Firm expectations and economic activity*

Zeno Enders, Franziska Hünnekes, Gernot J. Müller

August 27, 2021

Abstract

We assess how firm expectations about future production impact current production and pricing decisions. Our analysis is based on a large survey of firms in the German manufacturing sector. To identify the causal effect of expectations, we rely on the timing of survey responses and match firms with the same fundamentals but different views about the future. Firms that expect their production to increase (decrease) in the future are 15 percentage points more (less) likely to raise current production and prices, compared to firms that expect no change in production. In a second step, we show that expectations also matter even if they turn out to be incorrect. Lastly, we aggregate expectation errors across firms and find that they account for about 15 percent of aggregate fluctuations.

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

JEL-Codes: E32, D84, E71

*Enders: Heidelberg University and CESifo, zeno.enders@uni-heidelberg.de. Hünnekes: European Central Bank, franziska.huennekes@ecb.europa.eu. Müller: University of Tübingen, CEPR, and CESifo, gernot.mueller@uni-tuebingen.de. We thank Guido Lorenzoni (the editor), three anonymous referees and participants of various conferences and seminars for useful comments. We also gratefully acknowledge valuable suggestions by Nikolai Histrov in the early stage of the project, the technical support by the team of the LMU-ifo Economics & Business Data Center (EBDC) in Munich, and research assistance by Mark Brisske. This research has received financial support by the German Science Foundation (DFG) under the Priority Program 1859. Franziska Hünnekes thanks the Egon Sohmen Graduate Center at the University of Munich (LMU) for financial support while this research was conducted. The views expressed in this paper are the authors and do not necessarily reflect the views of the European Central Bank or the Eurosystem.

1 Introduction

To what extent do firm expectations affect current decision making? According to theory, expectations should have a first-order effect. Expectations take center stage in modern macroeconomic theory, which assumes that 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 firm expectations affect 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 exploiting a uniquely suited 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 to 2016. In each month, about 2000 firms in Germany's manufacturing sector report expectations regarding their future production in a qualitative manner: within the next three months it may increase, not change, or decrease. Similarly, firms report expectations about "business 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 variables for each firm. These include a large range of indicators that capture the economic and financial conditions under which firms operate.

We 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 based on 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 probit models in order to obtain an estimate for the probability that a firm expects production to rise or fall, respectively, given its fundamentals in a specific month. We then match firms that expect their production to increase or decrease with firms that expect no change using propensity scores (Rosenbaum and Rubin 1983). In this way, we interpret expectations that future production either increaeses or decreases as treatments that are randomly assigned across firms with the same fundamentals. We estimate for both treatments the average treatment effect on the treated by comparing the behavior of treated and non-treated firms with the same probability of being treated.

Identification relies on the timing of firms' survey responses: we estimate the effect of expectations, which firms report early in the month, on their production and price-setting decisions during the month, which they report in the following month. We find that expectations have a significant effect on production and prices. Firms that expect future production to increase are 15 percentage points more likely to raise production than firms that expect no change. They are also considerably more likely to raise prices. Firms that expect production to decline, in contrast, respond in the opposite way—they are more likely to reduce production and prices.

As we show formally below, these results are consistent with two distinct hypotheses on how expectations affect economic decision making. Under the first hypothesis, expectations that are orthogonal to current fundamentals are not necessarily orthogonal to future fundamentals. Put differently, expectations reflect genuine information ("news") about the future, which has not yet materialized 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; Görtz and Tsoukalas 2017; Schmitt-Grohé and Uribe 2012). Yet, these studies provide only indirect evidence on the role of expectations as news. In contrast to our analysis, they do not analyze expectations data explicitly.

Under the second hypothesis, changes in expectations are fully exogenous. Different labels are used to capture this notion, such as "noise," "sentiment," or "animal spirits".[1] A number of recent contributions have 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, these contributions also do not exploit expectations data directly. Instead, they show that noise helps quantitative business cycle models to account for key features of aggregate time-series data.

The unique nature of our data set allows us to test these two hypotheses directly. For not only do we observe firm expectations regarding future production and business conditions, we

[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).

also observe actual production and business conditions. We are thus able to construct a measure of firms' expectation errors and identify firms whose expectations turn out to be incorrect from an ex-post point of view. In the words of Pigou (1927), such expectations are "undue" with the benefit of hindsight. In the second step of our analysis, we therefore match firms that expect a change in production, which does not materialize, to firms that expect production to remain unchanged. Again, we find that expectations—even as they turn out to be incorrect—matter for output and price-setting decisions. Qualitatively the effect is the same as in the case when we do not condition on expectations being incorrect, quantitatively the effect is weaker.

In a third step, we aggregate expectation errors across firms in order to quantify their contribution to the business cycle. In this case, we no longer compare expectations on a firm-by-firm basis but rely on an ordered probit model to benchmark actual expectations against fundamentals. Since we are interested in the effect of expectations as such, we proceed as follows. First, we focus on those firms that expect an increases or a decrease of production even though the model suggests otherwise. Next, we classify expectations as either correct or incorrect based on actual outcomes as reported ex post. Finally, we aggregate across firms and project macro variables of interest on the aggregate expectation error. We find, in particular, that undue expectations of a production increase cause industrial production and prices to rise. We also find that the aggregate expectation error accounts for some 15 percent of aggregate fluctuations, in line with earlier estimates by Blanchard et al. (2013), Hürtgen (2014) or Enders et al. (2020).

This is remarkable because our approach differs fundamentally from those studies, which achieve identification in a classic time-series context—with or without a fully structural model. In the present paper, we exploit a large microdata set and put forward a new bottom-up approach: we start by identifying the effect of firm expectations on firm decisions, construct a measure of expectation errors and, eventually, trace out their aggregate implications. In doing so, our paper relates closely to two strands of research.

First, we build on recent work which uses survey data to shed light on the expectation formation process. In order to do so, most authors focus on surveys of professional forecasters following the influential work of Coibion and Gorodnichenko (2012, 2015). Surveys of firm expectations, in contrast, have received less attention. Coibion et al. (2018), for instance, study firm expectations based on a survey in New Zealand. Furthermore, there is work based on the ifo survey. An early study by Nerlove (1983) finds evidence in support of an adaptive expectations model. Bachmann and Elstner (2015) show that at most one-third of the firms in the ifo survey make systematic forecast errors. Massenot and Pettinicchi (2018), in turn, identify various factors which account for forecast errors of firms in the ifo sample. In Enders et al. (2019), we show that firm expectations respond systematically to monetary policy announcements.[2]

Second, we follow a number of recent papers that take up the challenge of directly estimating the effect of firm expectations on firm decisions. While a number of studies have highlighted the role of expectations regarding macroeconomic developments, such as GDP or inflation for firms' decisions (Coibion et al. 2020d; Tanaka et al. 2020), our focus is fundamentally different as we investigate the role of firms' expectations regarding their own, firm-specific developments. We share this focus with Gennaioli et al. (2015), Boneva et al. (2020) and Bachmann and Zorn (2020). Relative to these studies, our contribution is to put forward a new identification strategy in order to isolate the role of expectations, both as a transmission channel of news and as an exogenous trigger of economic decisions.

Last, we note that while a systematic empirical assessment of firm expectations based on survey data is still in its infancy, recent work has already established important insights into how households form economic expectations. These studies point to important heterogeneity in the expectation formation process across the population. D'Acunto et al. (2020) highlight the role of IQ, while other authors stress the role of distinct information channels (Coibion et al. 2020c; Conrad et al. 2021). The effect of household expectations on consumption and saving decisions has been explored using either natural experiments as in D'Acunto et al. (2021) or by running randomized control trials as in Coibion et al. (2020a). By now it appears that while firm and household expectations share certain characteristics, as found by Coibion et al. (2020b), they also display important differences (Candia et al. 2021; Link et al. 2021).

The remainder of the paper is structured as follows. The next section offers a brief formal exposition of the basic idea which underlies our empirical strategy. Section 3 provides details on our dataset as well as descriptive statistics. Section 4 describes the estimation approach and the results of the first step of our analysis. In Section 5, we zoom in on the transmission channels of firm expectations by distinguishing between firms with and without expectation errors. Afterwards, we quantify the aggregate effects of firm expectations using local projections. Section 7 concludes.

[2]Born et al. (2021b) survey recent work on firm expectations regarding their own variables.

2 Fixing ideas

Firm expectations depend on firm fundamentals. In order to identify the effect of firm expectations on firm decisions, we rely on variation in expectations that is not accounted for by fundamentals. In what follows, we provide a formal—if highly stylized—exposition of the issue at hand. In this way, we merely intend to fix ideas in order to guide our empirical analysis below, rather than to make an original point.

Let $a_{i,t}$ denote (the log of) a fundamental of a generic firm i at time t, say, its productivity or any other relevant variable. To simplify the exposition, we stick to the univariate case but our argument also goes through if there is a vector of fundamentals. We assume the following law of motion:

$$a_{i,t+1} = \rho a_{i,t} + u_{i,t+1}, \tag{2.1}$$

where $0 < \rho \leq 1$ and $u_{i,t+1} \sim N(0, \sigma_u^2)$ is an unforseen innovation to the fundamental. Generally, $a_{i,t}$ may depend on the decisions of firm i, but we abstract from this aspect since it is not essential for our argument. In period t, the firm receives a noisy signal about $u_{i,t+1}$:

$$s_{i,t} = u_{i,t+1} + \nu_{i,t}, \tag{2.2}$$

where $\nu_{i,t} \sim N(0, \sigma_\nu^2)$ is noise. Optimal signal-processing implies

$$E_{i,t}(a_{i,t+1}) = \rho a_{i,t} + \delta s_{i,t}, \tag{2.3}$$

where $E_{i,t}$ is the expectation operator of firm i and $\delta = \sigma_u^2/(\sigma_u^2 + \sigma_\nu^2)$.

In the first step of our empirical analysis, we take an *ex-ante perspective* and match firms, which report different expectations but feature the same fundamentals. Let's say firm i reports higher expectations than firm j, even though they have the same current fundamentals: $a_{i,t} = a_{j,t}$. In this case, any difference in the reported expectations is due to the difference in the signal across firms:

$$\begin{aligned} E_{i,t}(a_{i,t+1}) - E_{j,t}(a_{j,t+1}) &= \rho a_{i,t} + \delta s_{i,t} - \rho a_{j,t} - \delta s_{j,t} \\ &= \delta \left(s_{i,t} - s_{j,t} \right). \end{aligned}$$

In our empirical analysis, we thus attribute difference in the behavior across firms with identical fundamentals to their expectations and, more specifically, to the difference in expectations that is not accounted for by fundamentals, that is, to the difference in the signal. The signal may reflect either genuine news or noise. Hence, if firm i reports higher expectations, (2.2) implies the following inequality:

$$\nu_{i,t} - \nu_{j,t} > -(u_{i,t+1} - u_{j,t+1}). \tag{2.4}$$

Intuitively, firm i has higher expectations either because it received more positive noise or "better" news than firm j.

In the second step of our analysis, we take an *ex-post perspective* in order to further characterize how expectations impact current decisions. Key to the ex-post perspective is the expectation error, which we can compute for firm i using equations (2.1) and (2.3). It amounts to the difference between the actual innovation and the forecast of the innovation based on the signal:

$$a_{i,t+1} - E_{i,t}(a_{i,t+1}) = u_{i,t+1} - \delta s_{i,t} = (1-\delta)u_{i,t+1} - \delta\nu_{i,t}. \tag{2.5}$$

We may now repeat our matching exercise in the same way as before, that is, we match firms with the same fundamentals but different expectations, but now we account for firms' expectation error. Consider again the case where firm i has higher expectations than firm j but now assume that firm i was wrong, that is, its expectation error is negative. Firm j makes no expectation error. Then equation (2.5) implies:

$$\delta(\nu_{i,t} - \nu_{j,t}) > (1-\delta)(u_{i,t+1} - u_{j,t+1}).$$

or, substituting for δ:

$$\frac{\nu_{i,t} - \nu_{j,t}}{\sigma_\nu^2} > \frac{u_{i,t+1} - u_{j,t+1}}{\sigma_u^2}. \tag{2.6}$$

This expression shows that firm i has received relatively more positive noise than negative news, compared to firm j. In principle this could be, with equal levels of noise for both firms, because firm j received positive news. However, inequality (2.4) tells us that if news are relatively positive for firm j, noise must be even more positive for firm i. Hence, as we match firms with different expectations but identical fundamentals in period i and observe that the firm with higher expectations experiences a negative expectation error ex post, we may conclude that the difference in current behavior is caused to a large extent by noise.

3 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). Each month the ifo conducts four different surveys that cover German firms in the following sectors: manufacturing, retail, construction, and services. The surveys include the same basic stock of questions for each sector, but the wording of the questions and possible answers differs at times. In our analysis, we focus on the manufacturing survey, which covers the longest time period and the largest number of firms. A caveat is that the responses to the survey and the balance sheet data come at different frequencies: while the survey is conducted each month, the balance sheet data is only available at annual frequency. We will use balance sheet data, in addition to survey variables, to predict firm expectations. In order to ensure that we do not use information that is not yet available when firms report expectations, we use the most recent balance sheet data at a given point in time.[3]

While the BEP sample starts in 1980, we compute treatment effects for the sample from 1991 to 2016 due to data availability. The unit of observation in the manufacturing survey is a product. As a result, some firms respond to several questionnaires each month or different plants of one firm respond separately. In our sample, however, this is the case for less than 10% of firms. 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. We conduct a corresponding robustness check later on. In our analysis below, we refer to the individual observation as a "firm" in order to ease the exposition.

We compile a number of basic statistics for the firms in our sample. For this purpose, we distinguish between the sample of all firms ("full sample") and the sample for which balance sheet data is available. The latter sample is smaller but still includes some 5,000 distinct firms and more than 300,000 firm-month observations, see Table A.1 in the Appendix. We find that the mean and standard deviation of responses are generally quite similar across samples, although firms for which balance sheet data are available tend to be somewhat larger (Table A.2). Firms with balance sheet data stay in the survey for 87 months (7 years) on average and provide answers in 74 months, implying that they respond in 83% of the months during the period in which they

[3]For example, if a firm's financial year ends in September, we use this data for all the following months until the next balance sheet is available. Hence, our specification is conservative as we neglect potential information known to firms in the months close to but preceding the closing date of the balance sheet. In Appendix D.1 and D.2, we pursue two alternative strategies to using the balance sheet data but find that the results obtained for the baseline specification are robust.

Table 1: Selected ifo survey questions

Label	Question	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]

Notes: Tables shows our translation of the most recent formulation of the question in the German questionnaire. Additional questions used are listed in Table B.1. Changes to the questions are listed in Table B.2

are in the sample (again see Table A.1). We also stress that independently of the starting period, sample attrition is moderate (Table A.3).

The BEP includes a large set of questions, but only a subset of those are asked regularly. In our analysis, we focus on four main questions, listed in Table 1. These questions permit for qualitative responses only and for the purpose of our quantitative analysis, 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. However, since quantitative answers are available for a subset of questions since 2005, we perform an extended analysis which accounts for quantitative differences across responses as well (see Section 4.5).

Some questions vary over time. Especially in 2002 many changes were implemented due to a harmonization of business and consumer surveys in the European Union. The changes relevant for our analysis are documented below. Within firms, the questions are typically answered by the top management. In more than 80% of small and medium-sized firms and more than 60% of large firms, the CEO or the owner responds. Otherwise, the response is typically provided by the head of the relevant department (for details, see Sauer and Wohlrabe 2019).

Our baseline measure of firm expectations are the responses to 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 2001. Since these are minor changes affecting all firms in the same way, they are unlikely to matter for our results.

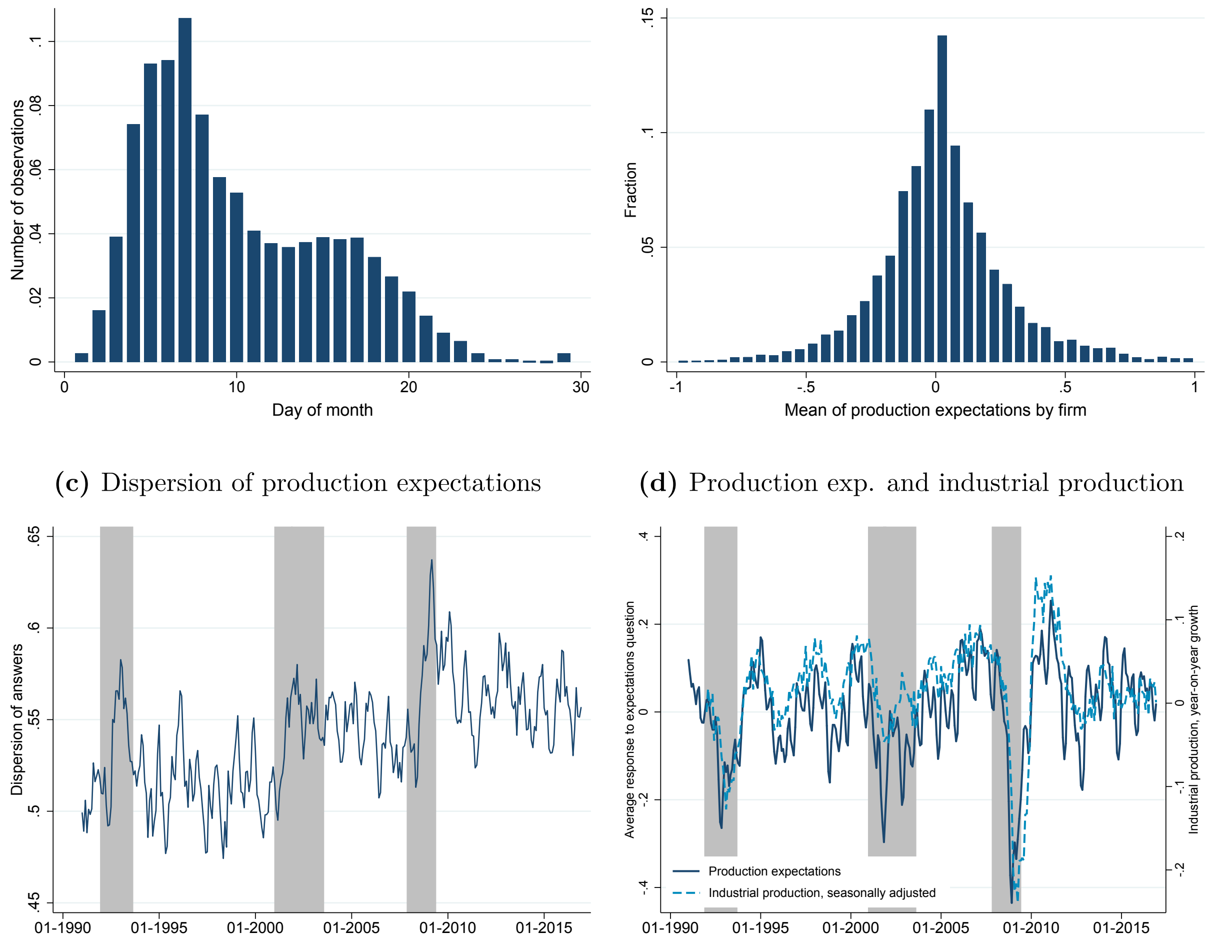
Q2 is a broader question about expectations of the state of business over the next six months. Combined with a question on the current state of business it provides the basis for the ifo business climate index. In a sensitivity analysis, we use this question and find similar results as in our baseline. Also, the answers to both questions tend to be highly correlated, see Figure A.1 in the appendix. In our baseline analysis, we use Q1, though, because its wording is more specific and the time horizon in question is shorter. Q2 also used to include an additional note to ignore seasonal fluctuations; it was dropped in 1997 (see Table B.2).

