



**COLLEGIUM OF ECONOMIC ANALYSIS
WORKING PAPER SERIES**

How do firms respond to demand and
supply shocks?

Michał Gradzewicz

How do firms respond to demand and supply shocks?

Michał Gradzewicz*

April 12, 2022

Abstract

The study aims to identify the granular demand and productivity shocks, their properties, and the responses of the important firm-level variables to these shocks. We use comprehensive data from the Polish enterprise sector that cover the 2002-2019 period. As the data do not include prices, the identification of the demand shocks relies on the information on inventory changes. We utilize the control function approach to estimate the parameters of the production function and to identify productivity shocks. We use projection methods with granular data to identify the dynamic impulse-response function. We show that the distributions of the two shocks differ: i.e., supply (productivity) shocks are symmetrically distributed, and the distribution of demand shocks is negatively skewed. Moreover, both distributions have fat tails. Productivity shocks have much more persistent effect on firms' outcomes than demand shocks. Following demand shocks, there are short-lived increases in output, market share, productivity, real wages and markups; whereas investment and employment demand remain elevated for a longer period. We also find a very limited transmission of productivity into wages and we showed that proxies for prices increase after demand shocks, and they decrease after the supply shock, in a theory-consistent way.

Keywords: demand shocks, supply shocks, granular impulse response function, granular local projections

JEL Codes: D22, D24, D4, J42, L11

Introduction and literature review

The aim of this study is to identify the dynamic responses of firms to granular supply (productivity) and demand shocks and to describe the properties of distributions of these shocks. Most of the existing literature on this topic has used macroeconomic data and macroeconomic identification schemes (like structural vector autoregressions or dynamic factor models) to derive the impulse responses. Following the routes pioneered by Cochrane (1994) and Gabaix (2011), we search for the rationale for macroeconomic shocks by measuring the granular reactions of firms to these shocks. The latter approach serves as a cross-check for the impulse responses identified using macroeconomic models. The need for such a consistency check was stressed by Buera et al. (2021). Our approach extends the existing literature by providing additional insights into the properties of shocks and of firms' adjustment mechanisms, in addition to performing a consistency check.

*Narodowy Bank Polski (NBP) & SGH Warsaw School of Economics, e-mail: michal.gradzewicz@nbp.pl The views expressed herein are those of the author and have not been endorsed by Narodowy Bank Polski. The author would like to thank Marek Góra, Jakub Mućk and Łukasz Postek; and participants of seminars at Narodowy Bank Polski and Department of Economics I (SGH Warsaw School of Economics) for helpful comments and discussions.

Because both demand and productivity shocks are unobserved and affect firms' outcomes simultaneously, separating them is challenging, as Klette and Griliches (1996) first noted. Various schemes for identifying demand shocks have been presented in the literature, and their applicability depends on the available data. For cases in which the data include information on prices, Foster et al. (2008) and Pozzi and Schivardi (2016) have offered a feasible solution based on the estimation of the demand schedules faced by firms. But if, as in our case, the prices are unavailable, indirect approaches need to be used instead. We use the identification scheme proposed by Kumar and Zhang (2019), as it is the most suitable approach for our analysis given our data. The approach builds on the assumption that the demand shocks are realized after the firm has chosen its inputs and output levels, whereas the supply shocks are realized before the inputs and output decisions are made (and thus should not affect inventory stock). The uncertainty introduced by demand shocks and timing assumptions create a gap between the expected and the realized sales. Thus, the firm-level deviations of inventories from the targeted level of inventory stock provide information about the unexpected shifts in demand. Note that the identification applied here is agnostic regarding the exact source of the demand shock – e.g., whether it is a result of a transitory shift in preferences or a change in monetary policy conditions – provided it is unobserved when decisions about inputs and output are made.

The identification of supply (productivity) shocks is related to the estimation of the production function. Here, we utilize here the control function approach pioneered by Olley and Pakes (1996), as it addresses the endogeneity problem that can arise when productivity shocks that are not observed by the econometrician are known to the firm (as is assumed in our identification scheme), and it affects both output and input choices. Moreover, the lack of firm-level prices in our data combined with the imperfect competition on the output markets implies that additional problems can arise when using revenue rather than the quantity production function, as was recently stressed by Bond et al. (2020), Doraszelski and Jaumandreu (2021) and Ridder et al. (2021) in the context of measuring markups.

Kumar and Zhang (2019) proposed an adjustment to the production function estimation procedure that incorporates the demand for the firm's products into the estimation of the revenue production function in a way that consistently allows for demand shocks to be controlled for in the estimation of the firm level productivity. This approach addresses some of the issues raised in the literature. First, it changes the first stage of the estimation of the production function, as inventories enter the production function as a state variable. Moreover, demand shocks may affect the investment decisions of firms due to factors such as the credit constraints. Thus, our identification assumes that demand shocks are part of the investment policy function used to control for productivity¹. Second, demand shocks are a source of independent variation between investments decisions and labor (and material) choices, which resolves the collinearity problem described by Akerberg et al. (2015). Third, Doraszelski and Jaumandreu (2021) argued that controlling for the heterogeneity in demand conditions across firms under imperfect competition is necessary to obtain unbiased estimates of the parameters of the production function and productivity, and, thus, the related measures of markups.² They showed in simulations that controlling for the demand shifts in the estimation procedure improves the estimates. In the empirical part of their study they used a proxy for demand shifts taken from the survey data. In our study, we directly identify a shock to the demand schedule.

We additionally refined the Kumar and Zhang (2019) identification scheme and added a measure of the firm's market share to the investment control function. Both Bond et al. (2020) and Doraszelski and Jaumandreu (2021) stressed that in cases in which firms are price-setters, the variables in the control function for productivity should include marginal costs. While these costs

¹We use investment as a control variable, as in Olley and Pakes (1996).

²Identified using a method proposed by De Loecker and Warzynski (2012).

are unobserved, they are closely related to markups. Moreover, many theoretical models (see, e.g., Atkeson and Burstein, 2008; Burstein et al., 2020) predict that firm-level markups are closely related to market shares.

In the empirical part of the paper we derive the demand and supply shocks using a comprehensive panel of financial statements from firms based in Poland, that covers more than 80% of employment and output in the enterprise sector. This rich dataset allows us to generate findings that are representative from a macroeconomic perspective. We show that although the standard deviations of the distributions of the two shocks are not large, their tails are extremely fat. The two distributions also differ in terms of skewness, as the productivity shocks are approximately symmetrically distributed, whereas the demand shocks are negatively skewed. Moreover, the demand shocks are persistent, while the autocorrelation of the supply shocks is essentially zero.

Having derived the productivity and demand shocks, we use the local projection method pioneered by Jordà (2005) to estimate the dynamic responses of variables related to firms' reactions to these shocks. Local projection is a method that allows for a direct estimation of impulse responses without the need to estimate and invert the full structural decision model. This characteristic of the method makes it especially well-suited for use in microanalyses based on panel data, as in such analyses, the estimation of a fully-blown simultaneous equations model is challenging. Nonetheless, to our knowledge, this paper is the first to use local projections based on a large microeconomic dataset, given that in the existing literature the method has been applied only in multi-country macroeconomic contexts, such as in Auerbach and Gorodnichenko (2013) or Jordà et al. (2015). Moreover, to our knowledge our study is the first attempt to identify dynamic impulse responses using a fully granular identification scheme.

Our empirical results indicate that the changes in firms' outcomes are much more persistent in response to productivity shocks than to demand shocks. Demand shocks result in short-lived increases in output, market share, productivity, real wages, and markups; and in increases in investment and employment for a couple of periods. Firms' reactions to supply shocks are qualitatively similar, but they are more persistent. Moreover, regardless of the nature of the shock, the resulting increases of labor productivity only partially translate into higher wages. We also use the estimated production function to construct the measures of prices and find that prices tend to increase temporarily after demand shocks, and to decrease temporarily after the supply shock, as predicted by theoretical models.

The studies that are the closest to ours are Pozzi and Schivardi (2016) and Carlsson et al. (2021). The authors of the former study also used granular measures of demand and supply shocks (using data for Italian manufacturing firms that include prices, which implies a different identification scheme), and found a limited pass-through of these shocks into firm's output growth for the current period. We extend their analysis to cover dynamic responses, which is important, given that many of the firms' characteristics are affected for a longer period of time. Moreover, our analysis covers a wider set of variables of interest. The authors of the latter study used data from Swedish manufacturing firms that include prices. They applied a set of long-run restrictions in a firm-level panel SVAR model to identify and examine the effects of demand and supply shocks. While their analysis included the reactions of output and prices, they concentrated on the responses of labor market flows to these shocks, and found that demand shocks have a larger impact on employment, whereas technology shocks have a larger impact on prices. Our work extends their analysis by using a more comprehensive set of variables of interest and by broadening the sectoral coverage of the analysis. Moreover, we use local projections as a more robust way to measure dynamic short-run and long-run responses to firm-specific shocks.

The related literature has also measured the effects of supply and demand shocks on market selection and firm turnover. The pioneering study of Foster et al. (2008) found that both productivity and demand shocks affect firm survival. As we do not have direct information on firm

bankruptcies, and the firm exits we observe in our data may have occurred for other reasons, we do not include this variable in the main analysis (although it still serves as a control variable in the identification of technology shocks). Roberts et al. (2018) focused on the role of firm-level demand and supply shocks in firms' exports, whereas Bachmann and Zorn (2020) used survey data to investigate the relative role of demand and supply shocks in firms' investment and output growth. Coviello et al. (2021) found that Italian firms tend to adjust investments rather than employment after experiencing negative demand shocks (identified as reductions in local government spending). The reactions of employment and wages were also studied by Cho (2018), who used matched firm-level financial data with a dataset of transactions from the American Recovery and Reinvestment Act.

