Firm expectations and economic activity*

Zeno Enders, Franziska Hünnekes, Gernot J. Müller

February 2019

Abstract

In this paper we assess empirically the role of expectations for firms' production and price-setting decisions. Our analysis is based on data for German manufacturing firms included in the EBDC Business Expectations Panel. To identify the causal effect of firms' expectations on their behavior, we match firms on the basis of fundamentals and compare decisions of firms that have the same fundamentals but differ in their views about the future. We find that optimistic (pessimistic) firms are about 17 percent more likely to raise (lower) production relative to neutral firms. Similarly, we find that optimistic (pessimistic) firms are more likely to raise (lower) prices. In the second step of our analysis we construct forecast errors and match, in turn, correctly and incorrectly optimistic to neutral firms. We find that incorrect optimists and pessimists do behave differently from untreated firms in the impact period. In a third step, we quantify the contribution of undue optimism and pessimism to aggregate fluctuations.

Keywords: Expectations, Firms, Survey data, Propensity score matching, Business cycle, News, Noise, Undue optimism

JEL-Codes: E32

*Enders: University of Heidelberg and CESifo, zeno.enders@uni-heidelberg.de. Hünnekes: University of Munich, and Kiel Institute for the World Economy, franziska.huennekes@econ.lmu.de. Müller: University of Tübingen, CEPR, and CESifo, gernot.mueller@uni-tuebingen.de

1 Introduction

To what extent do firms' expectations impact current decision making? According to theory, expectations should have a first-order effect. Expectations about the business cycle take center stage in modern macroeconomic theory as firms decide on production, investment and hiring as well as on prices in a forward-looking manner (e.g., Kydland and Prescott 1982; Lucas 1973; Mortensen and Pissarides 2009; Woodford 2003). This, in turn, is essential for why and how cyclical impulses propagate and how policy announcements shape economic outcomes (e.g., Del Negro et al. 2012; Eggertsson and Woodford 2003). Yet, at an empirical level the systematic exploration of how expectations impact economic decisions and hence economic outcomes is still in its infancy. Arguably, two major difficulties are to blame. First, expectations are not directly observable. Second, expectations are responsive to changes in the economic environment; identifying a causal effect of expectations on economic decisions is therefore challenging.

In this paper, we take up the issue by relying on a particular data set and a novel identification strategy. Specifically, our analysis is based on the EBDC Business Expectations Panel (BEP), maintained by the LMU-ifo Economics & Business Data Center (EBDC) in Munich. Our sample comprises monthly observations for the period 1991–2016. In each month, several thousand German firms report their expectations regarding future production in a qualitative manner: it may increase, not change, or decrease. Similarly, firms report expectations about business cycle conditions. The survey is the basis for the ifo business climate index, a widely-observed leading indicator for economic activity in Germany (Becker and Wohlrabe 2008). In addition, the BEP contains a rich set of observations for each firm. These include a large range of measures that capture the economic and financial conditions under which firms operate.

We also exploit these data in order to identify the causal effect of firm expectations on their behavior, notably in terms of production and price setting. For this purpose, we match firms on the basis of fundamentals and compare price-setting and production decisions of firms that have the same fundamentals but differ in their views about the future. Formally, we estimate a probit model and match optimistic and pessimistic firms, in turn, with neutral firms on the basis of their propensity scores (Rosenbaum and Rubin 1983). Intuitively, we consider "optimism" and "pessimism" as a treatment that is randomly assigned across firms with the same fundamentals: we estimate the average treatment effect on the treated by comparing the behavior of treated and non-treated firms with the same probability of being treated.

We find that expectations have a significant and lasting effect on production and prices. In the impact period optimistic firms are about 17 percent more likely to raise production than neutral firms. Similarly, we also find that optimistic firms are considerably more likely to raise prices. For pessimistic firms we find an effect of the same magnitude—they are more likely to reduce production and prices.

These results are consistent with two distinct hypotheses regarding how expectations impact economic decision making. Under the first hypothesis, expectations that are orthogonal to current fundamentals are not necessarily orthogonal to future fundamentals. Put differently, expectations represent genuine information ("news") about the future that is not yet reflected in current fundamentals. Under this interpretation, expectations matter as a transmission channel, but not as an exogenous source of variation. A number of influential contributions suggest that news are indeed an important source of business cycle fluctuations (Barsky and Sims 2012; Beaudry and Portier 2006; Schmitt-Grohé and Uribe 2012). Note, however, that these studies provide only indirect evidence on the role of expectations as news. In contrast to our analysis, they do not exploit expectations data directly.

Under the second hypothesis, changes in expectations are fully exogenous and different labels are used to capture this notion, such as "noise", "sentiment" or "animal spirits".[1] In this spirit, a number of recent contributions has put forward modern models of the business cycle in which "noise shocks" play a key role (Angeletos and La'O 2013; Lorenzoni 2009). But, again, also these contributions do not explore expectations data directly.

The unique nature of our data set allows us to test the two hypotheses directly. For not only do we observe firms expectations regarding future production and business conditions, we also observe actual production and business conditions. We are thus able to construct a measure of firms' forecast errors and identify firms whose optimism or pessimism turns out to be "undue" from an ex post point of view (Pigou 1927). In the second step of our analysis we match, in turn, unduely optimistic and pessimistic firms to ex-ante neutral firms. We find that on impact unduely optimistic firms tend to raise output and prices relatively more than neutral firms on impact. These decisions tend to be reversed afterwards, however. The behavior of unduely pessimistic firms mirrors those of unduely optimistic firms.

[1] According to Keynes, animal spirits are "a spontaneous urge to action rather than inaction", which drive economic decisions beyond considerations based "on nothing but a mathematical expectation" (Keynes 1936, pp. 161–162).

In a third step, we quantify the contribution of undue optimism and pessimism to aggregate fluctuations. For this purpose we compute an aggregate measure of undue optimism and pessimism in our population of firms. Specifically, we use an ordered probit model to measure the extent of optimism and pessimism at the firm level and classify such sentiment as undue whenever we observe a forecast error ex post. Finally, we aggregate across firms and project macro variables of interest on the resulting time series for undue optimism and pessimism. We find that optimism in particular causes industrial production and prices to rise.

Our paper relates to studies that focus on the expectation formation process. In this regard, recent work by Coibion and Gorodnichenko (2012, 2015) documents the presence of information rigidities. They use data from professional forecasters. There is also work on the expectation formation process that uses the ifo survey. An early study by Nerlove (1983) finds evidence in support of an adaptive expectations model. More recently, Bachmann and Elstner (2015) show that at most one-third of the firms in the ifo survey systematically over- or underpredict their production growth one-quarter ahead. Massenot and Pettinicchi (2018), in turn, identify various factors which account for forecasting errors of firms in the ifo sample.

Very few studies investigate empirically how expectations measured by survey data impact economic decision making. An exception is Boneva et al. (2018). They study firms' expectations in a panel compiled for the Confederation of British Industry (CBI) in the United Kingdom. Their focus is on price-setting decisions. They find that survey expectations about inflation feature significantly a variant of the New Keynesian Phillips curve. Coibion et al. (2018) use an Italian firm survey to study the effects of exogenously changing decision makers' inflation expectations on firm decisions. Bachmann and Zorn (2018) seek to establish the drivers of investment and find, among other things, a role for firm expectations. Their analysis also relies on the ifo survey. Gennaioli et al. (2015), instead, analyze the Duke University quarterly survey of Chief Financial Officers and show that firm investment is explained by CFOs' expectations of earnings growth.

The remainder of the paper is structured as follows. The next section details our data and provides some descriptive statistics. Section 3 describes the estimation approach and the results of the first step of our analysis. In section 4 we zoom further into our results as we distinguish between firms with and without forecast errors. Afterwards we document the aggregate effects of firms' expectations. Section 6 concludes.

2 Data

The EBDC Business Expectations Panel (BEP) combines monthly survey data from the ifo institute and annual balance sheet data from the Amadeus and Hoppenstedt databases (EBDC-BEP 2017). The survey data is collected through four different surveys that cover German firms in four sectors: manufacturing, retail, construction, and services. The surveys include the same basic stock of questions for each sector, but the wording of these questions and answers may differ at times. In our analysis we focus on the manufacturing survey which includes the largest number of firms. Also, the wording of the survey questions is particularly suitable for the purpose of our investigation in this case.

One caveat that the responses to the survey and the balance sheet data come at different frequencies: while the survey is conducted monthly, balance sheet data is only available annually. We will use balance sheet data to predict firm expectations. To ensure that we do not use information that is not yet available when firms report expectations, we only use the most recent balance sheet data at a given point in time. For example, if a firm publishes balance sheet data every September, we will use this data for all the following months until the next balance sheet is published.[2]

The BEP sample period starts in 1986 with the manufacturing survey. In our analysis we use data from 1991 to 2016 because some of the variables we rely on are available only since 1991. The unit of observation in the manufacturing survey may either be a product or a plant, depending on the firm. As a result, many firms provide several responses per month. We conduct our analysis at the product/plant level and do not explicitly account for whether a product/plant is part of a multi-product firm. Still, in our analysis below we refer to the individual observation as a “firm” in order to ease the exposition. Table A.1 in the appendix provides some basic information on our sample. Firms stay in the survey for 66 months (5.5 years) on average and provide answers in 56 months, implying that they respond in 86% of the months in the sample. The total number of firm-month observations amounts to more than 320,000 and there are 5922 different respondents.

The BEP contains answers to large set of questions, but only a subset of those are asked regularly. In our analysis we focus on four main questions. They are listed in Table 1. Some questions change over time, especially in 2002 many changes were implemented due to a har-

[2]Our baseline specification is conservative and as we neglect potential information known to firms in the months close to but preceding the publication of the balance sheet. In appendix D.1 we pursue an alternative strategy but find that our results are robust.

Table 1: Question used from the survey

Label	Question[1]	Possible answers
Q1	Expectations for the next 3 months: Our domestic production activity regarding good XY will probably ...	increase [1] not change [0] decrease [-1]
Q2	Expectations for the next 6 months: Taking economic fluctuations into account our state of business will be...	rather more favorable [1] not changing [0] rather less favorable [-1]
Q3	Tendencies in the previous month: Our domestic production activities with respect to product XY have ...	increased [1] not changed [0] decreased [-1]
Q4	Tendencies in the previous month: Taking changes of terms and conditions into account, our domestic sales prices (net) for product XY have ...	increased [1] not changed [0] decreased [-1]

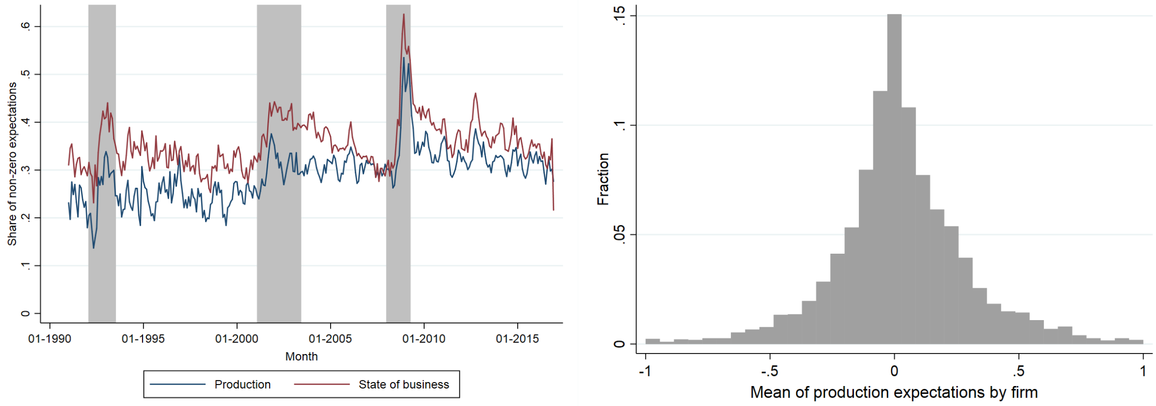
[1] Authors' translation of most recent formulation of question in German according to the EBDC Questionnaire manual.

monization of business and consumer surveys in the European Union. When necessary, we will report the changes in the questions. Details can also be found in Table B.2. Our measure of firm optimism will be based on question Q1 which refers to expectations about production activity in the next three months. The wording of this question has changed over time. Since July 1994 firms can additionally report that they have no significant domestic production. These firms are not included in our analysis. Furthermore, the question contained a note to ignore seasonal fluctuations until the end of 2000. Since these are minor changes which affect all firms in the same way, we believe they do no not affect our results.