Questions Q3 and Q4 refer to our outcome variables: changes in production and prices. 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 these variables in the previous month. We adjust the data to account for this change in the survey: we make sure that we always relate changes of production and prices in a given month to expectations regarding the developments in the following months. Actual developments are reported either in the same month as the expectations (before 2002), or in the following month (after 2002). In addition, we consider a reduced sample which starts in 2002 only in our sensitivity analysis. The results are very similar to those for the full sample (see Section 4). Further details on the wording of the questions can be found in Table B.2 in the appendix.

For our analysis below it is important to note that most firms provide answers early in the month. To see this, consider panel (a) of Figure 1 which shows the distribution of responses over time. Here, our sample is based on responses to the online version of the survey which is available since 2004 and by now used by the majority of firms. Note that about 60% percent of firms file their responses during the first 10 days of the month.

In panel (b) of the same figure, we show how the average response to question Q1 (over time) is distributed across firms. The distribution is close to normal. The same holds true when we consider the distribution of the average expectation error over time across firms. It is roughly symmetrically centered around zero (see Figure A.1). There is considerable and time-varying dispersion of responses, as panel (c) of Figure 1 illustrates. Here we plot a common measure of dispersion based

Figure 1: Survey responses



(a) Distribution of responses over time

(b) Average response by firm

(c) Dispersion of production expectations

(d) Production exp. and industrial production

Notes: Full sample. Panel (a) shows distribution of responses within the month for online responses, available since 2004; panel (b): average response of individual firm (only includes firms which respond at least 10 times); panel (c): dispersion of production expectations; shaded areas indicate recession periods as defined by the German Council of Economic Experts. panel (d): average production expectations and industrial production (German Statistical Office).

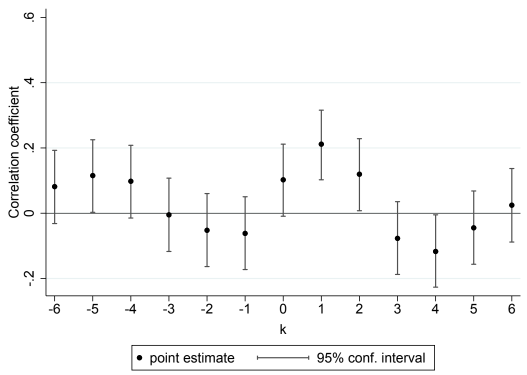
on the shares of positive and negative responses in a given month.[4] Dispersion generally increases during recessions (indicated by the shaded area) and tends to decrease afterwards.

At a fundamental level, production expectations co-move strongly with economic activity. To illustrate this, we plot average production expectations jointly with an index of industrial production in panel (d) of Figure 1. To investigate this issue more systematically, we plot the cross-correlation function of average firm expectations and the monthly growth rate of industrial production in panel (a) of Figure 2. The two time series are strongly correlated for small leads of

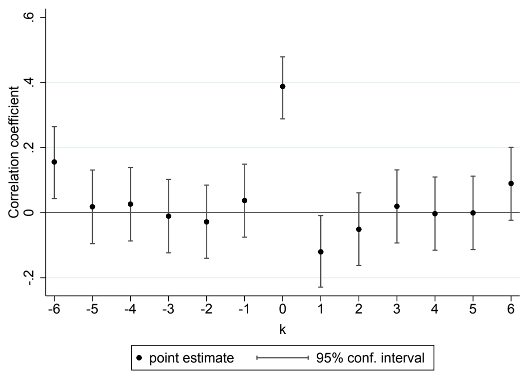
[4]Dispersion of expectations based on qualitative survey data is measured as $\sqrt{\text{frac}^+ + \text{frac}^- - \left(\text{frac}^+ - \text{frac}^-\right)^2}$, where frac^+ and frac^- are the fraction of positive and negative responses in each month, respectively, see e.g. Bachmann et al. (2013).

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

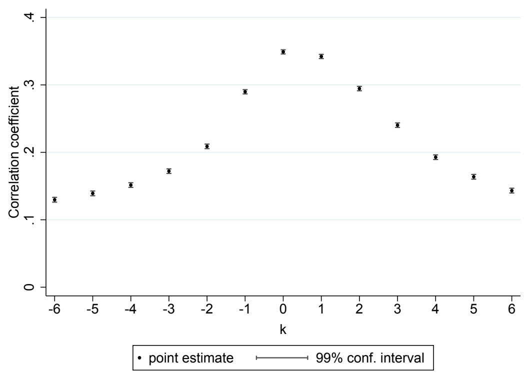
(a) Mean expected production in t and industrial production in $t+k$, monthly



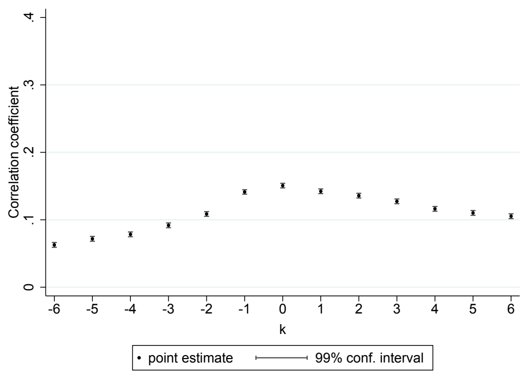
(b) Mean reported production in t and industrial production in $t+k$, monthly



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



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



Notes: Full sample. Data on industrial production in the manufacturing sector is taken from the German Statistical Office and measured as the month-on-month change in the not seasonally adjusted index.

industrial production, consistent with the well established fact that the ifo business climate index is a leading indicator of economic activity in Germany (Abberger and Wohlrabe 2006; Henzel and Rast 2013). Next, the strong contemporaneous correlation of reported production with industrial production, shown in panel (b), suggests that on average firms accurately report production.

In the bottom panels of the figure we consider the correlation of production expectations and actual production and prices at the firm level. Panel (c) of Figure 2 displays the average cross-correlation function of production expectations and production across firms, again for leads and lags of six months. Panel (d) reports the same statistic for prices. The contemporaneous correlation of production and prices, on the one hand, and production expectations, on the other hand, is particularly strong. In what follows, we seek to establish the causal effect of production expectations on production and price-setting decisions.

4 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 aim to assess to what extent firms' production and price-setting decisions depend on their production expectations. To this end, we compare the behavior of firms that expect an increase (decrease) of production to firms that expect production to remain unchanged. As the discussion in Section 2 above makes clear, a key challenge in this regard is to identify variation in expectations that is 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 since we are dealing with qualitative data on expectations: firms may either expect an increase, no change, or a decrease. Hence, in our analysis, if firms receive a treatment, they are treated either with expectations of an increase or with expectations of a decrease. Of course, we do not require expectations to be literally assigned in a random way. We merely assume that conditional on fundamentals, the assignment is random. Note also that our analysis does not require expectations to be unrelated to future fundamentals. We take up this issue in more detail in Section 5.

Generally, the matching approach offers several advantages over conventional regression analysis. First, it ensures that the distribution of control variables is 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 a significant bias when estimating treatment effects (Heckman et al. 1998). Second, the matching approach disciplines the analysis since 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).

Nevertheless, one may still be worried about reverse causality: rather than measuring the effect of firm expectations on production and pricing, we may pick up the effect of production on expectations. To see why this might be a problem, consider a scenario where actual production and expectations about future production increase simultaneously in a given month because, say, demand for a firm's product picks up. In light of this concern, two observations are key. First, we consider a large set of control variables ("fundamentals") in our matching routine. Importantly, this includes a number of lagged variables and non-linear interaction terms, but also variables which are contemporaneously observed, such as current orders and the reported state of business. Second, the timing of survey responses is the key to our identification strategy: because the large majority of responses to the survey is filed early in the month, we effectively sample expectations *before* actual production has taken place. As Figure 1 above shows, 50% of firms answer within the first eight days and another 25% answer in the following week. We also show below that our results are robust once we restrict the sample to those firms which respond within the first 10 days of the month. For these reasons we are confident that our estimate of the effect of expectations on production and price setting is not contaminated by reserve causality.

4.1 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 expects production to increase/decrease, 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:

$$\mathrm{E}[Y(1)|D = 1] - \mathrm{E}[Y(0)|D = 0] = \theta + \mathrm{E}[Y(0)|D = 1] - \mathrm{E}[Y(0)|D = 0]. \tag{4.1}$$

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 the 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 consider all variables 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 (4.1). 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 on the propensity to be treated, $p(X) \equiv Pr(D = 1|X)$, or to condition 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:

$$0 < p(X) = Pr(D = 1|X) < 1. \tag{4.2}$$

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: production may be expected to increase or decrease. To

Table 2: Control variables in the propensity score model

Variable	Description	Frequency	Reference period
debt share[1]	total debt over assets	annual[2]	$t-11$ to t
financing coefficient[1]	liabilities minus provisions divided by equity plus provisions	annual[2]	$t-11$ to t
employees	no. of employees	annual[3]	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 three lags are also 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 quarter round the question was asked.
[3] The number of employees is a three-year rolling average of the firm's response to reduce breaks in the series.

establish the effect of a treatment, we compare firms in each case to firms which do not expect production to change at all.

In order to estimate the propensity score, we pursue two alternative approaches. Since we are dealing with two treatments, we first estimate an ordered probit model where an expected increase and an expected decrease are outcomes of a common model. Alternatively, we consider two distinct probit models for both "treatments."

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

$$Pr(y_{it} = j) = Pr(a_{j-1} < y^*_{it} \leq \alpha_j) = \Phi(\alpha_j - X'_{it}\beta) - \Phi(\alpha_{j-1} - X'_{it}\beta), \qquad (4.3)$$

where $j = \{-1, 0, 1\}$ corresponds to the three possible answers to Q1. We collect the control variables in the 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). We provide more details on the selection criteria below. For the sectors, we consider the 2-digit breakdown of the German system of industry classification (WZ08). 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 $m = \{\text{expected production increase}, \text{expected production decrease}\}$, equals the conditional probability of the treatment given the alternative of no treatment, that is, no expected change of future production:

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

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:

$$Pr(D_{it}^m = 1) = Pr(X_{it}'\beta) = \Phi(X_{it}'\beta), \tag{4.4}$$

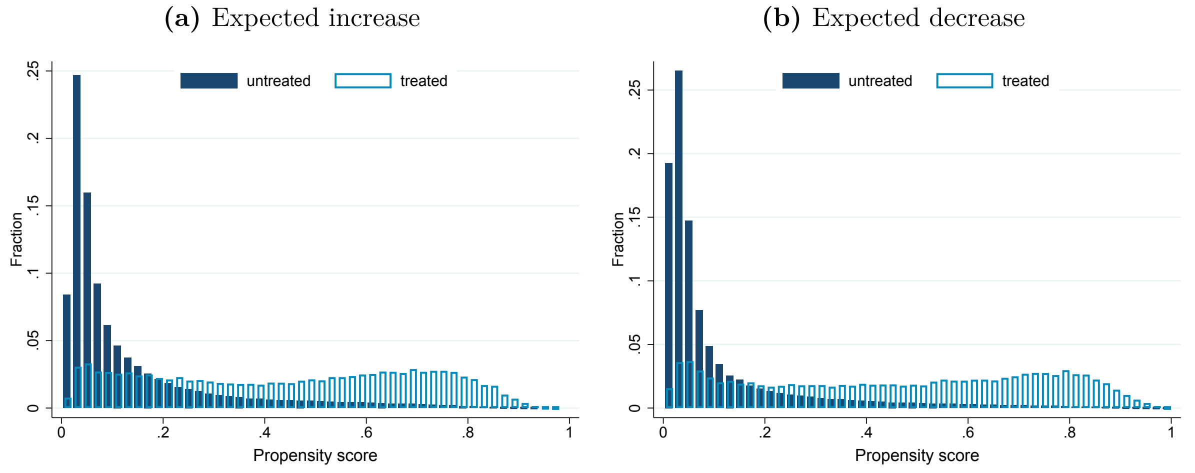
where D_{it}^m is a dummy variable, which equals 1 in the event of treatment, and 0 otherwise. We again collect the same control variables in vector X_{it}. Since the sample only includes the specific treatment group and the untreated, the estimated probability is a direct estimate of the propensity 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 treatment options. 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 indeed a slight advantage as it yields better balancing statistics. We therefore use it in our baseline. Results based on the ordered probit, however, do not differ much from results using the two probit regressions, see Table D.1.

We choose the control variables following a procedure by Imbens and Rubin (2015). First, we select a set of variables as the baseline group, which we want to include in the model in any case. These variables are production expectations, realized production, realized prices, demand, and capacity utilization, all from the previous month. In addition, we include the number of employees, the current order situation, the average state of business in the firm's sector, time and sector fixed effects in the baseline group. We include these variables because all of them should matter jointly for expectations and outcomes. Then we check whether to include additional variables, namely the

Figure 3: Histogram of the density of the propensity scores



Notes: Propensity scores for treated and untreated firms respectively, estimated as described by equation (4.4). Untreated firms are firms which report that they expect no change of production (Question Q1). Panel (a): treated firms expect an increase; panel (b): treated firms expect a decrease.

state of business, foreign orders, and the balance sheet variables. We decide on whether to include these additional variables on the basis of a log-likelihood test. Finally, we also check whether to include additional lags and interaction terms. This procedure yields the specification described above.

After computing the propensity scores, we match treated and untreated observations using a variant of caliper or radius matching (Caliendo and Kopeinig 2008).[5] We match each treated observation i to all untreated observations k within the same month which satisfy

$$p(X_{it}) - 0.02 \leq p(X_{kt}) \leq p(X_{it}) + 0.02. \tag{4.5}$$

Here we allow for a radius of 0.02. This corresponds to about a tenth of the standard deviation of the estimated propensity score.[6] In case a treated observation is matched to several untreated observations, the untreated observations receive equal weights which add up to unity. Note that an untreated observation may feature in several matches.

Figure 3 displays the distribution of the propensity scores. The left panel contrasts the distribution for firms that are treated with expectations of a production increase (light blue, transparent bars) with those for untreated firms (dark blue, solid bars). The right panel shows the analogous distribution for expectations of a production decrease. In each instance, we find

[5]We also test an alternative matching procedure proposed by Lechner et al. (2011). The results are very close to our baseline results. Details can be found in Appendix D.3.

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

Table 3: Number of matched observations

	Expected increase		Expected decrease	
	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

Notes: Panel (a) shows results for matching as discussed in this section. Panels (b) and (c) show results for matching based on more specific treatments as discussed in Section 5.

that there is considerable overlap of the distribution (common support), although the mass of untreated firms is more concentrated at lower propensity scores.[7] Panel (a) of Table 3 reports basic statistics regarding our matches. We are able to find matches for about 93% (90%) of all firms treated with expectations of a production increase (decrease). This is due to the large overlap in propensity scores between treated and untreated firms.