The rest of the paper is organized as follows. In the next section, we present our identification strategy. We then discuss our data sources and their properties. In the following section, we report the results of our estimation: i.e., the targeted levels of inventories and the properties of the distributions of demand and supply shocks. In the next section, we present our main results, which indicate how firms respond to these shocks. In the final section, we offer some concluding remarks and comments.

1 Identification

Our approach to the firm-level identification of demand shocks closely follows that of Kumar and Zhang (2019), who used information on inventory changes to derive the unexpected demand shocks faced by firms, and used the estimated production function to measure supply (productivity) shocks. We modify their approach to estimating the production function in order to address the identification issues recently discussed in the literature on firm-level markup estimation.

Assume the production function of firm i at time t is

$$Q_{it} = \exp(\omega_{it}^* + \epsilon_{it}^*) F(\mathbf{V}_{it}, K_{it}), \quad (1)$$

where Q_{it} , \mathbf{V}_{it} , K_{it} are the quantity produced, a vector of variable factors (e.g., labor, materials) and capital, respectively. ω_{it}^* is the firm's structural (observed) productivity, and ϵ_{it}^* is an unobserved, non-structural shock that contemporaneously affects production and productivity. We use lower-case letters to denote the logarithms of the respective variables. We assume $F(\cdot)$ is Cobb-Douglas (in Appendix B.1, we show our main results for the translog production function).

We assume the demand the firm faces follows a standard CES form:

$$Q_{it}^s = P_{it}^\eta \exp(z_{it}), \quad (2)$$

where Q_{it}^s is the quantity demanded (and sold), P_{it} is the price firm i charges in period t , η is the demand elasticity, and $z_{it} \sim \mathcal{N}(0, \sigma^2)$ is an unexpected demand shock.

We assume that when making decisions in period t , the firm knows its capital stock K_{it} , its productivity ω_{it} and enters the period with some initial inventory stock N_{it}^b . At the beginning of the period, the firm chooses variable factors (\mathbf{V}_{it}) and output prices P_{it} that maximize its expected profits conditional on unobserved productivity, ϵ_{it}^* , and demand shocks, z_{it} . The firm may choose to produce a different amount of output than it expects to sell, while aiming to achieve some level of targeted inventory stock λ_{it} . After production, the demand and productivity shocks are observed. These shocks determine the firm's sales and output levels, and lead to end-of-period inventory stock N_{it}^e and to the firm's profits. Thereafter, the firm chooses whether to exit, and (conditional on the exit decision) how much to invest to build capital for the next period.

The timing implies that the firm faces uncertainty when deciding how much to produce, and the optimal choice may generate too much or too little inventory compared to the targeted level.

As inventories cannot be negative, output shortages may occur. Thus, inventories contain useful information on demand shocks, and we use this information for the identification of demand shocks. Moreover, demand shocks can help us pin down productivity in the estimation of the production function, which we will discuss later.

1.1 Demand shocks

In each period, the production sold needs to correspond to the output and inventory changes, which implies:

$$Q_{it} + N_{it}^b = Q_{it}^s + N_{it}^e, \quad (3)$$

where Q_{it} and Q_{it}^s are output and sales, respectively. The demand function (2) can be decomposed into an expected and an unexpected part: $Q_{it}^s = E(Q_{it}^s | \mathcal{I}_{it}) \exp z_{it}$, where \mathcal{I}_{it} is the firm's beginning of the period information set. The literature on inventory behavior (see, e.g., Ramey, 1991; Kahn, 1992, among others) stresses that due to the factors like production smoothing and/or stockout avoidance in the presence of demand uncertainty, it is optimal for firms to maintain a targeted level of inventory stock. We assume that each firm i at time t targets the end-of-period inventory stock, which is proportional to its expected sales: $\lambda_{it} = \lambda_i E(Q_{it}^s | \mathcal{I}_{it})$. Kumar and Zhang (2019) stressed that this assumption is satisfied in a large class of inventory models that predict a fixed stockout rate as their optimal production strategy.

From the firm's perspective, the available output $Q_{it} + N_{it}^b$ need to be equal to the expected sales plus the targeted inventory stock:

$$Q_{it} + N_{it}^b = E(Q_{it}^s | \mathcal{I}_{it}) + \lambda_{it}. \quad (4)$$

Equation (4) can be solved for expected sales. Our assumption regarding the targeted inventory level then allows us to express equation (3) as:

$$\log \left(\frac{Q_{it}^s}{Q_{it} + N_{it}^b} \right) = -\log(1 + \lambda_i) + z_{it}. \quad (5)$$

Equation (5) can be used to identify both the targeted inventory level and demand shocks, provided we have information on quantities, which is often absent in the available data. However, if inventories are priced in line with output prices (which is the rule in practical accounting), the ratio in the LHS of equation (5) can be measured in nominal terms:

$$\log \left(\frac{R_{it}^s}{R_{it} + RN_{it}^b} \right) = -\log(1 + \lambda_i) + z_{it}, \quad (6)$$

where R_{it}^s is the value of the production sold in period t , R_{it} is the value of the output, and RN_{it}^b is the value of the initial inventories. Equation (6) shows that the cross-firm variation of the average ratio of sales to the available output (which is observed, and can be easily computed in the data) identifies the targeted inventory level λ_i , whereas its within-firm (time) variation identifies the demand shock z_{it} .

In the empirical application, we approximate the firm-specific constant $-\log(1 + \lambda_i)$ with a function of firm characteristics, like in Kumar and Zhang (2019). We use a third order polynomial of a firm's size (measured as a log of employment), the interaction of its size with its ownership status (state, private and foreign), and a regional indicator. We also experimented with a firm's fixed effects in a panel regression, but it did not have a qualitative effect on the results. The regression in (6) is estimated separately for 2-digit NACE sectors (in a limited number of cases the sectors are aggregated to avoid convergence problems with Tobit estimation that can occur in small samples). Moreover, in 7.4% of the cases, the estimated measures of λ_i are outside the

$[0, 1]$ interval, and we therefore fix them at the boundary values. Finally, the demand shocks z_{it} are calculated as the difference between the LHS of equation (6) and the firm-level constant $-\log(1 + \lambda_i)$.

There is an additional issue with the above identification scheme. The non-negativity of inventories implies that with a sufficiently high positive demand shock, firms may stockout. Hence, we cannot identify the full distribution of demand shocks with equation (5). This leads to two different kinds of problems. First, the non-negativity of inventory implies that the LHS of (5) is truncated at zero. Thus, to get consistent estimates of λ_j and the demand shocks one would have to include the censoring in the estimation of (6), and apply a Tobit estimation. Second, the truncation implies that we cannot recover the exact magnitude of the demand shocks. When $N_{jt}^e = 0$, then sales and available output coincide; thus, the LHS of equation (5) is zero, and in these censored (bounded) cases $z_{it} = \log(1 + \lambda_i)$. Kumar and Zhang (2019) proposed the use of the normality assumption of the demand shock and the conditional expectation as a measure of the magnitude of the demand shocks in censored cases. As $E(z_{it}|N_{jt}^e = 0) = \int_{\log(1+\lambda_j)}^{+\infty} z_{it} \frac{\phi(z_{it})}{1-\Phi(\log(1+\lambda_j))} dz_{it}$, it follows that the conditional demand shock (we call it \hat{z}_{it}) is:

$$\hat{z}_{it} = \begin{cases} \log\left(\frac{R_{it}^s}{R_{it} + RN_{it}^b}\right) + \log(1 + \lambda_i) & \text{for } N_{jt}^e > 0 \\ \frac{\sigma}{\sqrt{2\pi}} \frac{1}{1-\Phi(\log(1+\lambda_i))} \exp\left(\frac{-\log^2(1+\lambda_i)}{2\sigma^2}\right) & \text{for } N_{jt}^e = 0 \end{cases} \quad (7)$$

The logic and the approach applied here is similar to the Heckman correction commonly used in quantitative social sciences to tackle the problem of selection into the sample. We used the bounded shocks z_{it} in the production function estimation (the rationale is given in Kumar and Zhang, 2019). The conditional shocks \hat{z}_{it} are used in the discussion on the properties of the shocks and the impulse-response analysis.

1.2 Production function and supply shocks

The estimation of the production function is subject to a well-known endogeneity problem (first noticed by Marschak and Andrews, 1944): namely, that the productivity shock, which is unobserved by the econometrician, affects both output and production factors, and introduces a bias into the parameter estimates. We address this problem using the so-called control function, which was pioneered by Olley and Pakes (1996). Our approach to estimating the production function and productivity also addresses the problems of unobserved firm-level prices and imperfect competition, which were recently discussed by Bond et al. (2020), Doraszelski and Jaumandreu (2021), and Ridder et al. (2021).