Q2 is a broader question regarding expectations for the state of business over the next six months. Combined with a question on the current state of business it provides the basis for the ifo business climate index. In our baseline we rely on the narrower question Q1 and consider Q2 in the sensitivity analysis. Furthermore, we show in the following that both questions are highly correlated, see Figure A.1 in the appendix. Until 1997, question Q2 included the additional clarification "after elimination of purely seasonal fluctuations". Questions Q3 and Q4 measure our outcome variables: changes in production and prices. Also these questions changed in 2002. Before 2002 both questions asked about the change in production and prices in the current month compared to the previous month. Since 2002 both questions ask about the change in the variable in the previous month. We adjust the data to account for this change in timing. Also, to make sure that this adjustment does not affect results, we consider in our sensitivity analysis a reduced sample which starts in 2002 only. The results are very similar to those in the full sample (see Section 3).

Figure 1: Distribution of expectations, 1991-2016

(a) Share of non-zero answers (b) Average response by firm



Notes: Shaded areas mark recession periods as defined by the German Council of Economic Experts. Panel (b) only includes firms which respond at least 10 times.

In what follows, we compute a number of descriptive statistics for the variables of interest. For this purpose we assign a value of 1 to positive responses (*increase/improve*) and a value of -1 to negative responses (*decrease/worsen*) and a value of 0 otherwise. Figure 1 shows the distribution of responses over time and across firms. Distributional statistics are not straightforward to compute with qualitative data. We present two different approaches here. In panel (a) of Figure 1 we plot the sum of the non-zero responses, that is, the number of firms which respond that they expect a change in either direction. The figure reveals that the number of non-zero responses peaks during crises and tends to decrease before them. Furthermore, more firms report no change for production expectations than for state of business expectations. This might partially be due to the longer horizon of the state of business question. Panel (b) focuses on differences across firms. It plots the average response of each firm to Q1 during the period the firm is in the sample.[3] We find that the share of notorious optimists and pessimists is limited. Most firms' average response is zero. The distribution is close to normal.

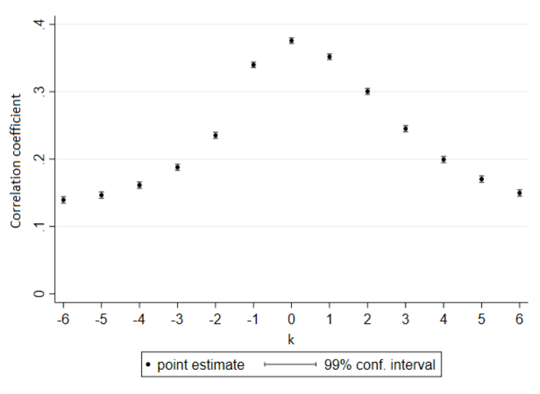
Figure 2 plots the correlation of expected production with six leads and lags of realized production and prices. The correlation between the expectations and reported outcomes is positive in both cases, across all leads and lags. The contemporaneous correlation of current production and prices, on the one hand, and production expectations, on the other hand, is particularly strong. In our analysis below we seek to establish the causal effect of production expectations on production and price setting decisions.

We also consider how average firm expectations within a month co-move with aggregate

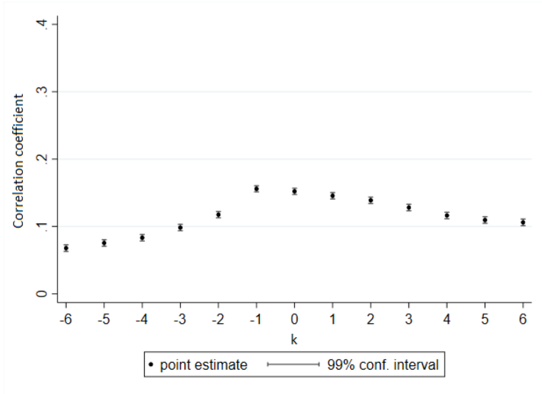
[3]Here we consider only firms which respond at least 10 times in total.

Figure 2: Correlations of expected changes in production with changes in realized in production and prices in the manufacturing sector

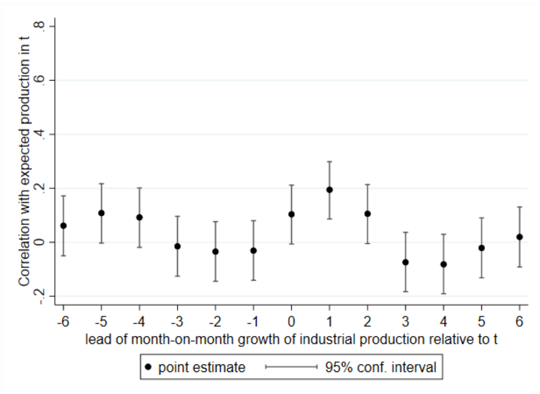
(a) Expected production in t and reported production $t+k$, firm level



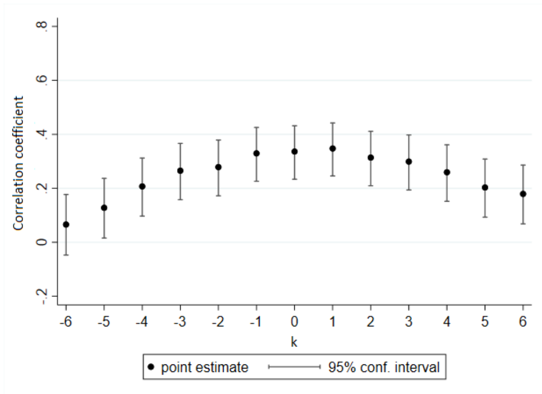
(b) Expected prices in t and reported prices $t+k$, firm level



(c) Mean expected production in t and industrial production in $t+k$, monthly



(d) Mean expected prices in t and producer prices in $t+k$, monthly



Notes: Survey data from the BEP. Production and Producer Price Index (both manufacturing) from the German Statistical Office and the Bundesbank respectively.

production and prices.[4] Since the average production and expectations co-move, it is no surprise that also realized industrial production and the average response to our main question are strongly correlated. Panel (c) of Figure 2 shows the cross correlation between the monthly average across firms of production expectations and leads and lags of the annual growth in industrial production. The two time series are strongly correlated especially for small leads of industrial production. This is in line with ifo business climate index being a leading indicator of economic activity. Analogously to the survey measures we also consider the cross correlation between month-on-month changes in the producer price index and average productions expectations. Panel (d) shows that also these measures are correlated albeit to a lesser degree.

[4]Recall that the ifo survey is the basis for the ifo business climate index, a widely watched leading indicator of the German business cycle (Abberger and Wohlrabe 2006; Henzel and Rast 2013).

3 Do firm expectations matter?

The main purpose of our analysis is to identify the effect of firm expectations on firm decisions. Specifically, in our baseline specification we assess to what extent firms' production and price-setting decisions depend on production expectations. For this purpose we compare the behavior of firms that expect an increase (decrease) of production to firms that expect production to remain unchanged. A key challenge in this regard is to identify variations in expectations which are orthogonal to current fundamentals. For only to the extent that firms are comparable in terms of fundamentals, we may think of expectations as a "treatment" into which some firms are randomly selected and others are not.

Put differently, as we compare the behavior of firms with different views about the future we face a selection problem because firms with better fundamentals are also more likely to enjoy a more favorable outlook. In order to address this selection problem we rely on propensity score matching (see e.g. Caliendo and Kopeinig 2008; Imbens and Rubin 2015). The idea is to mimic randomized control trials where treatment is actually assigned in a random fashion and hence orthogonal to observable characteristics. The matching approach is particularly suited for the purpose of our analysis because we are dealing with qualitative data on expectations: firms may be either optimistic, neutral or pessimistic. Hence, in our analysis, to the extent that firms receive a treatment, they are either treated with "optimism" or with "pessimism". Of course, our analysis does not require optimism or pessimism to be literally assigned in a random way. We merely require the assignment to be orthogonal to current fundamentals. Note also that we do not require optimism/pessimism to be unrelated to future fundamentals. We take up that issue in more detail in the Section 4 below.

In general, the matching approach offers several advantages over conventional regression analysis. First, it ensures that the distribution of control variables are similar across treated units and the control group (Dehejia and Wahba 2002; Imbens and Rubin 2015). This is important because differences in the distribution of controls can lead to significant bias when estimating treatment effects (Heckman et al. 1998). Second, the matching approach disciplines the analysis because the control group is specified prior to and independently of the estimation of the treatment effect (Imbens and Rubin 2015). Lastly, after matching, the treatment effect is estimated by a simple mean difference, thus allowing for a non-parametric estimation (Dehejia and Wahba 1999; Heckman et al. 1998).

3.1 Propensity score matching

We now briefly outline our approach following Caliendo and Kopeinig (2008). Inference is based on estimating the potential outcome of a treated firm under non-treatment, that is, the (unobserved) counterfactual outcome, had the treated firm not been treated. Formally, the object of interest is the average treatment effect on treated (ATT) firms:

$$\theta = \mathrm{E}[Y(1) - Y(0)|D = 1] = \mathrm{E}[Y(1)|D = 1] - \mathrm{E}[Y(0)|D = 1],$$

where $D = 1$ indicates treatment, $Y(1)$ the outcome of a treated firm, that is, a firm which is optimistic (pessimistic), and $Y(0)$ the counterfactual outcome of a treated firm in the absence of treatment. Since we do not observe the latter, we can only estimate the following relationship:

$$\mathrm{E}[Y(1)|D = 1] - \mathrm{E}[Y(0)|D = 0] = \theta + \mathrm{E}[Y(0)|D = 1] - \mathrm{E}[Y(0)|D = 0]. \tag{3.1}$$

This is equivalent to the ATT only if

$$\mathrm{E}[Y(0)|D = 1] - \mathrm{E}[Y(0)|D = 0] = 0,$$

that is, the potential outcomes are independent of treatment assignment. In randomized control trials this holds true due to the random assignment of treatment. In observational studies, additional assumptions are required. One approach is to assume that treatment is assigned randomly given a set of relevant covariates X:

$$Y(1), Y(0) \perp D|X.$$

Covariates are relevant if they affect both the (potential) outcome and the probability of being treated. In our case, this means that we need to include all information that matter for firms' expectation formation as well as for their production and price-setting decisions. We describe these variables below.

Since we are only interested in the effect on the treated, we merely need $Y(0)$ to be independent of treatment status, see equation (3.1). In this case the required conditional independence assumption simplifies to

$$Y(0) \perp D|X.$$

In the expressions above we condition on the whole set of control variables. This can be challenging when the number of observable controls is large. In our analysis we include 4 continuous variables and 18 categorical variables with three outcomes each. If we were to split the sample by the categorical variables only, we would already have 3^{18} potential bins. This makes accounting for controls by creating sub-samples of identical observations infeasible even with a large data set. We therefore rely on a result established by Rosenbaum and Rubin (1983): asymptotically, it is equivalent to condition either on the propensity to be treated, $p(X) \equiv Pr(D = 1|X)$ or directly on X. The conditional independence assumption can thus be stated as follows:

$$Y(0) \perp D|p(X).$$

Conditioning on the propensity score requires the additional assumption of common support, that is, treatment is not fully determined:

$$0 < p(X) = Pr(D = 1|X) < 1. \tag{3.2}$$

In what follows, we estimate the ATT by comparing the outcome of each treated observation to one or several untreated units with the same (or very similar) propensity score. In our analysis there are two possible treatments: optimism and pessimism. To establish the effect of a treatment we compare firms in each case to firms which do not expect production to change at all ("neutral firms").

