Bank of Israel **Research Department**

Sentiment Indicators Based on a Short Business Tendency Survey[1]

Daniel Roash* and Tanya Suhoy**

Discussion Paper 2019.11
September 2019

Bank of Israel - http://www.boi.org.il

[1] We thank Prof. Danny Pfeffermann, Prof. Moshe Pollak, Avigail Levy (Central Bureau of Statistics) and Ben Z. Schreiber, Ariel Mantzura and the participants of the Bank of Israel research seminar for helpful comments and discussion.

* Central Bureau of Statistics and Bar-Ilan University, Israel. Email: danielr@cbs.gov.il

** Bank of Israel. Email: tanya.suhoy@boi.org.il

חטיבת המחקר, בנק ישראל ת״ד 780 ירושלים 91007
Research Department, Bank of Israel. POB 780, 91007 Jerusalem, Israel.

Sentiment Indicators Based on a Short Business Tendency Survey

Daniel Roash and Tanya Suhoy

Abstract

The monthly frequency of the Business Tendency Survey, launched in 2011, sectoral representativeness, and early availability have created new opportunities for nowcasting. However, Israeli data confirm growing concerns that the aggregate balance of opinions has become less correlated with macroeconomic indicators in the post-crisis period.

We test this relationship using firm-level and macro (time-series) data.

At the firm-level, logit checks of qualitative evaluations of past domestic sales in the manufacturing, retail trade and services sectors in 2013–17 revealed significant cross-sectional correlations with corresponding revenue data, matched from administrative records; however, comovement between the qualitative evaluations and the aggregate sectoral index was documented only since the questionnaire wording was changed to focus on the specific month, instead of a three-month evaluation. Although this change has amplified seasonal variation in the categorical answers, correlations between qualitative and quantitative data remain (weakly) significant even after seasonal effects are controlled for. We find also that firms' heterogeneity has an effect on the reliability of qualitative evaluations, particularly in the services industry.

At the macro level, we are looking for a composite sentiment indicator that aggregates sectoral balances of opinions and tracks real growth at a monthly frequency. We suggest an indicator with time-varying weights, evaluated through Partial Least Squares regression with respect to GDP growth. As GDP is measured quarterly, we simulate intra-quarter GDP-changes from monthly interpolated and bootstrapped seasonally-adjusted GDP-levels. This sentiment indicator performs better than an overall balance of opinions, calculated as a composition of sectoral balances with predefined weights based on industrial GDP-shares. In most (about 85%) simulations the short-term forecasts outperform the benchmark of mean growth. The out-of-sample error is larger when the sentiment indicator forecast is compared to later GDP estimates published by the CBS than with the first estimate.

Keywords: Business tendency survey, Sentiment indicator, Partial Least Squares, monthly GDP

תקציר

סקר מגמות בעסקים, שנערך על ידי הלמ"ס החל מ-2011 על בסיס חודשי, פתח אפשרויות חדשות "לחיזוי ההווה" (nowcatsing) של הצמיחה הריאלית. זאת עקב התדירות הגבוהה בו הוא נערך, הזמינות המוקדמת של תוצאותיו והייצוג הענפי של המגזר העסקי. יחד עם זאת, נתונים ישראלים תומכים בממצאים המצביעים על היחלשות המתאם בין מאזן הדעות המצרפי מסקרי עסקים איכותניים מסוג זה לבין אינדיקטורים מאקרו-כלכליים בתקופה שלאחר המשבר העולמי בשנים 2008-2009.

אנו בוחנים קשר זה על בסיס תשובות פרטניות ברמת החברה ועל בסיס סדרות עתיות של מאזני התשובות המצרפיים.

על בסיס תשובות הפירמות לגבי מכירות שלהן בשוק המקומי, נבדק קשר מסוג logit בין נתונים קטגוריאליים אלה בענפי התעשייה, המסחר הקמעונאי והשירותים לבין נתוני הפדיון בפועל, אשר דווחו (בדיעבד) בתקופה 2013-2017 לרשויות המס. בניתוח זה נמצא קשר סטטיסטי מובהק בחתכי רוחב; אולם, מתאם בדינאמיקה מצרפית (ברמה סקטוריאלית) נמצא רק החל מ-2016, כאשר הוכנסו בשאלון הסקר שינויי ניסוח שנועדו לפשט אותו ולמקד אותו בהערכת מצב בחודש אחד, במקום הערכה לגבי שלושה חודשי כפי שהיה בניסוח הקודם. למרות ששינוי הניסוח הגביר תנודות עונתיות בהערכות הקטגוריאליות, השיפור במתאם הושג גם מעבר לאפקט זה.

אנו מוצאים גם שטיב הניבוי של ההערכות הקטגוריאליות מושפע מההטרוגניות של חברות בתוך הענף, במיוחד בענף השירותים.

על בסיס סדרות עתיות של מאזני הנטו הענפיים, אנו מציגים בעבודה את מדד הסנטימנט המצרפי, שנועד לספק הערכה בזמן אמת לגבי קצב הצמיחה החודשי. מדד זה משוקלל באמצעות רגרסיה של Partial Least Squares המאפשרת לעדכן את המשקולות על בסיס חודשי ולכוון אותן למתאם מרבי עם צמיחת התוצר. מאחר והתוצר נמדד בתדירות רבעונית, אנו מבצעים סימולציות של רמות התוצר החודשיות, מנוכות עונתיות, באמצעות אינטרפולציה ו-bootstrapping. מדד סנטימנט זה מנבא את מחזור הצמיחה טוב יותר בהשוואה לשקלול של מאזני הנטו במשקולות קבועים שנקבעו בהתאם להרכב הענפי של התוצר. בתחזית הסתברותית (density nowcast) מבוססת, כ-85% מהסימולציות עומדות במבחן של cross-validation התומך בכך שבמדד הסנטימנט נמצא מידע מוביל בהשוואה להנחת נאיבית המבוססת על קצב הצמיחה ארוך הטווח בלבד.

טעות התחזית גדלה במעבר מאומדן ראשון של התוצר לאומדן המעודכן המתקבל אחרי מספר רביזיות של פרסומים.

1. Introduction

Business Tendency Surveys have become a popular tool of nowcasting with "ragged-edge" data, due to earlier availability of their data and positive correlations between the balances of opinions statistics and GDP growth. Over the past two decades, special attention has been paid to mixed-frequency models, which allow the inclusion of monthly survey-based predictors in regressions of quarterly GDP growth, as well as factor models, which hold mutually correlated balances of opinions (Hansson et al. 2005; Frale, et al. 2010; Banbura and Rünstler, 2011; Österholm, 2014; Kaufmann and Scheufele, 2017; Mogliani, et al. 2017).

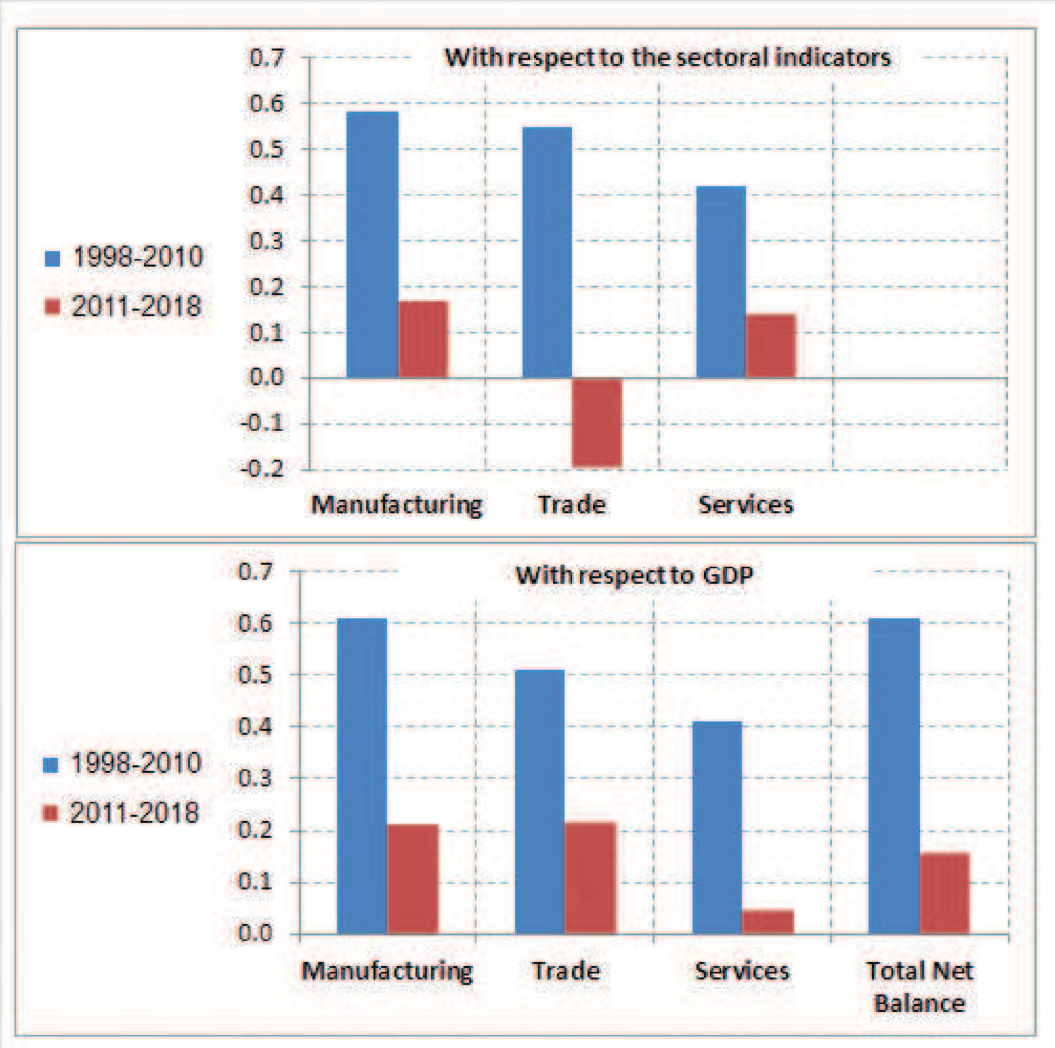
The Bank of Israel has conducted the Companies Survey since 1983 on a quarterly basis, and firms' participation in the survey is voluntary. In 2011, the Central Bureau of Statistics (hereinafter: CBS) began conducting a mandatory Business Tendency Survey on a monthly basis, in accordance with the OECD's methodological standards and reporting requirements.

Compared with the Companies Survey's balances of opinions, which display high historical correlations with macroeconomic variables, the CBS survey has appeared to fit the real indicators markedly worse.

"A clear drop" of correlations between the survey balances of opinions and real trends was also pointed out in the eurozone in the aftermath of the 2008–09 crisis (Malgarini, 2011; Tresor-Economics, No 125, 2014; Bruno, et al. 2016). Analysts have explained it by non-linearities in the relationship between the soft and hard data stemming from agents modifying over time their perceptions of long-run growth (a "new normal" situation), sufficient level of capacity utilization, and other settings. Another issue is growing sectoral heterogeneity—particularly in services—which requires a larger sample for its coverage. Regarding the poor performance of the Israeli Business Tendency Survey, CBS analysts have suggested low representativeness of the balances of opinions stemming from a large share of firms reporting "no-change".

Figure 1 depicts a decline in correlations between the balances of opinions and macroeconomic indicators, based on the quarterly Companies Survey and recorded after 2011, when the monthly Business Tendency Survey began.

Figure 1. Correlations between Companies Survey balances of opinions and seasonnally adjusted macroeconomic quarterly indicators, by period



It appears that high ex-post correlations documented in the period 1998–2010 stem from great cyclical fluctuations (dot-com boom of 1999–2000, recession 2001–03, recovery of 2004–07 and financial crisis of 2008–09). In the aftermath of the financial crisis the cyclical variance of real growth declined greatly, while the noise kept its magnitude, as seen from the X-12-Arima decomposition of two main macroeconomic series (Table 1): the industrial production index, measured at monthly frequency, and GDP in fixed prices, *measured at quarterly frequency.*

Table 1. Standard deviations of the trend-cycle and irregular components[1] of the Industrial Production index and GDP time series, by period

Period	Industrial Production Index (monthly % changes)			GDP (quarterly % changes)		
	Trend-cycle	Irregular	Signal-to-noise	Trend-cycle	Irregular	Signal-to-noise
1995-2005	0.51	9.12	0.056	1.12	2.58	0.434
2001-2010	0.68	8.89	0.076	1.03	2.24	0.460
2011-2018	0.23	8.68	**0.026**	0.53	2.33	**0.227**

[1] Trend cycle, seasonal and irregular components are available from seasonal adjustment processing, which is conducted monthly and quarterly with X-12-ARIMA.

Source: Industrial production – CBS data, GDP – authors' calculations.