4.2 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 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:

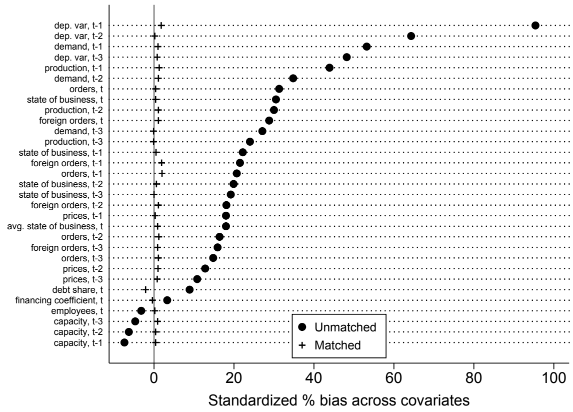
$$SB = 100 \frac{\bar{x}_1 - \bar{x}_0}{\sqrt{0.5(s_1^2 + s_0^2)}}, \tag{4.6}$$

where $\bar{x}_1$ is the mean of the control variable for the treated observations, $\bar{x}_0$ is the mean of the control variable for the untreated observations, s_1 is the standard deviation of the treated observations and s_0 the standard deviation of all untreated observations. Panels (a) and (b) of Figure 4 show that as a result of matching observations, we achieve a sizeable reduction of the

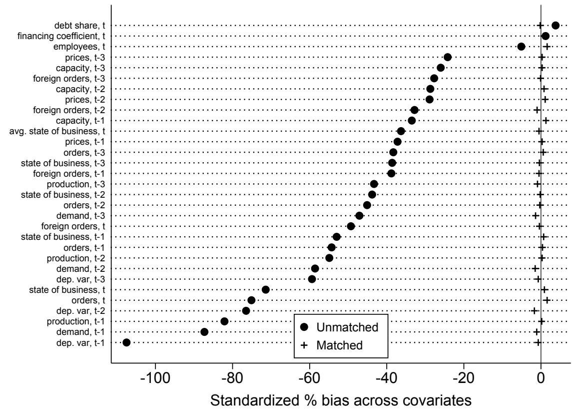
[7]There are also some treated observations with a larger propensity score than the largest propensity score of all untreated observations. We drop these observations in what follows. This trimming ensures that only suitable observations are matched.

Figure 4: Standardized bias and variance ratio, before and after matching

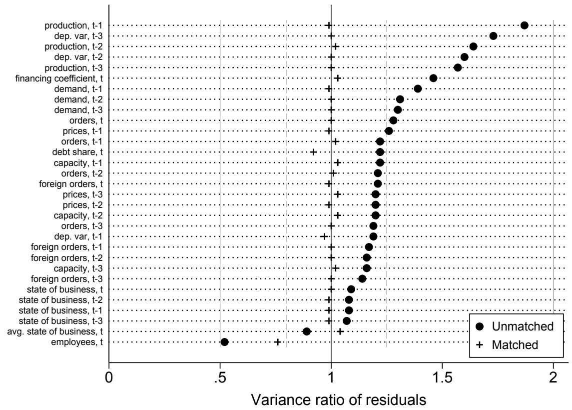
(a) Standardized bias, expected increase



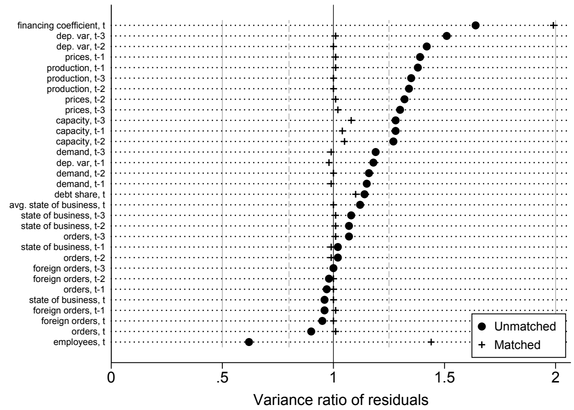
(b) Standardized bias, expected decrease



(c) Variance ratio, expected increase



(d) Variance ratio, expected decrease



Notes: Figure shows diagnostics statistics for the matching of firms expecting a production increase or decrease, respectively. The standardized bias measures the mean difference of each variable in the treated and untreated groups, as described by equation (4.6). The variance ratio measures the difference between the variances orthogonal to the propensity score. Variance ratios below 0.8 and above 1.25 (dashed lines) are considered "of concern"; ratios below 0.5 and above 2 (gray solid lines) are considered "bad", according to Rubin (2001).

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).[8] We meet this criterium in all instances, see also Table C.1 in the appendix.

Rubin (2001) suggests a second measure of balancing, arguing 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. Panels (c) and (d) of Figure 4 plot the variance ratios before and

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

after matching. Again, we find that matching firm-month observations ensures that treated and non-treated firms appear well balanced in terms of covariates. Only the ratio for the number of employees and the financing coefficient (for firms expecting a decrease) falls in the "of concern" area (outside dashed lines).

4.3 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. For both outcome variables, we compute the ATT as the mean difference, across all matches, of treated and untreated firms.

The computation of standard errors for the estimate of the ATT based on 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). He shows that in case of variants of nearest neighbor matching, as in our case, the variance of the ATT, $\hat{\tau}_{ATT}$, is given by:

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

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

4.4 Results

We now turn to the question that motivates our analysis: to what extent do firm expectations affect current decision making? Table 4 provides a first answer. In the upper part of the table, we report the ATT for expectations of a production increase, in the lower part we report the ATT for expectations of a production decrease. In each instance, we focus on production and price-setting decisions in the current month, while expectations pertain to production within the next three months. In each column, we consider alternative specifications.

The left-most column (1) reports results for our baseline specification. We find a significant positive treatment effect for production (panel (a)). This positive effect may reflect a stronger tendency among treated firms to raise production or a reduced tendency to lower production, or both. We disentangle these effects below. For prices, we also find a significant positive effect, although in this case the effect is much smaller, see panel (b). Taken at face value, such an

Table 4: Average treatment effect on the treated

	(1) Baseline	(2) Radius 0.01	(3) Sample 2002-2016	(4) Sample excl. fin. crisis[1]	(5) Match in sector	(6) Response in first 10 days[2]
Panel (a): Expected production increase – Effect on production						
ATT	0.172***	0.170***	0.181***	0.170***	0.165***	0.200***
	(30.43)	(29.34)	(30.22)	(28.52)	(23.30)	(19.20)
Observ.	129812	120335	108660	113690	52961	31722
Panel (b): Expected production increase – Effect on prices						
ATT	0.025***	0.025***	0.024***	0.025***	0.026***	0.032***
	(5.97)	(5.80)	(5.30)	(5.52)	(5.00)	(3.98)
Observ.	129858	120367	108691	113734	52962	31732
Panel (c): Expected production decrease – Effect on production						
ATT	-0.173***	-0.170***	-0.169***	-0.172***	-0.164***	-0.174***
	(-27.77)	(-26.47)	(-25.00)	(-25.37)	(-20.48)	(-13.81)
Observ.	125458	113992	104275	106764	47320	28855
Panel (d): Expected production decrease – Effect on prices						
ATT	-0.031***	-0.033***	-0.026***	-0.035***	-0.028***	-0.025**
	(-6.13)	(-6.41)	(-4.76)	(-6.53)	(-4.52)	(-2.46)
Observ.	125530	114050	104337	106821	47341	28877

Notes: Tables shows treatment effects on prices and production for different specifications. Treatment is the expectation of future production. Outcomes refer to the current month, t, while expectations refer to the next 3 months ($t+1$ to $t+3$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[1] Excluding the years 2008 and 2009.
[2] Results are based on online responses only, sample starts in 2004.

apparently small effect is consistent with the notion that prices are adjusted only infrequently in the short run. Note, however, that the effect is the outcome either of more frequent upward adjustments or less frequent downward adjustments of prices among treated firms. Last, we note that the effect on production and prices is quite symmetric for expectations of increasing and decreasing production, even though we estimate separate models. The estimate for the latter is shown in panels (c) and (d) of Table 4.

Table 4 also reports results for alternative specifications in columns (2) to (6). We stress upfront that across all specifications the estimate of the ATT is close to that for the baseline and significant throughout. Balancing statistics for the four sensitivity specifications that require changes in the matching procedure are summarized in Tables C.2 to C.5 in the appendix. 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 shorter sample period. It starts in 2002 rather than in 1991 because, as discussed in Section 3 above, the questionnaire has slightly changed in 2002. In column (4) we report results for a sample which excludes observations from the financial crisis, that is, the years 2008 and 2009.

Next, we repeat our estimation but allow only for matches within the same 2-digit NACE sector (out of 24). In this way, we address concerns that results are driven by sectoral demand shocks rather than expectations as such. Column (5) shows the results. They are very similar to the baseline. Lastly, we consider a restricted sample and include only firms which respond to the survey within the first 10 days of the month. The specific response date is known only for online responses. Hence, we consider a restricted sample in this case, as online responses are possible only since 2004 and a majority, but not all firms use the online option by now. By limiting the sample to early respondents, we make sure that production expectations are not driven by actual production in the current month. Results are shown in column (6). They are not weaker than the results for the baseline. The same holds if we limit the sample to firms which respond within the first week of the month, see column (2) of Table D.1 in the appendix.

Finally, we run a number of additional robustness tests and report results in Table D.1 in the appendix. Specifically, we first estimate an ordered probit model instead of two separate probit models to compute the propensity scores, as discussed in Section 4.1. Next, we limit the sample to firms that report only on one product in the survey. Third, we only include firm observations up to three months after the balance sheet closing date to avoid relying on outdated data.[9] Fourth, we also include firm age as a control variable given that this may be a further relevant factor which determines firms' fundamentals and expectation formation. The reason for excluding firm age in the baseline specifications is the limited data availability. Overall, we obtain similar ATTs for the alternative specifications. Only in the case of limiting the sample to firms responding within the first 7 days of the month, the effect of an expected production decline on prices is no longer significant. Lastly, we consider an alternative to propensity score matching by conditioning, using OLS, the ATTs parametrically on all the matching variables in the first stage. Results are provided in Appendix D.4 and are similar to those found in the baseline specification.

As noted above, our results regarding the response of production and prices may reflect more upward adjustments or fewer downward adjustments, or both. In order to disentangle the overall effect, we transform the dependent variable such that we obtain two binary variables for, in turn,

[9]There is a general issue when combining annual balance sheets with monthly survey data. In the baseline case, we always chose the most recently available data. Thus, the likelihood of the balance sheet information no longer reflecting the current situation increases the more time has passed between the balance sheet closing date and the survey month. We therefore double-check the robustness of our results by distributing the balance sheet data differently over the year, as discussed in Appendix D.2.

Table 5: Average treatment effect on the treated, increases and decreases in production and prices

	Dependent variable: change in production/prices			
	(1) Prod. increase	(2) Prod. decrease	(3) Price increase	(4) Price decrease
Panel (a): Expected production increase				
ATT	0.149***	-0.022***	0.018***	-0.008***
	(36.93)	(-6.86)	(5.35)	(-3.20)
Observations	129812	129812	129858	129858
Panel (b): Expected production decrease				
ATT	-0.024***	0.149***	-0.005	0.025***
	(-7.11)	(31.35)	(-1.63)	(7.23)
Observations	125458	125458	125530	125530

Notes: Table shows treatment effects for binarized production and price indicators, i.e., separately considering increases and decreases. Treatment is the expectation of future production. Outcomes refer to the current month, t, while expectations refer to the next 3 months ($t+1$ to $t+3$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

production and prices. This is a frequently used approach when dealing with survey data. We then compute the probability of treated firms to raise (lower) prices or production as the mean difference in the newly defined variable across treated and non-treated firms.

Table 5 reports the results for the baseline specification. Columns (1) and (2) show that firms which expect production to increase stand out by their increased probability of raising production in the same month. Specifically, the probability of a production increase is 14.9 percentage points higher for those firms compared to untreated firms. This accounts for the bulk of the overall effect discussed above. The probability of a production decrease, in turn, is reduced by 2.2 percentage points. Firms which expect production to go up are also more likely to raises prices by 1.8 percentage points and less likely to lower them by 0.8 percentage points, see columns (3) and (4) of Table 5.

Likewise, a treatment with expectations of a production decline increases the probability of a cut of current production by 14.9 percentage points, while the probability of a production increase falls by 2.4 percentage points. The response of prices to expectations of a production decline is somewhat larger than the one to expectations of a production increase. The probability of a price decline increases by 2.5 percentage points. The probability of a price increase is not affected significantly.

4.5 Additional evidence based on quantitative responses

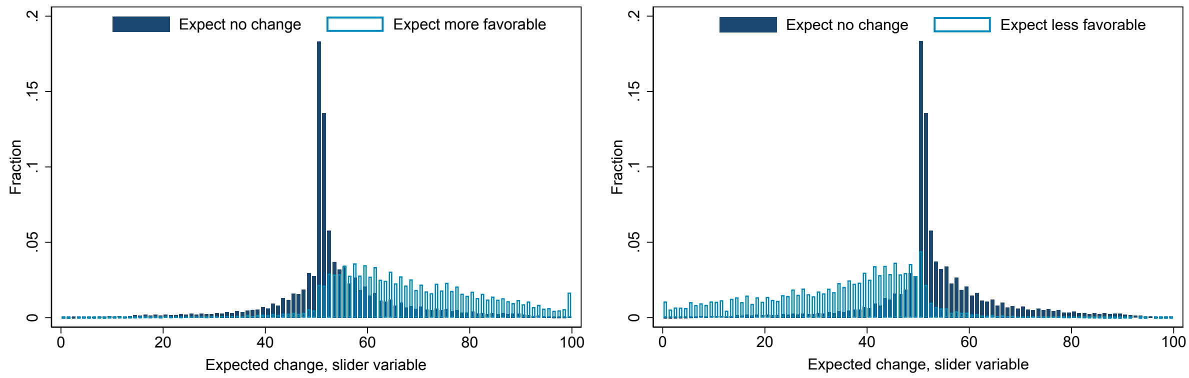
Our results so far are based on production expectations, measured qualitatively on a 3-point answer scale (decrease/no change/increase). We now take up the concern that the information content in our variable may be limited because of its qualitative nature. For this purpose, we exploit the fact that since August 2005 the ifo survey also asks for a quantitative response when it comes to the "expected state of business 6-months ahead" (but not for our baseline question regarding production expectations). In addition to a qualitative response regarding the state of business (see Q2 in Table 1) respondents use a graphical tool: they place a slider on a horizontal line which represents a scale that ranges from 0 to 100, with 0 corresponding to "rather less favorable", 50 corresponding to "stay about the same", and 100 corresponding to "rather more favorable". The endpoints and the middle of this scale are indicated by markers on the line.

Figure 5 shows the distribution of firm responses over the entire range of possible answers. In panel (a), we contrast the distribution of quantitative answers by firms that expect the state of business to become more favorable according to the 3-point answer scale (transparent bars) and the answers by firms which expect no change on that scale (solid bars). In panel (b), we repeat this exercise for firms which report a less favorable outlook. A consistent picture emerges across both panels. The mass of firms which report that they expect no change on the 3-point answer scale is centered around 50. Instead, the mass of firms which expect more or less favorable conditions is concentrated to the right or left of that value, respectively. This suggests that the qualitative responses provide a comprehensive and meaningful summary statistic of firms' quantitative expectations. Still, in both instances firm responses which suggest a change are distributed fairly equally to either the right or the left of those that suggest no change. This, in turn, implies that we lump together a wide range of expectations when working with the 3-point answer scale in our baseline specification.

Against this background, we further explore the importance of quantitative differences in firm responses and consider two alternative specifications for which we restrict our assessment to a subset of firms based on their quantitative responses. This allows us to focus on the effect of small and large expected changes, respectively. First, we limit our sample to those firms which report an expected change on the 3-point answer scale, but are close to reporting no change on the quantitative scale ("small-changes specification"). Specifically, for firms expecting more favorable conditions, we only include those with a slider value in the first quartile of all slider responses

Figure 5: Distribution of expectations about state of business

(a) No change vs. more favorable　　**(b)** No change vs. less favorable



Notes: Horizontal axes represent possible answers to question about the expected state of business where responses range from 0 to 100, with 0 corresponding to "rather less favorable", 50 corresponding to "stay about the same", and 100 corresponding to "rather more favorable". Bars show fraction of firms which tick a specific value. Panel (a)/(b): firms which respond "rather more/less favorable" in response to the qualitative question (transparent) vs. those which respond "no change" (solid), see Q2 in Table B.1.

in that group (value below 58). Accordingly, for those expecting less favorable conditions, we consider only those in the top quartile of that group (value above 45). We restrict the reference group of firms that report no change to those firms which report a quantitative response between the median of that group (51) and the limit for the respective treated group. In each instance, we thus constrain the estimation to match firms which are in a similar range on the slider, yet report qualitatively different expectations.

Second, to assess the effect of large expected changes, we restrict the sample to those firms which report a change on the 3-point scale and are also at the far end on the quantitative scale. We label this case the "large-changes" specification. In this case, we restrict the sample of treated firms to the first or to the top quartile in terms of the quantitative responses. Here we only keep treated firms with response values below 25 or above 76, for firms which expect less or more favorable conditions, respectively. At the same time, we do not restrict the sample of firms without treatment.

We report results in Table 6. First, in column (1) we simply estimate our baseline specification using the *expected state of business* (as reported on 3-point answer scale) rather than *production expectations* as the treatment variable, since there are no quantitative responses for production

Table 6: Average treatment effect on the treated, expected state of business

	(1) Full sample	(2) Sample w/ slider	(3) Small changes	(4) Large changes
Panel A: Expected more favorable – Effect on production				
ATT	0.080***	0.087***	0.046***	0.109***
	(13.42)	(10.30)	(3.39)	(7.51)
Observations	124196	54138	16075	66314
Panel B: Expected more favorable – Effect on prices				
ATT	0.018***	0.016***	0.015	0.013
	(4.07)	(2.58)	(1.48)	(1.19)
Observations	124262	54135	16061	66349
Panel C: Expected less favorable – Effect on production				
ATT	-0.083***	-0.098***	-0.041***	-0.137***
	(-12.85)	(-10.03)	(-2.97)	(-8.04)
Observations	121337	57924	19081	56723
Panel D: Expected less favorable – Effect on prices				
ATT	-0.027***	-0.024***	-0.011	-0.047***
	(-5.32)	(-3.13)	(-1.07)	(-3.33)
Observations	121442	57929	19089	56756

Notes: Table shows treatment effects on prices and production for different specifications. Treatment is the expectation of the future state of business. Outcomes refer to the current month, t, while expectations refer to the next 6 months ($t+1$ to $t+6$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For specification (3) and (4) sample is restricted based on quantitative responses, see main text for details.

expectations.[10] Not only does the concept differ, but also the time horizon: for the state of business it is six months, for production expectations three months. And yet, column (1) shows that results are in the same ballpark as those shown in Table 4 above. This holds both for the effect on production in the top panels and for the effect on prices in the bottom panels. Next, in column (2) we show results based on qualitative responses only, but for the sample for which quantitative responses are available. Again, we find results very similar to the baseline.