In order to estimate the propensity score we pursue two alternative approaches. Since we are dealing with two treatments we estimate an ordered probit model where optimism and pessimism as outcomes of a common model. Alternatively, we consider distinct probit models for optimists and pessimists.

In the first case we estimate the probability of the latent variable, y_i^*, falling between two thresholds α_{j-1} and α_j for treatment j

$$Pr(y_i = j) = Pr(a_{j-1} < y_i^* \leq \alpha_j) = \Phi(\alpha_j - X_{it}'\beta) - \Phi(\alpha_{j-1} - X_{it}'\beta), \tag{3.3}$$

where $j = \{1, 0, -1\}$ corresponds to the three possible answers to question Q1. We collect the control variables in vector X_{it}. It includes time and sector fixed effects, the sector average of the reported state of business in each month, three lags of the dependent variables, and all firm specific variables listed in Table 2 (including three lags for each of the survey variables). More detailed information on the survey variables is provided in Table B.1 in the appendix.

The ordered probit does not directly yield the propensity score. In this case the propensity score, $p^m(X_{it})$ for treatment $m = \{optimism, pessimism\}$, equals the conditional choice probability of the treatment given the alternative of no treatment, that is, if firms expect production to remain unchanged:

$$p^m(X_{it}) = \frac{Pr(y = m|X_{it})}{Pr(y = m|X_{it}) + Pr(y = 0|X_{it})},$$

see again Caliendo and Kopeinig (2008).

The second approach involves two separate probit regressions, one for each treatment. The specification is the same as for the ordered probit model:

$$Pr(D_{it}^m = 1) = Pr(X_{it}'\beta) = \Phi(X_{it}'\beta), \tag{3.4}$$

where $D_{it}^m = 1$ is a dummy variable which is 1 for an observation responding *increase* in the case of the optimism treatment, 1 for an observation responding *decrease* in the case of the pessimism treatment, and 0 for an observation responding *stay the same* in both cases. We again collect the same control variables in vector X_{it}. Since the sample only includes the specific treatment group and the untreated, the estimated probability is a direct estimate of the propensity score:

$$p^m(X_{it}) = Pr(D_{it}^m = 1).$$

Caliendo and Kopeinig (2008) discuss the use of serial probit estimation compared to multinomial models in the case of multiple treatments. They argue that generally authors found no difference or a slight advantage of using separate probit models. It turns out that also in our case the serial probit estimation has a slight advantange as it yields improved balancing statistics. We therefore use it as our baseline approach. However, the results using the ordered probit do not differ much from results using the two probit regressions (see results in Section 3.4.)

As mentioned above we include all potentially relevant variables as controls in order to capture the fundamentals of the firm, both current values as well as lags. However, we only consider realizations which are available at the time the survey is conducted. In this regard it is important to note that most firms respond to the survey in the first two weeks of the months. Figure 3 shows the distribution of participation days within the month for firms which respond to the survey online. 50% answer within the first eight days of the month and another 25% answer in the following week.[5]

[5]These statistics pertain to firms that answer the survey online. They represent more than 60% of the

Figure 3: Distribution of days firms respond to survey within month, 2004 to 2016

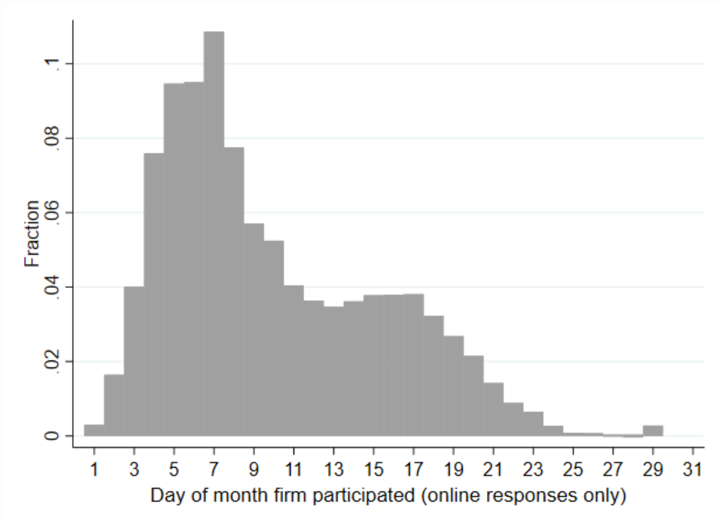


Table 2: Control variables in the propensity score model, measured in month t

Variable	**Description**	**Frequency**	**Reference period**
debt share[1]	total debt over assets	annual	$t-11$ to t
financing coefficient[1]	liabilities minus provisions divided by equity plus provisions	annual	$t-11$ to t
employees	no. of employees	annual[2]	October/November
state of business	answer to question on state of business (values: 1, 0, −1)	monthly	t
orders	answer to question on state of orders (values: 1, 0, −1)	monthly	t
foreign orders	answer to question on state of foreign orders (values: 1, 0, −1)	monthly	t
production	answer to question on change in production (values: 1, 0, −1)	monthly	$t-1$
prices	answer to question on change in prices (values: 1, 0, −1)	monthly	$t-1$
capacity utilization	utilization of existing capacity in %	quarterly[2]	$t-1$
demand	answer to question on demand in previous month (values: 1, 0, −1)	monthly	$t-1$

Notes: For all variables with monthly frequency also three lags are included. In addition various interaction terms are included (based on a log-likelihood ratio test).

[1] To ensure outliers and measurement error do not affect our results, we exclude the 99.99 percentile of observations for the debt share and the 0.02 and 99.98 percentiles for the financing coefficient.

[2] In months with no reporting we use data from the most recent balance sheet/most recent quarter the question was asked (if available).

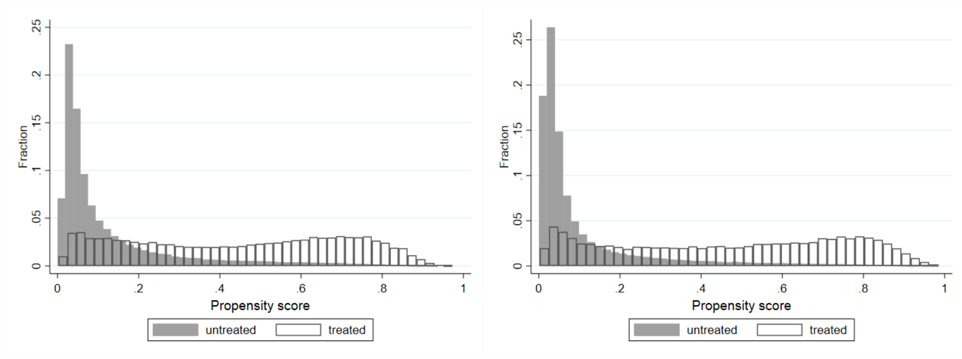
After computing the propensity scores, we match treated and untreated observations using a variant of caliper or radius matching (Caliendo and Kopeinig 2008).[6] We match each treated observation i (optimistic or pessimistic) to all untreated observations k (neutral) within the same

firms since 2004.

[6] We also test an alternative matching procedure proposed by Lechner et al. (2011). The results are very close to our results. Details can be found in appendix D.2.

Figure 4: Histogram of the density of the propensity scores

(a) Optimists　　**(b)** Pessimists



month which satisfy

$$p(X_{it}) - 0.02 \leq p(X_{kt}) \leq p(X_{it}) + 0.02. \tag{3.5}$$

Here we allow for a radius of 0.02. This corresponds to about a tenth of the standard deviation of the estimated propensity score.[7] All untreated observations to which a treated observation is matched are given equal weights: the inverse of the number of untreated observations in each match. Note that the untreated observations can be matched more than once to different treated observations.

Figure 4 displays the distribution of the propensity scores. The left panel contrasts the distribution for firms which receive an optimism treatment (white bars) with those for untreated (gray bars) firms. The right panel reports results for pessimism. In each instance, we find that there is considerable overlap of the distribution (common support), although the mass of untreated firms is more concentrated for lower propensity scores.[8]

Panel A of Table 3 reports basic statistics regarding our matches. We are able to find matches for about 93% of all treated optimists and for 90% of treated pessimists. This is due to the large overlap in propensity scores between treated and untreated firms.

3.2 Diagnostics

Before turning to the results, we report some diagnostics of the matching exercise. We compute balancing statistics in order to assess how similar the samples of treated observations and

[7] Alternative values for the radius give rise to similar results or, if not, fail to deliver satisfying balancing statistics (see next section).

[8] There are also some treated observations with a larger propensity scores than the largest propensity score of all untreated observations. We drop these observations in what follows. This trimming ensures that only suitable observations are matched. Specifically, we remove 15 optimists and 0 pessimists.

Table 3: Number of matched observations

	Optimism treatment		Pessimism treatment	
	Total	Matched	Total	Matched
Panel A: All firms				
Treated observations	26 974	25 050	23 327	20 947
Untreated observations	114 843	111 027	114 809	110 625
Panel B: Correct firms				
Treated observations	12 366	9 995	12 123	9 493
Untreated observations	82 317	73 321	82 519	72 762
Panel C: Incorrect firms				
Treated observations	10 634	9 671	7 641	6 614
Untreated observations	82 505	76 349	82 497	74 357

untreated observations are. The main statistic of interest is the standardized bias between the treated and untreated sample for each control variable. Following Rosenbaum and Rubin (1983), this is computed as follows:

$$SB = 100\frac{\bar{x}_1 - \bar{x}_0}{\sqrt{0.5(s_1^2 + s_0^2)}}, \tag{3.6}$$

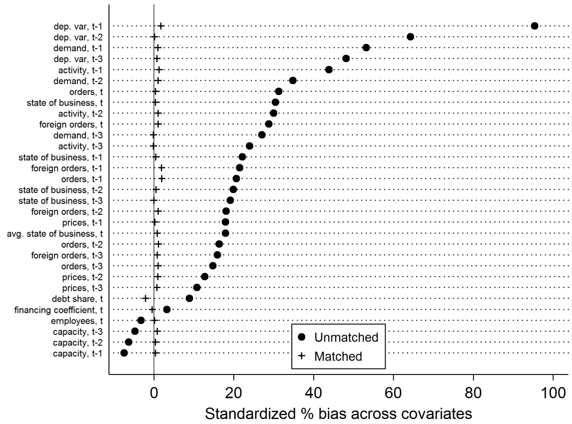
where $\bar{x}_1$ is the mean of the control variable among the matched treated observations, $\bar{x}_0$ is the mean of the control variable for all matched untreated observations, s_1 is the standard deviation of the treated observations and s_0 the standard deviation of all untreated observations. Figure 5 shows that as a result of matching observations, we achieve a sizeable reduction of the standardized bias. According to a widely used rule of thumb, the matched sample is regarded as well balanced when all standardized biases are below 5% (Caliendo and Kopeinig 2008).[9] We meet this standard in all instances (see also Table C.1 in the appendix).

Rubin (2001) suggest a second measure of balancing. He argues that the variance of the part of each covariate that is orthogonal to the propensity score (the residual of a regression of the covariate on the propensity score) should be similar for treated and untreated firms. Specifically, the ratio of the variances should not be below 0.5 or above 2. Ratios between a range of 0.8 and 1.25 are considered acceptable. Figure 6 plots the variance ratios before and after matching. Again, we find that matching firm-month observations ensures a that treated and non-treated firms appear well balanced in terms of covariates. Only for pessimists the ratio for the financing coefficient falls in the "of concern" area (dashed lines).

[9] Imbens and Rubin (2015) suggest that also 10% can be considered a satisfactory value, especially when the initial bias is large.