The change in the cyclical pattern and questionable reliability of the quarterly Companies Survey motivated our analysis of the usefulness of the monthly CBS survey in real-time monitoring.

We conduct the analysis with firm-level data and with the balances of opinions series. The Hölzl survey (2015) points out differences between the results obtained at the micro level and the results obtained from the balances of opinions, which suggests that idiosyncratic fluctuations are canceled out as a result of aggregation. In this context, Nieuwstad (2005) reported interesting results from the manufacturing survey in the Netherlands showing that only one-third of the respondents provide coherent and unbiased retrospective evaluations of the production change, and roughly 20 percent of firms respond in a completely irrational manner. Another important finding was that firms with seasonal production cycles are likely to provide more accurate evaluations.

We examine the consistency of the survey data by regressing (using an ordinal logistic model) qualitative evaluations by quantitative firm-level changes in the revenue and the aggregate index of the sectoral dynamics, adjusted for seasonality. In addition, we control for the monthly seasonal effects. This analysis revealed significant idiosyncratic effects uncorrelated with the macroeconomic variables. We were able to isolate the overall sectoral effect only in the samples of 2016–17 relating to the new questionnaire, which focuses on one month, rather than three consecutive months as the old questionnaire. Our firm-level analysis confirms significant seasonal variance in qualitative evaluations since this change was made. We also track the sensitivity of the likelihood statistics to incorporate factors of observed heterogeneity, like sub-sectoral affiliation, differences in reporting, export profile and company size.

In a time-series dimension, we consider monthly GDP projections by survey variables. Official GDP data are produced at quarterly frequency and the first estimates become available about 45 days after the end of the quarter of interest. The Bridge-equation nowcasts of the Bank of Israel based on two-month averages of monthly indicators are also produced once a quarter and are sensitive to fluctuation in consumer imports (car purchases).

Monthly GDP models with monthly predictors were developed by Mittnik and Zadrozny (2004); Mitchell, et al. (2005); and Frale, et al. (2010), which emphasized better accuracy of short-term forecasts compiled at monthly frequency, beyond the importance of intra-quarter monitoring for policy makers.

Here we exploit the concept of common sentiment, which drives mutually correlated balances of opinions and can show the direction of the growth cycle. The European Commission's[1] Sentiment Indicator summarizes multiple balances of opinions with fixed weights. As an alternative to ad hoc weights, Gelper and Croux (2010) proposed a Partial Least Squares (PLS) model, which extracts the sentiment indicator with regression-based weights.

---

[1] This index is based on the Business and Consumer confidence surveys, as the sectors covered are industry (with a weight of 40 percent), services (30 percent), consumers (20 percent), retail (5 percent) and construction (5 percent).
For details see:
European Commission 1st Quarter 2017 TECHNICAL PAPER 015 | APRIL 2017 European Business Cycle Indicators ISSN 2443-8049 (online):
https://ec.europa.eu/info/publications/economy-finance/european-business-cycle-indicators-1st-quarter-2017_en
European Business Cycle Indicators 4th Quarter 2018 Technical Paper 029 | January 2019 European Commission ISSN 2443-8049.

We are looking for the best composition of sectoral balances of opinions, derived with respect to monthly GDP growth, by taking advantage of the ability of PLS to handle mutually correlated explanatory series in short samples. To overcome the problem of low-frequency GDP series, we apply the bootstrap aggregation procedure by Bergmeir, Hyndman and Benitez (2016) for monthly interpolated series, thus enabling density nowcasts.

The rest of the paper is organized as follows. Section 2 briefly describes the Business Tendency Survey and the main macro-level correlations. Section 3 discusses the relationship between firm-level qualitative and quantitative data. Section 4 presents explanatory survey-based variables used in the nowcasting equations. Section 5 describes monthly GDP nowcasts, and Section 6 concludes.

2. Characteristics of the Business Tendency Survey and correlations with macroeconomic indicators

Table 2 shows the average number of companies surveyed since 2011, by industry and by periods that the same questionnaire phrasing was preserved. Appendix A presents the survey questions used for this study that refer to various aspects of business activity and different time aspects: the current situation, past activity or future expectations.

We begin by briefly describing the nature of the changes that occurred in the survey questionnaire until it stabilized in the current formulation beginning in 2015. The first change, in 2013—timed with the transition to the new classification of industries—expanded the Likert Scale of possible responses from 3 to 5 points to allow companies to better discriminate their assessments (as was previously adopted in manufacturing), thus reducing a high share of neutral responses that led to a negative bias in balances of opinions compiled from the 2011–12 survey data.

Subsequent checks carried out in 2015 documented weak performance of the survey with respect to quantitative indicators and balances of opinions that were still biased towards zero; as a result it was decided to simplify the formulation of the questionnaire by focusing on the month-of-interest activity compared to the previous one, rather than three-by-three months' comparison as before, which was found to be confusing. For even more simplicity, the requirement to adjust the response for seasonality was removed from the phrasing, except the questionnaire for hotels.

Table 2. The average number of companies[1] in the Business Tendency Survey, by industry and questionnaire version[2]

Industry \ Questionnaire	2011-2012	2013-2014	2015 (Old)	2015 (New)	since 2016
	(1)	(2)	(3)	(4)	(5)
Manufacturing	340.4 [12.7]	354.6 [35.5]	229.3 [10.5]	172.5 [12.7]	421.1 [28.7]
Retail Trade	153.6 [11]	178.9 [18.3]	112.1 [5.1]	78.4 [8.8]	209.6 [15.5]
Construction	202.5 [13]	204.8 [26.2]	111.8 [5.5]	134.0 [11.6]	272.3 [20.2]
Hotels	71.8 [5.9]	56.5 [4.6]	41.3 [2.9]	19.2 [2.3]	57.9 [5.1]
Services	404.7 [18.5]	433.0 [18.5]	288.3 [12.4]	202.5 [14.4]	521.2 [38.1]

[1] Monthly, standard deviation is given in brackets

[2] Methodological changes in the BTS questionnaire were made in 2013 and 2015. The change of 2015 concerned retrospective evaluations and was conveyed by a follow-up experiment, which required a split into treatment and control sub-samples.

To enable the follow-up of this change, the samples in each sector were divided into two parts: the companies that received a questionnaire in a new format and companies that continued to respond in a previous format; the average number of firms in treatment and control sub-samples managed from April to December 2015 is shown in columns (3) and (4) of Table 2, respectively.

Appendix B provides summary statistics from the follow-up experiment. A new formulation of the questionnaire is likely to make responses more optimistic and to reduce the negative bias in balances of opinions, except for the construction and hotels sectors. The share of firms reporting "no-change" has declined slightly in manufacturing and construction, but remained high in the services industry.

Table 3 presents the correlation coefficients between the retrospective balances of opinions related to various aspects of business activity and the corresponding changes in the macroeconomic indicators, adjusted and unadjusted for seasonality, by industry and questionnaire version. As shown, the correlations calculated from the new questionnaire (2016–18) are much higher than those based on the old one (2013–15). The table also shows that the balances of opinions have become much more correlated with unadjusted indices, thus indicating an increased seasonal variance in qualitative responses reported to the new questionnaire. The next section provides more details based on logit checks of firm-level data.

Table 3. Correlations[1] between selected net balances and industrial reference series[2], by questionnaire, industry, and question

Industry	Question	Old questionnaire: 2013:01-2015:12 (N=36)				New questionnaire: 2016:01-2018:12 (N=36)			
		Unadjusted reference series		Seas.adj. reference series		Unadjusted reference series		Seas.adj. reference series	
Manufacturing	Output	-0.14		0.22		0.84	***	0.26	
Manufacturing	Sales	0.08		0.14		0.91	***	0.27	
Construction	Ongoing activities	0.20		0.29	*	0.40	**	0.33	*
Trade	Sales	-0.09		0.14		0.74	***	0.39	**
Hotels	Local tourists	0.20		0.30		0.60	***	0.45	**
Services	Local sales	0.11		0.05		0.55	***	0.39	**
Services	Exports	0.29		0.00		0.21		0.47	**

[1] *, **, *** indicate significance at a level of 1%, 5%, and 10%, respectively.

[2] Reference series are monthly CBS series in real terms, seasonally adjusted and log-differenced: for manufacturing – industrial production index; for construction – housing starts; for hotels – number of Israeli tourists; for retail trade and services – the trade and services revenue indices, respectively. For the new questionnaire, the reference is compiled as the log difference between each monthly level and the previous one; for the old questionnaire, as a log difference between moving averages of three subsequent months, lagged by three months.

3. Firm-level relationship between qualitative and quantitative data

The panel datasets constructed for this analysis merge qualitative evaluations of past local sales provided by each firm to the survey with the corresponding monthly change (in log-difference terms) in the revenue recorded in the business register. Each dataset, constructed for the manufacturing, trade and services sectors, covers 24 months from the period 2013–14 when reported according with the old questionnaire and 24 months from the period 2016–17 when the questionnaire phrasing has changed. We excluded data of 2015 collected during the follow-up experiment in order to avoid an issue of unbalanced panel due to half sample size.

Using logistic regression of qualitative answers by quantitative changes and other controls we examined the following questions:

- Do quantitative sectoral data have a contemporaneous effect on qualitative answers? Can we isolate the aggregate sectoral effect particularly relevant for the use of the balance of opinions series?
- To what extent are the qualitative evaluations seasonally dependent?
- Is the relationship between qualitative and quantitative data affected by firms' characteristics, not captured by the sectoral balance statistics, like sub-sectoral differences in activities, differences in firm size based either on employment or on the revenue, reporting features ;
- How has the information content of the qualitative data changed as a result of the change in the wording of the questionnaire in 2015?

According to the questionnaire wording, the quantitative firm-level data in the new-questionnaire panel were processed as month-to-month changes in revenue in the month preceding the survey month and derived from the corresponding administrative records; whereas the changes in the revenue for the old questionnaire were calculated based on the mean revenue in the three-month period prior to the survey month relative to the mean revenue of the preceding three months.

To capture the overall sectoral dynamics we control for seasonally adjusted macroeconomic indicators, appropriately differenced and lagged. As detailed in Table 3, there are industrial production indices and retail trade and services revenue indices. We also map seasonal effects into 11 dummy variables.

For easier interpretation of the parameters, we aggregate "greatly increase" and "increase" responses, as well as “greatly decrease” and “decrease” responses, and move to a 3-point ordinal scale. We also winsorize extreme changes in firm-level revenue data, by setting the 97.5% and 2.5% percentiles of corresponding distributions as thresholds.[2]

We estimate the relationship between the qualitative and quantitative data through the cumulative logit link, as follows:

$$\log[\frac{\Pr(Y_{it} \leq j)}{\Pr(Y_{it} > j)}] = \alpha_j + \beta\Delta_{it} + \varphi\rho_t + \sum_{m=1}^{11} \gamma_m d\, \gamma_m + u_{it}\ ; \quad j = 1,2 \qquad (1)$$

Where the explanatory variables are denoted as following:

Y_{it} – qualitative response given by the firm i regarding domestic sales in month t;

Δ_{it} - change in the revenue in month t calculated for the firm i from the administrative records in appropriate terms, as described above;

ρ_t - macroeconomic series, seasonally adjusted and appropriately differenced;

d_m $(m = 1,\ldots 11)$ - seasonal dummies;

u_{it} - residuals;

and the estimated parameters of the relationship (1) are denoted as:

a_j $(j = 1,2)$ – two intercepts estimated for the probabilities of reporting “increase” and “no-change”, relative to reporting “decrease”;

ϕ - the parameter of the aggregate sectoral effect;

β - the parameter of the firm's idiosyncratic effect;

γ_m $(m = 1,\ldots 11)$ - seasonal effects.

---

[2] For details on the advantages of winsorization over filtering observations with outliers in similar datasets, we refer to :
Lui, S., J. Mitchell and M. R. Weale (2009), “Qualitative Business Survey: Signal or Noise”, National Institute of Economic and Social Research, London, September 2009.

Note that due to repeated measures in our datasets, the responses are not independent, so we use the GEE-method. This method delivers standard errors of the parameters much greater than would be obtained by the maximum likelihood assuming uncorrelated responses.

We estimate parameters by four different versions of (1)[3]: model (010) evaluates only cross-sectional effects, model (011) evaluates both cross-sectional and aggregate sectoral effects; while model (111) and model (211) account for the firm's size as well by using observation weights defined in terms of the number of employed or the volume of revenue, respectively. Appendices C1—C3 present estimated parameters, by industry.

In the manufacturing industry (Appendix C1), the parameters β and φ indicate statistically significant correlations with the quantitative data, while the overall sectoral effect is more significant in the new-questionnaire panel. The regression weighted by the number of employees (111) does not contribute much compared with the unweighted regression (011), and revenue-based weights (211) reduce the parameter of aggregate sectoral index.