As Kumar and Zhang (2019) pointed out, the explicit tackling of the demand helps to resolve the problem of unobserved prices. Most firm-level datasets include information on revenues, not on the quantities produced or sold. The observed (log) revenues are given by $r_{it} = q_{it}^s + p_{it}$ and together with the inverse demand function from (2) can be expressed as

$$r_{it} = \left(1 + \frac{1}{\eta}\right) q_{it}^s - \frac{1}{\eta} z_{it}. \quad (8)$$

Moreover, equation (3) gives the relationship between production and sold production: $q_{it} = q_{it}^s + \log(1 + x_{it})$ where $x = \frac{RN_{it}^e - RN_{it}^b}{R_{it}^s}$ is the ratio of the change in inventory value relative to sales. By plugging this, together with the definition of production function (1), into (8), we get:

$$r_{it} = \left(1 + \frac{1}{\eta}\right) \log F(\mathbf{V}_{it}, K_{it}) - \left(1 + \frac{1}{\eta}\right) \log(1 + x_{it}) - \frac{1}{\eta} z_{it} + \omega_{it} + \epsilon_{it}, \quad (9)$$

where (log) revenues r_{it} , (bounded) demand shock z_{it} , a term related to inventories $\log(1 + x_{it})$ and production factors are all observed (or consistently estimated, as in the case of demand shocks),

whereas structural productivity $\omega_{it} = (1 + \frac{1}{\eta})\omega_{it}^*$, non-structural productivity $\epsilon_{it} = (1 + \frac{1}{\eta})\epsilon_{it}^*$, the production function parameters, and η are unobserved, and thus need to be estimated.

Notice that the LHS of equation (9) are revenues rather than quantity produced, which solves the problem with unobserved individual prices (p_{it}) that was described by Bond et al. (2020). In other words, using the explicit assumption regarding the demand the firm faces allows us to estimate output elasticities rather than revenue elasticities.

A direct estimation of equation (9) yields biased estimates, as factor demands $\mathbf{V}_{it}(\omega_{it}, \cdot)$ in any feasible timings of shocks and decisions depend on productivity. We use the control function approach to get consistent estimates of equation (9). We use labor, materials, and outsourcing as the variable factors: $\log \mathbf{V}_{it} = [l_{it}, m_{it}, o_{it}]'$ (and capital as a state variable) and a Cobb-Douglas production function $F(\cdot)$.³ The timing we discussed above implies that the investment i_{it} depends on productivity, capital stock and the end-of-period inventories. Moreover, it also depends on profits, which are, given a non-competitive environment and a firm-specific residual demand (as in our case), a function of markups. While the estimation of markups is not our primary concern, as Doraszelski and Jaumandreu (2021) pointed out, the non-degenerate distribution of markups across firms implies that the estimation of (9) yields biased estimates of output elasticities.

Assuming a monotone relationship between investment and productivity and scalar observability (discussed extensively in Akerberg et al., 2007), we can invert the investment function and express productivity as: $\omega_{it} = \omega_{it}(i_{it}, k_{it}, rn_{jt}^e, z_{it}, \cdot)$. Doraszelski and Jaumandreu (2021) showed that productivity should depend on the planned output (i.e., the expected output before the factor decisions are made), and used a proxy for demand conditions to control for it. Our approach allows us to measure demand shocks directly from the data, and we use these shocks to identify productivity $\omega_{it}(\cdot)$. It is also consistent with the timing of decisions discussed above. Moreover, Bond et al. (2020) and Doraszelski and Jaumandreu (2021) showed that if firms are price-setters, one of the control variables in $\omega_{it}(\cdot)$ should be the firm's (log) marginal costs. These costs are usually non-observable, but they can be proxied by markups. Unfortunately, markups are also unobserved, and production function estimation is often used to identify them. We use market share as a proxy for marginal costs, given that in many economic models market share is correlated with markups.⁴

Summing up, we assume $\omega_{it} = \omega(i_{it}, k_{it}, rn_{jt}^e, z_{it}, ms_{it}, t)$, where ms_{it} is a measure of market share. Plugging this into (9) yields the following equation, which is usually referred to in the literature as the first-stage equation:

$$\begin{aligned} r_{it} &= \beta_l^* l_{it} + \beta_m^* m_{it} + \beta_o^* o_{it} + \beta_k^* k_{it} - (1 + \frac{1}{\eta}) \log(1 + x_{it}) - \frac{1}{\eta} z_{it} + \omega(i_{it}, k_{it}, rn_{jt}^e, z_{it}, ms_{it}, t) + \epsilon_{it} \\ &= \beta_l^* l_{it} + \beta_m^* m_{it} + \beta_o^* o_{it} - (1 + \frac{1}{\eta}) \log(1 + x_{it}) + \varphi(i_{it}, k_{it}, rn_{jt}^e, z_{it}, ms_{it}) + \epsilon_{it}, \end{aligned} \quad (10)$$

where $\beta_i^* = \frac{1+\eta}{\eta} \beta_i$ and $\varphi(\cdot)$ are unknown, and we approximate them using a polynomial of order three. Equation (10) gives consistent estimates of η , β_l^* , β_m^* and $\beta_o^* o_{it}$. As was pointed out by Kumar and Zhang (2019) the introduction of demand shocks into the first stage estimation resolves the collinearity problem between factors and investment discussed in Akerberg et al. (2015). The timing of our model implies that demand shocks affect investment only, and introduces independent variation between the two variables.

Similar to Olley and Pakes (1996), we can calculate: $\hat{\varphi}_{it} = \hat{r}_{it} + (1 + \frac{1}{\eta}) \log(1 + x_{it}) - \beta_l^* l_{it} - \beta_m^* m_{it} - \beta_o^* o_{it}$, where \hat{r}_{it} are fitted values from the estimation of (10). To identify β_k , we assume

³In the empirical analysis we used capital, labor, materials and outsourcing services as production factors. We use the translog production function as a robustness check, see Appendix B.1.

⁴See, e.g., Atkeson and Burstein (2008). Moreover, market share is a sufficient statistics for market power in a Cournot model. A fixed effect panel regression of market share on a proxy of marginal costs, which is discussed in section 1.4, yields a statistically significant coefficient, equal to 0.12.

that productivity follows a first-order Markov process: $\omega_{it} = g(\omega_{it-1}, \chi_{it-1}) + \xi_{it}$, where χ_{it} is an exit probability. The inclusion of exit probability addresses the attrition bias in the measurement of productivity, discussed in Olley and Pakes (1996), among others. The χ_{it} is measured as a predicted exit probability from an estimated non-parametric logit function of investments, demand shocks, and third-order polynomials of production factors as regressors (without the cross terms), all lagged by one period. The above considerations imply the following relationship, referred to as the second-stage equation:

$$\hat{\varphi}_{it} = \beta_k^* k_{it} - \frac{1}{\hat{\eta}} z_{it} + g\left(\varphi_{it-1} - \beta_k^* k_{it-1} + \frac{1}{\hat{\eta}} z_{it-1}, \chi_{it-1}\right) + \xi_{it}, \quad (11)$$

where $\hat{\eta}$ is an estimate derived from the first stage equation (10).

The $g(\cdot)$ function in equation (11) is also unknown and is approximated non-parametrically using a third-order polynomial. We estimate equation (11) using an iterative non-linear sum of squares estimation technique, which results in a consistent estimate of β_k^* . This allows us to calculate productivity as follows:

$$\omega_{it} = \varphi_{it} - \hat{\beta}_k^* k_{it} + \frac{1}{\hat{\eta}} z_{it}, \quad (12)$$

from which we calculate supply shocks as $\Delta\omega_{it}$

In empirical settings we estimate the production function parameters separately for 2-digit NACE sectors.⁵ As Akerberg et al. (2015) observed, firm-level datasets usually have considerable attrition, which affects productivity. In our case, the exit rate for each year varies between 6% and 12%. We control for the exit probability, approximated by fitted values of logit, which is estimated using production factors, the proxy variable (i.e., investments), and demand shock as regressors, all lagged by one period. In both the first and the second stage regressions, we apply the within transformation to take account of the panel structure of the data.

1.3 Firm-level responses to supply and demand shocks

So far, we have discussed the identification scheme for demand and supply shocks. Here, we focus on the identification of the firms' responses to these shocks. We use the local projections method pioneered by Jordà (2005) in the vector autoregressions context, which is flexible enough to accommodate non-linearities and the panel structure of the data (see e.g., Auerbach and Gorodnichenko, 2013; Jordà et al., 2015). This approach is especially well-suited to our granular data, as the estimation of a full microeconomic decision model, which is needed to calculate impulse responses, is often impossible when using such data. The approach is also very simple when using shocks that are already identified, as is the case in our analysis.

We use the following set of regressions to derive impulse responses from the granular data:

$$y_{it+h} = \alpha_{ih} + \gamma_t + \beta_h shock_{it} + \epsilon_{it+h}, \text{ for } h \in 0, 1, \dots, H-1, \quad (13)$$

where h denotes the horizon of a response, y_{it+h} is the response variable, and α_{ih} is the fixed effect. We include the time effects γ_t in the specification, as they control for the business cycle, generating common movements of response variables across a large fraction of firms. β_h measures an impulse response at a horizon h . It is directly estimated from the data; and, importantly, it makes it easy to quantify the accuracy of the estimator. Moreover, this method is more robust to misspecification, because it estimates a separate regression for each horizon h , instead of using

⁵We aggregated some sectors to address the convergence problems in the second step in equation (11) and to avoid the problem of negative point estimates (insignificant) of capital elasticity in a limited number of less capital-intensive sectors.

the same set of coefficients from the assumed and potentially incorrect autoregressive specification for y in calculating the responses analytically. In practice, to consistently estimate the covariance matrix, it is crucial to account for heteroskedasticity and autocorrelation in panel models. Thus, we 1) cluster errors at the firm level, and 2) use a consistent covariance matrix estimator first introduced by Driscoll and Kraay (1998). Appendix B.2 shows the main results of our paper when the set of estimation equations in (13) is augmented with a lagged dependent variable, as suggested in Olea and Plagborg-Møller (2021).