Figure 5: Standardized bias, before and after matching

(a) Optimists



(b) Pessimists

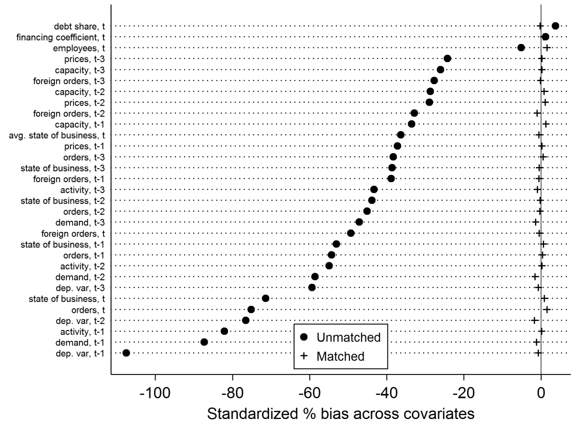
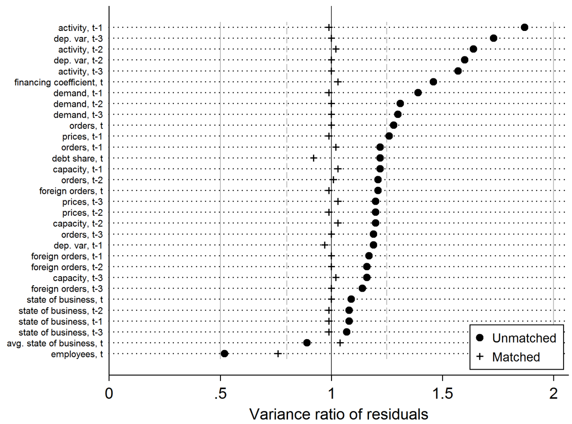
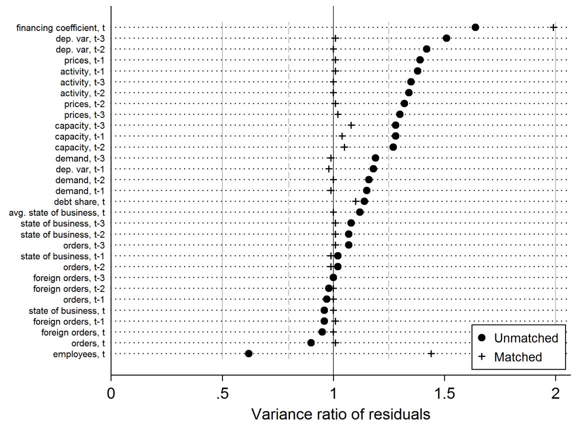


Figure 6: Variance ratio of residuals, before and after matching

(a) Optimists



(b) Pessimists



Notes: Ratios below 0.8 and above 1.25 (dashed lines) are considered "of concern"; ratios below 0.5 and above 2 (grey solid lines) are considered "bad" according to Rubin (2001).

3.3 Computation of the treatment effect

In what follows, we focus on the average treatment effect on the treated (ATT) in terms of production and price-setting decisions. The ATT is computed for each outcome variable as the difference of the mean of all treated and all untreated firms. In this part we focus on the ATT in the impact period. Later we will also look the outcome in each of the six months following the treatment.

The computation of standard errors for estimates of the ATT computed after matching is not straightforward. One can use analytical variances or bootstrapping. Since bootstrapping has sometimes been shown to be invalid (Caliendo and Kopeinig 2008), we use the methodology of Lechner (2001). The author shows that in case of variants of nearest neighbor matching, as in our case, the variance of the ATT, $\hat{\tau}_{ATT}$, is:

$$Var(\hat{\tau}_{ATT}) = \frac{1}{N_1} Var(Y(1)|D = 1) + \frac{\sum_{j\in\{D=0\}}(w_j)^2}{(N_1)^2} Var(Y(0)|D = 0),$$

where $Y(1)$ and $Y(0)$ refer to a variable of interest given the treatment indicator D equals 1 or 0. N_1 is the number of matched treated individuals and w_j is the weight of untreated individual j (see above).

3.4 Results

We now turn to the question which motivates our analysis: to what extent do firms' expectations impact current decision making? Table 4 provides a first answer. In the upper part of the table we report the ATT of optimism regarding future production, in the lower part we report the ATT of pessimism. In each instance, we focus on production and price-setting decisions. In each column we consider an alternative specification.

The left-most column (1) reports results for our baseline. We find that among optimists the fraction of firms which increase production is 17.2 percent higher than among neutral firms. Invoking the law of large numbers we thus conclude that optimism regarding future production raises the probability of an increase of current production by 17 percent. The effect is highly significant. It is also significant for prices, although in this case the effect is much smaller. Optimism regarding future production raises the probability of a price increase by 2.6 percent. These findings are consistent with the notion that prices are adjusted only infrequently in short run. Results are quite symmetric across optimism and pessimism, even though we estimate a

Table 4: Average treatment effect on the treated, optimistic and pessimist firms

	(1) Baseline	(2) Radius 0.01	(3) Sample 2002-2016	(4) Sample excl. fin. crisis	(5) Ordered probit	(6) Exp. state of business
Panel A: Optimists – Production (change in current month)						
ATT	0.172***	0.170***	0.181***	0.170***	0.152***	0.182***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
Observations	129812	120335	108683	113690	128932	129706
Panel B: Optimists – Prices (change in current month)						
ATT	0.025***	0.025***	0.027***	0.025***	0.016***	0.020***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Observations	129858	120367	108715	113734	128977	129759
Panel C: Pessimists – Production (change in current month)						
ATT	-0.173***	-0.170***	-0.169***	-0.172***	-0.198***	-0.204***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)
Observations	125458	113992	104490	106764	123941	125091
Panel D: Pessimists – Prices (change in current month)						
ATT	-0.031***	-0.033***	-0.029***	-0.035***	-0.038***	-0.029***
	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)
Observations	125530	114050	104551	106821	124014	125169

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

separate model for optimists and pessimists. Pessimism regarding future production increases the probability of a cut of current production by 17.2 percent as well. The probability of a price decline in response to pessimism is 3.1 percent.

Table 4 also reports results for alternative specifications. In column (2) we show results for a smaller radius in the matching procedure (0.01 instead of 0.02). In column (3) we consider a short sample period (start date 2002 rather than 1991). In column (4) we report results for a sample which excludes observations from the financial crisis, that is the years 2008 and 2009. Column (5) shows results for when we estimate an ordered probit model, jointly for pessimism and optimism. Column (6) shows results for when we use expectations for the future state of business rather than future production. Recall that in this case the "forecasting horizon" is 6 months rather than 3 months only. We observe that in all instances the estimate of the ATT is very close to that for the baseline. For instance, across specifications the probability of a production increase due to optimism ranges from 15.2 to 18.2 percent, the probability of a price increase ranges from 1.6 to 2.7 percent. Also, the effect is almost always estimated to be highly significant.

4 News or noise?

In the previous section we have established that firm expectations impact firm decisions about current production and pricing. This raises the question of why this is the case. Two alternative hypotheses appear plausible. According to the first hypothesis, firms may have information about the future that has not yet materialized in current fundamentals. While our set of fundamentals includes forward looking variables such as orders, one cannot rule out the possibility that firms have additional information beyond what is already reflected in current fundamentals. According to this "news" hypothesis, firms have therefore a good reason to be either optimistic or pessimistic. It is only that these reasons are not yet observable to the econometrician. Instead, according to the second hypothesis, optimism and pessimism are just "noise", that is, misperceptions about the future that are fundamentally unwarranted or "undue" (Pigou 1927). Of course, our estimate of the ATT may also reflect a mixture of news and noise.

In what follows, we seek to disentangle to what extent the expectations which, in turn, govern firms' decisions about production and prices reflect news and noise. We do so on the basis of firms' forecast errors. Intuitively, if a given firm appears particularly optimistic relative to its current fundamentals, but later reports that actual production is unchanged or down, its view about the future appears—with the benefit of hindsight—to have been misperceived. We are thus able to classify optimism and pessimism as incorrect or undue from an ex post perspective.[10]

In order to measure the forecast errors of firms we follow Bachmann et al. (2013). We interpret the qualitative responses to questions Q1 and Q3 as pertaining to the same latent variable and seek to measure whether or not there was an error, while remaining agnostic about its size. Table 5 provides on overview of all possible cases for which we define an error. According to our classification scheme, firms that reported to expect a change of their production were correct if they report at least once a change in the expected direction and no change in the opposite direction in the following three months. Firms which expected no change are considered to have made no error if they only report only one change in either direction or two changes in opposite directions.

Based on this classification scheme, we define a treatment with "correct optimism". It refers

[10] In principle, it is conceivable that firms have been optimistic about some aspect of the future and correctly so. Actual production may still fall short of the expected level because of some other unforeseen development. Still, since in our case optimism pertains to *future production as such*, we classify such firms are incorrect optimists.

Table 5: Computation of the forecast error for optimistic firms

expected production, t	average production, $t+1$ to $t+3$	error
1	> 0	0
	≤ 0	-1
	$> \frac{1}{3}$	1
0	$\leq \frac{1}{3}$ and $\geq -\frac{1}{3}$	0
	$< -\frac{1}{3}$	-1
-1	≥ 0	1
	< 0	0

to optimistic firms (answer 1 to question Q1) without forecast error. The control group are all neutral firms (answer 0 to Q1) without forecast error. The second treatment we consider is "incorrect optimism". Here we consider firms that are optimistic but make a forecast error from an ex post point of view. The control group is the same as in the first case. The third and fourth treatment are defined analogously for pessimists. Using these four new treatment indicators, we perform the same matching procedure as described in Section 3.1.

Before turning to the results, we again consider some diagnostics statistics to ensure the matching methodology works also in this case. The statistics are the same as described in Section 3.2. Panels B and C of Table 3 provide an overview of the number of observations for which a propensity score can be computed as well as the number of observations which can be matched. Since we now restrict our sample, the number of matches is smaller than in Section 3.2. Despite the sample reduction, however, common support is not an issue (see Figure C.1 in the appendix) and we still have a sufficiently large amount of observations to compute ATTs. We display balancing statistics in figures C.2 and C.3, the appendix. For all four treatments balancing is achieved and no bias is above 5% (Table C.1 again shows the bias in detail). Also the variance ratios are generally within the defined bounds.

We report results in Table 6. The structure is analogous to Table 4 and we focus again on how firms' current production and price setting decisions depend on optimism (upper part of the table) and on pessimism (lower part of the table). Now, however, we distinguish between correct and incorrect optimists/pessimists. Panels A and B show the result for production and prices for correct optimists, respectively, while panels C and D show the results for incorrect optimists.

We report results for the baseline in column (1), but as before results are robust across alternative specifications (columns 2-4). We find that the effect of optimism on firms current decisions

Table 6: Average treatment effect on the treated, correct and incorrect optimism and correct and incorrect pessimism

	(1) Baseline	(2) Radius 0.01	(3) Sample 2002-2016	(4) Sample ex. crisis
Panel A: Correct optimists – Production (change in current month)				
ATT	0.302***	0.298***	0.314***	0.297***
	(0.008)	(0.009)	(0.009)	(0.009)
Observations	81254	68946	68785	71391
Panel B: Correct optimists – Prices (change in current month)				
ATT	0.035***	0.034***	0.040***	0.033***
	(0.006)	(0.007)	(0.007)	(0.007)
Observations	81254	68945	68778	71392
Panel C: Incorrect optimists – Production (change in current month)				
ATT	0.063***	0.060***	0.072***	0.063***
	(0.007)	(0.008)	(0.008)	(0.008)
Observations	84029	74232	69715	73973
Panel D: Incorrect optimists – Prices (change in current month)				
ATT	0.016***	0.015**	0.017***	0.011**
	(0.006)	(0.006)	(0.006)	(0.006)
Observations	84032	74232	69714	73978
Panel E: Correct pessimist – Production (change in current month)				
ATT	-0.307***	-0.300***	-0.307***	-0.303***
	(0.009)	(0.010)	(0.010)	(0.009)
Observations	80282	66948	67112	68156
Panel F: Correct pessimist – Prices (change in current month)				
ATT	-0.030***	-0.021***	-0.028***	-0.044***
	(0.008)	(0.008)	(0.009)	(0.008)
Observations	80285	66941	67109	68158
Panel G: Incorrect pessimist – Production (change in current month)				
ATT	-0.086***	-0.093***	-0.074***	-0.086***
	(0.009)	(0.009)	(0.009)	(0.009)
Observations	79026	68414	65323	68835
Panel H: Incorrect pessimist – Prices (change in current month)				
ATT	-0.003	-0.008	0.000	-0.008
	(0.007)	(0.007)	(0.008)	(0.007)
Observations	79033	68420	65326	68842

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

is stronger for correct optimists than for incorrect optimists. The probability of production increase for correct optimists is about 30 percent, but about 6.5 percent for incorrect optimists. The probability of a price increase for correct optimists is about 3.5 percent, but 1.5 percent for incorrect optimists.