In retail trade (Appendix C2), the parameter of aggregate sectoral index, obtained from the old-questionnaire data is insignificant or has a negative sign, but the cross-sectional effect is positive and statistically significant. For the new-questionnaire data, a significant parameter for the macro variable was obtained only for unweighted data.

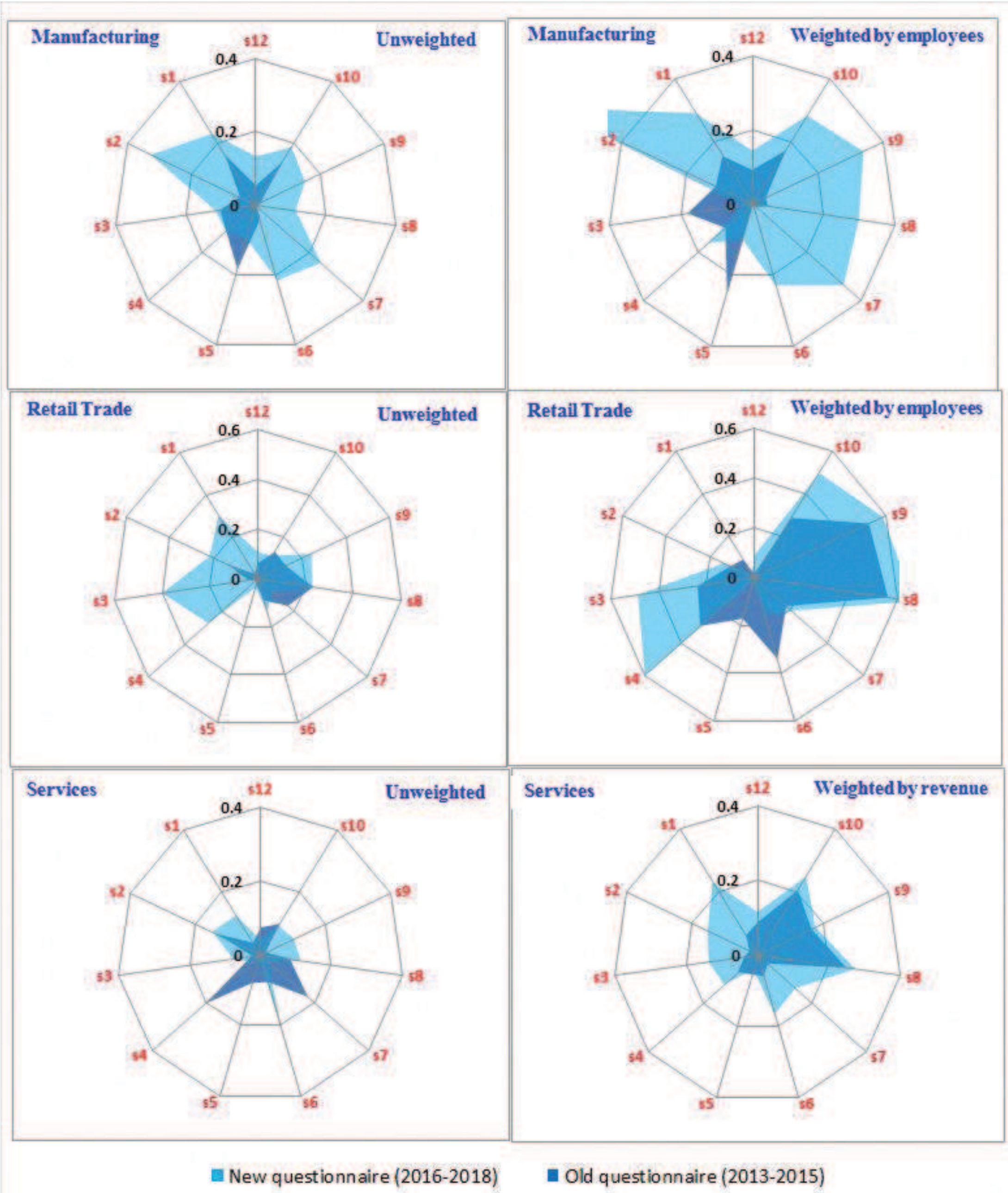
In services (Appendix C3), we have isolated statistically significant cross-sectional and overall sectoral effects from the new-questionnaire data, by using revenue-based firms' weights. Unweighted regression or employment-weighted regression failed to identify a statistically significant effect of aggregate sectoral dynamics.

Another result that stands out is an amplified seasonality of qualitative data detected in all new-questionnaire panels, which can be explained by the change in the questionnaire wording, placing a focus on a specific month of firm's activity instead of a cumulative three-month assessment.

Figure 2 depicts this result visually with help of the absolute values of the regression coefficients obtained for seasonal dummies in (1), according to various specifications.

[3] Numbers in parentheses correspond to model notation shown in Appendices C1-C3.

Figure 2. Seasonal effects[1] (in absolute terms) in the qualitative answers, by industry and questionnaire type



[1] Seasonal factors were estimated through logistic equation (1) and presented here in absolute values.

Table 4 shows that the relationship between the qualitative evaluations and qualitative data depend on additional variables of the observable firms' heterogeneity incorporated into our logistic regression (for only new-questionnaire data). We assess significant effects from an improvement (a decrease) in the quasi-likelihood statistics in response to inclusion various structural characteristics.

Table 4. Values of the quasi-likelihood statistics of panel regressions[1)] in response to incorporated factors of firms' heterogeneity, by industry

Model	Additional explanatory variables [2)]	Old questionnaire (2013-2014)		New questionnaire (2016-2017)	
		QIC	QICu	QIC	QICu
Panel A. Manufacturing					
010	Change in the firm-level revenue	17222.9	17196.1	19957.8	19934.0
011	Overall sectoral index, seas. adjusted	17225.1	17197.6	19931.1	19908.1
011 +	Technology-level category			19889.0	19852.3
011 ++	Export profiles			19888.7	19848.1
Number of firms		470		524	
Number of panel observations		8733		9909	
Panel B. Retail Trade					
010	Change in the firm-level revenue	7009.0	6988.7	8364.5	8340.5
011	Overall sectoral index, seas. adjusted	7010.6	6990.7	8359.1	8335.4
Number of firms		193		228	
Number of panel observations		3515		4196	
Panel C. Services					
010	Change in the firm-level revenue	18245.8	18215.2	21968.1	21932.8
011	Overall sectoral index, seas. adjusted	18239.3	18209.2	21942.9	21906.8
011#	Common reporting to tax authorities			21940.1	21897.0
011##	Sub-sector division			21824.4	21725.3
011###	Export profile			21711.7	21678.6
Number of firms		564		717	
Number of panel observations		10525		13009	

[1)] Variables of observed heterogeneity were tested only on new-questionnaire data.

[2)] The heterogeneity factors listed within each panel were added one by one, so a corresponding quasi-likelihood was recorded.

In manufacturing sector we examine the effect of different technology levels (high, medium-high, medium-low and low, model 011+) and export profiles, categorized as “low”, “medium” or “high” according to the percentiles of firms’ distribution by export shares in revenue in past years (model 011++). In the services industry we check whether differences in the way that revenue data have been reported to the tax authorities[4] may have some effect (model 011#). Additional dummies were assigned to specify sub-sectoral division, as banks, business services, accommodations, IT-services, transportation and other (model 011##), as well as export profile (model 011###).

As can be seen, accounting for sub-sectoral heterogeneity in the services industry led to the greatest improvement in QIC/QICu-statistics, decreasing by 0.5%/0.8%. Differences in the level of technology between manufacturing firms have a smaller effect, 0.2%/0.3%. In addition, differences in the export profile have a greater impact in services (a decrease of 0.5%/0.2% in QIC/QICu-statistics) than in manufacturing (0.1%/0.1%).

4. Nowcasting of sectoral indices by survey variables

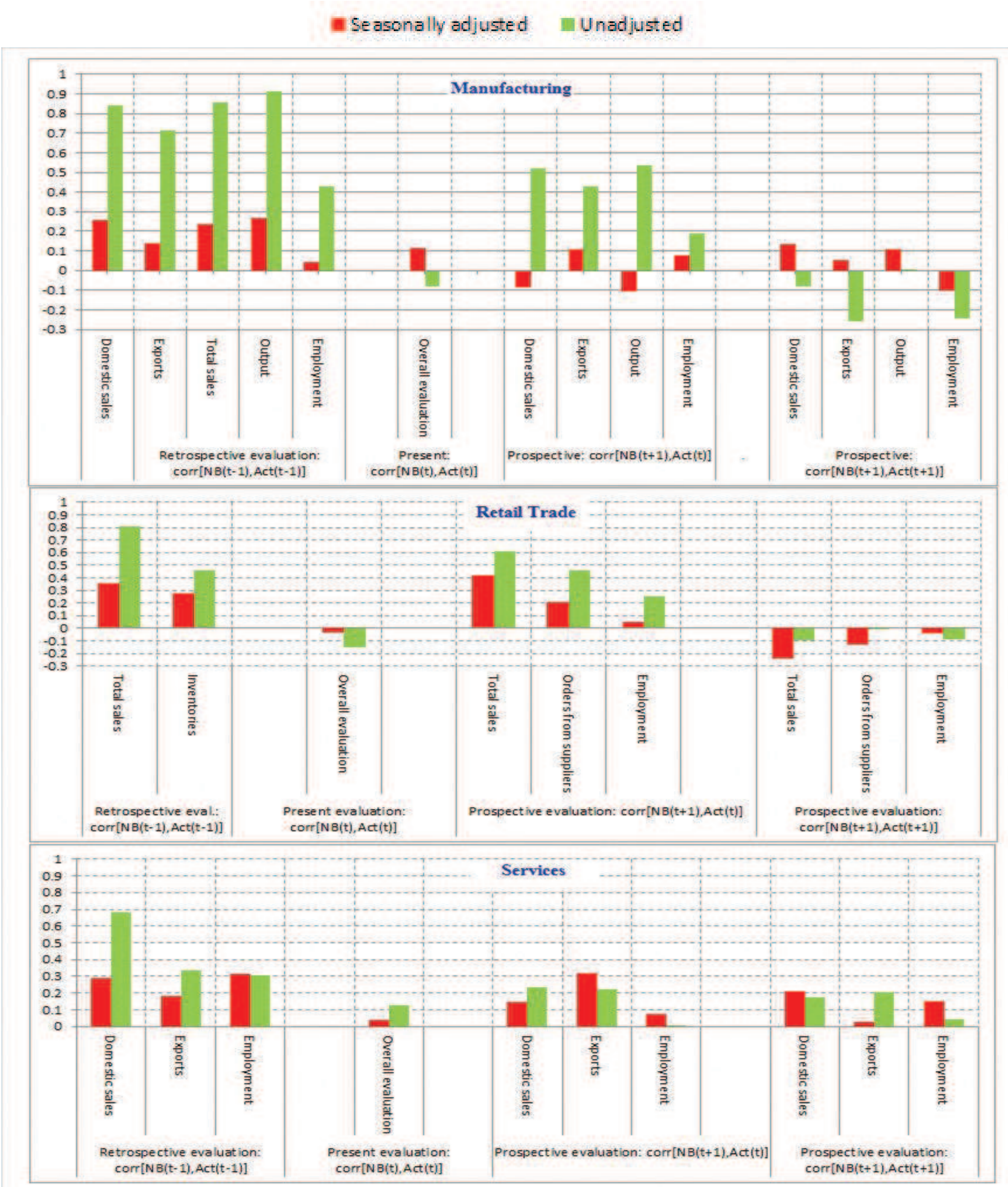
This section deals with the survey data, converted into balances of opinions series that could be used for short-term forecasts.

Figure 3 depicts the correlations between the new-questionnaire balances of opinions and contemporaneous month-to-month changes in manufacturing, retail trade and services indices—unadjusted as well as adjusted for seasonality—by various aspects of the activity (sales in the domestic market, exports, employed persons, etc.). The correlation coefficients are represented by colored columns and grouped by the time horizon required for providing an evaluation—retrospective, present situation or prospective. According to the notation of the month-of-interest on which a survey focuses, we denote the balances of opinions as $NB(t-1)$, $NB(t), NB(t+1)$ and the corresponding changes in macro-variables as $Act(t-1)$, $Act(t), Act(t+1)$.

---

[4] About 30 percent of firms are allowed to submit a consolidated report to the tax authorities and some others (small firms) report on a two-month basis.

Figure 3. Correlations between the balances of opinions and corresponding seasonally adjusted/unadjusted sectoral indicators, by industry, activity aspect and time perspective, based on the new questionnaire series (2016:01 – 2018:12) [1)]



[1)] $NB(t)$ denotes the balance of opinions related to the month t , $Act(t)$ denotes the sectoral index of month t.