1.4 Markups and proxies for prices

We identify markups using the methodology proposed by De Loecker and Warzynski (2012). We choose materials as a base factor, thus the markup μ_{it} is defined as:

$$\mu_{it} = \beta_i^M \left(\frac{w_{it}^M M_{it}}{R_{it}} \cdot \exp(\epsilon_{it}) \right)^{-1}, \quad (14)$$

where the expression in the parenthesis is a share of material costs $w_{it}^M M_{it}$ in the firm's revenues, corrected by the residual from the first-stage regression given in equation (10).

Our dataset does not allow us to calculate the responses of prices to the demand and supply shocks, but we construct two proxies for prices, and use them in the calculations of the impulse responses. We use the relationship $p_{it} = \log \mu_{it} + mc_{it}$, where μ_{it} is defined in equation (14). The first proxy for marginal costs mc_{it} is unit labor costs, (see, e.g., Sbordone, 2002). Thus, the proxy for a firm-level price becomes $p_{it} = \log \mu_{it} + \log W_{it}^L - \log \left(\frac{Q_{it}}{L_{it}} \right)$. Second, we use the cost function derived in the firm's cost minimization problem with the production function given in equation (1), and with a Cobb-Douglas specification of the $F(\cdot)$ function. When we define the economies of scale of a firm i as $\gamma_i = \sum_{j \in \{K, L, M, O\}} \beta_i^j$, then the cost function (measured up to a constant) is:

$$C_{it}(Q, \Omega, W_{it}^K, W_{it}^L, W_{it}^M, W_{it}^O) \propto \left(\frac{Q_{it}}{\Omega_{it} \exp(\epsilon_{it})} \right)^{\frac{1}{\gamma_i}} \prod_{j \in \{K, L, M, O\}} \left(W_{it}^j \right)^{\frac{\beta_i^j}{\gamma_i}} \quad (15)$$

where W_{it}^j is a price of a production factor j , and β_i^j are the estimates of equations (10) and (11). We measure $W_{it}^M = W_{it}^O = \frac{P_{it}^{IC}}{P_{jt}^{GO}}$, i.e., the relative price of intermediates to the price of output in sector's j (in which the firm operates). The firm-level wages W_{it}^L are defined as an average compensation of employees.⁶ The price of capital W_{it}^K is calculated following the seminal work of Jorgenson (1963) as a user cost of capital:

$$W_{it}^K = UCC_{it} = \frac{P_{jt}^{INV}}{P_{jt}^{GO}} (b_{it} roe_{it} + (1 - b_{it}) i_{it} - \Delta \log P_{jt}^{INV} + \delta), \quad (16)$$

where b_{it} is the share of equities in the sum of the firm's equities and liabilities, roe_{it} is the firm's after tax return on assets, and i_{it} is the unit cost of liabilities, and δ is set at 0.07. Thus, we used the average costs on the firm's liabilities and the returns of the firms' assets to measure the firm's rate of return. The firm's (log) marginal costs is calculated as $mc_{it} = \log \frac{\partial C_{it}(\cdot)}{\partial Q_{it}}$.

2 Data and measurement

Our annual data cover an 18-year period (2002-2019), and include the financial reports and balance sheets of almost all Polish enterprises with more than nine employees (in full-time equivalent

⁶We also calculate responses of a simpler version of the proxy for price, which is based on wages, output and productivity only, hence we assume that prices of capital and materials are not affected by the shock. The results are qualitatively similar.

Table 1: Data trimming and coverage

	initial	demand	final
No. of observations	928,168	855,924	772,275
Observation share	1.00	0.92	0.83
No. of firms	131,580	111,817	104,161
Firms share	1.00	0.85	0.79
Average spell	7.05	7.65	7.41
Output coverage	0.85	0.83	0.69
Employee coverage	0.83	0.79	0.78
Firm coverage	0.56	0.52	0.46

Source: own calculations and Eurostat data

positions). The data are collected by the Central Statistical Office, and comprise non-financial enterprises from the agriculture, mining, manufacturing, construction, and market and non-market service sectors (the latter covers only units from the enterprise sector), as well as a limited number of enterprises in the agriculture sector.

The original data⁷ is an unbalanced panel of over 0.928 million observations – i.e., more than 131,000 firms were observed for an average of seven years (see Table 1) – while also containing missing observations. The original data cover 56% of all firms with 10+ employees that are registered⁸, and they cover 83% of all employees of all registered firms with 10+ employees and 85% of the total economy’s output. Thus, smaller firms are under-represented in our dataset.⁹

We have trimmed the original data to render them usable for further analysis. Around 15% of the firms (8.2% of observations) did not have any inventories during the observation period. In these cases, it is impossible to apply the identification scheme for the demand shocks. The properties of the trimmed data are presented in the demand column in Table 1. The estimation of the production function and the supply shocks is more demanding, and not all of the firms in the sample reported usable data. The data on capital stocks and investments are especially problematic, with relatively high shares of zeros or missing values. Zeros are problematic, as we need to log most of the variables. Thus, we decided to replace zeros with ones in the analysis of the production function¹⁰ in order to use as much information from the data as possible.

The properties of the final sample for which both demand and supply shocks are available are presented in final column in Table 1. The final dataset has 0.772 million records based on observations of more than 104,000 firms over an average of 7.4 years. The firms in our final dataset represent 46% of all firms in Poland with 10+ employees, 78% of the country’s employees and 69% of the economy’s global output.

Nominal output is measured as total revenues from production and from the sale of purchased goods. The dataset contains information on the stock of inventories at both the beginning and the end of the period. The dataset also allows us to use relatively disaggregated production factors: i.e., labor input is measured as the firm’s employees, capital is measured as the firm’s tangible fixed assets (buildings, land, machinery, and vehicles), and materials are measured as the sum of

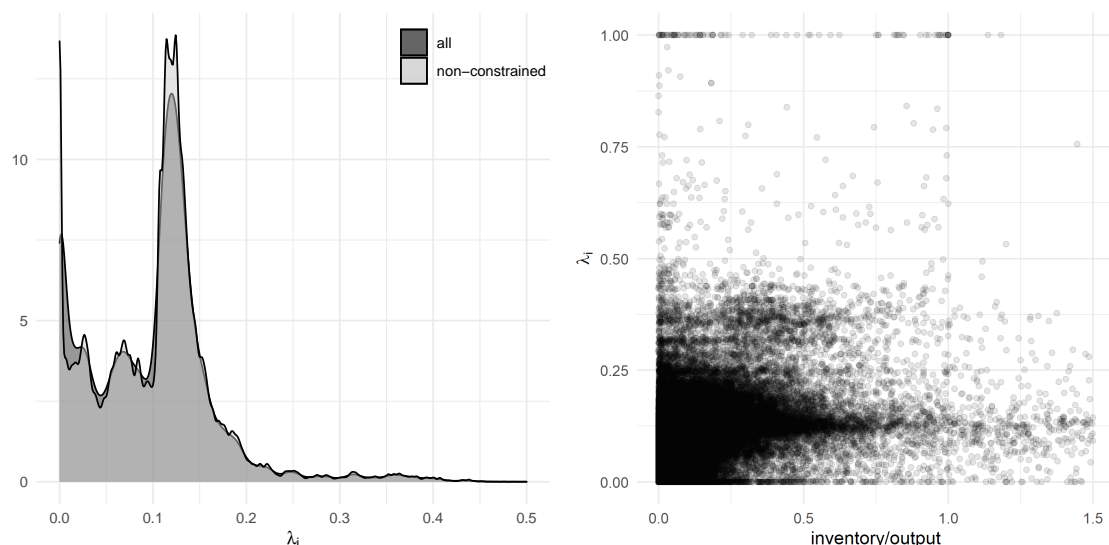
⁷A similar dataset with shorter time coverage and slightly lower firm coverage was used by Gradzewicz (2021) in the analysis of investment spikes.

⁸The firm registry also contains firms that are not active. Moreover, some smaller firms, mainly those employing fewer than 50 employees, decided not to fill out the compulsory form, which is a statistical base for our data.

⁹The information on the total economy is taken from Eurostat databases with the acronyms: ‘bd_9bd_sz_cl.r2’ and ‘nama_10_a64’.

¹⁰The unit in the data is 1000 PLN – approximately 255 USD – a very small amount compared to the scale of even smaller enterprises.

Figure 1: Density of targeted inventories λ_i and its relationship to the inventory-to-output ratio



Notes: inventories/output is defined as a mean of the firm-level ratio of the end-of-period inventories and the output produced, both measured in nominal terms.

purchases of materials used in the production process and purchases of commodities for resale. We also use outsourcing outlays as a separate production factor, observing that the share of this factor in total costs is rising, and these costs are functionally different from materials. In order to construct real values from nominal values (when needed, as our approach addresses the problem of a lack of price data for the firm's output), we use the respective deflators from the sectoral national accounts provided by Eurostat. Capital is deflated using the Eurostat sectoral data on values of fixed assets in the current and previous period replacement costs, while investments (used as a proxy variable in the control function approach) are deflated using the Eurostat sectoral data on gross capital formation deflators. Market shares are calculated as a fraction of industry sales, with the latter defined as a 4-digit NACE aggregation of sectoral sales (calculated using the original data, before trimming).

