefragungen Verzerrungseffekte auftreten können. Harris-Kojetin und Tucker (1999) konnten zeigen, dass ein Zusammenhang zwischen Rücklaufquoten in Individualbefragungen und verschiedenen ökonomischen und politischen Indikatoren besteht. Insbesondere in guten ökonomischen Zeiten sind die Rücklaufquoten im Allgemeinen niedriger. Für Unternehmensbefragungen, welche die Messung der konjunkturellen Entwicklung zum Ziel haben, ist gerade dieser Aspekt interessant. Inwieweit dies auch auf den ifo Konjunkturtest zutreffend ist, war Teil der Untersuchung.

Die Analysen in Seiler (2010) bestätigen, dass auch im ifo Konjunkturtest eine – wenn auch sehr geringe – Abhängigkeit vom Konjunkturzyklus besteht. Dies muss aber nicht zwangsläufig zu einer Verzerrung der Ergebnisse führen, da zunächst nur weniger Unternehmen in Boomzeiten eine Meldung abgeben. Es ist lediglich ein Hinweis auf mögliche Verzerrungen und sagt noch nichts darüber aus, ob strukturelle Unterschiede bezüglich ihrer eigentlichen »Antwortintention« vorhanden sind, und falls ja, wie groß die Verzerrung ist. Eine Untersuchung dessen wird im folgenden Kapitel durchgeführt.

## Robustheitsanalyse

### Eine Simulationsstudie

Mit Hilfe einer Simulationsstudie lässt sich eine Vielzahl verschiedener Szenarien erstellen und die berechneten Indikatoren auf ihre Robustheit überprüfen. Die meisten Fragen im ifo Konjunkturtest werden auf einer 3er-Skala (»gut«, »gleichbleibend«, »schlecht«) erhoben. Aus diesen Antworten wird dann ein sogenannter *Saldo* (Anteil der »Gut«-Antworten abzüglich des Anteils der »Schlecht«-Antworten) berechnet, der auch dem ifo Geschäftsklimaindikator zugrunde liegt. Wir abstrahieren im Folgenden von einer konkreten Variable (Geschäftslage, Preisänderungen, Lagerbestände etc.) und wollen untersuchen, wie sich Ausfallmuster generell auf die aggregierten Salden auswirken. Die formale Darstellung der Salden wird in Box 1 gezeigt.

In unserer Simulationsstudie wird nun eine Vielzahl verschiedener Ausfallmuster erstellt und deren Korrelationen mit vier verschiedenen Konjunkturzyklen berechnet.<sup>1</sup> Die Ausfallmuster sind kategorisiert als *fixe*, *zufällige*, *monotone* und *zyklusabhängige* Funktionen in  $t$ . Darüber hinaus wird eine *Mischung* dieser fünf Grundmuster untersucht. Für jedes dieser Muster werden nun 1 000 verschie-

<sup>1</sup> Hierbei handelt es sich um eine einfache Sinusfunktion, vier überlappende (Kitchin-, Juglar-, Kuznets- und Kondratiev-Zyklus-) Sinusfunktionen, einem AR(1)-Prozess sowie gefilterten Wachstumsraten der US-Industrieproduktion.

#### Box 1

##### Formale Darstellung der Salden

Vor Beginn der Simulationsstudie definieren wir eine Funktion  $g(t)$ , die den Konjunkturzyklus darstellt. In idealisierter Form hat dieser einen sinusförmigen Verlauf (vgl. Schirwitz, Seiler und Wohlrabe 2009). Die Antwortbildung wird nun von dieser Funktion, einer individuellen Abweichung  $\varepsilon_i$  sowie »Schranken«  $\tau^s$  bedingt. Die Schranken  $\tau^s$  geben die Werte an, welche über- bzw. unterschritten werden müssen, damit der Antwortende die entsprechende Kategorie ankreuzt.<sup>a)</sup> Der (erwartete) Saldo zum Zeitpunkt  $t$  ist dann definiert als