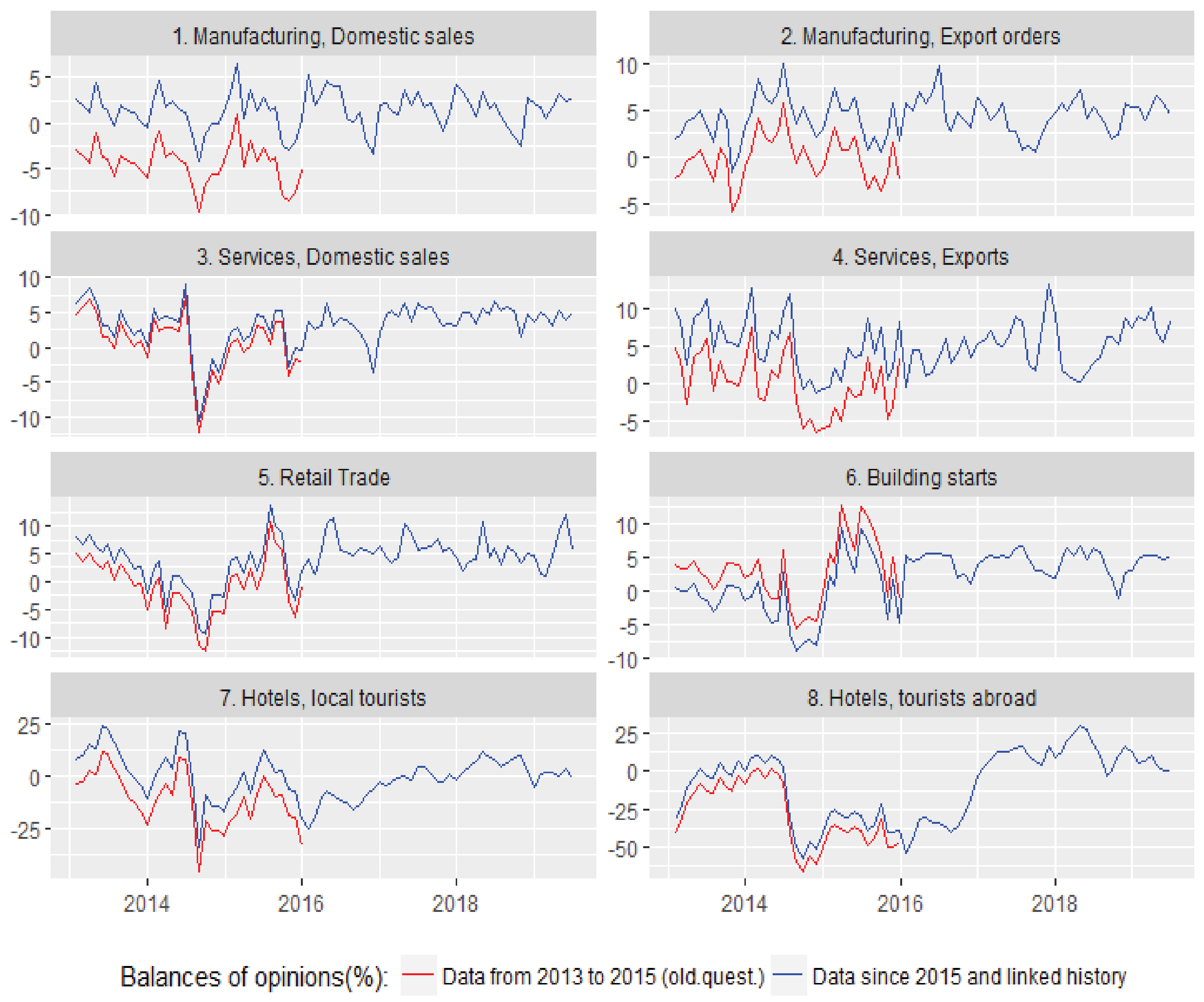
As shown, the retrospective balances of opinions correlate with the contemporaneous macroeconomic indicator, while higher correlations were recorded for seasonally unadjusted data. In contrast, the balances of the present situation do not correlate with the corresponding sectoral indicator.

The prospective balances do not correlate with the actual data for the month to which they relate; however, some correlations were documented with the activity in the survey month. The correlations of retrospective and prospective balances of opinions with quantitative data for two consecutive months creates the possibility of smoothing seasonal volatility used below.

The main issue of forecasting by survey variables is discontinuity of balances of opinions over a relatively short period. We try to eliminate the structural break that occurred in the balances of opinions between 2015 and 2016 using the series from the treatment and control groups. These data are available over 9 months of the follow-up period before the new questionnaire was introduced. Since the series are short and not adjusted for seasonality, we apply the geometric-mean conversion of the retrospective and prospective balance of opinions obtained in the same survey, which should smooth seasonal effects in adjacent months.

Keeping in mind the transition from the three-month to one-month evaluation and implied differences in the volatility of the transformed series, we additionally smooth the new-questionnaire series by weighted three-month moving averages. Then, we estimate the bias and reconstruct the explanatory series starting from 2013. Appendix D describes it in more detail. Figure 4 demonstrates the time series of the old-questionnaire series alongside the chained series, used in the nowcasting equations.

Figure 4. Geometric-mean conversion of retrospective and prospective balances of opinions: old-questionnaire vs. reconstructed series for the period 2013:01–2019:05, by sector and business activity aspect



5. Monthly GDP nowcasting by survey variables

To obtain historical GDP data at monthly frequency, we perform linear-spline interpolation of quarterly seasonally adjusted GDP series, while assigning the known GDP level to the first month of the quarter and missing values to the remaining months. Then, we apply the bagging procedure suggested in Bergmeir, Hyndman and Benítez (2016) which derives the trend component, bootstraps the remainder and adds it back. To ensure enough observations for bootstrapping, we have used the GDP series since 1995.

Figure 5. Ten monthly interpolated and bagged GDP series (seasonally adjusted): upper left—in levels, upper right—log-differenced, bottom—implied quarterly changes compared to original ones (1995:Q1–2018:Q1)

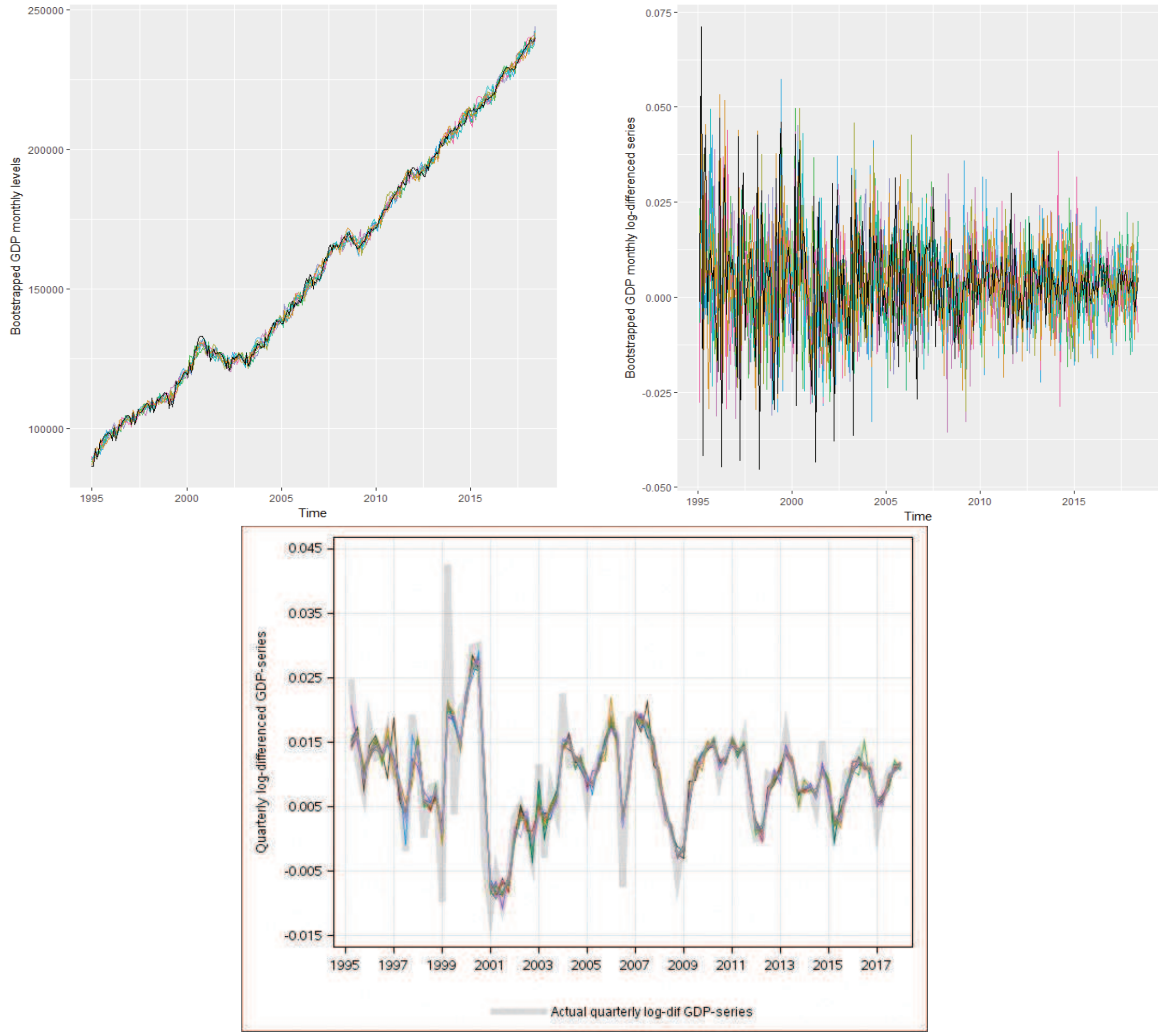


Figure 5 (the upper graph, on the left) depicts 10 monthly interpolated and bagged GDP series. The upper graph, on the right, represents these series in terms of log-differences, since they are designed for the regression. The bottom graph shows changes in quarterly aggregated bagged monthly levels. Thus, the bagging creates multiple replications of the "noised" monthly GDP trend, although the quarterly and intra-quarterly volatility are not fully addressed.

Using the bagged GDP-series (in log-difference terms) as a left-side variable, the PLS regression runs with five survey-based predictors, calculated as the geometric-mean conversion of the retrospective and prospective balance of opinions related to domestic sales and export orders in manufacturing, domestic sales and exports in services industry and total retail sales in trade. Three additional candidates presented in Figure 4—construction activity and the number of local and foreign tourists in the hotel sector—have been filtered out because of very low VIP-scores.[5]

Having a total of M bagged GDP series (in log-difference terms) we get M different forecasts for the month of interest t, some of which, say $\left\{y_t^{(1)}, y_t^{(2)}, \dots y_t^{(M_{val})}\right\}$, $(M_{val} < M)$ passed the test set validation. This enables a density nowcast for a given month t, average (point) nowcast $\hat{y}_t$, as well as 5% and 95% distribution percentiles for estimating the confidence interval.

We allow various definitions of the dependent variable, like GDP at market prices, excluding import taxes, at basic prices, as well as the business-sector GDP at market and basic prices.

Due to the different frequency of data and unsynchronized publications of the Business Tendency Survey and National Accounts, the forecast horizon in our GDP equations varies from two to four months: the maximum forecast horizon occurs in February, May, August and November as the survey data are published 10–12 days earlier than the first GDP estimate of a new quarter. In contrast, in March, June, September and December, the model provides forecasts for up to two months.

Recent experience with this model can be summarized as follows.