The results in column (2) provide the benchmark to assess the results for the restricted samples. In column (3), labeled "small changes", we report the ATT for firms that expect a change according to the 3-point answer scale but are still close to the untreated firms according to their quantitative response. In this way, we seek to measure the effect of a small change in expectations. In column (4), we report the ATT for large changes. And indeed, we find that the

[10] All results are based on new matching procedures based on the same specification as in the baseline, except that we replace the lags of the dependent variable accordingly. Balancing statistics are reported in Table C.6. The matching still works well on the restricted sample and, with the exception of three values, all standardized biases remain below 5%.

ATT is considerable weaker for small changes than the effect for the unrestricted sample (reported in column (2)): it is about half, but still significant for production. For prices, the effect ceases to be significant. For large changes the ATT is considerably stronger (with the exception of prices in case of an expected increase in the state of business). Taken together, these results lend additional support to our approach. While our baseline results are based on qualitative responses only, it seems that these responses are meaningful even if the quantitative difference between treated and untreated firms is small, see column (3). At the same time, comparing results in columns (3) and (4) shows that the size in the expected change also matters. A more systematic exploration of the quantitative differences in expectations is thus a promising venue for further research. In the present paper, however, we focus on the qualitative responses because only these are available for production expectations and a much longer time span.

5 News or noise?

In the previous section, we established that firm expectations affect decisions on current production and pricing at the firm level. This raises the question of why this is the case. As shown in Section 2, there are two distinct possibilities. According to the first, firms may have information about future developments that is unrelated to 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 a good reason to expect production to in- or decrease, even if their current fundamentals do not differ from their peers that expect no change—it is only that this reason is not yet observable to the econometrician. Instead, according to the second hypothesis, changes in expectations about the future that are fundamentally unwarranted may simply be "undue", in the language of Pigou (1927). Or, put differently, they are basically misperceptions about the future or "noise" (Lorenzoni 2009). Of course, our estimate of the ATT may also reflect a mixture of news and noise.

5.1 Effect on Production and Prices

In what follows, we seek to determine to what extent the expectations that impact current decisions about production and prices reflect news or noise. We do so on the basis of firms' expectation errors. Importantly, we are agnostic about whether expectations are rational or not from an

Table 7: Ex-post classification of expectations

Expectation	Realization	Classification
expected *increase* in t	realization in t+1 to t+3 > 0	correct
expected *increase* in t	realization in t+1 to t+3 ≤ 0	incorrect
expected *no change* in t	realization in t+1 to t+3 $> \frac{1}{3}$	incorrect
expected *no change* in t	$-\frac{1}{3} \geq$ realization in t+1 to t+3 $\leq \frac{1}{3}$	correct
expected *no change* in t	realization in t+1 to t+3 $< -\frac{1}{3}$	incorrect
expected *decrease* in t	realization in t+1 to t+3 ≥ 0	incorrect
expected *decrease* in t	realization in t+1 to t+3 < 0	correct

Notes: classification of expectations based on reported expectation and reported realization. The latter is the simple average of the responses of a firm to the question regarding realized production in periods *t+1*, *t+2* and *t+3* (Q3).

ex-ante point of view. We simply classify them as correct or incorrect based on actual outcomes. For instance, a firm may entertain expectations of a production increase and correctly so given all available information. Yet, actual production may still fall short of the expected level because of some other unforeseen development. Since expectations pertain to future production as such rather being conditional on a specific development, such firms make an expectation error under our classification scheme. In our terminology, the firm has, rationally or irrationally, responded to noise.

More specifically, as in Bachmann et al. (2013), we interpret the qualitative responses to questions about expected and realized production (Q1 and Q3, respectively) as pertaining to the same latent variable. We say that a firm expecting a change in production makes an expectation error whenever it reports an average realization in the following three months that differs in sign from the expected change. If a firm expects no change, we classify expectations as correct if the firm reports on average at most one change in either direction. Table 7 provides an overview of our classification scheme.

Based on this classification scheme, we define a treatment with a "correct expectation of a production increase" if a firm correctly expects production to increase (answer "1" to Q1). The control group are firms that correctly report that production will not change (answer "0" to Q1). The second treatment we consider is an "incorrect expectation of a production increase." Also in this case we use as a control group those firms that correctly report that production will not change. The third and fourth treatments are defined analogously but for firms that expect a production decline. Using these four new treatment indicators, we perform the same matching procedure as in Section 4.1. In this way, we make sure that we a) isolate the effect of expectations that cannot be fully accounted for by current fundamentals and b) distinguish the effect of these expectations depending on whether they turn out to be correct or incorrect from an ex-post point of view.

Before turning to the results, we again consider some diagnostic statistics to ensure that the matching works reasonably well. The statistics are the same as those for the baseline, described in Section 4.2. Panels (b) and (c) in Table 3 above report the number of observations for which a propensity score can be computed as well as the number of observations that can be matched. Even though the number of matches is now smaller than before, there is still common support (see Figure C.1). Also, for all four treatments balancing is achieved, no bias is above 5%, and the variance ratios are generally within the defined bounds with similar exceptions as before (see Figures C.2 and C.3 as well all as Table C.1).

We report the ATTs in Table 8. We focus again on how firms' current production and price-setting decisions depend on expectations of a production increase (upper part of the table) and on expectations of a production decrease (lower part of the table), but we now distinguish between correct and incorrect expectations. Panels (a) and (b) show the results for production and prices for correct expectations of a production increase, while panels (c) and (d) display the results for incorrect expectations of a production increase.

We present estimates for the baseline specification in column (1) and stress that results are, as before, robust across alternative specifications, reported in columns (2) to (6). We find that the effect of expectations on firms' current decisions is stronger if they are correct. Still, also for incorrect expectations, we find a significant effect, except for prices in case of incorrect expectations of a production decrease. As before, this effect may reflect a mixture of more upward or fewer downward adjustments compared to untreated firms. For expectations of a production decrease it may reflect more downward adjustments and fewer upward adjustments.

In order to shed some light on this point, we rely once more on the transformation of the dependent variables into two binary variables indicating increases and decreases, respectively. Table 9 shows the results. We find that the probability of a production increase is 27.6 percentage points higher for correct expectations of a production increase compared to untreated firms, and 7.2 percentage points higher for incorrect expectations. The probability of reducing production, instead, does not change much in both cases. A similar picture emerges for prices. By and large,

Table 8: Average treatment effect on the treated, correct and incorrect firms, production and prices

	(1) Baseline	(2) Radius 0.01	(3) Sample 2002-2016	(4) Sample excl. fin. crisis[1]	(5) Match in sector	(6) Response in first 10 days[2]
Panel A: Correctly expected production increase – Effect on production						
ATT	0.302***	0.298***	0.313***	0.297***	0.290***	0.331***
	(36.89)	(34.85)	(35.95)	(34.26)	(25.37)	(22.75)
Observ.	81254	68946	68597	71391	20644	18040
Panel B: Correctly expected production increase – Effect on prices						
ATT	0.035***	0.034***	0.037***	0.033***	0.034***	0.033***
	(5.40)	(5.18)	(5.24)	(4.90)	(4.03)	(2.83)
Observ.	81254	68945	68587	71392	20635	18044
Panel C: Incorrectly expected production increase – Effect on production						
ATT	0.063***	0.060***	0.075***	0.063***	0.082***	0.081***
	(8.58)	(7.94)	(9.55)	(8.13)	(8.42)	(5.90)
Observ.	84029	74232	69659	73973	26203	18716
Panel D: Incorrectly expected production increase – Effect on prices						
ATT	0.016***	0.015***	0.014**	0.011*	0.012	0.006
	(2.92)	(2.58)	(2.26)	(1.89)	(1.61)	(0.58)
Observ.	84032	74232	69656	73978	26205	18723
Panel E: Correctly expected production decrease – Effect on production						
ATT	-0.307***	-0.300***	-0.302***	-0.303***	-0.281***	-0.304***
	(-33.71)	(-30.52)	(-30.13)	(-32.00)	(-22.03)	(-17.14)
Observ.	80282	66948	66312	68156	18875	15243
Panel F: Correctly expected production decrease – Effect on prices						
ATT	-0.030***	-0.021**	-0.024***	-0.044***	-0.044***	-0.048***
	(-3.83)	(-2.52)	(-2.76)	(-5.66)	(-4.23)	(-3.08)
Observ.	80285	66941	66303	68158	18859	15250
Panel G: Incorrectly expected production decrease – Effect on production						
ATT	-0.086***	-0.093***	-0.077***	-0.086***	-0.116***	-0.075***
	(-9.99)	(-10.29)	(-8.34)	(-9.15)	(-10.12)	(-4.34)
Observ.	79026	68414	65304	68835	22376	16195
Panel H: Incorrectly expected production decrease – Effect on prices						
ATT	-0.003	-0.008	-0.003	-0.008	-0.019**	-0.004
	(-0.36)	(-1.07)	(-0.38)	(-1.08)	(-2.04)	(-0.32)
Observ.	79033	68420	65305	68842	22375	16209

Notes: Table shows treatment effects on prices and production for different specifications. Treatment is the expectation of future production, separately for firms which turn out to be correct and incorrect, respectively. Outcomes refer to the current month, t, while expectations refer to the next 3 months ($t+1$ to $t+3$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[1] Excluding the years 2008 and 2009.
[2] Results are based on online responses only, sample starts in 2004.

the effect of an expected production decrease mirrors that of an expected production increase. Estimates are shown in the bottom panel of Table 9. There is, however, some asymmetry in the

Table 9: Average treatment effect on the treated, correct and incorrect firms, increases and decreases in production and prices

	Dependent variable: change in production/prices			
	(1) Prod. increase	(2) Prod. decrease	(3) Price increase	(4) Price decrease
Panel (a): Correctly expected production increase				
ATT	0.276***	-0.026***	0.028***	-0.007*
	(44.90)	(-5.87)	(5.45)	(-1.93)
Observations	81254	81254	81254	81254
Panel (b): Incorrectly expected production increase				
ATT	0.072***	0.009**	0.012***	-0.004
	(13.82)	(2.11)	(2.84)	(-1.24)
Observations	84029	84029	84032	84032
Panel (c): Correctly expected production decrease				
ATT	-0.033***	0.274***	0.000	0.030***
	(-6.62)	(39.14)	(0.01)	(5.57)
Observations	80282	80282	80285	80285
Panel (d): Incorrectly expected production decrease				
ATT	-0.002	0.084***	0.004	0.007
	(-0.46)	(12.66)	(0.96)	(1.42)
Observations	79026	79026	79033	79033

Notes: Table shows treatment effects for binarized production and price indicators, i.e., separately considering increases and decreases. Treatment is the expectation of future production, separately for firms which turn out to be correct and incorrect, respectively. Outcomes refer to the current month, t, while expectations refer to the next 3 months ($t+1$ to $t+3$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

response to incorrect expectations of a production in- or decrease: firms that incorrectly expect a production decrease do not respond by adjusting prices in a significant way. The average effect is roughly zero and the effects on the binary outcome variables are close to zero and insignificant.

This observation lends support to the view that downward price rigidities prevent an adjustment of prices unless the need for adjustment is particularly strong. Arguably this is the case if expectations are correct but not if they are incorrect, perhaps because the expected decrease in production is more moderate or firms are less certain about it. In this case firms appear more responsive in terms of quantities instead of prices.

In sum, we find that while firm expectations matter for firm decisions, this holds not only for expectations that turn out to be correct ex post. It also holds for incorrect expectations. Hence, the role of expectations for today's decisions is not limited to the transmission channel of news. Our results show that expectations also have a noise component, that is, they cause firms to adjust prices and production even though there is no fundamental reason for firms to do so.

5.2 Further evidence

By now we have established that firms respond to expectations—both to correct and incorrect ones—by adjusting current prices and production. In what follows, we turn to additional variables for which we may also expect an effect in light of our main results. In each instance, we use the same framework as above, but report the ATT on variables other than production and prices.

Table 10: Average treatment effect on the treated, correct and incorrect firms, inventories and profits

	Correct expectations			Incorrect expectations		
	(1) Baseline	(2) Radius 0.01	(3) Sample excl. fin. crisis[1]	(4) Baseline	(5) Radius 0.01	(6) Sample excl. fin. crisis[1]
Panel (a): Expected production increase – Effect on inventories						
ATT	0.034***	0.030***	0.032***	0.027***	0.026***	0.027***
	(3.97)	(3.36)	(3.58)	(3.42)	(3.19)	(3.27)
Observ.	11150	9384	9543	11137	10188	9457
Panel (b): Expected production increase – Effect on profits +2						
ATT	0.141***	0.149***	0.136***	-0.070***	-0.074***	-0.076***
	(5.16)	(5.33)	(4.67)	(-2.91)	(-3.01)	(-2.96)
Observ.	11930	10661	10220	12149	11034	10410
Panel (c): Expected production increase – Effect on profits +3						
ATT	0.229***	0.240***	0.207***	-0.077***	-0.066***	-0.088***
	(8.49)	(8.79)	(7.21)	(-3.18)	(-2.63)	(-3.38)
Observ.	11403	9931	9712	11680	10562	9940
Panel (d): Expected production decrease – Effect on inventories						
ATT	-0.067***	-0.063***	-0.072***	-0.018*	-0.026**	-0.022**
	(-5.98)	(-5.25)	(-6.34)	(-1.78)	(-2.50)	(-2.05)
Observ.	10586	9353	8768	10200	8865	8526
Panel (e): Expected production decrease – Effect on profits +2						
ATT	-0.122***	-0.141***	-0.148***	0.068**	0.032	0.088**
	(-3.44)	(-3.83)	(-4.16)	(2.24)	(1.01)	(2.56)
Observ.	11403	9756	9426	11267	10137	9491
Panel (f): Expected production decrease – Effect on profits +3						
ATT	-0.240***	-0.201***	-0.290***	0.047	0.035	-0.004
	(-6.59)	(-5.02)	(-7.57)	(1.52)	(1.09)	(-0.13)
Observ.	10356	8256	8295	10441	9087	8743

Notes: Tables shows treatment effects on inventories and demand for different specifications. Treatment is the expectation of future production, separately for firms which turn out to be correct and incorrect, respectively. Outcomes refer to the current month, *t*, while expectations refer to the next 3 months (*t+1* to *t+3*). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
[1] Excluding the years 2008 and 2009.

In particular, we conjecture that firms that expect production to increase (decrease) consider current inventories as insufficiently low (too high), independently of whether expectations turn out to be correct or not. In the survey, firms can evaluate the current state of their inventories as "too small" [1], "sufficient" [0], or "too large" [-1], see Table B.1 in the appendix. Panels (a) and (d) of Table 10 show how expectations affect this assessment. As it turns out, both correct and incorrect expectations of a production increase induce firms to evaluate their inventories as too low (panel (a)). The opposite is true for expectations of a production decrease, either correct or incorrect (panel (d)).

Furthermore, taking decisions based on expectations that turn out to be correct or incorrect should also have a bearing on future profits, in addition to realized fundamentals. Panels (b) and (e) of Table 10 report the effect of expectations on profits in the second month and panels (c) and (f) in the third month after impact.[11] We find that profits increase in both months in case of correct expectations of a production increase. Given that firms anticipate higher production and act accordingly, this result appears plausible.[12] Likewise, we find that profits decline in case of correct expectations of a production decrease (panel (e)). However, so do profits in case of incorrect expectations of a production increase—suggesting that acting based on incorrect expectations can be costly. Note, however, that according to our estimates, incorrect expectations of a production decrease do not seem to lower profits significantly. This result appears to be consistent with our previous finding, according to which firms that incorrectly expect a production decrease do not adjust prices.

6 Noise and aggregate fluctuations

Up to now, we have focused on individual firms and, more specifically, we have documented that incorrect expectations cause firms to adjust prices and production. In what follows, we investigate whether noise at the firm level matters for aggregate outcomes. Intuitively, if a sufficiently large number of firms or a number of sufficiently large firms maintains and responds to incorrect expectations, economic activity at the aggregate level may change as a result.

[11] In the survey, the question about profits is asked twice a year, in May and in September. Our results are therefore based on different sets of firms. For results pertaining to profits two months after impact, we rely on expectations data from March and July. For profits three months after impact, we use responses from February and June. For details on the question see Table B.1.

[12] Profits in the first month after the price and production changes are less responsive to these changes.

Table 11: Correlation of expectation errors

Sector	Correlation with same sector	Correlation with all firms	Sector	Correlation with same sector	Correlation with all firms
All sectors	0.1967	0.1310	Rubber&plastic prod.	0.1902	0.1513
Food	0.1558	0.0383	Glass prod.	0.1889	0.1266
Beverages	0.2669	0.0186	Basic metals	0.2735	0.1977
Tobacco	0.6281	-0.0207	Fabricated metal prod.	0.1646	0.1465
Textiles	0.1985	0.1018	Computer&electronic prod.	0.1700	0.1339
Wearing apparel	0.2185	0.0397	Electrical equipment	0.1775	0.1460
Leather&related prod.	0.2965	0.0893	General-purpose machinery	0.1568	0.1333
Wood&cork products	0.2161	0.1361	Motor vehicles&trailers	0.2592	0.1966
Paper products	0.2130	0.1687	Other transport equi.	0.3299	0.1413
Printing	0.1731	0.0989	Furniture	0.2245	0.1081
Coke&refined petrol.	0.4659	0.0865	Other manufacturing	0.1969	0.1060
Chemical products	0.2126	0.1697	Repair&installation	0.3821	0.0881
Pharmaceuticals	0.3073	-0.0134			

Notes: Correlation of firms' individual forecast error with average of the forecast error in the same 2-digit WZ08 sector and the whole economy, shown separately for each 2-digit sector. Error computed following the approach of Bachmann et al. (2013): the error is 0 if the firm is correct, that is, if the sign of the expectation and the average realization is the same. If the firm is incorrect, the error equals the difference between the sum of realized production in *t+1* to *t+3* and the expectation in *t*, divided by 3.