Results for pessimists are quantitatively comparable. Correct pessimists are 30 percent more likely to lower production, while the probability for incorrect optimists is 8.5 percent. We thus note that there is some asymmetry in the response to incorrect optimism/pessism with incorrect pessimists responding somewhat stronger than incorrect optimists. Similarly, correct pessimist tend to lower the price to the same extent as correct optimists tend to raise them. Incorrect optimists, instead, do not respond by adjusting prices in a significant way. This observation lends support to the view that downward price rigidities prevent an adjustment of prices unless the need for adjustment is strong. Arguably this is the case if pessimism is correct, but not if it is incorrect, which might be related to a lower degree of pessimism. In the latter case, it turns out, firms appear more responsive in terms of quantities instead of prices.

5 Undue optimism and aggregate fluctuations

So far we have focused on individual firms and, more specifically, we have documented that undue optimism causes firm to raise prices and production. Likewise, undue pessimism causes firms to lower prices and production. In what follows, we ask whether undue optimism and pessimism at the firm level matters for aggregate outcomes. Intuitively, if a sufficiently large number of firms or a number of sufficiently large firms are unduely optimistic (pessimistic) this may cause economic activity to rise (fall) at the aggregate level. Our analysis proceeds as follows. In a first step, we construct a time series for undue optimism and undue pessimism at the aggregate level. For this purpose we measure for each month in our sample the fraction of unduely optimistic and pessimistic firms. In a second step, we rely on local projections to estimate the impact of undue optimism and pessimism for industrial production and prices.

5.1 Undue optimism and pessimism at the macro level

In order to identify undue optimism and pessimism at the aggregate level we simply aggregate the measures established at the firm level in the previous sections. We proceed as follows. First, we rely on the ordered probit model estimated in Section 3. Now, however, rather than matching firms based on their propensity score, we measure the extent of optimism and pessimism by computing the difference between a firm's response and the prediction of the ordered probit model as in equation (3.3). Note that we use the ordered probit, rather than distinct models for optimists and pessimists, because we need to be able to account for all outcomes simultaneously.

Recall that the ordered probit model includes as control variables time and sector fixed effects, the sector average of the reported state of business in each month, three lags of the dependent variables, and all firm specific variables listed in Table 2 (including three lags for each of the survey variables). On the basis of the ordered probit model we classify firms as optimists (pessimists) whenever their respond with "+1" ("−1") even though the model predicts otherwise. We further narrow down our aggregate measure as we focus on undue optimism and pessimism. For this purpose we only count firms for which we additionally observe a non-zero forecast error, as defined in Section 4 above. Finally, we compute the shares of firms that turn out either unduely optimistic or unduely pessimistic relative to all firms in the corresponding month.

We compute three measures. One is the unweighted share. For the second measure, we compute weighted shares using firm employees (without the top 5%).[11] Finally, we weight firms in line with the approach by the ifo institute for aggregating answers to the business climate index (Sauer and Wohlrabe 2018). In a first step all firms within a 2-digit WZ08 sector are weighted using the number of employees in production as reported in the survey.[12] Instead of using the number of employees directly, the weight is a logarithmic transformation of employment.[13] Then the sector averages are aggregated using data on gross value added by sector from the German Statistical Office.

Figure 7 displays the unweighted and the employee-weighted time series for undue optimism and pessimism (using ifo weights results in a very similar plot, such that we present only two series per plot for better readability). The computation of these shares requires firms to be in the survey for at least eight consecutive months because we need three lags for the estimation of the ordered probit model and four leads for the computation of the forecast error.[14] This leads to a gap in our time series from August 2001 to March 2002 because the ifo survey was not conducted in December 2001. In addition, it reduces the number of observations in the last five months of 2016 since currently not all data for 2017 is available to us.

[11]We exclude the largest 5% of firms because of known issues with employment figures in the BEP data. By excluding the largest firms we avoid single observations driving our results.

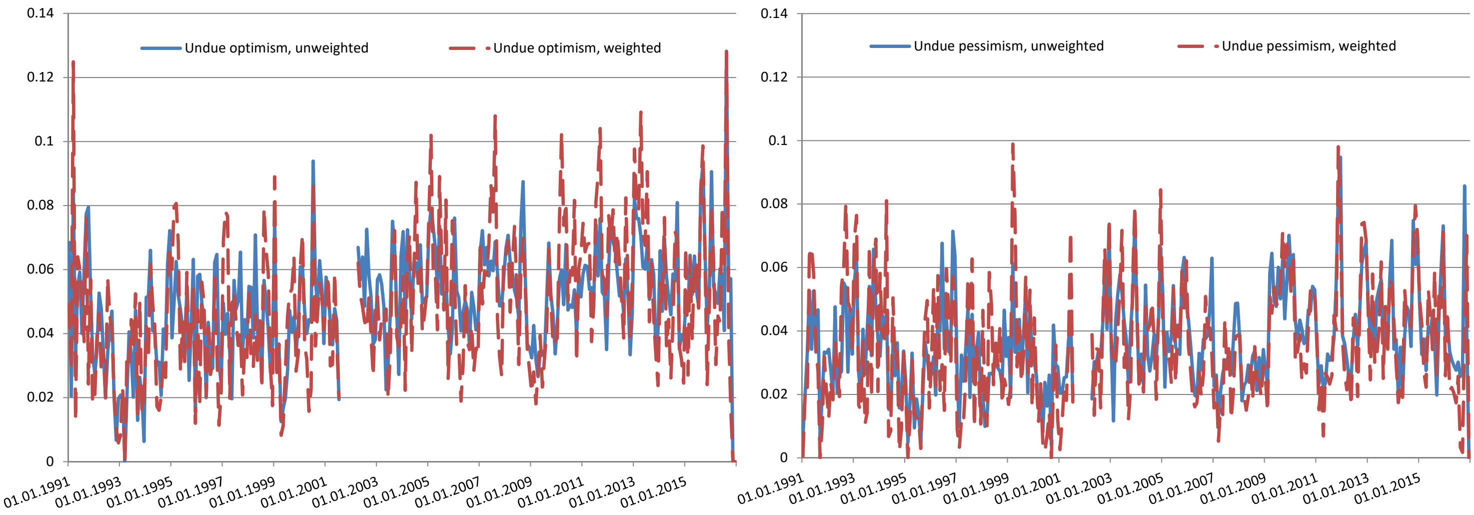
[12]WZ08 is the German system of industry classification.

[13]Specifically, the weight is $w = \log_{10}(N)^e$, with N being the number of employees, see the EBDC Questionnaire Manual. This transformation ensures that very large firms do not distort the averages.

[14]That is, we need production for the next three months, but since production is reported only for the previous month we need four leads of the survey.

Figure 7: Undue optimism and pessimism

(a) Undue optimism **(b)** Undue pessimism



Notes: aggregate time series for undue optimism and pessimism, unweighted and using employee weights.

5.2 Aggregate effects of undue optimism

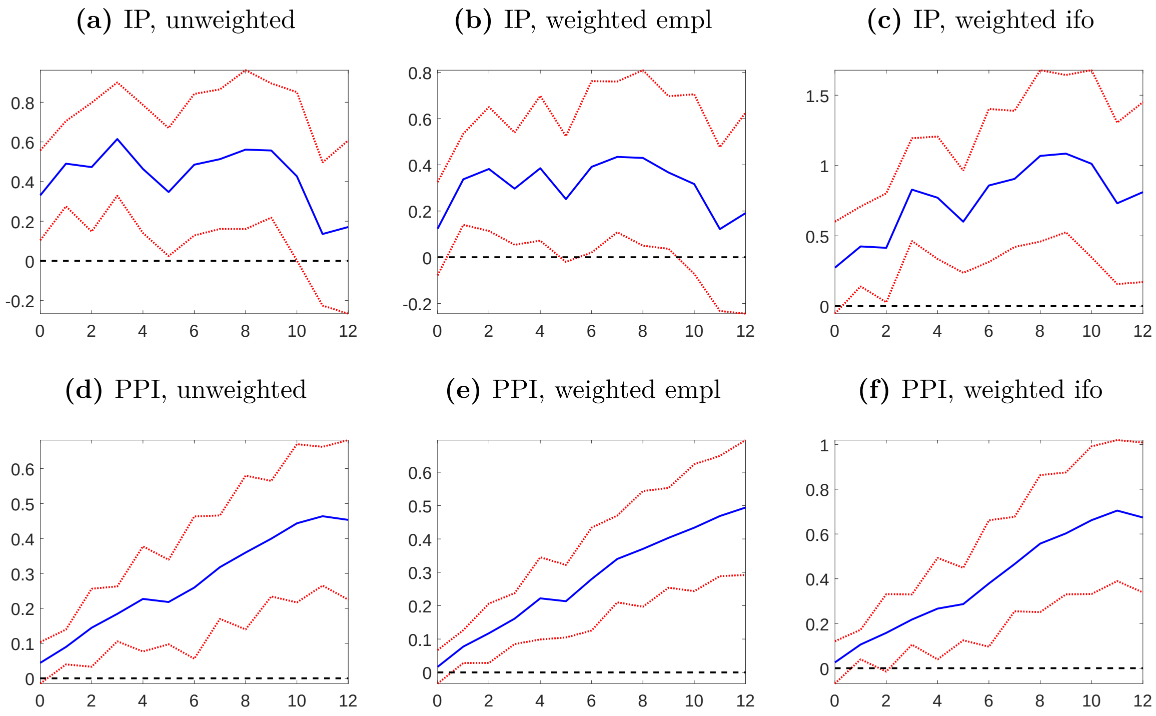
Our measure of undue optimism and pessimism is an aggregate of firm-level responses. In computing our measure we have allowed for time-fixed effects. As such, aggregate optimism and pessimism is unlikely the to be caused by macroeconomic shocks. Similarly, to the extent macroeconomic shocks impact different firms differently, it is important that we focus on unduely optimistic/pessimistic firms, that is, firms whose expectations turn out to be unwarranted from an ex-post perspective. For these reasons we treat our aggregate measure as an explanatory variable of macroeconomic outcomes that is not itself caused by macroeconomic developments.