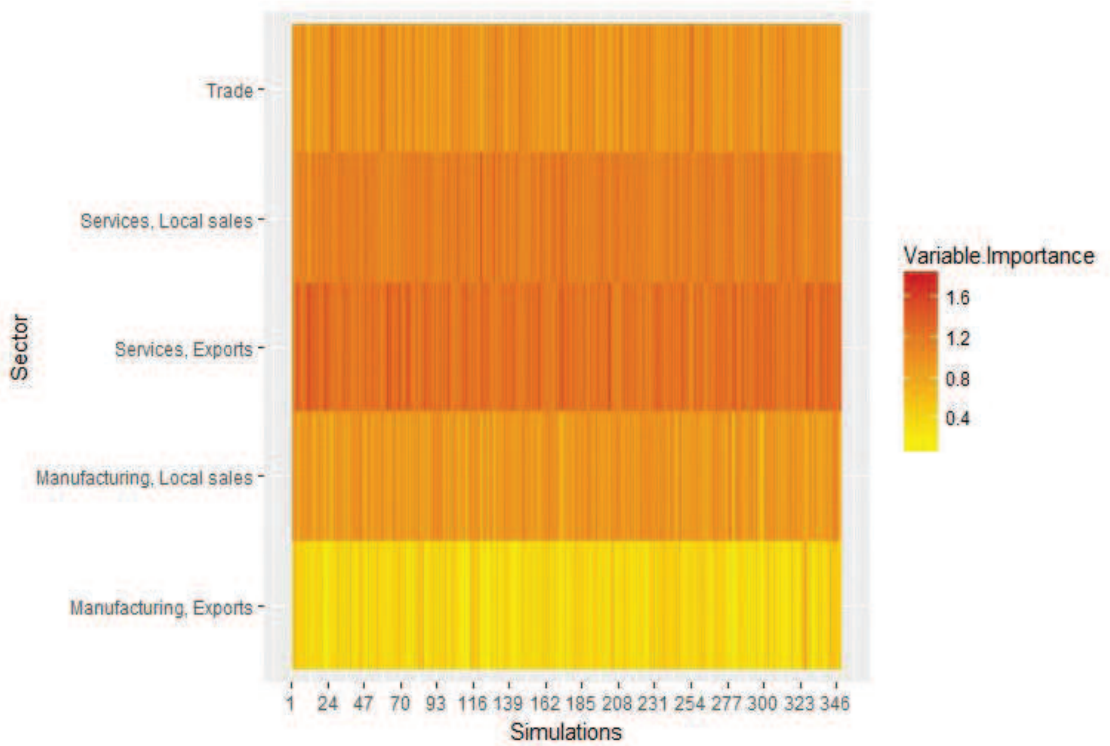
The fraction of filtered-out predictions varies between 15 percent and 20 percent, depending on the data and the type of dependent GDP series. The nowcast estimates of

[5] See Appendix E for more details. We filtered out variables whose VIP-scores remained below 0.8 in all simulations.

monthly GDP growth obtained so far explain between 12 percent and 30 percent of the variance in the target variable. The derived sentiment factors explain between 43.5 percent and 56.2 percent of the variance in the survey-based explanatory series. Mostly, only one sentiment factor has been derived: the hazard rates are about 87 percent for the nowcasts obtained with respect to the GDP at market prices and business-sector GDP, 76 percent for the GDP excluding import taxes and 67 percent for the GDP at basic prices.

Figure 6 depicts the VIP-scores of the explanatory series. It can be seen that the relative importance of the service sector is the largest, although these variables were found to provide a poor fit with respect to the monthly revenue index of services. Retail sales also show considerable relative importance, and no less than industrial sales in the domestic market, despite its smaller weight in the industrial GDP-composition.

Figure 6. The heatmap of VIP-scores[1] of the explanatory variables recorded from the simulations in January, 2019 with respect to the GDP at market prices



[1] The definition of the relative importance measure VIP is given by (E.4) in Appendix E.

The relative importance of export orders in manufacturing is the lowest. We do not remove this variable because in some simulations the corresponding VIP-scores exceed 0.8.

Table 5 presents a summary of the model coefficients, by type of the dependent GDP variable. These statistics are calculated over cross-validated simulations, whose number (out of 350 for each type of the dependent variable) is also shown.

Table 5. Summary of PLS-parameters of monthly GDP regressions, by type of dependent variable

		Employee-weighted series				Revenue-weighted series[1]			
Sector	Aspect of business activity	Mean	Median	St.dev.	Number[2] of simulations	Mean	Median	St.dev.	Number[2] of simulations
Panel A: Dependent variable - GDP at market prices									
Retail Trade	Total sales	0.036	0.041	0.033		0.038	0.039	0.017	
Services	Domestic sales	**0.061**	0.067	0.042		**0.093**	**0.094**	0.012	
Services	Exports	**0.017**	0.013	0.030		**0.005**	**0.013**	0.014	
Manufacturing	Domestic sales	0.171	0.164	0.052	288	0.158	0.158	0.022	334
Manufacturing	Exports	0.092	0.094	0.018		0.089	0.090	0.016	
Panel B: Dependent variable - GDP at market prices, excluding import taxes									
Retail Trade	Total sales	0.057	0.059	0.018		0.072	0.072	0.010	
Services	Domestic sales	**0.099**	0.101	0.027		**0.105**	**0.106**	0.012	
Services	Exports	**0.062**	0.061	0.025	294	**0.024**	**0.037**	0.012	332
Manufacturing	Domestic sales	0.138	0.135	0.032		0.147	0.148	0.016	
Manufacturing	Exports	0.066	0.067	0.016		0.041	0.041	0.015	
Panel C: Dependent variable - Business-sector GDP at market prices									
Retail Trade	Total sales	0.074	0.074	0.010		0.054	0.055	0.012	
Services	Domestic sales	**0.089**	0.089	0.011		**0.127**	**0.127**	0.014	
Services	Exports	**0.009**	0.009	0.015	312	**0.041**	**0.041**	0.014	336
Manufacturing	Domestic sales	0.152	0.151	0.014		0.126	0.126	0.017	
Manufacturing	Exports	0.044	0.044	0.014		0.062	0.062	0.015	
Panel D: Dependent variable - GDP at basic prices									
Retail Trade	Total sales	0.064	0.063	0.028		0.187	0.182	0.044	
Services	Domestic sales	0.065	0.084	0.082		**0.172**	**0.173**	0.025	
Services	Exports	0.155	0.142	0.068	294	**0.098**	**0.099**	0.033	178
Manufacturing	Domestic sales	0.142	0.131	0.061		0.126	0.126	0.031	
Manufacturing	Exports	0.089	0.090	0.034		0.029	0.025	0.018	
Panel E: Dependent variable - Business-sector GDP at basic prices									
Retail Trade	Total sales	0.100	0.099	0.026		0.166	0.163	0.056	
Services	Domestic sales	0.082	0.100	0.079		**0.150**	**0.155**	0.056	
Services	Exports	0.156	0.140	0.069	296	**0.144**	**0.142**	0.038	255
Manufacturing	Domestic sales	0.146	0.131	0.071		0.078	0.076	0.041	
Manufacturing	Exports	0.015	0.024	0.053		0.028	0.022	0.020	

[1] An experimental version with two revenue-weighted explanatory variables of the service industry; the data have been replaced for the months from 2016:01 to 2019:02.

[2] The number of simulations supported by the test set validation, out of a total of 350 simulations.

In the services industry, there is some evidence that the estimated parameters have strengthened since the balances of opinions are weighted in terms of revenue rather than employment. However, it is too early to draw conclusions, because the revenue-weighted series are currently available only from 2016 and the results are not completely comparable.

Figure 7. Statistical densities of the parameters, by type of dependent GDP variable

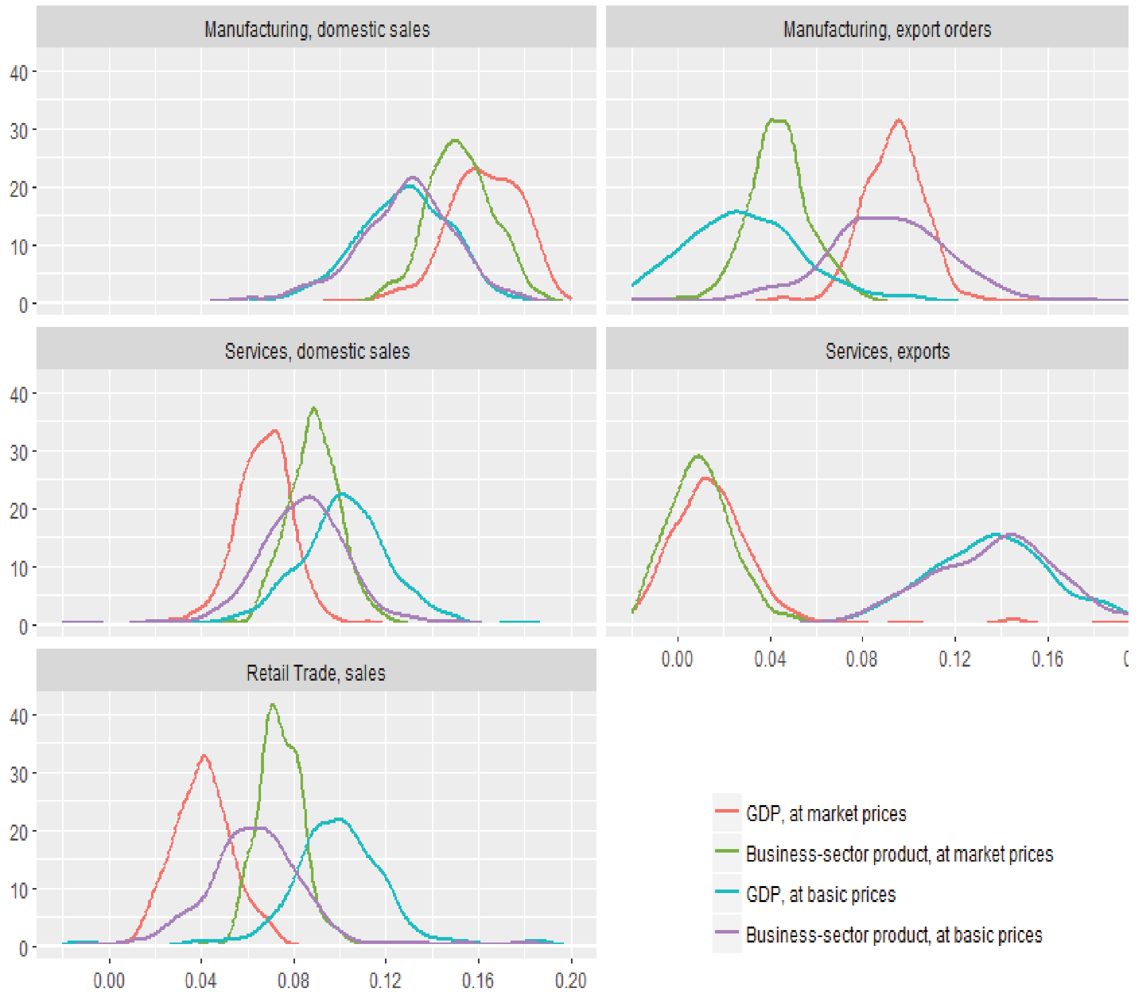


Figure 7 depicts statistical density of the parameters, by explanatory variable and the type of the dependent GDP variable. In this figure, the effect of the left-side variable is especially noticeable for exports in the services sector, since the densities of the corresponding coefficients are quite different. This difference is consistent with other simulation results showing better performance of the model with respect to the GDP at basic prices than at market prices.

Figure 8 depicts the in-sample fit alongside the actual quarterly growth of business-sector GDP, spread uniformly over the quarter. For comparison, Figure 9 depicts the aggregated balance of opinions (calculated with fixed weights) alongside the monthly simulated GDP growth.

Figure 8. In-sample fit of the PLS-sentiment with respect to GDP growth, at market prices (2013:01–2019:05)