3 Estimation results

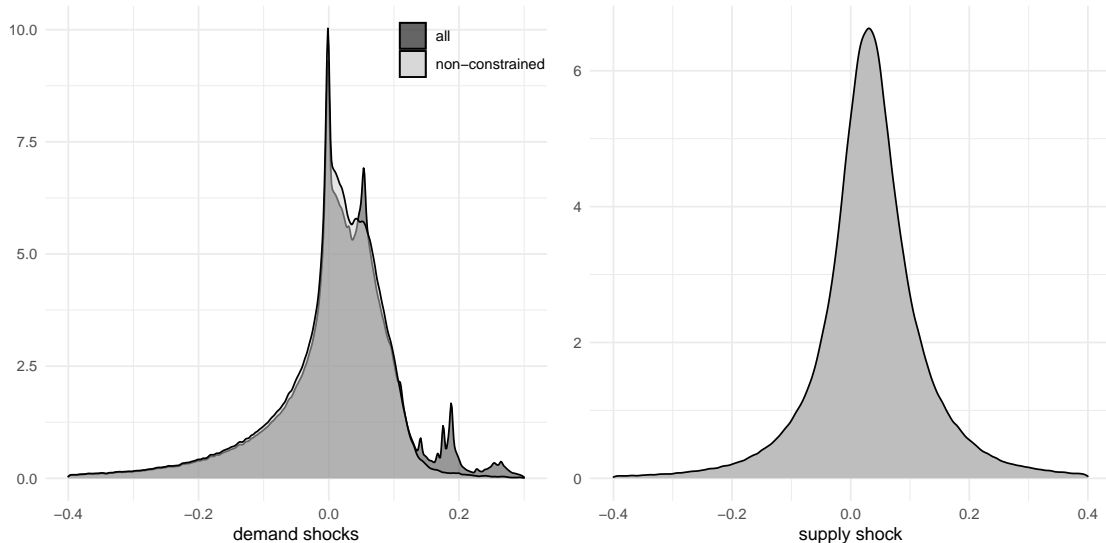
Before we proceed to our main findings, we present some intermediate results from the estimation that are worth highlighting. They concern the targeted inventory levels, the production function parameters, and the properties of the shocks.

3.1 Targeted inventory levels

The estimation of equation (6) yields $\hat{\lambda}_i$, which is the targeted inventory level (measured as a fraction of expected sales). The left panel of Figure 1 displays the density of λ_i . For over 95% of the firms, the targeted inventory level is below 20% of expected sales, and the dominant values are in the 12-13% range. The graph presents the kernel densities of all firms, and specially of the non-constrained firms (and thus only of the firms that did not fail to meet the demand for their products). The left panel of Figure 1 suggests that the approximation to the truncation problem presented in equation (7) does not affect the density of λ_i .

The right panel of Figure (1) shows the scatterplot of λ_i and its closest (albeit imperfect) directly observed counterpart: i.e., the mean of the firm-level ratio of the end-of-period inventories

Figure 2: Densities of demand and supply shocks



and the output produced, both measured in nominal terms. There is no clear pattern between the two, and the correlation is significant but small (0.052). This indicates that the information that λ_i holds is unique, and cannot be directly (and simply) derived from the data.

3.2 Production function parameters

Our baseline calculations assume the Cobb-Douglas production function with four factors, estimated separately for the 2-digit NACE aggregations. The cross-sector average labor elasticity is $\beta_l = 0.2$, ranging from 0.013 in agriculture to 0.4 in the computer repairs. This range is commonly found in the literature. The average elasticities of materials and outsourcing, $\beta_m = 0.26$ and $\beta_o = 0.15$, respectively, sum up to 0.4, which is roughly in line with the values reported in empirical studies. The capital elasticity $\beta_c = 0.023$ is relatively small, on average, although the variation is quite large, ranging from 0.001 in financial services¹¹ to 0.11 in coal mining. The graph with the densities of production function parameters is presented in Figure A.1 in the appendix.

3.3 Properties of demand and supply shocks

The identification procedure outlined in chapter 1 allows us to simultaneously derive demand and supply shocks. As we mentioned in the introduction, demand and supply shocks are usually identified at the macroeconomic level, assuming some distribution; usually a normal distribution. In the identification scheme, we assume that at the firm level, shocks are (log)normally distributed. As the distribution of firms approximately follows the power law, Gabaix (2011) showed that the distributions of firm groupings and aggregate outcomes can have properties that differ from the properties of firm-level distributions. Below, we will characterize the properties of the distributions of the demand and supply shocks identified for all firms in our data.

Figure 2 presents the kernel densities, and Table 2 shows the moments and the positional statistics of demand and supply shocks. The density of the demand shocks is pictured for all firms (for which the calculation is feasible), and for the subgroup of non-constrained firms. Interestingly, the shapes of the distribution of the two shocks differ considerably. The distribution of the demand

¹¹As the dataset excludes firms classified into the sector of financial institutions, all of the firms in this sector are classified into the sector of enterprises.

Table 2: Properties of demand and supply shocks

	mean	sd	skewness	kurtosis	q10	q25	q75	q90
demand shock z_{it}	0.00	0.17	-12.55	411.08	-0.12	-0.03	0.06	0.11
supply shock $\Delta\omega_{it}$	0.04	0.13	2.37	179.48	-0.07	-0.01	0.08	0.14

shocks exhibits negative skewness, whereas the supply shock distribution is almost symmetric. The distribution of the demand shocks is centered at zero while the supply shocks are centered at 4%, which indicates that demand shocks have redistributive effects whereas supply shocks have both redistributive and growth effects. Moreover, significant differences are also observed in both the dispersion and the thickness of tails. The distribution of demand shocks is much more dispersed, less in terms of standard deviation (which is only 30% higher) than in terms of kurtosis, which is more than 2.3 times higher than in case of supply shocks. The mixture of both negative skewness and high kurtosis of demand shocks is apparent in the positional statistics: i.e., while the 75th and 90th percentiles are quite similar across the distributions (and are even higher for supply shocks), the 10th and 25th percentiles are significantly lower for demand shocks.¹²

There are also noticeable differences in the persistence of the two shocks, measured as an autoregression coefficient in a pooled regression. The persistence of supply shocks is low and is even negative, at -0.05; although it is significantly different from zero (with a standard error equal to 0.001).¹³ By contrast, demand shocks are much more persistent: i.e., the autoregression coefficient is 0.70 (0.001), and is thus slightly lower than that reported in Foster et al. (2008). Additionally, the demand and supply shocks are almost orthogonal, as the correlation coefficient between the two is -0.04. These findings show our assumptions regarding the different timings of the shocks are reasonable (a high correlation would suggest that demand shocks as well as supply shocks affect inventory changes). The correlation between the shocks calculated for each period separately varies in time, and the time pattern of correlation is slightly anticyclical: e.g., the correlation with the HP-adjusted log of sales is -0.24.

The next section will focus on the responses of various variables, most notably sales, to the identified shocks. Here, we investigate the relationship between shock volatility and sales volatility. Our methodological approach does not allow us to perform the variance decomposition exercise. Instead, we use a simple statistic: namely, the ratio of a standard deviation of a shock to a standard deviation of sales.¹⁴ Figure 3 presents the cross-firm distributions of these measures. In most firms, the volatility of demand shocks constitutes a relatively small fraction of the overall sales volatility: i.e., in half of the firms, the volatility of demand shocks is at most 3% of sales volatility; and in only 13% of the firms, the volatility of demand shocks exceeds 10% of sales volatility. Combined with a relatively large dispersion of demand shocks, these findings suggest that there needs to be a substantial correlation between the volatility of sales and the volatility of demand shocks. This is indeed the case here, as the slope of a simple regression¹⁵ is 1.3 with a standard error equal to 0.009.

The relative volatility of supply shocks is generally higher (as presented in Figure 3). In half of the firms, the volatility of supply shocks constitutes at most 7.5% of sales volatility; whereas for

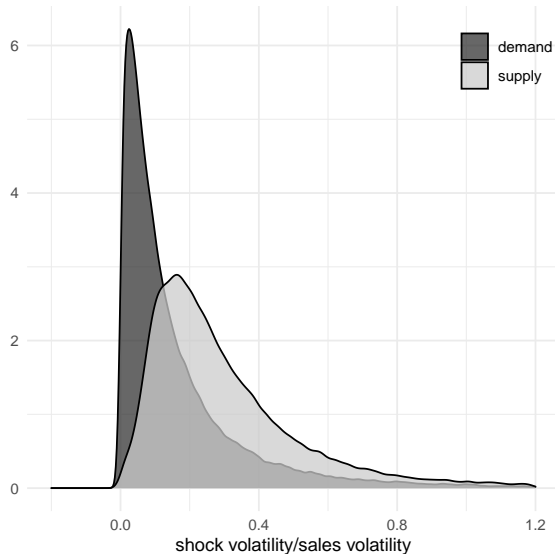
¹²It is worth emphasizing that the magnitude of the demand shocks is relatively large, as the size of about 10% of the negative demand shocks exceeds 12% of sales. A similar pattern is found for the positive demand shocks.

¹³It is important to keep in mind that supply shocks are defined as $\Delta\omega_{it}$, and, in turn, that the productivity ω_{it} is a highly persistent variable, with a persistence coefficient equal to 0.99. For reference, the persistence of log real sales is 0.95. Foster et al. (2008) also found high persistence in the levels of productivity.

¹⁴The volatility of shocks and the volatility of sales are positively related, as presented in Figure A.2 in the appendix.

¹⁵It can be interpreted as a sort of pass-through coefficient that reflects how much output volatility is, on average, generated by shock volatility.