Against this background, we first assess the degree to which the expectation errors of firms are correlated, both within sectors and across the entire economy. For this purpose, we now not only classify expectations as correct or incorrect, but also quantify the extent to which they are incorrect, following the approach of Bachmann et al. (2013).[13] We report descriptive statistics and the serial correlation pattern of expectation errors in Tables A.4 and A.5 in the appendix. Table 11 shows that errors are generally positively correlated within sectors and more strongly so than across all firms. While expectation errors within a sector may be caused by (sector-specific) fundamentals, we are interested in identifying the economy-wide effect of expectations *as such*, that is, changes in expectations that are not caused by fundamentals, neither from an ex-ante nor from an ex-post point of view.

[13]Specifically, the error is 0 if the expectation is correct, that is, if the sign of the expectation and the average realization is the same. If the expectation is incorrect, the error equals the difference between the sum of realized production in periods $t+1$ to $t+3$ and the expectation in t, divided by 3.

6.1 An aggregate measure of noise

For this purpose, we develop an aggregate measure of noise that builds on our analysis of firm-level expectations above. Importantly, to make sure our aggregate expectation error captures noise, we control for current fundamentals and consider ex-post outcomes. First, in order to account for current fundamentals, we rely on the ordered probit model introduced in Section 4. Now, however, rather than matching firms based on their propensity score, we compute the difference between a firm's response and the prediction of the ordered probit model, given in equation (4.3) above.[14] We select those firms which expect an increase or a decrease of production even though the model suggests otherwise. In this way, we capture the extent to which expectations are not accounted for by current fundamentals.[15] Second, among those firms, we only consider firms for which expectations turn out to be incorrect from an ex-post perspective as defined in Section 5 above. Our measure of the aggregate expectation error is then given by the share of firms that fulfill these two criteria relative to all firms in a given month.

In computing the aggregate error, we consider three alternative weights. First, we compute the share while giving equal weight to each firm. For the second measure, we use the number of employees as weights. We drop the largest 5 percent of our observations to ensure that results are not driven by individual firms. Finally, we weigh firms using the approach of the ifo institute for aggregating answers to the business climate index (Sauer and Wohlrabe 2018). This approach weighs all firms within a 2-digit WZ08 sector (the German system of industry classification) using the number of employees in production as reported in the survey. Instead of using the number of employees directly, however, the weight is a logarithmic transformation of employment.[16] The sector averages are then aggregated using data on gross value-added by sector from the German Statistical Office.

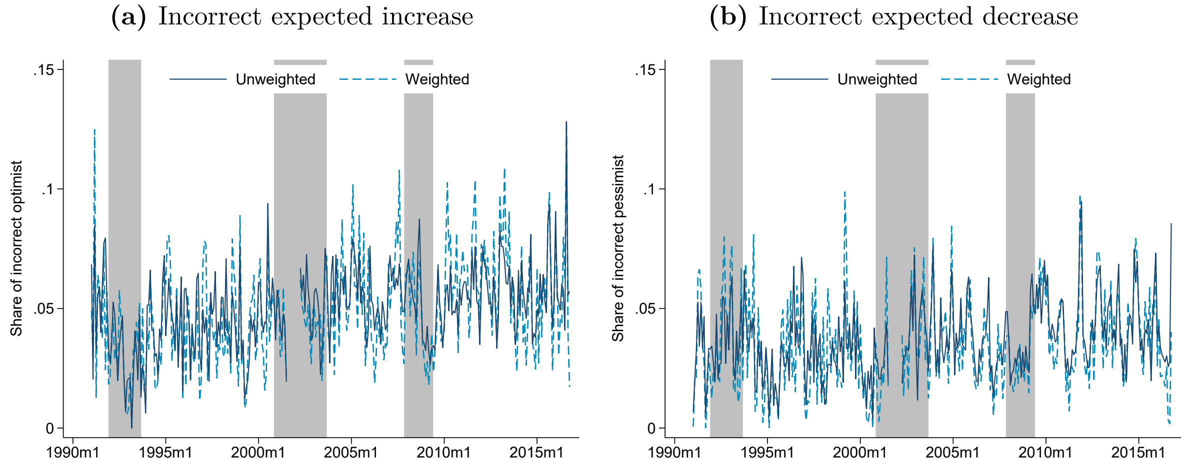
Figure 6 displays the unweighted and the employee-weighted time series for the aggregate expectation error (using ifo weights results in a very similar time series). For firms to enter this

[14]We use the ordered probit because we seek 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).

[15]The predicted response is the response to which the ordered probit model assigns the highest probability.

[16]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.

Figure 6: An aggregate measure of noise, German manufacturing sector 1991–2016



Notes: Aggregate time series for incorrect expectations, unweighted and weighted by employees. Shaded areas indicate recession periods as defined by the German Council of Economic Experts.

measure, we require them to be in the survey for at least eight consecutive months.[17] 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 four months of 2016. In our analysis below, we consider only those time periods for which we have all observations, including the required lags. The main takeaway of Figure 6 is that there is considerable variation of incorrect expectations over time, even at the aggregate level—expectation errors do not wash out. In addition, we note that the time series exhibit little persistence.

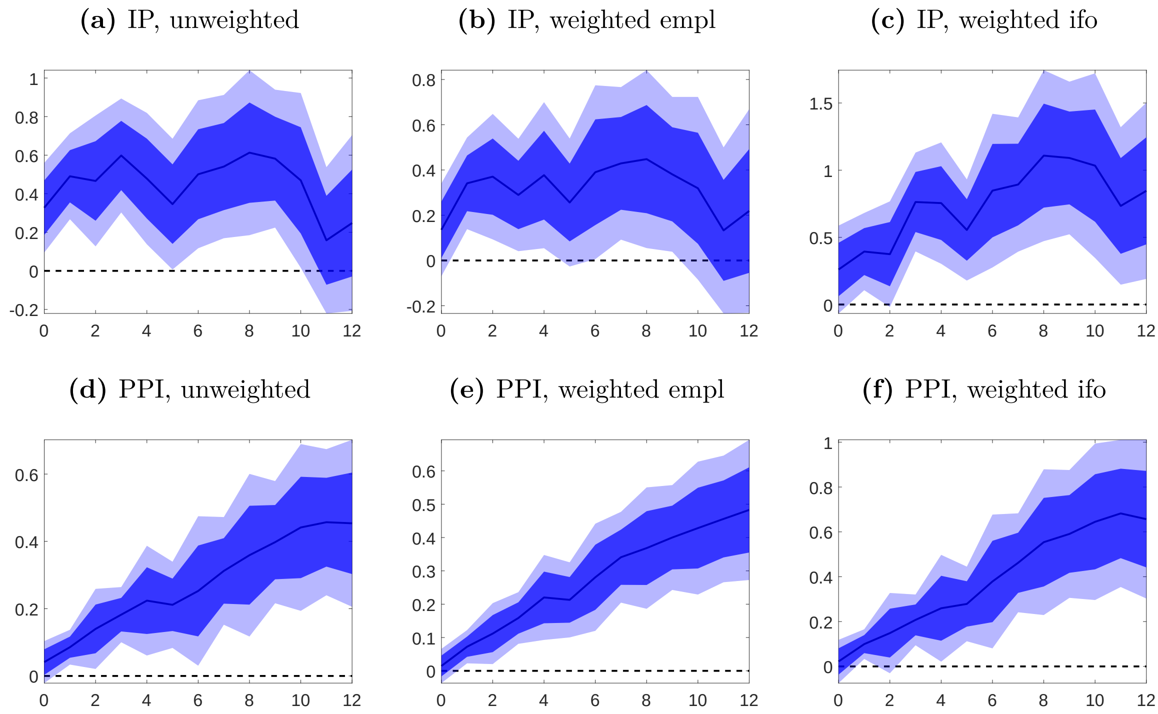
6.2 The effect of aggregate expectation errors

Our noise measure is an aggregate of firms' expectation errors. We compute this measure while controlling for current fundamentals, conditional on ex-post outcomes. In addition, we allow for time-fixed effects as we model expectations at the firm level. As a result, aggregate expectations are unlikely to be caused by macroeconomic shocks. In what follows, we assess to what extent macroeconomic variables are driven by the aggregate expectation error.

For this purpose, we rely on local projections. Formally, using e_t^i and e_t^d to denote the time-series observations for incorrect expectations of a production increase and a production decrease, respectively, and letting x_t denote a macroeconomic variable of interest, we estimate the following

[17]We need three lags for the estimation of the ordered probit model and four leads for the computation of the expectation error. Since production is reported only for the previous month, we need four leads of the survey.

Figure 7: Effects of a noise shock, expected production increase



Notes: Responses of industrial production (IP) and producer price index (PPI), both in manufacturing, to incorrect expectations of production increase (one standard deviation shock). Local projections with constant, linear trend, one lag of dependent variable and 12 of the shocks. Shaded areas indicate 68 and 90 percent confidence intervals. IP data from the German Statistical Office, PPI data from the German Bundesbank.

model:

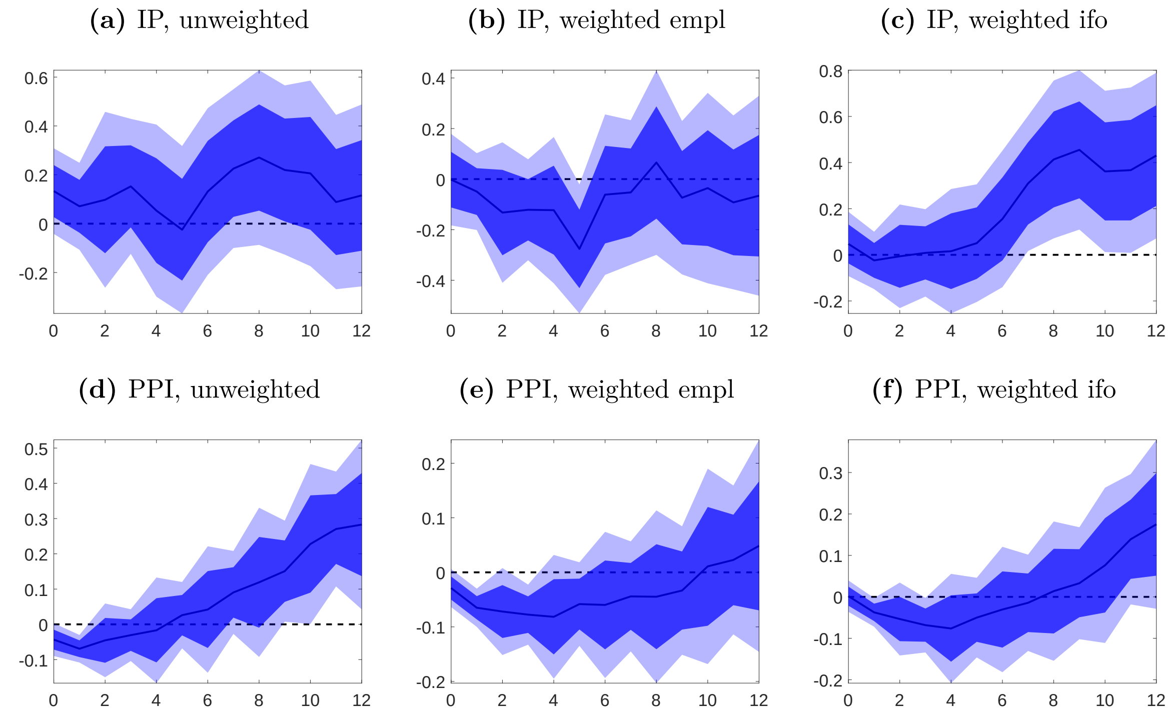
$$x_{t+h} = c^{(h)} + \sum_{j=1}^{J} \alpha_j^{(h)} x_{t-j} + \sum_{k=0}^{K-1} \beta_k^{(h)} e^i_{t-k} + \sum_{k=0}^{K-1} \gamma_k^{(h)} e^d_{t-k} + \varepsilon_{t+h}, \tag{6.1}$$

where $c^{(h)}$ is a (horizon-specific) constant.[18] In addition, we include a linear time trend. To enhance efficiency, we also include the residuals of the previous horizon when increasing the horizon by steps of one (Jordà 2005). For the estimation, we include 1 lag of the dependent variable and 12 lags of both expectations errors. That is, we include incorrect expectations of both a production increase and a production decrease 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 shock in period t.

We show the effect of a noise shock in Figure 7. It displays the response to an increase of one standard deviation in the share of incorrect expectations of a production increase. The top panels show the response of industrial production in the manufacturing sector (IP), measured

[18] In order to account for non-stationarity, the dependent variable can be expressed relative to its pre-shock level (Stock and Watson 2018). Including lags of the dependent variable, as we do above, generally yields the same result.

Figure 8: Effects of a noise shock, expected production decrease



Notes: Responses of industrial production (IP) and producer price index (PPI), both in manufacturing, to incorrect expectations of production decrease (one standard deviation shock). Local projections with constant, linear trend, one lag of dependent variable and 12 of the shocks. Shaded areas indicate 68 and 90 percent confidence intervals. IP data from the German Statistical Office, PPI data from the German Bundesbank.

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, shaded areas indicate 68 and 90 percent confidence intervals. We find that industrial production responds strongly and significantly to the expectation error. The observed increase is temporary and becomes insignificant after approximately one year, except when we use ifo weights. Consistent with the results at the firm level, reported in Section 5 above, we also find a strong and significant increase in the price level in response to the aggregate expectation error regarding an increase of future production. This reaction is in line with the interpretation of noise shocks as a specific form of demand shocks (Enders et al. 2020; Lorenzoni 2009).

Table 12: Forecast error variance decomposition (one year horizon)

	Variable	Unweighted	Empl. weights	ifo weights
Expected increase	IP	15%	9.5%	19%
	PPI	20%	22%	22%
Expected decrease	IP	2.5%	1.3%	7.2%
	PPI	7.3%	1.2%	2.3%

Figure 8 displays the results for incorrect expectations of a production decrease. Here we find much weaker effects. Moreover, a direct comparison with Figure 7 suggests that positive and negative noise shocks affect the aggregate economy asymmetrically. Specifically, in response to an adverse noise shock industrial production does not respond significantly and the producer price index declines only marginally in the first month after the shock. This result is consistent with our findings for firm-level data which show that incorrect expectations of a production decrease do not seem to cause a significant downward adjustment of prices, while incorrect expectations of a production increase cause prices to increase (Table 8).

And indeed, downward nominal (wage) rigidity may go some way in accounting for the asymmetry in the response of the economy to positive and negative shocks documented elsewhere (Born et al. 2021a; Dupraz et al. 2021). However, in these accounts negative shocks tend to impact economic activity adversely—in contrast to what we observe in Figure 8. Against this background, it seems noteworthy that an adverse noise shock reflects the fact that a sizeable fraction of firm expects a decrease of production that does not materialize—either because things stay the same or because things actually improve. It is thus conceivable that an adverse shock to expectations merely dampens an upswing which is about to set in. The pattern observed in Figure 8 is consistent with this interpretation, not only in terms of industrial production but also in terms of the adjustment of prices over time, notably in panels (d) and (f).[19]

Finally, Table 12 displays a forecast error variance decomposition for a horizon of 12 months, using the methodology of Gorodnichenko and Lee (2020). We report results for incorrect expectations of a production increase and a production decrease for all three measures of noise. We find that incorrect expectations of a production increase account for 10-19% of aggregate fluctuations of industrial production at a one-year horizon. At the same horizon, around 20% of the PPI

[19]In case we weigh firms' expectation errors by employees, production remains low and the tendency of prices to increase over time is much weaker (panels (b) and (e)). This may reflect better forecasting abilities of large firms (Born et al. 2021b).

is driven by incorrect expectations of a production increase. These are sizeable contributions. Incorrect expectations of a production decrease, on the other hand, have much a smaller effect on IP and the PPI. The specification for the PPI with unweighted observations delivers a value of around 7%, similar to that with ifo weights for IP. Regarding the small impact on prices, these results are in line with those of firm-level effects, reported in Section 5 above.

7 Conclusion

In this paper, we ask to what extent firm expectations matter for economic activity. From a theoretical point of view, expectations should matter a great deal. To date, however, there is little direct evidence to support the theory. We aim at filling this gap by applying a new identification strategy to a particularly suited data set. 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, which differ in their expectations, but not in their fundamentals. To identify the effect of expectations on firm decisions in a given month, we rely on the fact that firms report expectations early in the month.

We find that expectations of a production increase induce firms to raise prices and production. This result can be explained in two ways. According to the "news view," firms simply have additional information about future developments that is not reflected in current fundamentals, but justified by future fundamentals. Instead, according to the "noise view," firms entertain certain expectations for no fundamental reason. They simply have wrong ideas about the future and this accounts—in part—for their actions today. In light of these considerations, we classify expectations as correct and incorrect based on actual outcomes and match firms with incorrect observations to neutral firms. In line with the noise view, we find that incorrect expectations impact current production and price-setting decisions, although the effect is considerably smaller than in case of correct expectations. In a last step, we show that firms' expectations errors contribute to aggregate fluctuations as well.

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 Peter Zorn (2020). "What drives aggregate investment? evidence from german survey data". *Journal of Economic Dynamics and Control 115*, 103873.