Against this background we may resort to local projections to estimate the effect of undue optimism on aggregate outcomes (Jordá 2005). Formally, using e_t^o and e_t^p to denote the time-series observations for undue optimism and pessimism, respectively, and x_t for the realization of a macroeconomic variable of interest, we estimate following model:

$$x_{t+h} = c^{(h)} + \sum_{j=1}^{J} \alpha_j^{(h)} x_{t-j} + \sum_{k=0}^{K-1} \beta_k^{(h)} e_{t-k}^o + \sum_{k=0}^{K-1} \gamma_k^{(h)} e_{t-k}^p + \varepsilon_{t+h}, \tag{5.1}$$

where c is a (horizon-specific) constant. Additionally, we include a linear time trend. To enhance efficiency, we also include the residuals of the previous horizon when increasing the horizon in steps of one (see Jordá 2005). For the estimation, we include 1 lag of the dependent variable and 12 lags of undue optimism and pessimism. In all estimations, we include both undue optimism and pessimism to account for a potential correlation between the two variables. The estimated coefficients β^h and γ^h provide a direct measure of the impulse response at horizon h, given a

Figure 8: Impulse response to undue optimism

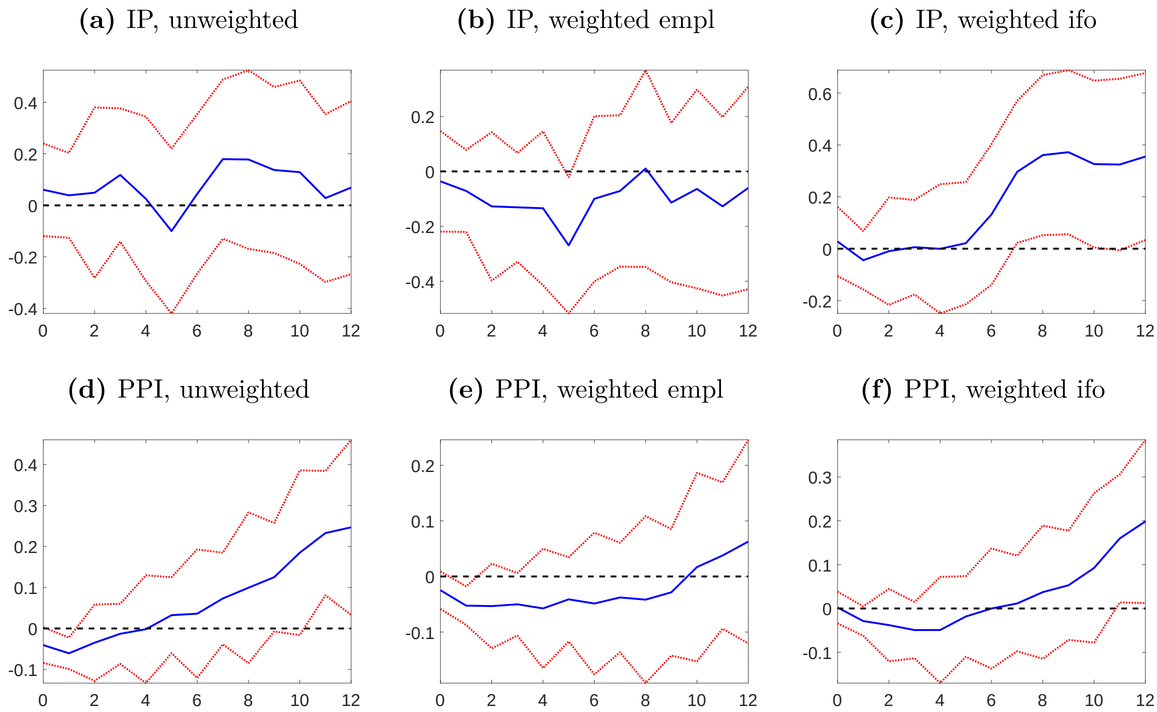


Notes: Responses of IP and PPI (both manufacturing) to undue optimism. Local projections with constant, linear trend, one lag of dependent variable and 12 of the shocks. 90% confidence intervals.

shock in period t. Figure 8 displays the responses to a one percentage point increase in the share of unduely optimistic firms. The top panels show the response of industrial production in the manufacturing sector (IP), measured in percentage deviations from trend, while the bottom panels show the response of the producer price index in the manufacturing sector (PPI), also measured in percentage deviations from trend. The left column displays results using the unweighted measure, the middle column is based on employee-weighted shares, while the right column shows responses for ifo weights.

In each instance, time is measured in months along the horizontal axis. The blue solid line represents the point estimate, dashed red lines are 90 percent confidence bounds. We find that industrial production responds strongly and significantly to an increase in undue optimism. The observed increase is temporary and turns insignificant after approximately one year (using ifo weights prolongs the time period of a significant increase somewhat). Similar to the firm-level results, reported in Section 4 above, we also find a strong and significant increase in the price level after an increase in undue optimism. This reaction is in line with the interpretation of optimism shocks as a specific form of demand shocks, see Lorenzoni (2009) and Enders et al. (2017). Figure 9 repeats the same exercise for undue pessimism. We find much weaker effects

Figure 9: Impulse response to undue pessimism



Notes: Responses of IP and PPI (both manufacturing) to undue pessimism. Local projections with constant, linear trend, one lag of dependent variable and 12 of the shocks. 90% confidence intervals.

compared to undue optimism. Specifically, industrial production remains insignificant throughout while the producer price index falls (marginally significant) only in period 1 after the shock.

Finally, Table 7 displays a forecast error variance decomposition (FEVD) for the horizon of 12 months, using the methodology of Gorodnichenko and Lee (2017). We present results for undue optimism and pessimism, both for the unweighted measure and for the time series based on employee or ifo weights. We find that undue optimism is responsible for 10-18% of aggregate fluctuations in industrial production at a one-year horizon. At the same horizon, around 20% of the PPI is driven by undue optimism. Pessimism, on the other hand, has much smaller impact on IP and the PPI. The specification for the PPI with unweighted observations and the one with ifo weights for IP deliver a FEVD of around 6%. Regarding the small impact on prices, these results for the PPI are in line with those of firm-level effects, reported in Section 4 above.

Table 7: Forecast error variance decomposition (one year horizon)

	Variable	Unweighted	Empl. weights	ifo weights
Optimism	IP	14%	10%	18%
	PPI	19%	23%	20%
Pessimism	IP	1.3%	1.7%	6.4%
	PPI	5.9%	0.9%	2.9%

6 Conclusion

In this paper, we ask to what extent firm expectations matter for firm decisions. From a theoretical point of view, the answer to the question seems obvious: expectations should matter a great deal. However, to date there is little direct evidence to support the theory. In this paper, we aim at filling this gap on the basis of a particularly suited data set and a new identification strategy.

We use a large survey of firms in the German manufacturing sector. Firms report on a monthly basis whether they expect production to increase, to remain constant or to decline. For each firm-month observation we also observe a large number of firm characteristics, including balance-sheet information. This allows us to match firms on the basis of fundamentals, that is, we compare firms which have very similar fundamentals but differ in their views about the future. Firms which expect production to go up are thus more optimistic than firms which do not—despite approximately identical fundamentals. The converse applies to pessimistic firms.

We find that optimistic firms tend to raise prices and production today. This can be explained in two ways. According to the news view, firms have simply additional information about future developments that are not reflected in current fundamentals. Their optimism is thus justified, even if the reason for the optimism are future fundamentals only. According to the noise view, firms are optimistic or pessimistic for no particular reason. They simply have wrong ideas about the future or, put differently, are driven by animal spirits.

We disentangle the effect of news and noise on our decision makers as we define a forecast error. Because we observe the actual developments ex posts we can see whether firms were right about the future or not. We match incorrectly optimistic and pessimistic firms, in turn, to neutral firms, as well as correctly optimistic/pessmistic firms. For both groups we find a positive (negative) effect of optimism (pessimism) on current production and prices, although the effect is relatively small in case of incorrect "sentiments". These results suggest that firms are to some extent guided by noise/animal spirits.

Finally, we turn to the aggregate effects of animal spirits. We aggregate firm-level expectations by determining the share of unduely optimistic and pessimistic firms. We run local projections to assess the effect of undue optimism and pessimism at the aggregate level. We find, in particular, that aggregate optimism, even if undue, causes industrial production and the producer prices to increase. This confirms the notion that animal spirits are a force which underly cyclical fluctuations.

References

Abberger, Klaus and Klaus Wohlrabe (2006). "Einige Prognoseigenschaften des ifo Geschäftsklimas - Ein Überblick über die neuere wissenschaftliche Literatur". *ifo Schnelldienst 59*, 19–26.

Angeletos, George-Marios and Jennifer La'O (2013). "Sentiments". *Econometrica 81* (2), 739–779.

Bachmann, R, Steffen Elstner, and Eric R Sims (2013). "Uncertainty and Economic Activity: Evidence from Business Survey Data". *American Economic Journal: Macroeconomics 5* (2), 217–249.

Bachmann, Rüdiger and Steffen Elstner (2015). "Firm optimism and pessimism". *European Economic Review 79*, 297 –325.

Bachmann, Ruediger and Peter Zorn (2018). "What Drives Aggregate Investment? Evidence from German Survey Data". Mimeo. CEPR Discussion Paper 12710.

Barsky, Robert B. and Eric R. Sims (2012). "Information, animal spirits, and the meaning of innovations in consumer confidence". *American Economic Review 52*, 1343–1377.

Beaudry, Paul and Franck Portier (2006). "Stock Prices, News, and Economic Fluctuations". *American Economic Review 96* (4), 1293–1307.

Becker, Sasha. O. and Klaus Wohlrabe (2008). "Micro Data at the Ifo Institute for EconomicResearch - The 'Ifo Business Survey' Usage and Access". *Schmollers Jahrbuch-Zeitschrift für Wirtschafts- und Sozialwissenschaften 182* (2), 307–319.

Boneva, Lena, James CLoyne, Martin Weale, and Tomasz Wieladek (2018). "Firms' Expectations of New Orders, Employment, Costs and Prices: Evidence from Micro Data". Mimeo. CEPR Discussion Paper 12722.

Caliendo, Marco and Sabine Kopeinig (2008). "Some practical guidance for the implementation of propensity score matching". *Journal of Economic Surveys 22* (1), 31–72.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy 120* (1), 116–159.

——— (2015). "Information rigidity and the expectations formation process: a simple framework and new facts". *American Economic Review 105* (8), 2644–2678.

Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele (2018). "Inflation expectations and firm decisions: new causal evidence". NBER Working Paper 25412.

Dehejia, Rajeev H. and Sadek Wahba (1999). "Causal effects in nonexperimental studies: reevaluating the evaluation of training programs". *Journal of the American Statistical Association 94* (448), 1053–1062. eprint: `https://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1999.10473858`.

——— (2002). "Propensity Score-Matching Methods For Nonexperimental Causal Studies". *The Review of Economics and Statistics 84* (1), 151–161.

Del Negro, Marco, Marc Giannoni, and Christina Patterson (2012). “The forward guidance puzzle”. Staff Reports 574. Federal Reserve Bank of New York.

EBDC-BEP (2017). “Business Expectations Panel 01/1980 – 06/2017, LMU-ifo Economics & Business Data Center, Munich”. *doi: 10.7805/ebdc-bep-2017*.

Eggertsson, Gauti B. and Michael Woodford (2003). “The zero bound on interest rates and optimal monetary policy”. *Brookings Papers on Economic Activity 2003* (1), 139–211.

Enders, Zeno, Michael Kleemann, Gernot J. Müller, and Almuth Scholl (2017). “Growth expectations, undue optimism, and short-run fluctuations”. Bundesbank Discussion Paper No. 11/2017.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer (2015). “Expectations and Investment”. *NBER Macroeconomics Annual 2015, Volume 30*. NBER Chapters. National Bureau of Economic Research, Inc, 379–431. URL: `https://ideas.repec.org/h/nbr/nberch/13589.html`.

Gorodnichenko, Yuriy and Byoungchan Lee (2017). “A note on variance decomposition with local projections”. NBER Working Paper No. 23998.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd (1998). “Characterizing Selection Bias Using Experimental Data”. *Econometrica 66* (5), 1017–1098.

Henzel, Steffen R. and Sebastian Rast (2013). “Prognoseeigenschaften von Indikatoren zur Vorhersage des Bruttoinlandsprodukts in Deutschland”. *ifo Schnelldienst 66*, 39–46.

Huber, Martin, Michael Lechner, and Andreas Steinmayr (2015). “Radius matching on the propensity score with bias adjustment: tuning parameters and finite sample behaviour”. *Empirical Economics 49*, 1–31.

Imbens, Guido W and Donald B Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Jordá, Òscar (2005). “Estimation and inference of impulse responses by local projections”. *American Economic Review 95* (1), 161–182.

Keynes, John M. (1936). *The general theory of employment, interest, and money*. Orlando, Florida: Edition Harvest/Harcourt Brace 1964.

Kydland, Finn E. and Edward C. Prescott (1982). “Time to Build and Aggregate Fluctuations”. *Econometrica 50* (6), 1345–1370.