Figure 3: The ratio of shock volatility to sales volatility



a large fraction (35%) of firms, the volatility of supply shocks exceeds 10% of sales volatility. The regression of the volatility of sales shocks on the volatility of sales yields a higher multiplier, 1.39 (0.01), suggesting a higher pass-through of supply shock volatility to sales volatility.

4 Firm responses to shocks

After discussing the properties of demand and supply shocks, we now focus on the dynamic responses of important variables to these shocks. As we discussed in section 1.3, we use linear projections to identify an impulse-response function (IRF) based on the data. We decided to calculate IRFs of a relatively large subset of variables in order to gain a comprehensive picture of firm-level adjustments to either a temporarily higher demand for the firm’s products or to an improvement in the firm’s productivity.

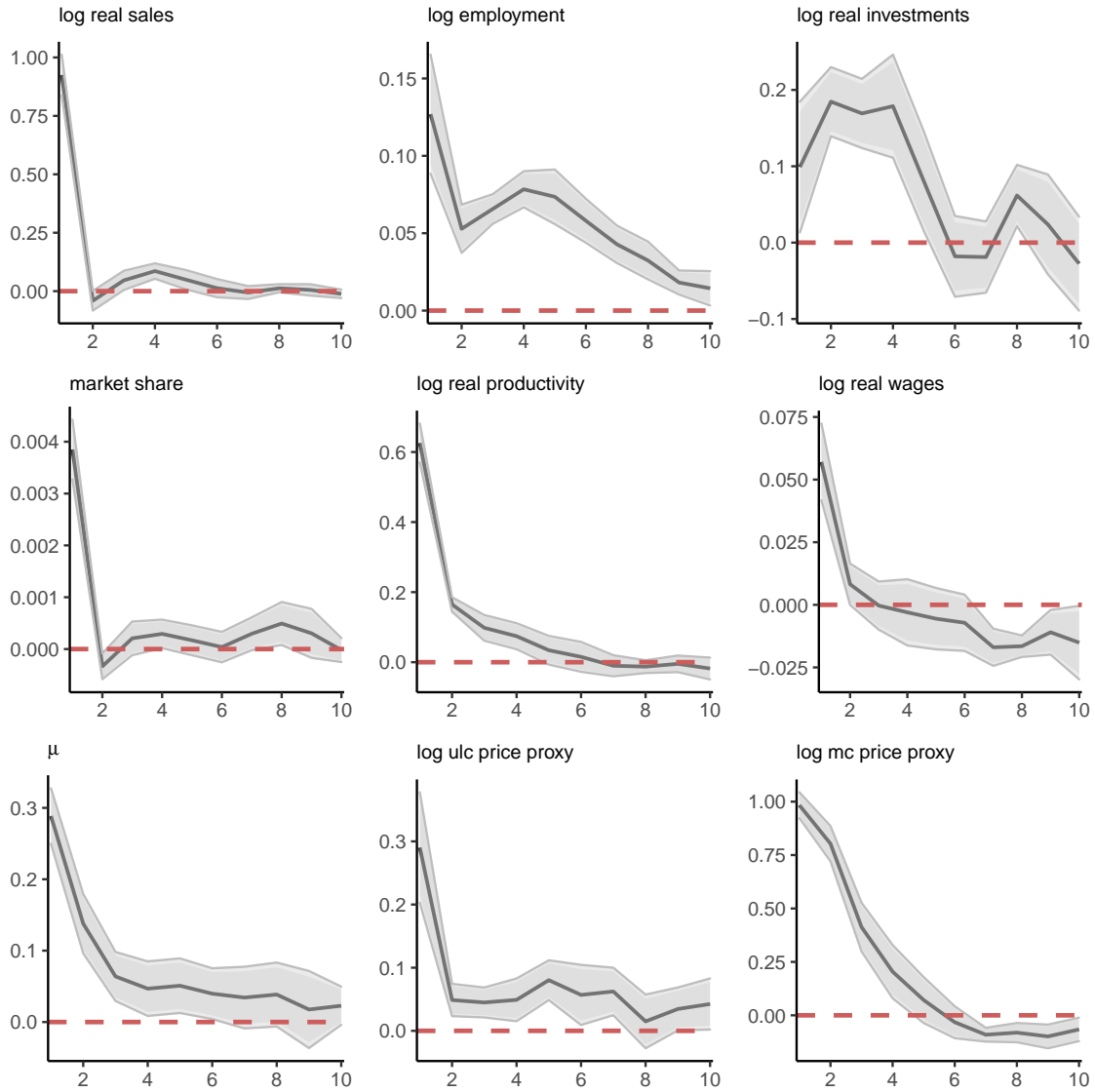
We have chosen nine indicators as response variables. In addition to real sales, we also present the results for market share, displaying the firm’s outcomes relative to the industry. Moreover, we calculate impulse responses for employment and investments to gauge the extent to which firms are adjusting their factors of production. We also show the behavior of labor productivity (log of value added per person employed), and how it translates into real average wages (calculated as the compensation of employees over employment). Additionally, we estimate the responses of markups in order to check if the shocks trigger the short-run changes in the firm’s monopoly power. As the shocks should induce the opposite reactions of prices, we also calculate the impulse responses of both proxies for prices, i.e., the ULC-based measure and the marginal-cost-based measure, which are described in section 1.4.

4.1 Demand shocks

Figure 4 displays the responses of a set of variables to a one-period firm-level demand shock. A positive demand shock triggers an increase in real sales. Interestingly, although the reaction of sales to the shock is large on impact (see the discussion in the previous chapter for a reference regarding the shock volatility), it exhibits almost no persistence. Although sales are significantly higher even five years after the shock, the size of the response is very small. The increase in sales lifts the

firm's output more than that of its competitors, and the firm's market share rises, but only in the period of the shock. Higher sales in response to the shock translate into higher employment. A substantial share of this additional employment is probably short-term, as the level of employment drops a year after the shock. Employment returns to its initial level very slowly, staying slightly elevated even 10 years after the shock. This pattern indicates that at least a portion of the initial increase in employment is permanent, most likely because employment protection, which is quite strong in Poland¹⁶, prevent firms from reducing their employment levels.

Figure 4: Responses to demand shocks



Notes: The dark and light gray areas represent 90% and 95% confidence intervals, respectively. The horizontal axis is expressed in years.

A positive demand shock triggers the firms to increase its investments. The firm's purchases of capital goods remain elevated levels for about four years, which increases the efficiency of its

¹⁶The various dimensions of the OECD's Employment Protection Legislation index usually rank Poland among the countries with relatively high labor protection.

production processes. Hence, the firm’s labor productivity (measured as real value added per employee) rises. It increases substantially in response to the shock (technically, sales – and hence the value added – increase more than employment), and stays at an elevated level for four years after the shock. But the increase in productivity is not permanent, and productivity eventually returns to the pre-shock level. An increase in labor productivity translates into higher real wages. However, higher productivity only partially translates into wage increases. The short-term pass-through of productivity into wages is close to 0.1 and declines thereafter. Higher wages persist for the first year only, then return to the pre-shock level, and eventually stay on a level slightly lower than in the period before the shock.

Following an expansion in the demand for a firm’s products, the markups that the firm charges rise. While the initially large increase in markups is short-lived, the markups are somewhat elevated three to six years after the shock, and then return to the pre-shock level. Our observation that markups increase in response to a positive demand shock is consistent with the findings of Nekarda and Ramey (2020) who used macroeconomic data to identify markups and their responses to demand shocks. Moreover, in line with sectoral oligopolistic models of the economy (see especially Burstein et al., 2020), we find a positive correlation between markups and market share, conditional on demand shocks.¹⁷

Finally, the evolution of both proxies for prices suggests that firm-level prices tend to rise after a demand shock. However, the two measures do not provide a consistent picture of the inertia of the price response. The ULC-based measure indicates a sharp and immediate increase of the price, but the proxy stays on a slightly elevated level up to seven years after the shock. In contrast, there is a higher inertia of an increase in price, measured using marginal costs. The mc-based proxy indicates prices are above the initial level up to five years after the shock. Summing up, prices tend to increase temporarily after the demand shock.

4.2 Supply shocks

We defined a supply shock as a one-period increase of the change in the TFP. Figure 5 displays the path of TFP (Ω_{it}) after the shock (estimated using the local projection method). There are two observations that we need to highlight. First, the shock does not permanently shift the path of TFP. Second, although the persistence of the shock is very low, TFP remains at an elevated level for about six years.

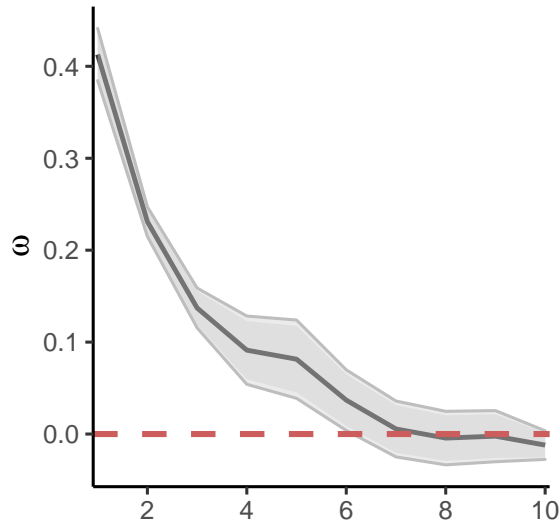
The responses of our variables of interest to supply shocks are presented in Figure 6. The first observation that emerges is that compared to the demand shocks, the firms’ responses to supply shocks are much more persistent.