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

Barsky, Robert B. and Eric R. Sims (2012). "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence". *American Economic Review 52* (4), 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 Economic Research - The 'Ifo Business Survey' Usage and Access". *Schmollers Jahrbuch-Zeitschrift für Wirtschafts- und Sozialwissenschaften 182* (2), 307–319.

Blanchard, Olivier, Jean-Paul L'Huillier, and Guido Lorenzoni (2013). "News, Noise, and Fluctuations: An empirical investigation". *American Economic Review 103* (7), 3045–3070.

Boneva, Lena, James Cloyne, Martin Weale, and Tomasz Wieladek (2020). "Firms' Price, Cost and Activity Expectations: Evidence from Micro Data". *The Economic Journal 130* (627), 555–586.

Born, Benjamin, Francesco D'Ascanio, Gernot J. Müller, and Johannes Pfeifer (2021a). "Mr. Keynes meets the Classics: Government Spending and the Real Exchange Rate". mimeo.

Born, Benjamin, Zeno Enders, Gernot J. Müller, and Knut Niemann (2021b). "Firm expectations about production and prices: Facts, determinants, and effects". mimeo.

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

Candia, Bernardo, Olivier Coibion, and Yuriy Gorodnichenko (2021). "The inflation expectations of u.s. firms: evidence from a new survey". CEPR Discussion Paper 16161.

Coibion, O., D. Georgarakos, Y. Gorodnichenko, and M. Weber (2020a). “Forward guidance and household expectations”. NBER Working Paper No. 26778.

Coibion, O., Y. Gorodnichenko, S. Kumar, and M. Pedemonte (2020b). “Inflation expectations as a policy tool?” *Journal of International Economics 124*, 103297.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020c). “Monetary policy communications and their effects on household inflation expectations”. NBER Working Paper No. 25482.

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 Saten Kumar (2018). “How Do Firms Form Their Expectations? New Survey Evidence”. *American Economic Review 108* (9), 2671–2713.

Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele (2020d). “Inflation Expectations and Firm Decisions: New Causal Evidence”. *The Quarterly Journal of Economics 135* (1), 165–219.

Conrad, Christian, Zeno Enders, and Alexander Glas (2021). “The role of information and experience for households’ inflation expectations”. Deutsche Bundesbank Discussion Paper No 07/2021.

D’Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber (2020). “IQ, Expectations, and Choice”. *Review of Economic Studies*, forthcoming.

D’Acunto, Francesco, Daniel Hoang, and Michael Weber (2021). “Managing households’ expectations with unconventional policies”. *Review of Financial Studies*, forthcoming.

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.

——— (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”. Federal Reserve Bank of New York Staff Reports No. 574.

Dupraz, Stéphane, Emi Nakamura, and Jón Steinsson (2021). “A Plucking Model of Business Cycles”. mimeo.

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, Franziska Hünnekes, and Gernot J. Müller (2019). "Monetary policy announcements and expectations: Evidence from German firms". *Journal of Monetary Economics 108*, 45–63.

Enders, Zeno, Michael Kleemann, and Gernot J. Müller (2020). "Growth Expectations, Undue Optimism, and Short-Run Fluctuations". Review of Economics and Statistics, forthcoming.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer (2015). "Expectations and Investment". *NBER Macroeconomics Annual 2015, Volume 30*. NBER Chapters. National Bureau of Economic Research, 379–431.

Gorodnichenko, Yuriy and Byoungchan Lee (2020). "A Note on Variance Decomposition with Local Projections". *Journal of Business and Economic Statistics*, forthcoming.

Görtz, Christoph and John D. Tsoukalas (2017). "News and Financial Intermediation in Aggregate Fluctuations". *The Review of Economics and Statistics 99* (3), 514–530.

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.

Hürtgen, Patrick (2014). "Consumer Misperceptions, Uncertain Fundamentals, and the Business Cycle". *Journal of Economic Dynamics and Control 40*, 279–292.

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

Jordà, 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, Michael (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, Michael, Ruth Miquel, and Conny Wunsch (2011). "Long-Run Effects of Public Sector Sponsored Training". *Journal of the European Economic Association 9* (4), 742–784.

Link, Sebastian, Andreas Peichl, Christopher Roth, and Johannes Wohlfart (2021). "Information frictions among firms and households". CESifo Working Paper No. 8969.

Lorenzoni, Guido (2009). “A Theory of Demand Shocks”. *American Economic Review 99* (5), 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 Christopher Pissarides (2009). “Job creation and job destruction in the theory of unemployment”. *The Review of Economic Studies 63* (2), 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, Donald 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.

——— (2019). “CEO or Intern – Who Actually Answers the Questionnaires in the ifo Business Survey?” *CESifo Forum 20* (2), 29–31.

Schmitt-Grohé, Stephanie and Martín Uribe (2012). “What’s news in business cycles?” *Econometrica 80* (6), 2733–2764.

Stock, James H. and Mark W. Watson (2018). “Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments”. *The Economic Journal 128* (610), 917–948.

Tanaka, Mari, Nicholas Bloom, Joel M. David, and Maiko Koga (2020). “Firm performance and macro forecast accuracy”. *Journal of Monetary Economics*, forthcoming.

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

A Additional descriptive statistics

Table A.1: Observations and average duration in panel

	Full sample	Sample with balance sheet data
Avg. duration in survey (months)	140.2	87.1
Avg. number of responses (months)	119.3	74.1
Response rate	82.5%	83.8%
Respondents	6625	4938
Respondents × months	620671	322839

Notes: Number of firms and duration of firms in the ifo survey. Separately for the full sample and the sample with balance sheet data.

Table A.2: Descriptive statistics for different samples

	Full sample			Sample with balance sheet data		
	Mean	Std. Dev.	Observ.	Mean	Std. Dev.	Observ.
Employees in production	505	3024	620411	579	3665	322823
Exp. production, t	0.00	0.55	607578	0.02	0.56	313788
Exp. state of business, t	-0.01	0.61	619148	0.01	0.61	321946
Production, t-1	-0.05	0.58	585075	-0.03	0.58	311828
Prices, t-1	0.00	0.42	597422	0.02	0.43	320617
Orders, t	-0.27	0.63	618229	-0.21	0.64	321226
Foreign orders, t	-0.28	0.57	616199	-0.22	0.58	319775
Capacity utilization, t	81.15	16.36	527754	81.10	16.13	279901
Demand, t-1	-0.02	0.64	598360	-0.00	0.65	321154
Inventories, t	-0.17	0.51	425005	-0.15	0.49	227742
State of profits, t	-0.12	0.69	51995	-0.11	0.69	42528
Change in profits, t	-0.06	0.71	51389	-0.05	0.71	42031
Increase in prices, t-1	0.09	0.28	597422	0.10	0.30	320617
Decrease in prices, t-1	0.09	0.28	597422	0.08	0.27	320617
Increase in production, t-1	0.15	0.35	585075	0.15	0.36	311828
Decrease in production, t-1	0.19	0.40	585075	0.18	0.39	311828

Notes: Descriptive statistics for the variables used in the micro-level analysis. Separately for the full sample and the sample with balance sheet data. For details on the related survey questions see Table B.1.

Table A.3: Sample attrition

Start date	Total	Fraction of firms surviving after										
		6m	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y
1991m1	1896	1	1	0.99	0.96	0.94	0.91	0.87	0.84	0.81	0.78	0.75
1992m1	2114	1	0.99	0.96	0.93	0.89	0.86	0.82	0.79	0.76	0.73	0.70
1993m1	2320	0.97	0.95	0.92	0.88	0.85	0.81	0.78	0.74	0.71	0.69	0.65
1994m1	2213	0.98	0.96	0.93	0.89	0.84	0.81	0.78	0.74	0.72	0.68	0.64
1995m1	2130	0.97	0.95	0.91	0.87	0.83	0.80	0.77	0.74	0.71	0.66	0.62
1996m1	2072	0.98	0.96	0.92	0.88	0.85	0.81	0.78	0.74	0.70	0.66	0.63
1997m1	2040	0.97	0.95	0.92	0.88	0.84	0.81	0.77	0.72	0.68	0.65	0.62
1998m1	1979	0.98	0.96	0.93	0.88	0.85	0.81	0.75	0.71	0.68	0.64	0.60
1999m1	2038	0.97	0.95	0.91	0.87	0.83	0.77	0.73	0.69	0.66	0.62	0.58
2000m1	2055	0.98	0.96	0.92	0.87	0.81	0.77	0.73	0.69	0.65	0.61	0.56
2001m1	2050	0.98	0.96	0.91	0.84	0.80	0.76	0.72	0.67	0.63	0.58	0.56
2002m1	2006	0.98	0.95	0.88	0.83	0.79	0.75	0.70	0.66	0.61	0.58	0.54
2003m1	1990	0.97	0.93	0.88	0.83	0.79	0.73	0.69	0.64	0.61	0.57	0.53
2004m1	2044	0.97	0.94	0.89	0.84	0.78	0.74	0.68	0.65	0.61	0.57	0.52
2005m1	1979	0.97	0.94	0.90	0.83	0.78	0.72	0.69	0.64	0.60	0.55	0.51
2006m1	1973	0.97	0.95	0.88	0.83	0.77	0.73	0.69	0.64	0.59	0.55	0.50
2007m1	1935	0.96	0.93	0.88	0.82	0.78	0.73	0.68	0.63	0.59	0.54	0
2008m1	1859	0.97	0.95	0.88	0.84	0.79	0.73	0.68	0.63	0.58	0	0
2009m1	1874	0.96	0.93	0.88	0.84	0.77	0.70	0.65	0.60	0	0	0
2010m1	1845	0.97	0.94	0.89	0.82	0.75	0.70	0.64	0	0	0	0
2011m1	1930	0.97	0.94	0.87	0.78	0.73	0.68	0	0	0	0	0
2012m1	2228	0.96	0.92	0.83	0.76	0.70	0	0	0	0	0	0
2013m1	2105	0.94	0.91	0.84	0.77	0	0	0	0	0	0	0
2014m1	1976	0.96	0.92	0.86	0	0	0	0	0	0	0	0
2015m1	1881	0.96	0.92	0	0	0	0	0	0	0	0	0

Notes: Table shows share of firms from initial period given by the row still in the sample after the time specified in the column; e.g., the last row shows that of the 1881 firms which were part of the sample in January 2015, 96% were still there after 6 months, 93% after 1 year, and 84% after two years. The zeros in the bottom right corner are due to our sample ending in December 2016. '6m' = 6 months, '1y' = 1 year.

Table A.4: Expectation error in the full sample

	Full sample				Excl. recession months			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
Expectation error, t	-0.038	0.362	-1.33	1.33	-0.022	0.358	-1.33	1.33
Exp. production, t	0.002	0.551	-1	1	0.028	0.545	-1	1
Realized prod., t+1 to t+3	-0.152	1.297	-3	3	-0.053	1.273	-3	3

Notes: Expected production based on response to Q1 and realized production based on Q3, see Table B.1 for wording of question and coding of answers. Expectation error computed following the approach of Bachmann et al. (2013): the error is 0 if the firm is correct, that is, if the sign of the expectation and the average realization is the same. If the firm is incorrect, the error equals the difference between the sum of realized production in *t+1* to *t+3* and the expectation in *t*, divided by 3. Right panel excludes months in which Germany was in a recession according to the German Council of Economic Experts.

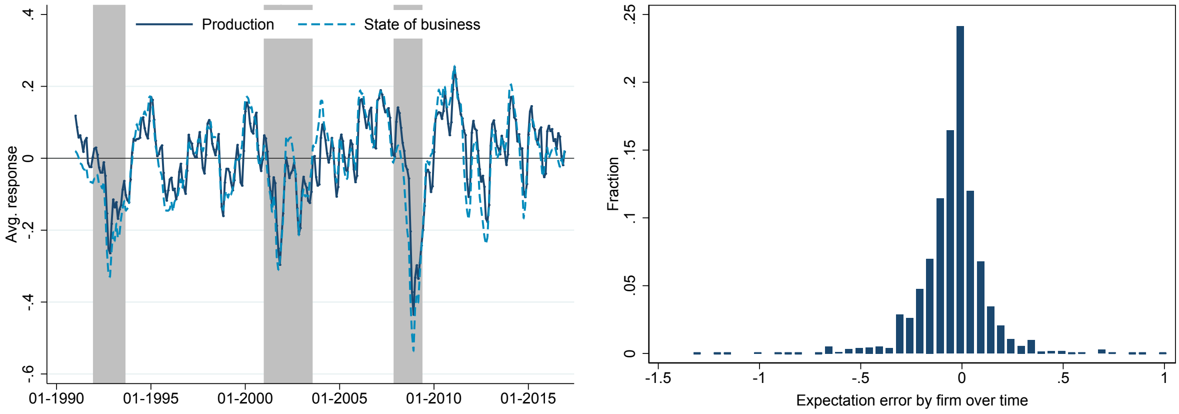
Table A.5: Serial correlation of the expectation error in full sample

	Correlation with forecast error in									
	t-1	t-2	t-3	t-4	t-5	t-6	t-7	t-8	t-9	t-10
Expectation error, t	0.580	0.341	0.152	0.128	0.120	0.109	0.095	0.087	0.082	0.095

Notes: Table shows correlation of forecast errors over time at the firm level. Expectation error computed following the approach of Bachmann et al. (2013): the error is 0 if the firm is correct, that is, if the sign of the expectation and the average realization is the same. If the firm is incorrect, the error equals the difference between the sum of realized production in *t+1* to *t+3* and the expectation in *t*, divided by 3.

Figure A.1: Average expectations

(a) Production and state of business **(b)** Distribution of expectation errors across firms



Notes: in panel (a), the shaded areas indicate recession periods as defined by the German Council of Economic Experts. Average response defined as share of positive responses (*increase/improve*) minus share of negative responses (*decrease/worsen*). Survey data from the BEP. Industrial Production from the German Statistical Office. Panel (b) shows the distribution of the average expectation error of firms over time.

B Details on survey questions

Table B.1: Complete list of survey questions used in our analysis

Label	Name	Question	Possible answers
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, of which x persons are for producing product XY.	x is the number of persons employed for XY
Q6	orders	We consider our order backlog for XY to be...	relatively high [1] sufficient [0] too small [-1] usually no backlog of orders [4]
Q7	foreign orders	We consider our order backlog for exports of XY 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]
Q10	inventories	Current situation: We evaluate our current stock of unsold finished goods of XY to be ...	too small [1] sufficient [0] too large [-1] no stock keeping [4]
Q11	change in profits	Profit situation and development: As compared to fall last year/the first quarter this year[1] the profit situation of our company measured by the operating results from customary business operations ...	improved [1] did not change [0] deteriorated [-1]

Notes: Authors' translation of the most recent formulation of the question in German according to the EBDC Questionnaire manual. We only show the answer possibilities that we consider. Specifically, we exclude "no production" or similar answers, which indicate that the question does not apply to the firm. [1]Questions asked biannually in May and September. In May question refers to "fall last year", in September it refers to "first quarter this year".

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

Label	Time period	Question
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.[1]
	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.

Notes: Authors' translation of the question in German according to the EBDC Questionnaire manual. **Bold font** highlights components which change from the initial formulation or drop out. *Italic font* highlights components which are added later on.
[1]In several months in 1980 the question was split into two parts, one covering regular and another additional orders.

C Balancing statistics

Table C.1: Standardized bias, baseline

	Expected prod. increase			Expected prod. decrease		
Variable	All	Correct	Incorrect	All	Correct	Incorrect
Exp.prod., *t-1*	1.8	1.6	1.8	-0.7	-0.7	-0.9
Exp.prod., *t-2*	0.2	-0.2	-0.3	-1.7	-3.5	-2.5
Exp.prod., *t-3*	0.8	0.7	0.1	-0.7	-2.8	-0.5
Production, *t-1*	1.3	0.1	-0.6	0.2	-0.3	-0.6
Production, *t-2*	1.1	0.3	-0.9	0.3	1.1	0.8
Production, *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

Notes: Table shows standardized bias which captures the difference in means between the treated and untreated groups. The standardized bias is the mean difference of each variable in the treated and untreated groups relative to the variances, details can be found in equation (4.6) in Section 4.

Table C.2: Standardized bias, radius r=0.01

	Expected prod. increase			Expected prod. decrease		
Variable	All	Correct	Incorrect	All	Correct	Incorrect
Exp.prod., *t-1*	1.3	2.4	1.9	-0.1	-1.1	-0.1
Exp.prod., *t-2*	0.1	-0.9	-0.3	-1.8	-2.8	-2.2
Exp.prod., *t-3*	0.1	0.8	0.5	-0.2	-2.6	0.6
Production, *t-1*	0.8	0.1	-1.6	0.3	0.1	-0.4
Production, *t-2*	0.8	-0.7	-0.3	0.9	2.3	2.4
Production, *t-3*	-1.0	-2.2	-0.8	-0.4	0.4	3.4
Prices, *t-1*	0.6	1.4	0.0	-0.2	3.6	-0.5
Prices, *t-2*	1.3	-0.6	2.0	0.8	4.7	1.0
Prices, *t-3*	0.7	1.2	0.1	-0.1	3.5	-0.5
Demand, *t-1*	0.9	1.6	-0.5	-1.3	-1.1	-1.6
Demand, *t-2*	0.9	0.8	-1.2	-0.3	-1.5	3.0
Demand, *t-3*	-0.4	-0.7	-0.5	-1.0	-1.6	3.5
Capacity, *t-1*	0.0	-0.7	1.2	2.8	4.3	1.4
Capacity, *t-2*	0.1	-0.6	1.5	2.3	3.7	1.1
Capacity, *t-3*	0.6	-0.6	0.8	1.7	3.8	0.5
Employees, *t*	0.7	-0.0	0.6	2.0	2.0	-0.3
Avg. state of business, sector, *t*	0.6	1.0	0.0	-0.7	-0.5	-1.0
State of business, *t*	0.8	-1.6	-0.8	1.9	3.6	1.5
State of business, *t-1*	0.5	-1.0	0.4	2.1	2.7	1.2
State of business, *t-2*	-0.0	-0.4	-0.2	1.4	1.7	1.9
State of business, *t-3*	-0.2	-1.4	0.7	0.7	0.4	1.0
Orders, *t*	0.7	-2.1	-0.8	2.0	3.6	0.5
Orders, *t-1*	1.5	0.5	-0.2	1.3	0.9	0.5
Orders, *t-2*	0.5	0.6	0.6	0.6	-0.1	-0.7
Orders, *t-3*	0.9	-0.2	1.1	1.8	0.5	-0.5
Foreign orders, *t*	1.3	-1.2	-0.9	0.5	-1.1	0.3
Foreign orders, *t-1*	1.6	-0.2	-0.4	0.3	-0.7	1.2
Foreign orders, *t-2*	0.9	-0.6	-0.9	0.0	-0.7	-1.4
Foreign orders, *t-3*	1.0	-0.4	0.1	1.1	-1.0	-0.9
Debt share, *t*	-1.7	-0.9	-1.2	0.1	-5.2	-2.2
Financing coefficient, *t*	-1.3	-0.6	-1.2	1.5	1.5	0.7

Notes: Table shows standardized bias which captures the difference in means between the treated and untreated groups. The standardized bias is the mean difference of each variable in the treated and untreated groups relative to the variances, details can be found in equation (4.6) in Section 4.