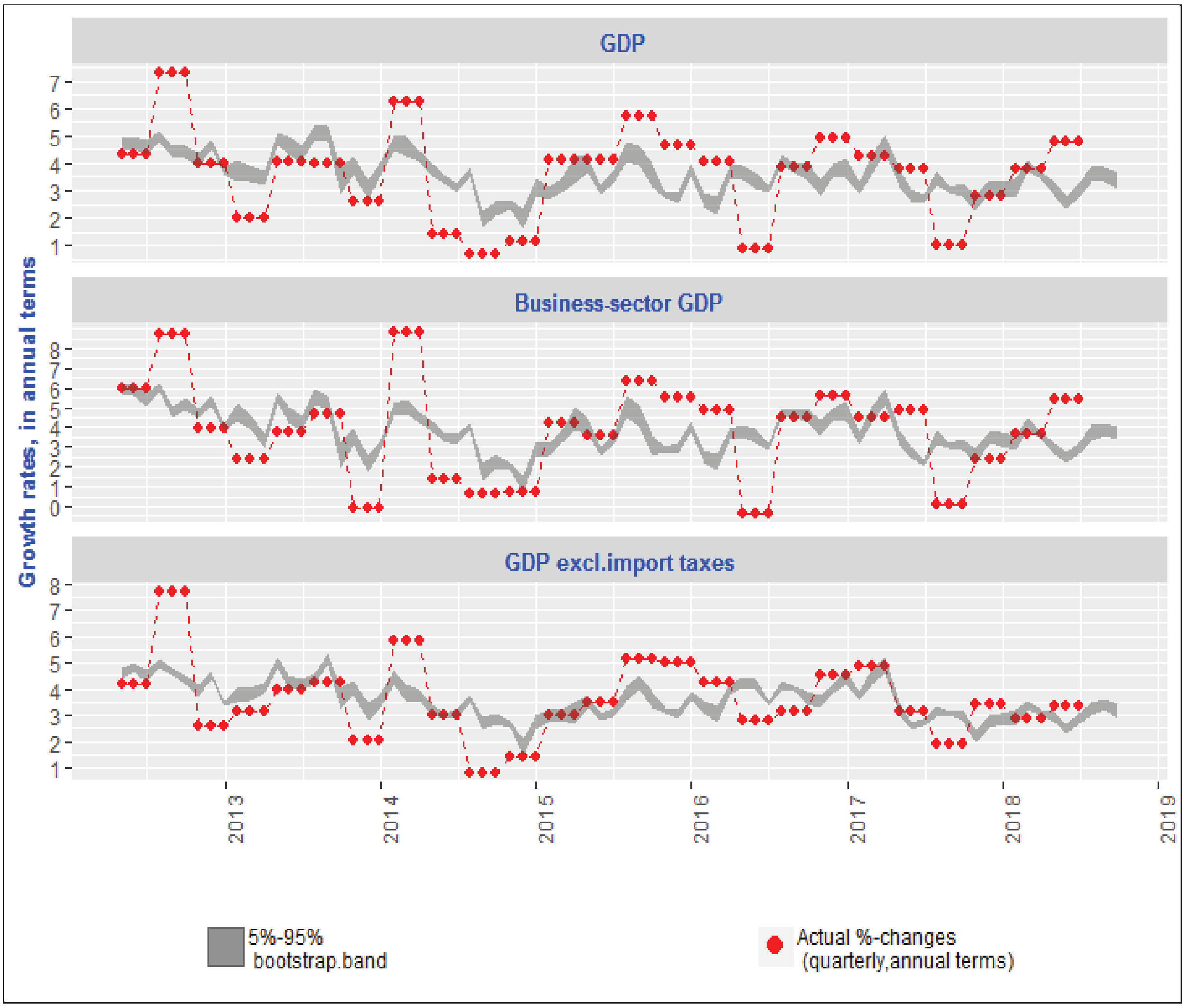
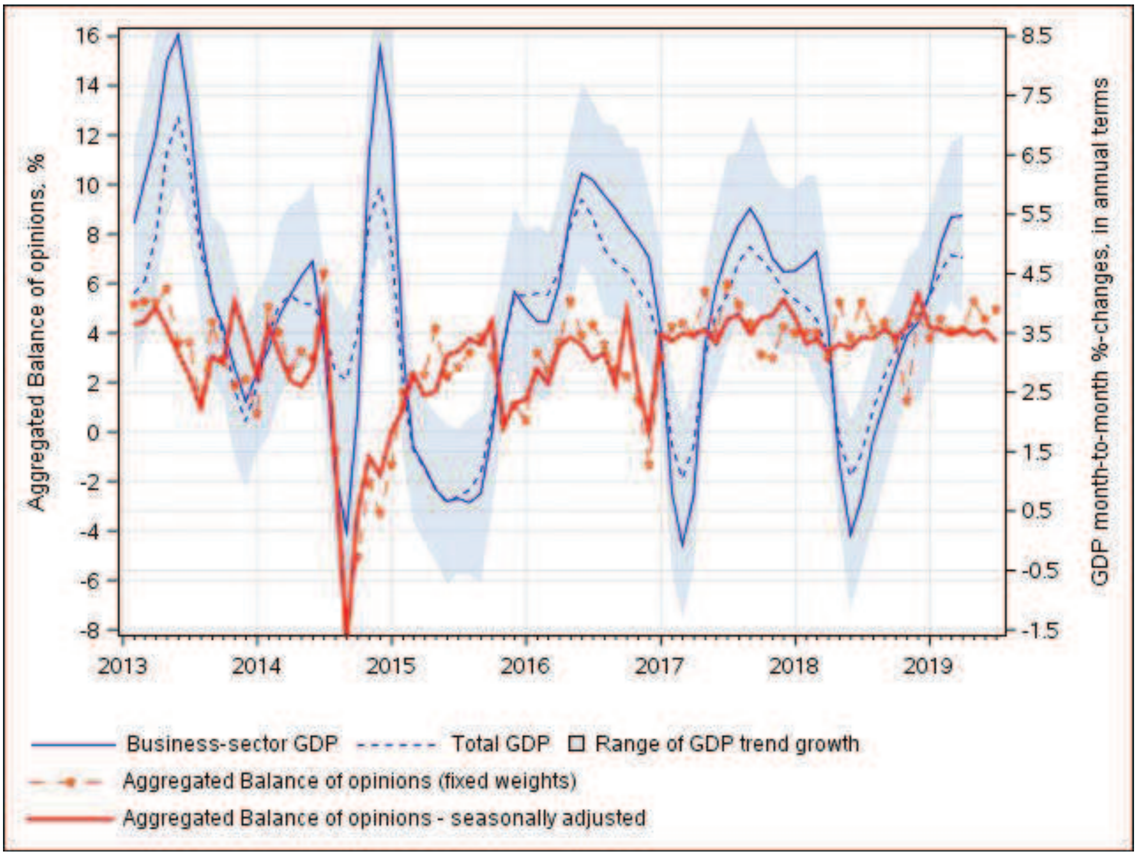


Figure 9. Aggregated Balance of opinions (with fixed weights)[1] alongside estimated trend[2] of monthly GDP-growth (in annual terms, 2013:01-2019:05)



[1] Seasonally adjusted balance has been calculated as a composite series, based on seasonally adjusted sectoral components;

[2] Monthly GDP growth estimates have been obtained through monthly interpolation and bootstrapping of quarterly seasonally adjusted GDP growth rates

Table 6 shows that the sentiment indicator with regression-based weights is more closely correlated with the target variable than the sentiment with fixed weights. As seen from the table, the PLS-sentiment provides a better fit for GDP, adjusted for import taxes, which is less affected by the volatility of consumer imports.

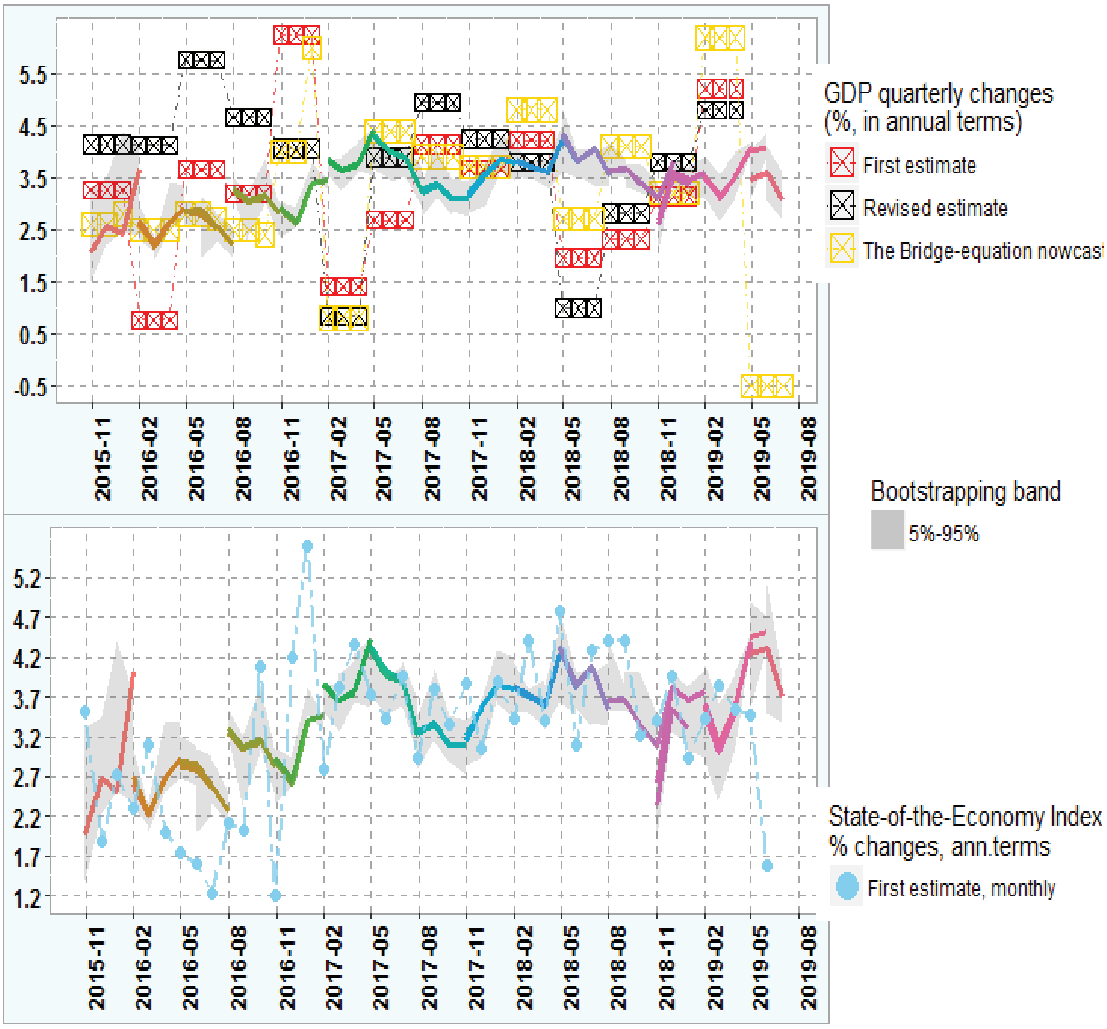
Table 6. In-sample correlations between the sentiment indicators and monthly GDP growth (at market prices), by type of target variable over the period 2013:01–2019:05

	Pearson correlations (Prob >|r| under H0:Rho=0)	
	with the target	between the Sentiment indicators
Panel A. The target variable is GDP		
PLS-Sentiment	**0.311**	**0.701**
	(0.007)	(<0.0001)
Aggregated balance of opinions[1)]	0.031	
	(0.791)	
Panel B. The target variable is GDP excluding import taxes		
PLS-Sentiment	**0.530**	**0.272**
	(<0.0001)	(0.016)
Aggregated balance of opinions[1)]	0.103	
	(0.378)	
Panel C. The target variable is business-sector GDP		
PLS-Sentiment	**0.285**	**0.867**
	(0.013)	(<0.0001)
Aggregated balance of opinions[1)]	0.087	
	(0.459)	

[1)] The aggregate balance of opinions is an alternative construction of the sentiment indicator which summarizes all available chained sectoral variables with fixed weights, assigned according to industrial GDP-composition.

Figure 10 shows real-time nowcasts of the monthly GDP (top graph) and business-sector GDP (bottom graph) growth, collected for the period 2015:11–2019:05, while different color lines denote the sequence of monthly nowcasts released within the same vintage. The official CBS figures of the GDP quarterly changes, first releases and revised estimates are shown in the upper graph. Substantial revisions of the 2016 data made retrospectively explain, to a large extent, bigger out-of-sample errors of the model in this period. For comparison, we also show nowcasts of the bridge equation model, which is estimated in the Bank of Israel once a quarter (first estimate) and leads the first CBS release by two months.

Figure 10. The out-of-sample monthly GDP nowcasts[1)] alongside the actual growth rates and other real-time estimates[2)] for the period 2015:11–2019:05



[1)] The upper graph presents the real-time forecasts simulated with respect to GDP, the bottom—with respect to business-sector GDP, at market prices. The colored lines denote real-time vintages of monthly GDP nowcasts, obtained at the beginning of each month.

[2)] The monthly nowcasts simulated by the PLS-model are shown alongside the quarterly nowcasts of the bridge equation (total GDP) and the monthly State-of-the-Economy Index (business-sector GDP).

The bottom graph show that the real-time PLS-nowcasts of monthly business-sector GDP growth rates are highly correlated with the Composite State-of-the-Economy Index[6], which summarizes ten monthly available macroeconomic indicators and runs around the 20th day of each month. Since the sentiment and the Composite indices vary in the same range and the latter—by construction—evaluates the real-time monthly business-sector GDP growth, we conclude that comparable monthly GDP nowcasts could be obtained at the beginning and at the end of each month, based on different datasets.

Due to short data span of survey-based explanatory series, the follow-up period which allows to calculate out-of-sample errors relative to actual growth rates includes only 14 quarters, from 2015:IV to 2019:I. This period is also characterized by large revisions of the 2016 GDP data which negatively affected the predictive accuracy of short-term forecasts.

Table 7. Mean absolute quarterly forecast error (%) of the PLS-sentiment model compared[1)] to the Nowcast and the average-growth assumption, calculated for the follow-up between 2015:IV and 2019:I, in annual terms

	Target variable		
	Total GDP	GDP excl	Business-
Average revison in growth rates	1.18	0.96	1.48
Panel A. Relative to the first GDP-estimate			
Nowcast model	0.78 **		
PLS-sentiment	0.83 **	0.95 *	1.19 **
Mean-growth baseline assumption	1.20	1.09	1.83
Panel B. Relative to the revised GDP-estimate			
Nowcast model	1.39		
PLS-sentiment	1.36	1.09	1.78
Mean-growth baseline assumption	0.79 **	0.87 *	1.11 **

[1)] * and ** denote significance at 5% and 1% level of MDM-statistic indicating whether the model (PLS or nowcast-equation) outperforms the baseline assumption of mean growth, known in real time and calculated over a rolling window of 52 quarters. For details on the MDM statistic for small sample see Harvey et al. (1998):

[6] For details see: https://www.boi.org.il/en/Research/Pages/ind.aspx.