$$E_t(s) = \Phi^+(t) - \Phi^-(t)$$

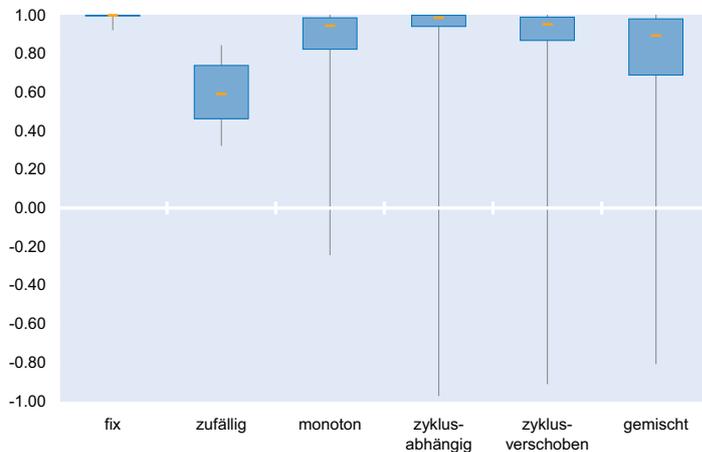
mit  $\Phi^+(t)$  als der prozentuale Anteil der »Positiv«-Antworten und  $\Phi^-(t)$  als der prozentuale Anteil der »Negativ«-Antworten. Fügt man nun die Wahrscheinlichkeit zur Beantwortung des Fragebogens (gegeben die Antwortintention) mit ein, so ändert sich der (erwartete) Saldo der *beobachteten Werte* zu

$$E_t(s^{obs}) = \frac{\pi^+(t)}{\bar{\pi}(t)} \Phi^+(t) - \frac{\pi^-(t)}{\bar{\pi}(t)} \Phi^-(t),$$

wobei  $\pi^+(t)$  die Wahrscheinlichkeit für eine Antwort, gegeben einen »positiven« Zustand;  $\pi^-(t)$  die Wahrscheinlichkeit für eine Antwort, gegeben einen »negativen« Zustand; und  $\bar{\pi}(t) = \sum \pi^s(t) \Phi^s(t)$  die gewichtete Summe der Antwortwahrscheinlichkeiten über alle Zustände ist. Vergleicht man  $E_t(s^{obs})$  mit  $E_t(s)$ , so ist deutlich erkennbar, dass der beobachtete Saldo  $E_t(s^{obs})$  mit den relativen Antwortwahrscheinlichkeiten  $\frac{\pi^+(t)}{\bar{\pi}(t)}$  bzw.  $\frac{\pi^-(t)}{\bar{\pi}(t)}$  gewichtet wird. Nähere Informationen zu den Herleitungen findet man bei Seiler (2012).

<sup>a)</sup> Weiterführende Literatur zu dem formalen Aufbau finden sich in den frühen Arbeiten von Anderson (1951; 1952) und Theil (1952) sowie in Carlson und Parkin (1975).

**Abb. 1**  
**Boxplots der Korrelationen der beobachteten Indikatoren mit dem Konjunkturzyklus nach verschiedenen Ausfallmustern**



Quelle: Berechnungen des ifo Instituts.

dene Konstellationen simuliert. Aus diesen wird dann der unter diesem Ausfallmuster beobachtete Indikator berechnet und dessen Korrelation mit dem Konjunkturzyklus gemessen.

Abbildung 1 zeigt die Verteilung der Korrelationen nach den verschiedenen Grundmustern. Es ist deutlich zu erkennen, dass die Korrelationen in nahezu allen Fällen weiterhin sehr hohe Werte nahe 1 besitzen. Lediglich im Falle zufälliger Ausfallmuster sinken die Korrelationen verhältnismäßig deutlich. Dies ist allerdings nicht primär auf strukturelle Unterschiede zurückzuführen, sondern darauf, dass der beobachtete Indikator stärkeren (aber unstrukturierteren) Schwankungen unterliegt. Glättungsverfahren können in diesem Fall Abhilfe schaffen. Die Ergebnisse zeigen, dass trotz des Auftretens sehr komplexer Ausfallmuster der Konjunkturzyklus weiterhin sehr gut abgebildet werden kann. Lediglich in wenigen Einzelfällen ist die Korrelation nahe 0 oder wird sogar negativ. Allerdings ist zu bemerken, dass diese Ausfallmuster Extremfälle darstellen, die in der Realität kaum auftreten dürften, insbesondere nicht über einen längeren Zeitraum.