Sales increase after the supply shock, remain higher for about five years, and then return to the pre-shock levels. The response of the market share closely mirrors that of sales. The shock we identified and used in the estimation of impulse responses does not affect permanently the level of TFP, thus sales and the market share return to the level before the shock. It suggests that the firm’s initial efficiency improvement vanishes after six years, driving down its market share and its sales.

A positive supply shock drives the demand for production factors, leading to increases in both employment and investment. In contrast to demand shocks, which triggered the highest employment response in the short run, there is almost no short-term change in employment after the supply shocks. However, employment rises substantially in the second period, and then slowly returns to the pre-shock level, reaching it seven years after the shock. It appears that given the labor market frictions and lags associated with finding new employees, firms adjust relatively slowly to supply shocks. The changes in investments are concentrated in the second and third

¹⁷The unconditional correlation of markups and market share is 0.005.

Figure 5: Response of TFP, Ω_{it} , to a supply shock



Notes: The dark and light gray areas represent 90% and 95% confidence intervals, respectively. The horizontal axis is expressed in years.

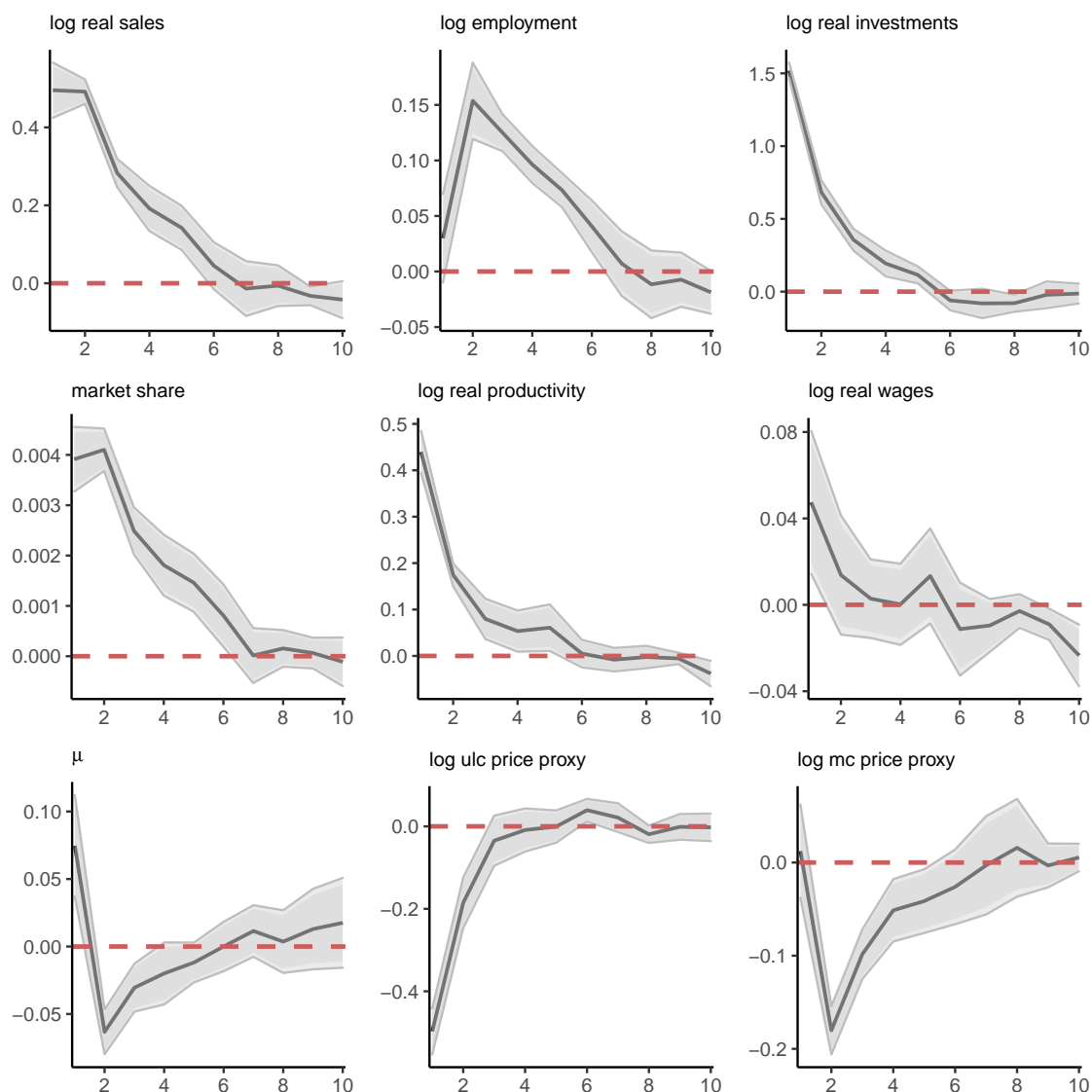
periods after the shock, which is also understandable given the time needed to adjust the firm’s capital stock. Importantly, both labor and investments rise after a productivity shock, indicating a conditional complementarity between labor and capital after a supply shock.

Labor productivity also rises after a supply shock, staying above the pre-shock level for about five years. Likewise in the case of a demand shock, the increase in labor productivity only partially translates into real wages, which are higher only immediately after the shock. Moreover, there is a high dispersion of the evolution of wages after the supply shock across firms, and the confidence bands associated with the estimate are wide. The pass-through of productivity into real wages is again close to 0.1. The incomplete pass-through of productivity shocks to wages was also found by Guiso et al. (2005), who attributed it to risk-sharing considerations.

Markups tend to rise immediately after the supply shock, and to decline shortly thereafter. The adjustment is short-lived and four years after the shock the markups return to the pre-shock level. Again, as predicted in the theoretical model of Burstein et al. (2020), markups tend to move in line with the market share, at least in the short run. Moreover, our short-run results are again consistent with the findings in Nekarda and Ramey (2020), who found using a macroeconomic identification scheme that markups are procyclical, conditional on TFP shocks. However, they found that the positive response of markups have relatively more inertia.

Finally, our results provide quite consistent evidence on the behavior of prices after the productivity shock. The ULC-based measure declines immediately after the shock, and then for the next two years it returns to the pre-shock level. The marginal-cost-based measure, which includes not only the information on markups and wages, but also the costs of capital and intermediates, does not change in the moment of the shock, and declines a year after the shock. The prices stay on a reduced level up to five years after the shock and then return to the pre-shock level. Summing up, prices tend to decrease temporarily after the supply shock.

Figure 6: Responses to supply shocks



Notes: The dark and light gray areas represent 90% and 95% confidence intervals, respectively. The horizontal axis is expressed in years.

Conclusions

The aim of our study was 1) to identify shifts in the demand and productivity schedules of individual firms; 2) to describe the properties of these shocks; and 3) to estimate the empirical granular dynamic responses of variables like sales, employment, and investments to these shocks. We used a comprehensive dataset from the Polish enterprise sector, that covers more than 80% of the whole output and employment of the country's enterprise sector. These data allowed us to draw conclusions that are representative from a macroeconomic standpoint. However, as our data do not contain information on physical output and prices, we used the identification scheme of demand shocks proposed by Kumar and Zhang (2019). It extracts the unexpected demand shocks from the time variation of inventories and sales. The direct observation of demand shocks also helped us to address the problematic issues in the production function estimation implied by the lack of

price data under imperfect competition. Thus, in the control function approach to the estimation – which is discussed extensively in, for example Akerberg et al. (2015) – we controlled for the heterogeneity of demand changes, as recommended by Doraszelski and Jaumandreu (2021). We identified the dynamic responses of a set of variables to temporary demand and supply shocks using a projection method pioneered by Jordà (2005). To our knowledge, our study is the first attempt to identify dynamic impulse responses using a fully granular identification scheme. Moreover, we use the information from the estimation of the production function to construct the proxies for individual prices and we measure their responses to the two shocks.

We showed that although the distributions of the two shocks have quite regular standard deviations, their tails are very fat. The two distributions also differ in terms of skewness; i.e., the productivity shocks are approximately symmetrically distributed, whereas the demand shocks are negatively skewed. Moreover, the demand shocks are persistent, while the supply shocks are not. The empirical analysis revealed that the changes in firms’ outcomes in response to productivity shocks are much more persistent than those in response to demand shocks. Demand shocks result in short-lived increases in output, market share, productivity, real wages, and markups; and to increases in investments, and employment for a couple of periods. Firms’ reactions to supply shocks are in most cases qualitatively similar, but they are more persistent. Our analysis showed that prices tend to increase temporarily after demand shocks, and to decrease temporarily after the supply shock, as predicted by theoretical models and found by Pozzi and Schivardi (2016) in a firm-level analysis. Moreover, regardless of the nature of the shock, the resulting increases of labor productivity only partially translate into higher wages.

The differences we found in the persistence of both shocks and responses to shocks are consistent with the results of many macroeconomic models (in which the persistence of the technology shocks is usually assumed in the calibration, like, e.g., in the canonical DGSE models, see Smets and Wouters, 2003). The positive reactions of output and demand for production factors we observed are also consistent with the findings of many macroeconomic models. In addition, other observations we made are consistent with assumptions that are frequently found in macroeconomic models. In particular, our finding that markups increase after a productivity shock suggests that prices are sticky, as rising productivity does not fully translate into prices. By contrast, the limited impact of both productivity and demand shocks on wages we found is consistent with nominal wage rigidities. Moreover, the sluggish response of employment to both types of shocks points to the importance of labor market frictions, which are, for example, emphasized in search models.