Table C.3: Standardized bias, sample 2002-2016

	Expected prod. increase			Expected prod. decrease		
Variable	All	Correct	Incorrect	All	Correct	Incorrect
Exp. prod., $t-1$	2.4	2.7	1.2	-1.1	-1.3	-2.3
Exp. prod., $t-2$	0.9	-0.3	0.8	-2.4	-2.3	-1.5
Exp. prod., $t-3$	1.0	-0.7	2.4	-1.0	-3.5	0.5
Production, $t-1$	1.4	0.5	-0.6	0.2	-0.7	1.9
Production, $t-2$	1.1	0.2	-0.4	-0.4	-0.8	1.3
Production, $t-3$	0.4	1.0	-0.3	-2.3	-3.0	0.5
Prices, $t-1$	-0.1	0.6	-0.0	0.3	3.9	1.4
Prices, $t-2$	0.3	-1.2	1.5	0.5	1.9	-0.2
Prices, $t-3$	0.8	1.1	0.1	-0.6	1.0	1.0
Demand, $t-1$	1.1	0.8	-0.1	-0.4	-0.7	-0.1
Demand, $t-2$	0.9	1.7	-0.4	-1.5	-2.6	1.3
Demand, $t-3$	0.7	0.9	0.1	-2.6	-4.7	1.3
Capacity, $t-1$	0.8	0.7	0.1	0.3	3.0	1.4
Capacity, $t-2$	0.3	0.6	0.4	-0.4	3.7	1.4
Capacity, $t-3$	0.8	1.0	0.0	-0.7	4.1	0.4
Employees, t	0.3	-0.3	0.4	1.4	2.0	1.5
Avg. state of business, sector t	0.8	1.7	-0.5	-0.2	0.1	-0.6
State of business, t	0.3	-0.2	0.3	-0.7	-0.9	-0.6
State of business, $t-1$	-0.0	-0.2	0.3	0.1	-0.3	-1.0
State of business, $t-2$	0.2	0.5	1.2	-1.2	0.1	-0.5
State of business, $t-3$	-0.0	-0.3	0.7	-1.1	0.8	-1.4
Orders, t	0.3	-1.0	-0.0	-0.1	-0.2	-0.0
Orders, $t-1$	1.6	0.4	0.2	-0.4	-1.5	-0.3
Orders, $t-2$	1.3	-0.0	0.8	-2.1	-3.1	-1.2
Orders, $t-3$	0.9	0.0	1.1	-0.9	-2.5	-0.8
Foreign orders, t	0.9	-0.2	0.1	-0.5	-1.8	-0.1
Foreign orders, $t-1$	1.6	0.6	1.1	-0.8	-2.2	-0.5
Foreign orders, $t-2$	1.1	0.3	1.2	-1.3	-2.6	-1.0
Foreign orders, $t-3$	1.0	0.0	1.6	-0.4	-2.1	-0.7
Debt share, t	-1.4	-3.7	-2.7	0.1	-5.4	-0.2
Financing coefficient, t	-0.0	-0.7	-0.4	0.9	2.6	1.3

Notes: Table shows standardized bias which captures the difference in means between the treated and untreated groups. The standardized bias is the mean difference of each variable in the treated and untreated groups relative to the variances, details can be found in equation (4.6) in Section 4.

Table C.4: Standardized bias, match in sector

Variable	Expected prod. increase			Expected prod. decrease		
	All	Correct	Incorrect	All	Correct	Incorrect
Exp. prod., $t-1$	-1.0	-0.1	-2.3	1.7	-0.1	-1.0
Exp. prod., $t-2$	1.5	2.5	2.9	0.4	-0.1	-2.8
Exp. prod., $t-3$	1.6	2.3	2.3	-0.3	-2.1	-1.0
Production, $t-1$	0.5	2.2	2.2	-1.3	0.8	-2.9
Production, $t-2$	1.1	1.0	-0.1	1.1	2.4	-1.2
Production, $t-3$	0.2	-0.6	-0.0	0.3	-0.9	-1.8
Prices, $t-1$	0.4	2.2	-0.5	-1.5	0.3	0.2
Prices, $t-2$	-0.3	-1.3	-0.2	-1.4	0.6	-0.6
Prices, $t-3$	0.9	1.0	-1.1	-1.1	0.6	-2.1
Demand, $t-1$	2.5	5.7	2.0	-3.7	-5.0	-2.9
Demand, $t-2$	0.6	1.7	0.1	0.0	-2.8	-0.2
Demand, $t-3$	2.6	1.4	0.0	0.7	-3.0	-1.4
Capacity, $t-1$	-0.6	0.7	0.7	1.9	3.4	-0.3
Capacity, $t-2$	-0.8	1.1	0.8	1.3	2.8	1.1
Capacity, $t-3$	-0.5	0.8	0.3	0.2	2.2	-0.4
Employees, t	-0.4	2.1	0.5	0.8	-1.4	-1.2
Avg. state of business, sector t	0.0	0.0	0.0	-0.0	-0.0	-0.0
State of business, t	1.2	-0.2	1.1	-0.5	0.5	-1.0
State of business, $t-1$	0.8	-1.2	0.2	0.5	1.8	0.7
State of business, $t-2$	0.7	-1.1	-0.2	0.3	2.7	2.4
State of business, $t-3$	-0.3	-2.1	-0.6	0.2	-0.1	1.4
Orders, t	0.7	0.6	1.2	0.3	1.3	-0.4
Orders, $t-1$	1.4	-1.3	-0.5	0.1	2.4	-0.5
Orders, $t-2$	1.0	-1.9	-0.6	0.4	1.1	1.5
Orders, $t-3$	-0.7	-3.0	-1.3	0.3	2.9	2.3
Foreign orders, t	-0.7	-2.4	-0.1	0.2	1.5	-0.2
Foreign orders, $t-1$	0.2	-2.9	-1.2	0.9	1.7	0.9
Foreign orders, $t-2$	0.5	-2.8	-0.7	0.3	1.7	1.5
Foreign orders, $t-3$	-0.5	-3.1	-0.7	0.5	2.4	2.7
Debt share, t	-0.3	1.6	-1.3	0.6	0.3	-1.1
Financing coefficient, t	0.9	0.2	3.1	0.4	-2.1	-0.2

Notes: Table shows standardized bias which captures the difference in means between the treated and untreated groups. The standardized bias is the mean difference of each variable in the treated and untreated groups relative to the variances, details can be found in equation (4.6) in Section 4.

Table C.5: Standardized bias, firms responding in first 10 days of month

	Expected prod. increase			Expected prod. decrease		
Variable	All	Correct	Incorrect	All	Correct	Incorrect
Exp. prod., $t-1$	2.3	2.1	5.0	-1.1	-0.2	0.5
Exp. prod., $t-2$	0.8	0.2	0.3	-1.9	-0.2	-4.9
Exp. prod., $t-3$	2.3	2.0	2.0	-0.4	-0.4	-0.2
Production, $t-1$	1.8	3.6	1.1	1.2	2.3	0.0
Production, $t-2$	0.2	0.6	1.5	-0.3	-2.3	0.8
Production, $t-3$	-0.9	-0.4	0.5	-0.8	-2.2	1.5
Prices, $t-1$	0.5	-1.8	0.3	1.6	0.8	-2.6
Prices, $t-2$	1.5	-1.4	1.6	1.7	0.6	-0.9
Prices, $t-3$	3.1	-1.0	0.5	0.2	0.7	0.2
Demand, $t-1$	1.1	0.7	0.4	-1.9	-3.1	-2.2
Demand, $t-2$	1.7	3.3	-1.2	-1.6	-3.4	1.7
Demand, $t-3$	1.1	0.2	-1.1	0.1	0.3	2.1
Capacity, $t-1$	0.2	-4.6	1.9	0.7	3.5	2.3
Capacity, $t-2$	0.1	-2.6	0.8	-0.6	4.5	0.8
Capacity, $t-3$	0.3	-2.0	1.5	-0.7	4.7	-1.4
Employees, t	0.6	0.4	-1.1	0.3	3.7	-1.3
Avg. state of business, sector t	1.3	2.1	2.7	1.8	0.3	0.2
State of business, t	0.4	-0.8	2.6	-0.2	2.1	-1.4
State of business, $t-1$	-0.3	1.8	2.8	0.4	2.2	1.7
State of business, $t-2$	0.3	-0.3	2.4	-1.2	-0.1	0.2
State of business, $t-3$	0.1	2.4	2.1	-0.4	1.1	-0.3
Orders, t	-0.8	-1.4	1.6	0.6	0.6	-2.3
Orders, $t-1$	1.3	-0.3	0.6	1.8	2.5	0.4
Orders, $t-2$	1.0	-1.1	1.8	0.0	-0.9	-1.5
Orders, $t-3$	0.5	0.6	3.0	1.1	3.1	-2.1
Foreign orders, t	-0.9	-1.9	3.7	1.4	-3.7	-4.6
Foreign orders, $t-1$	0.7	-1.4	3.2	1.7	-2.8	0.6
Foreign orders, $t-2$	1.0	-3.0	2.7	1.3	-4.7	-0.6
Foreign orders, $t-3$	0.8	-1.1	2.6	-0.2	-3.3	-0.0
Debt share, t	-2.6	-1.2	-0.9	0.7	-1.8	0.5
Financing coefficient, t	0.1	0.5	-4.2	1.5	0.7	3.0

Notes: Table shows standardized bias which captures the difference in means between the treated and untreated groups. The standardized bias is the mean difference of each variable in the treated and untreated groups relative to the variances, details can be found in equation (4.6) in Section 4.

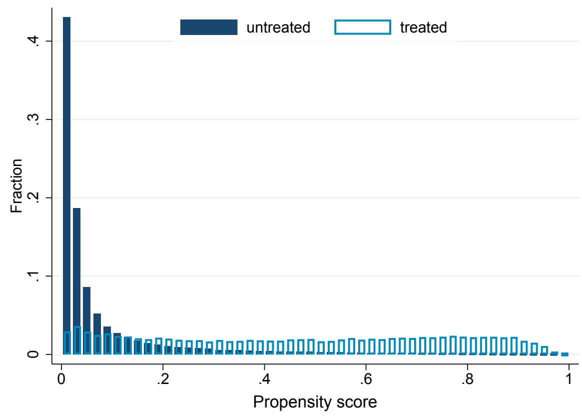
Table C.6: Standardized bias, models with expected state of business

Variable	Expected more favorable: Full sample	Sample w/ slider	Small changes	Large changes	Expected less favorable: Full sample	Sample w/ slider	Small changes	Large changes
Exp. state of bus., $t-1$	1.2	2.4	-1.0	2.7	-0.1	-0.4	1.8	-0.8
Exp. state of bus., $t-2$	1.2	-0.0	-2.5	1.9	-0.6	-0.9	1.2	-3.5
Exp. state of bus., $t-3$	1.8	-0.3	-1.4	3.0	-1.1	-0.7	-2.1	-6.2
Production, $t-1$	1.1	0.8	2.1	-0.9	-1.8	-0.8	-1.0	-3.2
Production, $t-2$	1.3	-0.4	-3.7	-0.8	-0.4	-2.1	1.5	-1.5
Production, $t-3$	1.2	0.8	-1.5	1.8	-0.6	-0.9	2.1	-5.5
Prices, $t-1$	1.2	0.0	-0.8	-2.2	-0.2	-0.1	0.6	-1.5
Prices, $t-2$	1.1	-1.2	0.6	-1.0	0.3	0.8	3.0	1.4
Prices, $t-3$	1.0	-1.0	0.1	-1.7	1.2	-0.6	2.5	-3.0
Demand, $t-1$	0.7	-0.0	2.6	1.4	-2.9	-1.9	-3.7	-5.5
Demand, $t-2$	1.0	-0.4	-2.2	-0.6	-1.9	-1.7	-0.3	-1.6
Demand, $t-3$	1.2	0.8	-1.4	0.8	-1.1	-0.5	2.2	-4.8
Capacity, $t-1$	2.4	1.9	-1.8	2.2	1.0	1.8	2.2	1.1
Capacity, $t-2$	2.6	2.2	-1.6	3.0	0.9	1.9	2.4	2.8
Capacity, $t-3$	2.5	1.5	-0.9	3.7	1.4	2.7	0.8	2.1
Employees, t	0.9	0.3	-1.9	0.0	1.0	-0.0	-0.6	-0.6
Avg. state of bus., sector t	-0.1	0.3	-0.3	0.2	-0.3	1.2	0.6	-3.4
State of bus., t	0.1	-0.6	-2.3	2.7	-1.0	-0.4	1.3	-2.1
State of bus., $t-1$	-0.4	-0.9	-4.2	2.3	0.0	0.5	1.6	-1.4
State of bus., $t-2$	-0.1	-0.9	-3.3	3.8	-0.8	0.7	1.1	-3.5
State of bus., $t-3$	0.0	-0.4	-2.5	5.0	0.1	0.7	0.6	-1.7
Orders, t	0.1	-1.3	-0.6	1.9	-0.5	-0.1	1.0	-2.6
Orders, $t-1$	1.4	0.7	-2.4	2.2	-1.2	0.2	1.5	-0.4
Orders, $t-2$	0.9	-0.4	-2.9	4.1	0.5	1.3	1.3	-0.9
Orders, $t-3$	1.3	0.8	-0.6	4.8	0.3	2.4	0.4	1.4
Foreign orders, t	0.7	-0.4	-0.1	2.0	-0.8	0.1	1.7	-1.1
Foreign orders, $t-1$	1.0	0.5	-1.4	3.5	0.0	1.6	3.1	-0.1
Foreign orders, $t-2$	1.1	0.8	-1.5	4.2	-0.6	0.5	1.9	-1.8
Foreign orders, $t-3$	1.3	0.7	0.1	6.5	-0.9	1.4	-0.4	-0.4
Debt share, t	-1.2	-1.2	0.9	-0.9	-3.2	-2.0	-2.4	-3.1
Financing coefficient, t	0.4	-1.2	-0.0	-0.7	0.5	0.8	0.8	0.6

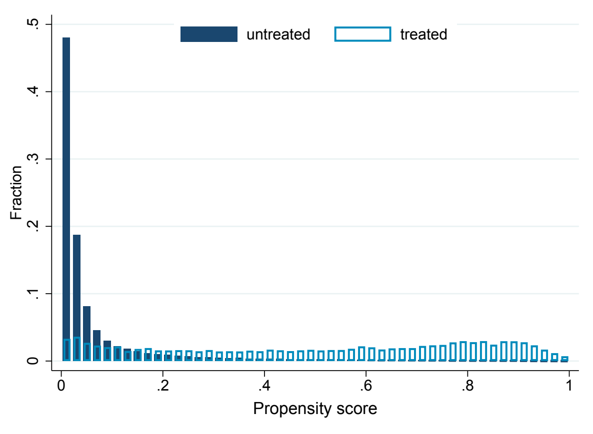
Notes: Table shows standardized bias which captures the difference in means between the treated and untreated groups. The standardized bias is the mean difference of each variable in the treated and untreated groups relative to the variances, details can be found in equation (4.6) in Section 4.