Lechner, M. (2001). “Identification and estimation of causal effects of multiple treatments under the conditional independence assumption.” *Econometric Evaluation of Labour Market Policies*. Ed. by M. Lechner and F. Pfeiffer. Heidelberg: Physica, 1–18.

Lechner, M., R. Miquel, and C Wunsch (2011). “Long-Run Effects of Public Sector Sponsored Training”. *Journal of the European Economic Association 9* (4), 742–784.

Lorenzoni, Guido (2009). “A theory of demand shocks”. *American Economic Review 99*, 2050–2084.

Lucas Robert E, Jr (1973). “Some International Evidence on Output-Inflation Tradeoffs”. *American Economic Review 63* (3), 326–334.

Massenot, Baptiste and Yuri Pettinicchi (2018). “Can firms see into the future? survey evidence from germany”. *Journal of Economic Behavior & Organization 145*, 66 –79.

Mortensen, Dale and Christoper Pissarides (2009). “Job creation and job destruction in the theory of unemployment”. *Review of Economic Studies 63*, 397–415.

Nerlove, Marc (1983). “Expectations, Plans, and Realizations in Theory and Practice”. *Econometrica 51* (5), 1251–1279.

Pigou, Arthur C. (1927). *Industrial fluctuations.* Macmillan.

Rosenbaum, Paul R. and Donald B. Rubin (1983). “The central role of the propensity score in observational studies for causal effects”. *Biometrika 70* (1), 41–55.

Rubin, D. B. (2001). “Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation”. *Health Services & Outcomes Research Methodology 2*, 169–188.

Sauer, Stefan and Klaus Wohlrabe (2018). “The New ifo Business Climate Index for Germany”. *CESifo Forum 19* (2), 59–64.

Schmitt-Grohé, Stephanie and Martín Uribe (2012). “What’s news in business cycles?” *Econometrica 80*, 2733–2764.

Woodford, Michael (2003). *Interest and prices.* Princeton University Press.

A Additional descriptive statistics

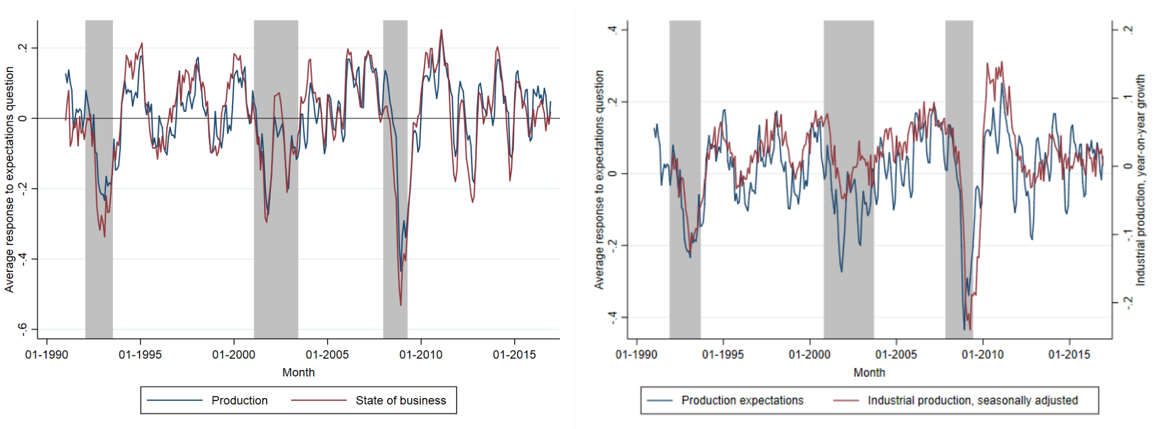
Table A.1: Observations and average duration in panel, 1991 to 2016

Observations		Months	
Respondents	5 922	Avg. number of responses	56.39
Respondents × months	323 823	Avg. duration in survey	66.52
		Response rate	85.71%

Figure A.1: Average expectations and industrial production, 1991-2016

(a) Production and state of business expectations

(b) Production expectations and industrial production



Notes: Shaded areas mark recession periods as defined by the German Council of Economic Experts. Average response defined as share of positive responses (*increase/improve*) minus share of negaative responses (*decrease/worsen*). Survey data from the BEP. Industrial Production from the German Statistical Office.

B Details on survey questions

Table B.1: All survey questions used

Label	Name	Question[1]	Possible answers[2]
Q1	expected production	Expectations for the next 3 months: Our domestic production activity regarding good XY will probably ...	increase [1] not change [0] decrease [-1]
Q2	expected state of business	Expectations for the next 6 months: Taking economic fluctuations into account our state of business will be...	rather more favorable [1] not changing [0] rather less favorable [-1]
Q3	production	Tendencies in the previous month: Our domestic production activities with respect to product XY have ...	increased [1] not changed [0] decreased [-1]
Q4	prices	Tendencies in the previous month: Taking changes of terms and conditions into account, our domestic sales prices (net) for product XY have been ...	increased [1] not changed [0] decreased [-1]
Q5	employees	Number of employees: In our company (domestic enterprises only) we employ [...] persons. Thereof x persons for producing product XY.	x is the number of persons employed for XY
Q6	orders	We consider our order backlog to be	relatively high [1] sufficient [0] too small [-1] no export of XY [4]
Q7	foreign orders	We consider our order backlog for exports to be	relatively high [1] sufficient [0] too small [-1] no export of XY [4]
Q8	capacity utilization	The current utilization of our capacities for producing XY (standard utilization = 100%) is currently x%.	x is a value between 30 and 100 divisible by 10
Q9	demand	Tendencies in the previous month: The demand situation with respect to product XY is ...	better [1] not changed [0] worse [-1]

[1] Authors' translation of the most recent formulation of the question in German according to the EBDC Questionnaire manual.
[2] Only those answers which we consider. Specifically we exclude answers like "no production" or similar answers which indicate the questions does not apply to the firm.

Table B.2: Main survey questions, changes over time

Label	Time period	Question[1]
Q1	01/1980-06/1994	Our domestic production activity regarding good XY in the next 3 months **taking economic fluctuations into account – i.e. after eliminating purely seasonal fluctuations** – will probably . . . increase/not change/decrease.
	07/1994-06/1997	Our domestic production activity regarding good XY in the next 3 months **taking economic fluctuations into account – i.e. after eliminating purely seasonal fluctuations** – will probably . . . increase/not change/decrease/*no substantial domestic production.*
	07/1997-11/2001	Our domestic production activity regarding good XY in the next 3 months **taking economic fluctuations into account** will probably . . . increase/not change/decrease/no substantial domestic production.
	Since 01/2002	Expectations for the next 3 months: Our domestic production activity regarding good XY will probably . . . increase/not change/decrease/no substantial domestic production.
Q2	01/1980-06/1997	Our state of business regarding good XY in the next 6 months taking economic fluctuations into account – **i.e. after eliminating purely seasonal fluctuations** – will be . . . rather more favorable/ not changing/rather less favorable.
	07/1997-11/2001	Our state of business regarding good XY in the next 6 months taking economic fluctuations into account will be . . . rather more favorable/ not changing/rather less favorable.
	Since 01/2002	Expectations for the next 6 months: Taking economic fluctuations into account our state of business will be . . . rather more favorable/ not changing/rather less favorable.
Q3	01/1980-06/1994	**In comparison to the previous month** our domestic production activities regarding good XY have . . . **been more lively/unchanged/weaker**.
	07/1994-11/2001	**In comparison to the previous month** our domestic production activities regarding good XY have . . . **been more lively/unchanged/weaker** */no substantial domestic production.*
	01/2002-02/2002	**In the last 2-3 months** our domestic production activities regarding good XY have . . . **been more lively/unchanged/weaker**/no substantial domestic production.
	Since 03/2002	Tendencies in the previous month: Our domestic production activities with respect to product XY have . . . increased/not changed/decreased/no substantial domestic production.
Q4	01/1980-11/2001	**Compared to the previous month** our domestic prices (net prices) of good XY – taking changes of terms and conditions into account – have been . . . increased/not changed/decreased.[2]
	01/2001-02/2002	**In the last 2-3 months** our domestic prices (net) of good XY – taking changes of terms and conditions into account – have been . . . increased/not changed/decreased.
	Since 03/2002	Tendencies in the previous month: Taking changes of terms and conditions into account, our domestic sales prices (net) for product XY have been . . . increased/not changed/decreased.

[1] Authors' translation of the question in German according to the EBDC Questionnaire manual.
[2] In several months in 1980 the questions was split into two parts, one covering regular and additional orders.

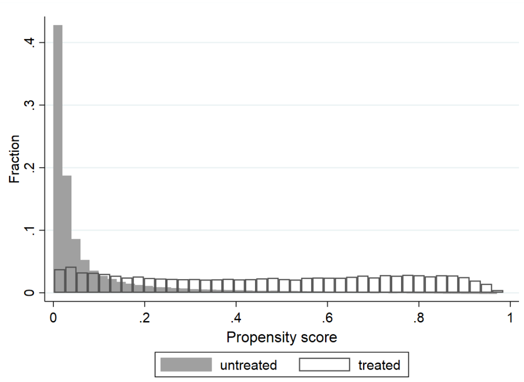
C Balancing statistics

Table C.1: Balancing statistics, standardized bias

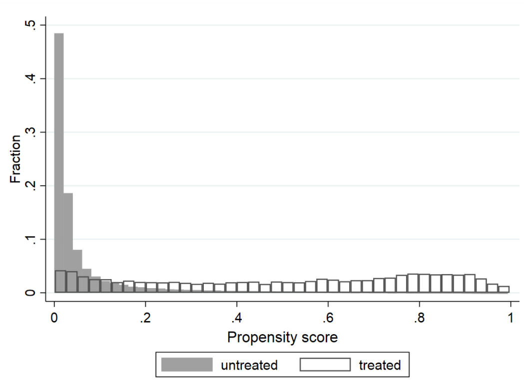
Variable	Optimists All	Optimists Correct	Optimists Incorrect	Pessimists All	Pessimists Correct	Pessimists Incorrect
dep. var, $t-1$	1.8	1.6	1.8	-0.7	-0.7	-0.9
dep. var, $t-2$	0.2	-0.2	-0.3	-1.7	-3.5	-2.5
dep. var, $t-3$	0.8	0.7	0.1	-0.7	-2.8	-0.5
activity, $t-1$	1.3	0.1	-0.6	0.2	-0.3	-0.6
activity, $t-2$	1.1	0.3	-0.9	0.3	1.1	0.8
activity, $t-3$	-0.1	-0.2	-0.8	-0.9	-1.3	1.6
prices, $t-1$	0.3	1.8	-0.0	0.3	1.5	0.1
prices, $t-2$	1.0	-0.8	0.8	1.1	3.1	-0.1
prices, $t-3$	0.8	1.5	0.1	0.3	2.0	-0.6
demand, $t-1$	1.0	1.1	0.6	-1.1	-0.6	-1.5
demand, $t-2$	1.1	2.0	-1.3	-1.5	-1.0	1.2
demand, $t-3$	-0.1	-0.4	-0.8	-1.4	-1.7	1.6
capacity, $t-1$	0.4	0.4	0.3	1.3	3.3	1.0
capacity, $t-2$	0.4	0.6	0.5	0.8	3.1	1.3
capacity, $t-3$	0.9	1.2	-0.1	0.3	3.1	0.6
employees, t	0.2	-0.8	0.1	1.6	1.0	1.2
Avg. State of business, sector t	0.9	1.8	-0.5	-0.5	-0.9	-0.3
state of business, t	0.4	-0.7	-0.5	0.9	2.0	1.2
state of business, $t-1$	0.5	-0.0	-0.7	0.7	1.0	1.0
state of business, $t-2$	0.6	0.6	-0.5	-0.2	0.9	1.8
state of business, $t-3$	0.0	0.1	-0.3	-0.4	-0.2	0.7
orders, t	0.4	-1.4	-0.4	1.6	1.8	-0.4
orders, $t-1$	2.0	2.1	-0.8	0.4	0.7	0.1
orders, $t-2$	1.2	1.4	0.0	-0.3	-1.4	-0.6
orders, $t-3$	1.1	0.9	0.0	0.6	-0.9	0.1
foreign orders, t	1.1	-0.4	-0.4	-0.4	-2.1	-0.0
foreign orders, $t-1$	1.9	1.0	-0.8	-0.5	-1.5	0.9
foreign orders, $t-2$	1.1	0.6	-0.5	-1.0	-1.7	-0.6
foreign orders, $t-3$	0.9	0.8	-0.1	-0.1	-2.1	-0.1
debt share, t	-2.1	-2.8	-1.2	-0.2	-4.5	-2.6
financing coefficient, t	-0.4	-0.7	-1.5	0.7	0.9	0.1