Table 7 summarizes out-of-sample errors, obtained using the quarterly (implied) sentiment indicator and compared – in terms of the mean absolute forecast error (MAFE) - with errors of the Bank of Israel's Nowcast which is currently in use (with respect to total GDP growth) and the baseline assumption of mean growth, calculated in real-time over a rolling window of 52 quarters.

As can be seen, the predictive ability of the sentiment regarding the first estimate of GDP growth rates is close to the Nowcast and both models deliver smaller out-of-sample errors than the average growth assumption. Furthermore, the sentiment index provides forecasts a few weeks earlier and updates its forecasts during a quarter, as soon as a new survey becomes available. However, out-of-sample errors calculated relative to revised GDP data for this period are much larger and both models failed to outperform a simple assumption of average growth.

Conclusion

Our analysis shows that the Business Tendency Surveys provide new information for nowcasting on a monthly basis through retrospective and prospective balances of opinions, related to specific aspects of business activity and focused on a specific month.

We find that retrospective balances of opinions are positively correlated with the corresponding monthly changes in sectoral macroeconomic indices, but this covariance is largely due to seasonal effects. Overall evaluations of present business situation do not correlate with industry indicators. Prospective balances of opinions are weakly correlated with the month of survey data, but not with the data of the next month to which they refer. Based on this, a geometric-mean conversion of retrospective and prospective balances of opinions was proposed for the linkage between the old and new-questionnaire series and for the smoothing until a seasonal adjustment procedure is feasible.

The firm-level data of the new questionnaire yield significant seasonal and cross-sectional variance and provide evidence that firms' heterogeneity significantly affects the reliability of qualitative answers, particularly in the services industry.

The monthly sentiment indicator calculated from five sectoral balances of opinions with regression-based weights is positively correlated with the GDP trend, derived ex post in terms of month-to-month changes. These real-time monthly GDP estimates are available much earlier than quarterly estimates of the bridge-equation and less affected by the volatility of consumer imports.

The out-of-sample GDP nowcasts vary in the same range as the Composite State-of-the-Economy Index, based on macroeconomic indicators of sectoral activity, exports, imports, employment and vacancy rate. This makes it possible to obtain comparable monthly estimates of GDP at the beginning and at the end of the month.

References

Armeanu, D., J.V.Andrei, L.Lache and M.Panait (2017). "A Multifactor Approach to Forecasting Romanian Gross Domestic Product in the Short Run"—URL: https://doi.org/10.1371/journal.pone.0181379

Banbura, M. and Rünstler, G. (2011). "A Look into the Factor Model Black Box: Publication Lags and the Role of Hard and Soft Data in Forecasting GDP", *International Journal of Forecasting*, 27, 333–46.

Bergmeir, C., R. J. Hyndman, and J. M. Benitez (2016). "Bagging Exponential Smoothing Methods using STL Decomposition and Box-Cox Transformation", International Journal of Forecasting 32, 303-312.

Bruno G., L. Crosilla and P. Margani (2016). "Inspecting the Relationship between Business Confidence and Industrial Production: Evidence Based on Italian Survey Data", ISTAT Rome, Italy, URL: https://ec.europa.eu/info/sites/info/files/file_import/g1_istat_it_new_normal_ciret2016_0.pdf

Chong, I. and C. Jun (2005). "Performance of Some Variable Selection Methods When Multicollinearity is Present", Chemometrics and Intelligent Laboratory Systems, 78 (2005), 103-112.

Frale, C., M. Marcellino, G.L.Mazzi and T.Proietti (2010) "Survey data as coincident or leading indicators" - *Journal of Forecasting*, 2010, vol. 29, issue 1-2, 109-131

Gelper, S. and C. Croux (2010). "On the Construction of the European Economic Sentiment Indicator", Oxford Bulletin of Economics and Statistics, 72.1 (2010) 0305–9049.

Hansson, J., Jansson, P. and Löf, M. (2005). Business Survey Data: Do They Help in Forecasting GDP Growth?", International Journal of Forecasting, 21, 377–89.

Harvey, D.L., S.J. Leybourne and P. Newbold (1998). "Tests for Forecast Encompassing", *Journal of Business and Economic Statistic,* Vol.16, No 2, 254-259.

Hölzl, W. (2015). "Business Cycle Dynamics and Firm Heterogeneity. Evidence for Austria Using Survey Data", Working Paper, December 2015, Austrian Institute of Economic Research.

Kaufmann, D and Scheufele, R. (2017). "Business Tendency Surveys and Macroeconomic Fluctuations", *International Journal of Forecasting*, vol. 33(4), p.p. 878-893.

Malgarini M. (2011). "Industrial Production and Confidence after the Crisis: What's Going On?", Fifth Joint EU-OECD Workshop on International Developments of Business and Consumer Tendency Surveys.

Martinsen K., F. Ravazzolo and F. Wulsberg (2013). “Forecasting Macroeconomic Variables Using Disaggregate Survey Data”, Norges Bank, 2013.

Mitchell J., R.J. Smith, M.R. Weale, S.Wright and E.L.Salazar (2005) "An Indicator of Monthly GDP and an Early Estimate of Quarterly GDP Growth" – *The Economic Journal*, 155 (February), F108-F129.

Mogliani, M., O. Darne and B. Pluyaud (2017) "The New MIBA Model: Real-time Nowcasting of French GDP Using the Banque de France's Monthly Business Survey", *Economic Modelling*, 64 (2017), 26-39.

Mittnik, S. and P. Zadrozny (2004) "Forecasting Quarterly German GDP at Monthly Intervals Using Monthly IFO Business Condition Data”, CESifo Working Paper no 1203.

Nieuwstad, C.H. (2005). “A Comparison on the Company-Level of Manufacturing Business Sentiment Survey Data and Turnover “, Statistics Netherland, Discussion Paper No 05007.

Österholm, P. (2014). “Survey Data and Short-Term Forecasts of Swedish GDP Growth”, *Applied Economics Letters*, Vol.21, 2014 - Issue 2, p.p.135-139.

OECD (2003). “Business Tendency Surveys: A Handbook”, OECD Publications, Paris, ISBN 92-64-19894.

Szori B. (2015). “A Survey-Based Measure of Slack”, in: Economic Bulletin, ECB, June, 2015.

Tresor-Economics (2014). “Are Business Surveys Equally Successful to Forecast Economic Activity in France?”, Tresor-Economics, No 125, March 2014.

Wold, S., Johansson A, Cochi M (editors) (1993). “PLS-Partial Least Squares Projections to Latent Structures”, Leiden: ESCOM Science Publishers; 1993, p.p. 523-550.

Appendix A. Balances of opinions[1)], by industry, time perspective, and question

Industry \ Time perspective	Present		Past		Expected	
	Question	No	Question	No	Question	No
Manufacturing	Overall	1				
Manufacturing			Local orders	3	Local orders	10
Manufacturing			Export orders	4	Export orders	11
Manufacturing			Output	5	Output	12
Manufacturing			Sales	6		
Manufacturing			Employment	9	Employment	14
Construction	Overall	1	Overall	2	Overall	4
Construction			Building starts	3		
Construction					Employment	5
Retail Trade	Overall	1	Sales	3	Sales	5
Retail Trade					New orders	6
Retail Trade					Employment	7
Services	Overall	1				
Services			Local sales	2	Local sales	7
Services			Export sales	3	Export sales	8
Services			Employment	4	Employment	6
Hotels	Overall	1				
Hotels			Local tourists	2	Local tourists	7
Hotels			Foreign tourists	3	Foreign tourist	8
Hotels			Revenue	6		
Hotels			Employment	4	Employment	9

[1] Only questions considered in our study are shown. The BTS also asks about output prices, capacity utilization, credit volume, credit, production and market limitations that we leave aside.

Appendix B. Statistical summary of retrospective evaluations made according the new and old questionnaires over follow-up period (April–December 2015), by industry and selected questions (unweighted statistics)

Question	N of firms		Share of non-response		Share of voting "No-change"		"Increase" (upper) and "Decrease"(bottom) relatively to "No-change"		Average balance of opinions	
	New	Old	New	Old	New	Old	New	Old	New	Old
Panel A. Manufacturing										
Sales (6)	175.9 (9.8)	227.4 (12.7)	0.6% (0.4%)	0.2% (0.3%)	50.1% (3.7%)	54.4% (2.8%)	0.54 0.47	0.29 0.55	2.0%	-9.6%
Output (5)	175.6 (9.7)	226.1 (12.6)	4.2% (1.3%)	1.9% (0.6%)	58.1% (5.6%)	62.9% (2.4%)	0.39 0.35	0.23 0.37	1.8%	-5.8%
Employment (9)	175.9 (9.7)	227.1 (12.7)	0.6% (0.8%)	0.7% (0.5%)	77.7% (2.3%)	73.4% (2.6%)	0.14 0.14	0.12 0.25	-0.2%	-5.0%
Panel B. Construction										
Overall activity (2)	136.8 (7.2)	109.5 (6.4)	0.1% (0.3%)	1.7% (0.3%)	59.2% (4.3%)	66.5% (3.9%)	0.34 0.36	0.27 0.24	-2.6%	-0.3%
Building starts (3)	135.4 (7.8)	107.8 (5.4)	5.6% (1.2%)	3.9% (1.4%)	57.2% (4.8%)	63.3% (5.9%)	0.41 0.35	0.36 0.23	-1.4%	2.7%
Panel C. Trade										
Sales (3)	82 (7.7)	110.3 (4.3)	0.3% (0.8%)	1.7% (0.6%)	50.2% (8.2%)	48.7% (8.0%)	0.53 0.51	0.54 0.57	0.1%	-1.2%
Panel D. Hotels										
Revenue (6)	17.1 (2.7)	39.5 (2.5)	0.0%	4.4% (1.7%)	30.3% (17.1%)	22.2% (2.9%)	1.24 1.90	0.50 3.08	-10.1%	-34.7%
Panel C. Services										
Local sales (2)	203 (14.2)	288.4 (13.5)	6.7% (0.6%)	7.4% (1.1%)	64.5% (4.1%)	65.5% (2.1%)	0.32 0.23	0.26 0.27	2.3%	-1.1%
Employment (4)	204.1 (14.3)	288.8 (13.3)	0.5% (0.4%)	1.7% (0.5%)	71.7% (2.5%)	72.3% (1.6%)	0.25 0.15	0.19 0.19	2.9%	-1.1%

Appendix C1. Parameters[1)] of firm-level logistic regression, estimated for the manufacturing industry, by model specification[2)] and questionnaire version

		Old questionnaire (2013-2014)				New questionnaire (2016-2017)			
Model	Parameter	Estimate	StdErr	Z	Pr>|Z|	Estimate	StdErr	Z	Pr>|Z|
010	α_1	**-1.474**	0.072	-20.39	<.0001	**-1.132**	0.056	-20.05	<.0001
	α_2	**1.066**	0.067	15.91	<.0001	**1.261**	0.059	21.48	<.0001
	β	**1.118**	0.114	9.80	<.0001	**0.810**	0.067	12.04	<.0001
011	α_1	**-1.475**	0.073	-20.36	<.0001	**-1.153**	0.056	-20.44	<.0001
	α_2	**1.064**	0.067	15.87	<.0001	**1.245**	0.059	21.16	<.0001
	β	**1.119**	0.114	9.80	<.0001	**0.781**	0.066	11.80	<.0001
	φ	-0.514	0.828	-0.62	0.535	**4.888**	0.788	6.20	<.0001
111	α_1	**-1.419**	0.139	-10.20	<.0001	**-1.127**	0.118	-9.57	<.0001
	α_2	**1.263**	0.136	9.28	<.0001	**1.435**	0.121	11.90	<.0001
	β	**1.577**	0.196	8.05	<.0001	**0.984**	0.124	7.91	<.0001
	φ	-1.333	1.456	-0.92	0.360	**5.416**	1.588	3.41	0.001
211	α_1	**-1.584**	0.269	-5.89	<.0001	**-1.006**	0.159	-6.63	<.0001
	α_2	**1.828**	0.245	7.47	<.0001	**1.588**	0.148	10.71	<.0001
	β	**2.186**	0.334	6.53	<.0001	**1.048**	0.134	7.80	<.0001
	φ	-2.244	2.165	-1.04	0.300	4.397	2.399	1.83	0.067

[1)] The significance of the GEE-parameters is estimated on the basis of Z-scores, instead of χ^2 for the maximum likelihood estimation.

[2)] Four model specifications presented here are: 010 – includes quantitative changes only in a firm-level dimension, 011 – includes firm-level quantitative changes, as well as the aggregate seasonally adjusted index of industrial production (log-differenced); 111 and 211 include quantitative changes in both firm-level and sectoral dimensions and also account for the firm's size by using employment-based or revenue-based observations weights, respectively.