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

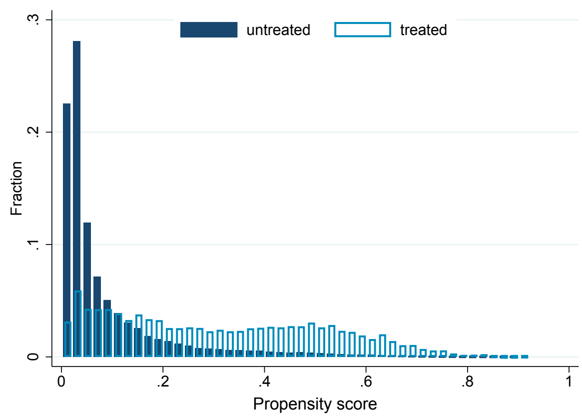
(a) Correctly expected production increase



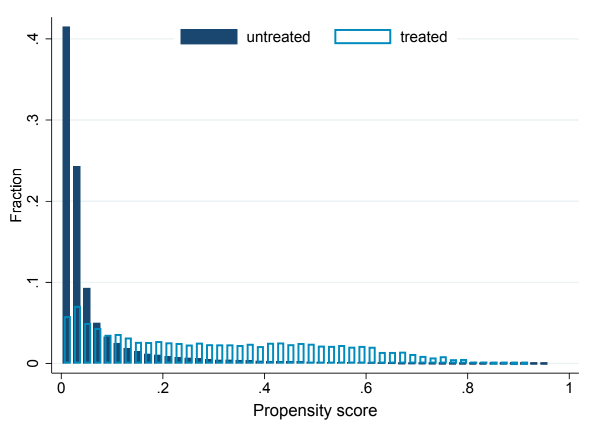
(b) Correctly expected production decrease



(c) Incorrectly expected production increase



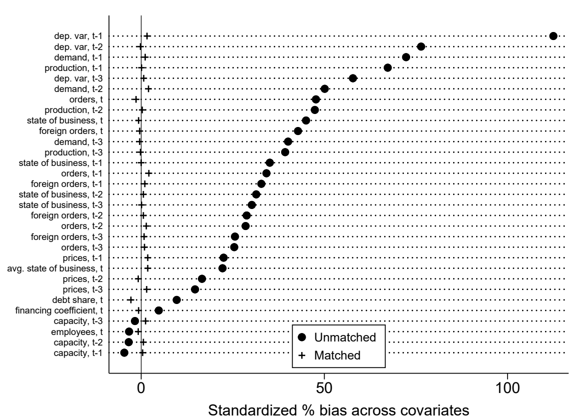
(d) Incorrectly expected production decrease



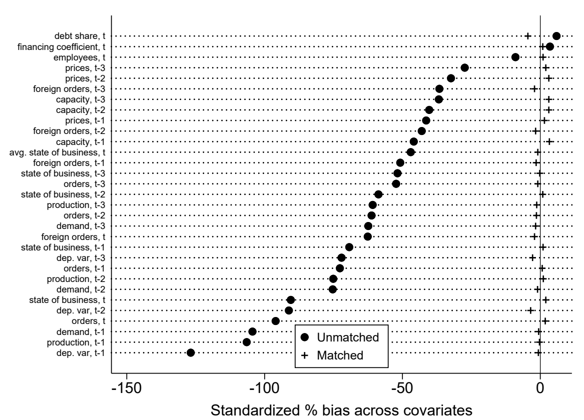
Notes: Histograms show the propensity scores for treated and untreated firms respectively, estimated as described by equation (4.4). In Panel (a) treated firms expect an increase and are correct, in panel (b) treated firms expect a decrease and are correct, in panel (c) treated firms expect an increase and are incorrect, and in panel (d) treated firms expect a decrease and are incorrect.

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

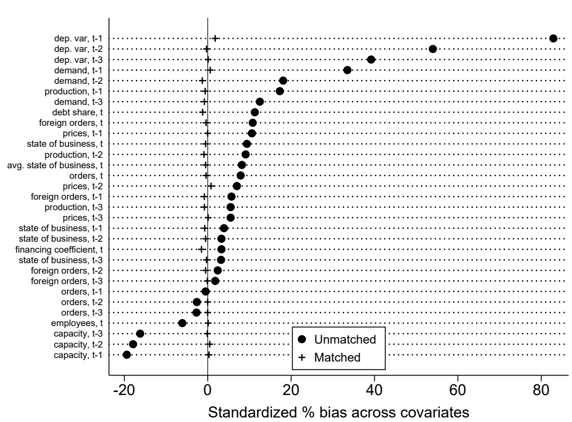
(a) Correctly expected production increase



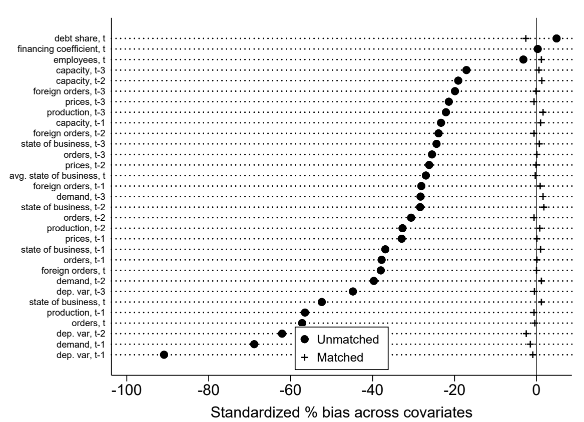
(b) Correctly expected production decrease



(c) Incorrectly expected production increase



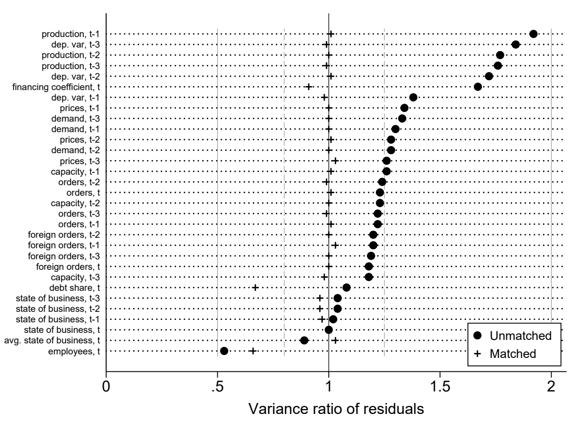
(d) Incorrectly expected production decrease



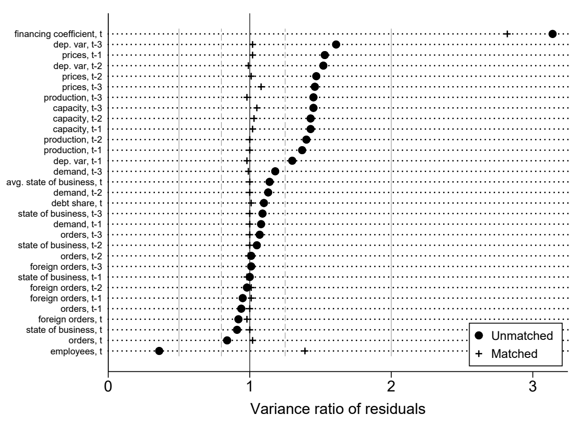
Notes: Figure shows standardized bias for the matching firms expecting a production increase or decrease and being correct or incorrect, respectively. The standardized bias measures the mean difference of each variable in the treated and untreated groups, as described by equation (4.6) in Section 4.2.

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

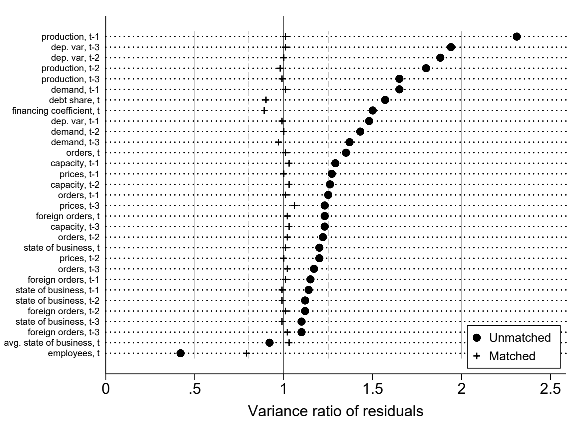
(a) Correctly expected production increase



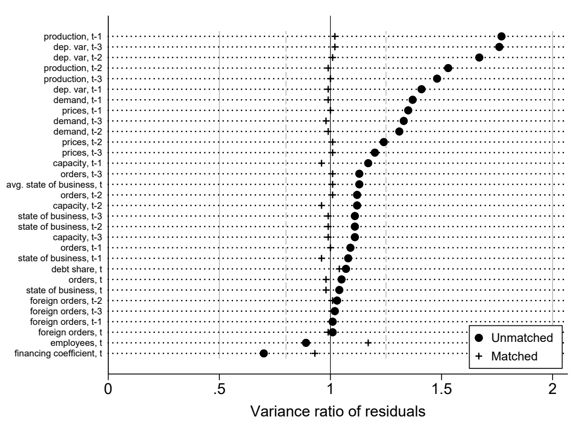
(b) Correctly expected production decrease



(c) Incorrectly expected production increase



(d) Incorrectly expected production decrease



Notes: Figure shows variance ratios for the firms expecting a production increase or decrease and being correct or incorrect, respectively. The variance ratio measures the difference between the variances orthogonal to the propensity score. Details on this measure can be found in Section 4.2. Variance 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 4 and 5

D.1 Additional specifications

In addition to the robustness checks included in the main text, we consider some more checks in Table D.1. First, we also use propensity scores resulting from an ordered probit estimation instead of the two separate probit estimations, as explained in Section 4.1. Column (1) shows that this does not change results materially: the ATTs are somewhat smaller for firms expecting a production increase and somewhat larger for firms expecting a production decrease but remain significant.

In column (2) of Table D.1, we restrict the sample to firms responding only within the first 7 days of the month. Effects remain similar compared to the baseline specification in all cases, except the ATT on prices for firms expecting a decline in production. Here, the ATT is smaller than in the baseline specification and no longer significant.

Column (3) restricts the sample in another way: here we focus on firms which report only on one product in the survey. As mentioned above, some firms fill out more than one questionnaire each round. The results in column (3) show that this does not seem to affect our baseline results.

A further issue is the frequency mismatch between survey and balance sheet data. In the baseline case, we always use the most recently available data based on the end of the accounting period. This means that the likelihood of the balance sheet information no longer reflecting the current situation increases the more in the past the closing date of the balance sheet is compared to the current period. Therefore, in column (4) of Table D.1, we only include firm observations from the three months after the balance sheet closing date. Again the results are not materially affected. Another way to consider the role of the frequency mismatch is to distribute the balance sheet data differently over the year. This approach is discussed in Appedix D.2 below.

Finally, firm age maybe a further relevant factor determining firms' fundamentals and expectation formation. Since firm age is not available for many firms, we did not include it in the baseline specification. Column (5) of Table D.1 shows the results from including firm age in the probit estimations used to compute the propensity scores. This again leads to similar results as the baseline specification.

Table D.2 shows the same robustness specifications for the matching procedure that further separates firms based on whether their expectations turn out to be correct or incorrect (except for the ordered probit approach which is not feasible with these treatment variables). Also here results remain qualitatively unchanged, only the size of the coefficients varies in some cases.

All robustness estimation in Tables D.1 and D.2 involve a new matching. Therefore, we again check the balancing statistics, which are overall similar to the baseline case with

Table D.1: Average treatment effect on the treated, additional robustness checks

	(1) Ordered probit	(2) Response in first 7 days	(3) Single-product firms	(4) 3 month after fiscal year end	(5) Firm age
Panel A: Expected production increase – Effect on production					
ATT	0.152***	0.203***	0.172***	0.163***	0.175***
	(27.14)	(16.21)	(28.64)	(16.29)	(27.73)
Observations	128932	21250	110817	37509	96597
Panel B: Expected production increase – Effect on prices					
ATT	0.016***	0.028**	0.024***	0.032***	0.024***
	(3.87)	(2.86)	(5.23)	(4.12)	(5.12)
Observations	128977	21253	110861	37529	96627
Panel C: Expected production decrease – Effect on production					
ATT	-0.198***	-0.177***	-0.177***	-0.188***	-0.173***
	(-32.53)	(-11.50)	(-26.69)	(-16.12)	(-24.46)
Observations	123941	19772	106468	34575	93038
Panel D: Expected production decrease – Effect on prices					
ATT	-0.038***	-0.018	-0.027***	-0.022**	-0.031***
	(-7.80)	(-1.47)	(-5.14)	(-2.27)	(-5.53)
Observations	124014	19785	106541	34604	93094

Notes: Table shows treatment effects on prices and production for different specifications. Treatment is the expectation of future production. Outcomes refer to the current month, t, while expectations refer to the next 3 months ($t+1$ to $t+3$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

few exceptions, in particular for the ordered probit model. For brevity, results are not reported but available upon request.

Table D.2: Average treatment effect on the treated, correct and incorrect firms, additional robustness checks

	(1) Response in first 7 days	(2) Single-product firms	(3) 3 month after fiscal year end	(4) Firm age
Panel A: Correctly expected production increase – Effect on production				
ATT	0.320***	0.300***	0.300***	0.308***
	(17.65)	(34.77)	(21.03)	(33.57)
Observations	11243	67250	23731	59039
Panel B: Correctly expected production increase – Effect on prices				
ATT	0.035**	0.040***	0.047***	0.035***
	(2.36)	(5.85)	(4.08)	(4.81)
Observations	11242	67249	23726	59042
Panel C: Incorrectly expected production increase – Effect on production				
ATT	0.088***	0.066***	0.081***	0.072***
	(5.23)	(8.45)	(6.22)	(8.72)
Observations	12196	69741	23987	60314
Panel D: Incorrectly expected production increase – Effect on prices				
ATT	0.029**	0.016***	0.020**	0.013**
	(2.23)	(2.72)	(1.93)	(1.98)
Observations	12195	69747	23983	60313
Panel E: Correctly expected production decrease – Effect on production				
ATT	-0.273***	-0.308***	-0.309***	-0.302***
	(-12.11)	(-31.79)	(-18.61)	(-29.05)
Observations	9914	65778	20471	57939
Panel F: Correctly expected production decrease – Effect on prices				
ATT	-0.044**	-0.024***	-0.045***	-0.030***
	(-2.26)	(-2.82)	(-3.05)	(-3.35)
Observations	9915	65773	20467	57927
Panel G: Incorrectly expected production decrease – Effect on production				
ATT	-0.073***	-0.087***	-0.124***	-0.076***
	(-3.53)	(-9.52)	(-7.74)	(-7.79)
Observations	10492	65386	21167	56662
Panel H: Incorrectly expected production decrease – Effect on prices				
ATT	0.019	-0.003	-0.012	-0.005
	(1.17)	(-0.47)	(-0.91)	(-0.58)
Observations	10500	65401	21166	56664

Notes: Table shows treatment effects on prices and production for different specifications. Treatment is the expectation of future production, separately for firms which turn out to be correct and incorrect, respectively. Outcomes refer to the current month, t, while expectations refer to the next 3 months ($t+1$ to $t+3$). T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Alternative use of balance sheet data

Table D.3: Aggregate results with alternative use of balance sheet data

	Expected prod. increase		Expected prod. decrease	
	(1)	(2)	(3)	(4)
Dep. variable:	Production	Prices	Production	Prices
ATT	0.170***	0.026***	-0.175***	-0.029***
	(28.85)	(5.82)	(-27.23)	(-5.57)
Observations	120754	120802	116470	116548

Notes: T-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

As discussed in Section 3, the survey data has a different frequency than the balance sheet data. In our baseline setting, we use the most recently available balance sheet data to estimate the debt share and financing coefficient in each month. This implies that in the months before the end of the accounting year, we use information which is almost one year old. We use this approach to avoid including any 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 closing date 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 closing date, we use the most recent report as before. However, for the next six months until the new balance sheet is available, we use the new data. This means that we always use the balance sheet data with the closing date closest to the respective month.

Table D.3 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.3 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 other variables (in the latter case using the Mahalanobis distance) without replacement. In our case, we use the propensity score from the simple probit regressions described in Section 4.1 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

Table D.4: Aggregate results with alternative matching procedure

	Expected prod. increase		Expected prod. decrease	
	(1) No bias corr.	(2) Bias corr.	(3) No bias corr.	(4) Bias corr.
Panel A: Effect on production in t				
ATT	0.172***	0.172***	-0.174***	-0.174***
	(0.007)	(0.007)	(0.007)	(0.007)
Observations	135170	135170	131656	131656
Panel B: Effect on prices in t				
ATT	0.027***	0.027***	-0.037***	-0.037***
	(0.005)	(0.005)	(0.006)	(0.006)
Observations	135170	135170	131656	131656

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

radius, additional matches are selected if they are within the radius around the respective observation. This matching step is done with 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 observations 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 the distance, matches can be from different months (albeit only from close months because we include the 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.4. 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: -0.037 compared to -0.031 in the baseline. Reassuringly the bias adjustment also does not have any effects up to three digits. This implies that using a more simple matching procedure with no bias correction is valid in our data set.

D.4 Parametric estimation of the ATT

We also consider a fully parametric estimation of the treatment effects without any matching. For this purpose, we run a simple OLS regression of the outcome variables (realized prices and production) on the firm's expected change in production and control variables. To keep results comparable to the ATTs obtained from the matching, we choose the same control variables as for the propensity score estimation, see model (4.3) in Section 4.1 and Table 2. In addition, we estimate the effect of an expected increase and an expected decrease in production separately. In both cases, the control group are the firms expecting no change. In line with the analysis in Section 4, we consider the increase and decrease in the outcome once jointly (column (1) and (4) in Table D.5) and once separately.

The results in Table D.5 qualitatively confirm the results obtained from the matching. In terms of size, the OLS coefficients are somewhat smaller in most cases, indicating a potential underestimation of the effect using OLS instead of propensity score matching.

Table D.5: OLS estimation of the effect of expected production on realized production and prices

	Change in production			Change in prices		
	(1) All changes	(2) Increase	(3) Decrease	(4) All changes	(5) Increase	(6) Decrease
Panel A: Expected production increase						
ATT	0.160***	0.140***	-0.020***	0.019***	0.014***	-0.005**
	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)
	(0.037)	(0.025)	(0.023)	(0.033)	(0.025)	(0.023)
Observations	135342	135342	135342	135389	135389	135389
Adjusted R^2	0.23	0.25	0.18	0.31	0.28	0.29
Panel B: Expected production decrease						
ATT	-0.178***	-0.024***	0.154***	-0.024***	-0.005*	0.019***
	(0.006)	(0.002)	(0.005)	(0.004)	(0.002)	(0.003)
	(0.037)	(0.027)	(0.022)	(0.032)	(0.027)	(0.020)
Observations	131831	131831	131831	131917	131917	131917
Adjusted R^2	0.26	0.15	0.30	0.33	0.26	0.35

Notes: Standard errors clustered at the firm level in parentheses. Each panel is based on separate OLS regressions with the same control variables as in model (4.3). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.