**Schätzung der fehlenden Werte**

Nachdem die Simulationsstudie bereits sehr positive Ergebnisse zeigte, wird nun im nächsten Schritt untersucht, wie sich der ifo Geschäftsklimaindikator ändert, wenn die fehlenden Werte mit Hilfe statistischer Imputationsverfahren geschätzt werden. Die Grundidee besteht darin, Strukturen in den beobachteten Daten zu finden, um damit die fehlenden Werte, gegeben diese Struk-

tur, zu schätzen. Hierzu werden üblicherweise Regressionsmodelle verwendet, wobei diese nur begrenzt auf *unit nonresponse* anwendbar sind. Im Falle des ifo Konjunkturtests muss man sich die Panelstruktur der Daten zunutze machen. Das Analogon zu einem autoregressiven Prozess in der Zeitreihenanalyse stellen auf diskreter Ebene Markov-Ketten dar, welche sich in unserem Fall gut als Imputationsverfahren eignen. Diese modellieren die Übergänge zwischen den drei Zuständen »gut«, »gleichbleibend« und »schlecht« in Abhängigkeit von den bis zu diesem Zeitpunkt beobachteten Antwortverläufen. Seiler und Heumann (2012) konnten zeigen, dass sich mit Hilfe zweistufiger Markov-Ketten mehr als 70% der Daten zur Geschäftslage korrekt geschätzt werden können. Bei den Geschäftserwartungen erreicht ein besonders

in der Biostatistik verbreiteter Ansatz namens *last observation carried forward*, bei dem man die zuletzt verfügbare Information verwendet, ebenfalls einen guten Wert von knapp 60%.

Schätzt man nun mit den genannten Methoden die fehlenden Werte und rechnet den ifo Geschäftsklimaindikator mit den geschätzten fehlenden Werten nach, so lässt sich in Abbildung 2 deutlich erkennen, dass die Unterschiede zwischen beiden Indikatoren nur marginal sind. Es lassen sich keine strukturellen Unterschiede zwischen dem originalen und dem imputierten Indikator feststellen. Auch ein Vergleich beider Indikatoren hinsichtlich ihrer Prognoseleistung des BIP-Wachstums führt zu einem erfreulichen Ergebnis: Der imputierte Indikator besitzt keine signifikant bessere Prognoseleistung als der originale Indikator.

**Abb. 2**  
**Originaler und imputierter ifo Geschäftsklimaindikator**  
 jeweils um Saisoneffekte bereinigt mit X12-ARIMA



Quelle: Berechnungen des ifo Instituts.

## Fazit

Die Ergebnisse der Studien zeigen, dass der ifo Geschäftsklimaindikator sehr robust gegenüber systematischen Antwortausfällen ist. Trotz einer geringfügigen Abhängigkeit der Teilnahmeentscheidung von dem Konjunkturzyklus konnte gezeigt werden, dass diese Ausfallmuster nur zu sehr geringen Verzerrungen führen. Dies ließ sich sowohl empirisch durch die Schätzung der fehlenden Werte und der anschließenden Neuberechnung des Indikators sowie mit Hilfe von Simulationsstudien, die eine Vielzahl verschiedener Ausfallmuster erzeugen, zeigen. Insbesondere die Prognoseleistung des ifo Geschäftsklimaindikators ist trotz des Auftretens fehlender Werte unvermindert hoch.

## Literatur

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 de l'Institut International de Statistique* 20, 1–17.

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. und M. Parkin (1975), »Inflation expectations«, *Economica* 42, 123–138.

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

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

OECD (2003), *Business Tendency Surveys – A Handbook*, Paris.

Schirwitz, B., C. Seiler und K. Wohlrabe (2009), »Regionale Konjunkturzyklen in Deutschland – Teil II: Die Zyklendatierung«, *ifo Schnelldienst* 62(14), 24–31.

Seiler, C. (2010), »Dynamic Modelling of Nonresponse in Business Surveys«, Ifo Working Paper No. 93.

Seiler, C. (2012), »On the Robustness of the Balance Statistics with respect to Nonresponse«, Ifo Working Paper No. 126.

Seiler, C. und C. Heumann (2012), »Microdata Imputations and Macrodata Implications: Evidence from the Ifo Business Survey«, Department of Statistics, Technical Report Nr. 119.

Theil, H. (1952), »On the shape of economic microvariables and the Munich business test«, *Revue de l'Institut International de Statistique* 20, 105–120.

Willimack, D.K. und E. Nichols (2010), »A Hybrid Response Process Model for Business Surveys«, *Journal of Official Statistics* 26(1), 3–24.