Appendix C2. Parameters[1)] of firm-level logistic regression, estimated for the retail trade industry, by model specification[2)] and questionnaire version

		Old questionnaire (2013-2014)				New questionnaire (2016-2017)			
Model	Parameter	Estimate	StdErr	Z	Pr>|Z|	Estimate	StdErr	Z	Pr>|Z|
010	α_1	**-1.577**	0.102	-15.42	<.0001	**-1.033**	0.084	-12.34	<.0001
	α_2	**0.889**	0.093	9.52	<.0001	**1.404**	0.083	16.92	<.0001
	β	**0.656**	0.148	4.42	<.0001	**1.597**	0.205	7.78	<.0001
011	α_1	**-1.578**	0.107	-14.78	<.0001	**-1.050**	0.083	-12.00	<.0001
	α_2	**0.887**	0.097	9.13	<.0001	**1.390**	0.084	16.59	<.0001
	β	**0.656**	0.149	4.41	<.0001	**1.553**	0.207	7.52	<.0001
	φ	0.133	2.515	0.05	0.958	**3.861**	1.386	2.79	0.005
111	α_1	**-1.360**	0.213	-6.38	<.0001	**-0.861**	0.147	-5.85	<.0001
	α_2	**1.041**	0.206	5.06	<.0001	**1.578**	0.141	11.16	<.0001
	β	**1.788**	0.458	3.91	<.0001	**3.169**	0.563	5.63	<.0001
	φ	**-6.235**	2.903	-2.15	0.032	-0.199	4.284	-0.05	0.963
211	α_1	**-1.957**	0.436	-4.48	<.0001	**-1.243**	0.332	-3.75	<.0001
	α_2	0.255	0.560	0.45	0.6491	**2.097**	0.287	7.31	<.0001
	β	0.517	0.717	0.72	0.4704	**3.620**	0.588	6.15	<.0001
	φ	-2.313	3.056	-0.76	0.449	-5.055	4.056	-1.25	0.213

[1)] The significance of the GEE-parameters is estimated on the basis of Z-scores, instead of χ^2 for the maximum likelihood estimation.

[2)] Four model specifications presented here are: 010 – includes quantitative changes only in a firm-level dimension, 011 – includes firm-level quantitative changes, as well as the aggregate seasonally adjusted index of retail trade revenue (log-differenced); 111 and 211 include quantitative changes in both firm-level and sectoral dimensions and also account for the firm's size by using employment-based or revenue-based observations weights, respectively.

Appendix C3. Parameters1) of firm-level logistic regression, estimated for the services industry, by model specification2) and questionnaire version

		Old questionnaire (2013-2014)				New questionnaire (2016-2017)			
Model	**Parameter**	**Estimate**	**StdErr**	**Z**	**Pr>|Z|**	**Estimate**	**StdErr**	**Z**	**Pr>|Z|**
010	α_1	**-1.695**	0.080	-21.27	<.0001	**-1.360**	0.067	-20.37	<.0001
	α_2	**1.538**	0.074	20.93	<.0001	**1.996**	0.075	26.68	<.0001
	β	**0.376**	0.072	5.23	<.0001	**0.111**	0.063	1.77	0.0771
011	α_1	**-1.728**	0.081	-21.45	<.0001	**-1.391**	0.067	-20.66	<.0001
	α_2	**1.507**	0.073	20.59	<.0001	**1.972**	0.075	26.25	<.0001
	β	**0.383**	0.072	5.29	<.0001	0.106	0.063	1.68	0.092
	φ	**4.196**	1.226	3.42	0.001	**5.345**	0.902	5.93	<.0001
111	α_1	**-1.747**	0.119	-14.64	<.0001	**-1.367**	0.146	-9.34	<.0001
	α_2	**1.677**	0.168	9.98	<.0001	**2.189**	0.181	12.10	<.0001
	β	-0.178	0.271	-0.66	0.5118	0.099	0.120	0.83	0.409
	φ	3.728	1.873	1.99	0.047	**6.821**	1.537	4.44	<.0001
211	α_1	**-1.708**	0.283	-6.03	<.0001	**-1.035**	0.191	-5.42	<.0001
	α_2	**1.397**	0.273	5.11	<.0001	**1.865**	0.263	7.10	<.0001
	β	**0.431**	0.143	3.01	0.0026	**0.595**	0.243	2.45	0.010
	φ	0.615	3.235	0.19	0.849	**8.348**	3.168	2.63	0.008

[1)] The significance of the GEE-parameters is estimated on the basis of Z-scores, instead of χ^2 for the maximum likelihood estimation.

[2)] Four model specifications presented here are: 010 – includes quantitative changes only in a firm-level dimension, 011 – includes firm-level quantitative changes, as well as the aggregate seasonally adjusted index of retail trade revenue (log-differenced); 111 and 211 include quantitative changes in both firm-level and sectoral dimensions and also account for the firm's size by using employment-based or revenue-based observations weights, respectively.

Appendix D. Linkage between the old and new questionnaire balances of opinions

Denote by t the month that the survey was carried out, s - sector, v - the related aspect of business activity; then $NB_t^{bkw(s,v)}$ and $NB_t^{frw(s,v)}$ are corresponding retrospective and prospective balances of opinions derived from this survey. The geometric-mean conversion of these data is calculated as follows:

$$G_t^{(s,v)} = \sqrt{(\frac{NB_t^{bkw(s,v)}}{100} + 1) \times (\frac{NB_t^{frw(s,v)}}{100} + 1)} - 1 \qquad \text{(D.1)}$$

The applied three-month moving average of this transformation looks like:

$$\dddot{G}_t^{(s,v)} = 0.6G_t^{(s,v)} + 0.3G_{t-1}^{(s,v)} + 0.1G_{t-2}^{(s,v)} \qquad \text{(D.2)}$$

Then, we can evaluate the average gap between (D.1) and (D.2) series over the follow-up period and reconstruct the series from 2013, as follows:

$$\hat{G}_t^{(s,v)} = \begin{cases} G_t^{(s,v)} + \frac{1}{9}\sum_{t=2015/04}^{t=2015/12}\left(G_t^{(s,v)} - \dddot{G}_t^{(s,v)}\right) & if\ 2013 \leq year(t) \leq 2015 \\ \dddot{G}_t^{(s,v)}\ otherwise \end{cases} \qquad \text{(D.3)}$$

Table D1. Summary of explanatory series[1)] in the follow up period (2015:04–2015:12) and estimated bias

Sector	Aspect of business activity	Average		Standard deviation		Estimated bias
		New	Old	New	Old	
Manufacturing	Domestic sales	-5.02	0.48	2.20	1.86	5.49
Manufacturing	Export orders	-0.89	3.30	2.09	2.23	4.19
Retail Trade	Total sales	1.45	4.49	5.17	4.18	3.04
Services	Domestic sales	0.60	2.16	2.76	1.80	1.56
Services	Exports	-0.84	3.70	3.05	3.94	4.54
Construction	Building activity	6.86	3.47	4.96	3.60	-3.39
Hotels	Local tourists	-13.36	-1.19	9.33	10.65	12.16
Hotels	Foreign tourists	-42.50	-32.86	6.51	8.93	9.64

[1)] Explanatory series are calculated through geometric-mean conversion of the retrospective and prospective balances of opinions of the old and new questionnaires, as defined in (D1) and (D2), respectively.

Table D1 presents summary statistics of transformed series and estimated bias used in (E.3) for construction of linked series starting in 2013.

Appendix E. The PLS-based sentiment index

Like the principal component regression, the PLS-method constructs uncorrelated linear combinations of the predictors via eigenvalue-decomposition of the correlation matrix. The difference is that the PLS-regression identifies each new combination of the original predictors with respect to the target variable.

We regress the dependent variable y_t (seasonally adjusted and log-differenced) by p explanatory balances of opinions $X_{\overline{t}}=\{X_{1t} \cdots X_{pt}\}$ combined into h $(h < p)$ unobservable sentiment factors $T_{\overline{t}}=\{T_{1t} \cdots T_{ht}\}$, as defined below:

$$\begin{aligned} X &= TP` + e \\ y &= Tb + u \end{aligned} \qquad \text{(E.1)}$$

where

$X(nxp)$ – matrix of p explanatory variables, with n monthly observations;

$T(nxh)$ – matrix of h derived and uncorrelated sentiment factors, with n monthly observations; $T'T = I$;

$P(pxh)$ – matrix of weights of the explanatory series in the derived sentiment factors;

$b(hx1)$ – vector of regression parameters of the dependent variable by the derived sentiment factors;

$e(nxp)$ and $u(nx1)$ are random residuals of the explanatory and dependent variables, respectively.

As all variables are standardized to zero mean and unit variance, the parameters do not include an intercept. The sentiment factors $T_1, \cdots T_h$ are derived one by one; we begin with $X_{(1)} = X$, $y_{(1)} = y$ by constructing the first factor $T_{(1)} = X_{(1)}W_{(1)}$, where $W_{(1)}$ is a vector of weights proportional to the first eigenvector of the matrix $(X'_{(1)}y_{(1)}y'_{(1)}X_{(1)})$ which maximizes $T'_{(1)}y$, i.e., the correlation between the first derived sentiment factor with the target variable. The second factor is derived in a similar way, but instead of X and y we take the residuals $X_{(2)} = X_{(1)} - T_{\{1\}}P'_{(1)}$ and $y_{\{2\}} = y_{\{1\}} - T_{(1)}b_{(1)}$, respectively. This procedure is repeated until the desired number of factors is created.

The number of sentiment factors is supervised through test set validation: the optimal h is determined through the minimization of the predicted residual error sum of squares (PRESS), calculated as $\sum_{j=1}^{n} u_{h,j}^2$, where $u_{h,j}$ denotes the prediction error for y_j obtained by the model of h factors, while y_j is excluded. The test ensures that the forecast outperforms (in terms of out-of-sample mean squared error) the benchmark of the mean growth, calculated as a conditional mean of the dependent variable with deleted observation to be predicted, as follows:

$$\bar{y}_{\in j} = (\sum\nolimits_{k=1,k\neq j}^{n} y_k)/(n-1) \qquad \text{(E.2)}$$

If the PLS-regression does not outperform the benchmark forecast, in terms of PRESS, we conclude that no sentiment factors related to the dependent variable can be extracted. In other words, the explanatory variables don't pass the test set validation, if:

$$\sum\nolimits_{j=1}^{n} u_{h,j}^2 \geq \sum\nolimits_{j=1}^{n} (y_j - \bar{y}_{\in j})^2 \qquad \text{(E.3)}$$

In this case, we need to make changes in the explanatory set. The variable selection may be based on the relative importance measure (Chong and Jun, 2005), calculated as follows[7]:

$$VIP_i = \sqrt{\frac{p}{\sum_{m=1}^{h} SS(b_m T_m)} \sum\nolimits_{m=1}^{h} w_{mi}^2 SS(b_m T_m)} \qquad \text{(E.4)}$$

where

p is the number of explanatory variables;

[7] According to Wold (1993), VIP scores below 0.8 indicate the low importance of explanatory variables and this threshold may be used for variable selection.

h is the number of derived sentiment factors;

w_{mi} denotes the loading of the standardized balance of opinions X_i $(i = 1, \ldots p)$ in the sentiment factor T_m $(m = 1, \ldots h)$;

$SS(b_m T_m) = b_m^2 T_m^` T_m$ evaluates the share of variance of the dependent variable y explained by the sentiment factor T_m .

Appendix F. Industrial composition weights used for the aggregated balance of opinions

Sector	Aspect of business activity	Share in GDP[1)]	Intra-sectoral division[2)]	Component weight
Manufacturing	Domestic sales		80%	17.9%
Manufacturing	Export orders	22.4%	20%	4.5%
Retail Trade	Total sales	6.5%	-	6.5%
Services	Domestic sales		80%	47.3%
Services	Exports	59.1%	20%	11.8%
Construction	Building activity	10.3%	-	10.3%
Hotels[3)]	Local tourists		50%	0.9%
Hotels	Foreign tourists	1.7%	50%	0.9%

[1)] The industrial composition of business-sector GDP is given in the "Statistical Abstract of Israel" (2018, Table 18.1), URL: https://old.cbs.gov.il/reader/shnaton/templ_shnaton_e.html?num_tab=st18_01x&CYear=2018

Since not all sectors of the business sector are covered by the Business Tendency Survey and some are partly covered (for example, manufacturing does not include mining and quarrying, commerce is represented by only retail trade, business services do not include education etc.) and the hotels are represented separately from the services industry, although originally it is a part of the food and accommodation sub-industry), the composition was adjusted according to VAT sources.

[2)] Exports shares are calculated according to perennial VAT composition.

[3)] The hotels are covered by the Survey as a separate sector, although it is a part of the services industry (the food and accommodation sub-industry) according to industrial CBS classification. The weight of this sector was calculated from the VAT sources