Figure C.1: Histogram of the density of the propensity scores

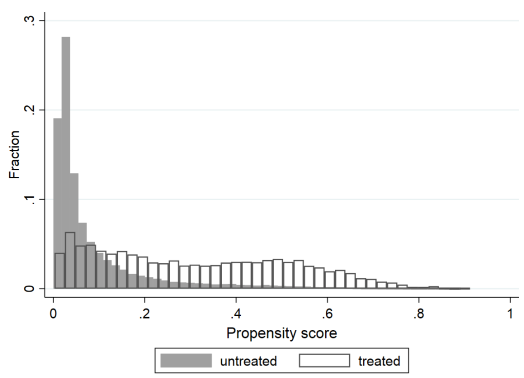
(a) Correct optimists



(b) Correct pessimists



(c) Incorrect optimists



(d) Incorrect pessimists

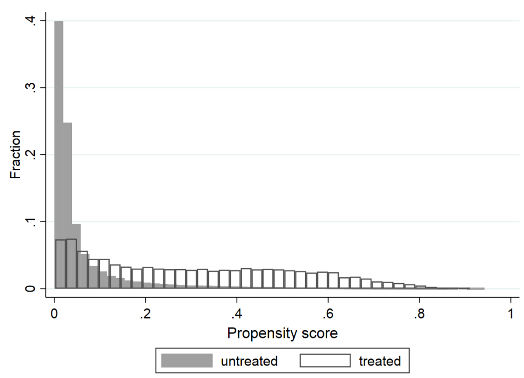
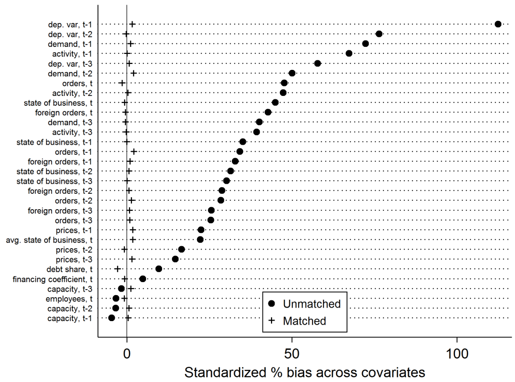
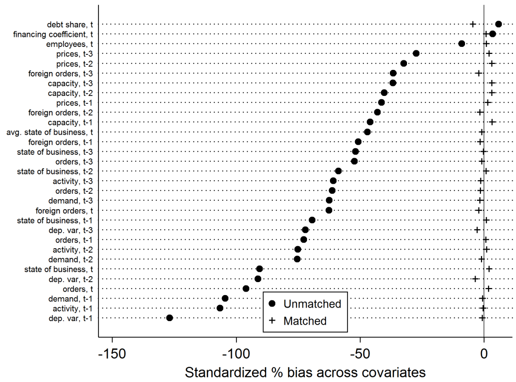


Figure C.2: Standardized bias, before and after matching, correct and incorrect treatments

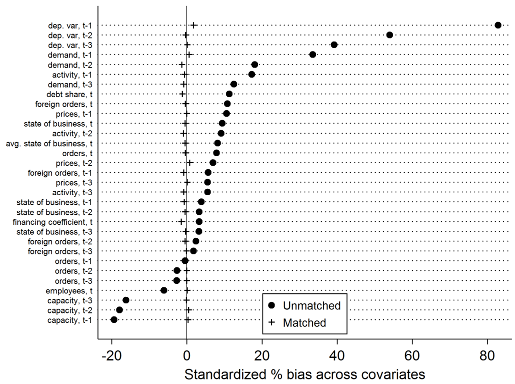
(a) Correct optimists



(b) Correct pessimists



(c) Incorrect optimist



(d) Incorrect pessimists

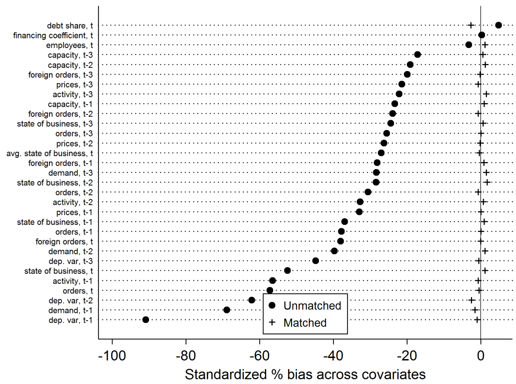
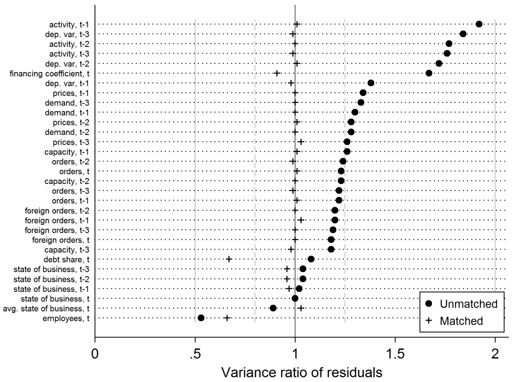
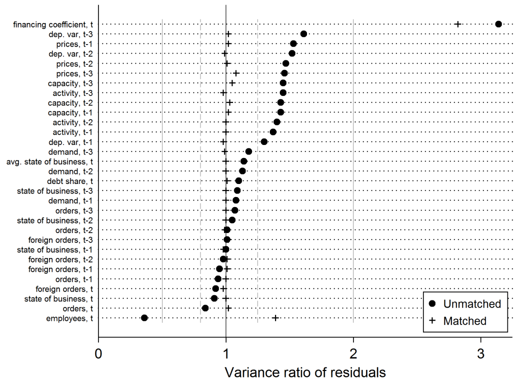


Figure C.3: Variance ratio of residuals, before and after matching

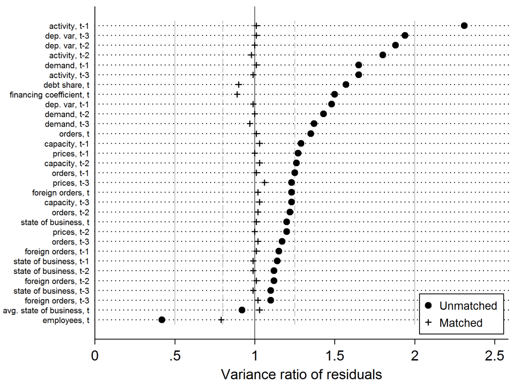
(a) Correct optimists



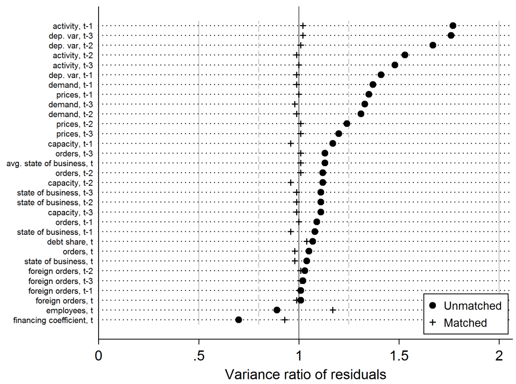
(b) Correct pessimists



(c) Incorrect optimists



(d) Incorrect pessimists



Notes: Ratios below 0.8 and above 1.25 (dashed lines) are considered "of concern"; ratios below 0.5 and above 2 (grey solid lines) are considered "bad" according to Rubin (2001).

D Sensitivity analysis for Sections 3 and 4

Table D.1: Aggregate results with alternative use of balance sheet data

	Optimists		Pessimists	
Dep. variable:	Production	Prices	Production	Prices
ATT	0.170***	0.026***	-0.175***	-0.029***
	(0.006)	(0.004)	(0.006)	(0.005)
Observations	120754	120802	116470	116548

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.1 Alternative use of balance sheet data

As discussed in section 2 the survey data has a different frequency than the balance sheet data. In our baseline setting we use the most recently published balance sheet data to estimate the debt share and financing coefficient in each month. This implies that in the months before a new balance sheet is published we use information which is almost one year old. We use this approach to avoid including any future information which is not yet available to firms at the time expectations are formed. However, one may argue that firms become aware of changing fundamentals already ahead of the publication of the new balance sheet. We therefore now propose an alternative method to link the annual balance sheet data to the monthly survey data. Specifically, for the six months following the publication of the balance sheet we use the most recent report as before. However, for the next six months until the new balance sheet is published we use the new data. This means that we always use the balance sheet data with the publication date closest to the respective month.

Table D.1 shows that changing the method for allocating the balance sheet data barely affects the results. Given that we only use two balance sheet variables in the probit regressions determining the propensity score this is not very surprising. Nevertheless, it is reassuring that our estimation is robust in this regard.

D.2 Alternative matching method

In order to ensure our results are not affected by our choice of matching algorithm, we implement an alternative algorithm as described in Lechner et al. (2011). These authors propose a radius (or caliper) matching procedure which includes weighting proportional to the distance of the match and a bias adjustment.

Specifically, the algorithm first selects all nearest neighbors in terms of the propensity score and potentially other variables (in the latter case using the Mahalanobis distance) without replacement. In our case we use the propensity score from the simple probit

Table D.2: Aggregate results with alternative matching procedure

	Optimists		Pessimists	
	No bias corr.	Bias corr.	No bias corr.	Bias corr.
Panel A: Production (change in current month)				
ATT	0.172***	0.172***	-0.174***	-0.174***
	(0.007)	(0.007)	(0.007)	(0.007)
Observations	135170	135170	131656	131656
Panel B: Prices (change in current month)				
ATT	0.027***	0.027***	-0.037***	-0.037***
	(0.005)	(0.005)	(0.006)	(0.006)
Observations	135170	135170	131656	131656

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

regressions described in section 3.1 and the month as an additional variable. The latter is done to ensure comparability to our matching procedure. In a next step the radius is computed as a function of the maximum distance within a matched pair in step one. Using this radius additional matches are selected if they are within the radius around the respective observation. This matching step is done without replacement, i.e. untreated observations can be matched to different treated observations. Weights are computed as the inverse of the distance between the untreated and treated observation in a match.

Finally, a regression bias adjustment is implemented by regressing the outcome variable on an intercept, the propensity score, the square of the propensity score, and any further variables used to define the distance. The regression is done only for the matched untreated observations using the weights obtained from matching. Using the regression coefficient one then predicts the potential outcome under no treatment for all observations. The difference between the weighted mean of the predicted outcome in the untreated group and the mean of the predicted potential outcome in the treated group is the estimated bias. This bias is then subtracted from the estimated ATT. The variance is computed analytically.

This approach differs from our matching algorithm because the radius is determined endogenously, the weights are proportional to distance, matches can be from different months (albeit only from close months because we include month as an additional distance measure), and finally there is regression adjustment. We implement this procedure using the STATA code provided by Huber et al. (2015). For simplicity we use their default settings. The results can be found in Table D.2. Using this alternative matching procedure does not affect our results substantially. Compared to our baseline specification in column 1 of Table 4 results only differ at the third digit. The largest difference is observed for prices of pessimists: -0.036 compared to -0.031 in the baseline. Reassuringly also the bias adjustment does not have any effects up to three digits. This implies that by using a more simple matching procedure with no bias correction is valid in our data set.