References

- ACKERBERG, D., L. BENKARD, S. BERRY, AND A. PAKES (2007): “Econometric Tools for Analyzing Market Outcomes,” in *The Handbook of Econometrics*, ed. by J. Heckman and E. Leamer, Amsterdam: North-Holland, vol. 6A, 4171–4276.
- ACKERBERG, D., K. CAVES, AND G. FRAZER (2015): “Identification Properties of Recent Production Function Estimators,” *Econometrica*, 83, 2411–2451.
- ATKESON, A. AND A. BURSTEIN (2008): “Pricing-to-Market, Trade Costs, and International Relative Prices,” *American Economic Review*, 98, 1998–2031.
- AUERBACH, A. J. AND Y. GORODNICHENKO (2013): “Output Spillovers from Fiscal Policy,” *American Economic Review*, 103, 141–146.
- BACHMANN, R. AND P. ZORN (2020): “What Drives Aggregate Investment? Evidence from German Survey Data,” *Journal of Economic Dynamics and Control*, 115, 103873.

- BOND, S., A. HASHEMI, G. KAPLAN, AND P. ZOCH (2020): “Some Unpleasant Markup Arithmetic: Production Function Elasticities and Their Estimation from Production Data,” Working Paper 27002, National Bureau of Economic Research.
- BUERA, F. J., J. P. KABOSKI, AND R. M. TOWNSEND (2021): “From Micro to Macro Development,” *Journal of Economic Literature*.
- BURSTEIN, A., V. M. CARVALHO, AND B. GRASSI (2020): “Bottom-up Markup Fluctuations,” Working Paper 27958, National Bureau of Economic Research.
- CARLSSON, M., J. MESSINA, AND O. NORDSTRÖM SKANS (2021): “Firm-Level Shocks and Labour Flows,” *The Economic Journal*, 131, 598–623.
- CHO, D. (2018): “The Labor Market Effects of Demand Shocks: Firm-Level Evidence from the Recovery Act,” Tech. rep.
- COCHRANE, J. H. (1994): “Shocks,” *Carnegie-Rochester Conference Series on Public Policy*, 41, 295–364.
- COVIELLO, D., I. MARINO, T. NANNICINI, AND N. PERSICO (2021): “Demand Shocks and Firm Investment: Micro-evidence from Fiscal Retrenchment in Italy,” *The Economic Journal*, ueab073.
- DE LOECKER, J. AND F. WARZYNSKI (2012): “Markups and Firm-Level Export Status,” *American Economic Review*, 102, 2437–2471.
- DORASZELSKI, U. AND J. JAUMANDREU (2021): “Reexamining the De Loecker & Warzynski (2012) Method for Estimating Markups,” CEPR Discussion Papers DP16027, CEPR.
- DRISCOLL, J. C. AND A. C. KRAAY (1998): “Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data,” *The Review of Economics and Statistics*, 80, 549–560.
- FOSTER, L., J. HALTIWANGER, AND C. SYVERSON (2008): “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, 98, 394–425.
- FRAUMENI, B. (1997): “The Measurement of Depreciation in the U.S. National Income and Product Accounts,” *Survey of Current Business*, 77, 7–23.
- GABAIX, X. (2011): “The Granular Origins of Aggregate Fluctuations,” *Econometrica*, 79, 733–772.
- GRADZEWICZ, M. (2021): “What Happens After an Investment Spike—Investment Events and Firm Performance,” *Journal of Business & Economic Statistics*, 39, 636–651.
- GUISSO, L., L. PISTAFERRI, AND F. SCHIVARDI (2005): “Insurance within the Firm,” *Journal of Political Economy*, 113, 1054–1087.
- JORDÀ, Ò. (2005): “Estimation and Inference of Impulse Responses by Local Projections,” *American Economic Review*, 95, 161–182.
- JORDÀ, Ò., M. SCHULARICK, AND A. M. TAYLOR (2015): “Betting the House,” *Journal of International Economics*, 96, S2–S18.
- JORGENSEN, D. W. (1963): “Capital Theory and Investment Behavior,” *The American Economic Review*, 53, 247–259.
- KAHN, J. A. (1992): “Why Is Production More Volatile than Sales? Theory and Evidence on the Stockout-Avoidance Motive for Inventory-Holding,” *The Quarterly Journal of Economics*, 107, 481–510.
- KLETTE, T. J. AND Z. GRILICHES (1996): “The Inconsistency of Common Scale Estimators When Output Prices Are Unobserved and Endogenous,” *Journal of Applied Econometrics*, 11, 343–361.
- KUMAR, P. AND H. ZHANG (2019): “Productivity or Unexpected Demand Shocks: What Determines Firms’ Investment and Exit Decisions?” *International Economic Review*, 60, 303–327.

- MARSHAK, J. AND W. H. ANDREWS (1944): “Random Simultaneous Equations and the Theory of Production,” *Econometrica*, 12, 143–205.
- NEKARDA, C. J. AND V. A. RAMEY (2020): “The Cyclical Behavior of the Price-Cost Markup,” *Journal of Money, Credit and Banking*, 52, 319–353.
- OLEA, J. L. M. AND M. PLAGBORG-MØLLER (2021): “Local Projection Inference Is Simpler and More Robust Than You Think,” *Econometrica*, 89, 1789–1823.
- OLLEY, G. S. AND A. PAKES (1996): “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 64, 1263–1297.
- POZZI, A. AND F. SCHIVARDI (2016): “Demand or Productivity: What Determines Firm Growth?” *The RAND Journal of Economics*, 47, 608–630.
- RAMEY, V. A. (1991): “Nonconvex Costs and the Behavior of Inventories,” *Journal of Political Economy*, 99, 306–334.
- RIDDER, M. D., B. GRASSI, AND G. MORZENTI (2021): “The Hitchhiker’s Guide to Markup Estimation,” Tech. Rep. 677, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.
- ROBERTS, M. J., D. YI XU, X. FAN, AND S. ZHANG (2018): “The Role of Firm Factors in Demand, Cost, and Export Market Selection for Chinese Footwear Producers,” *The Review of Economic Studies*, 85, 2429–2461.
- SBORDONE, A. M. (2002): “Prices and Unit Labor Costs: A New Test of Price Stickiness,” *Journal of Monetary Economics*, 49, 265–292.
- SMETS, F. AND R. WOUTERS (2003): “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area,” *Journal of the European Economic Association*, 1, 1123–1175.

Appendices

A Additional graphs

Figure A.1: Cobb-Douglas elasticities (cross-sectional dispersion)

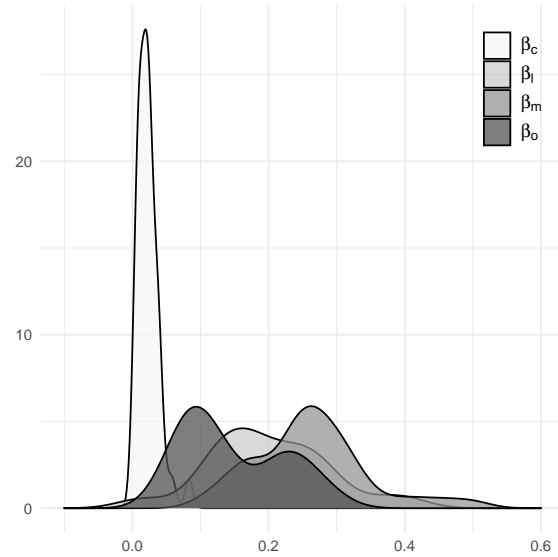
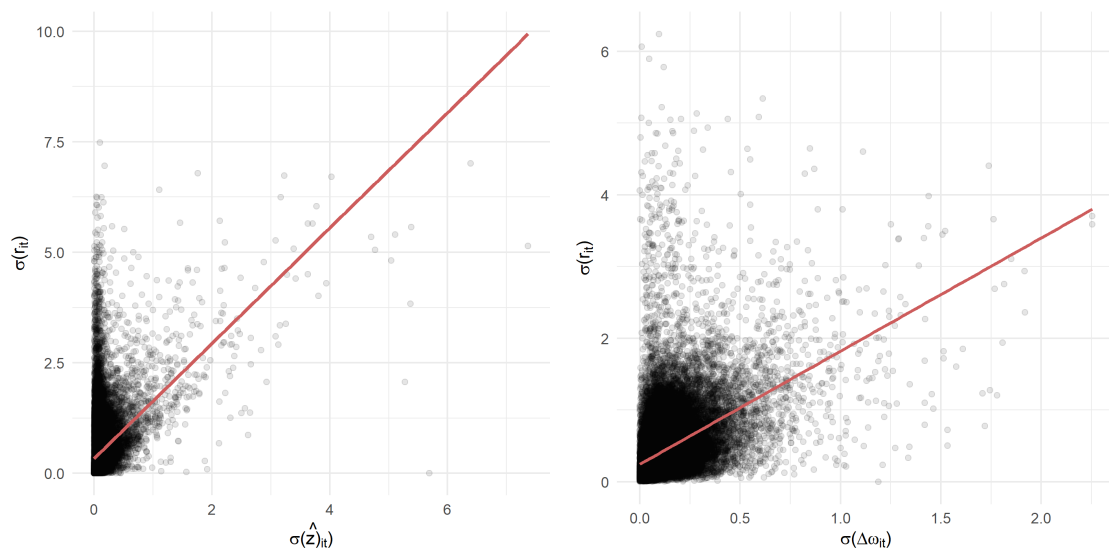


Figure A.2: Sales volatility and demand (right panel) and supply (left panel) shock volatility

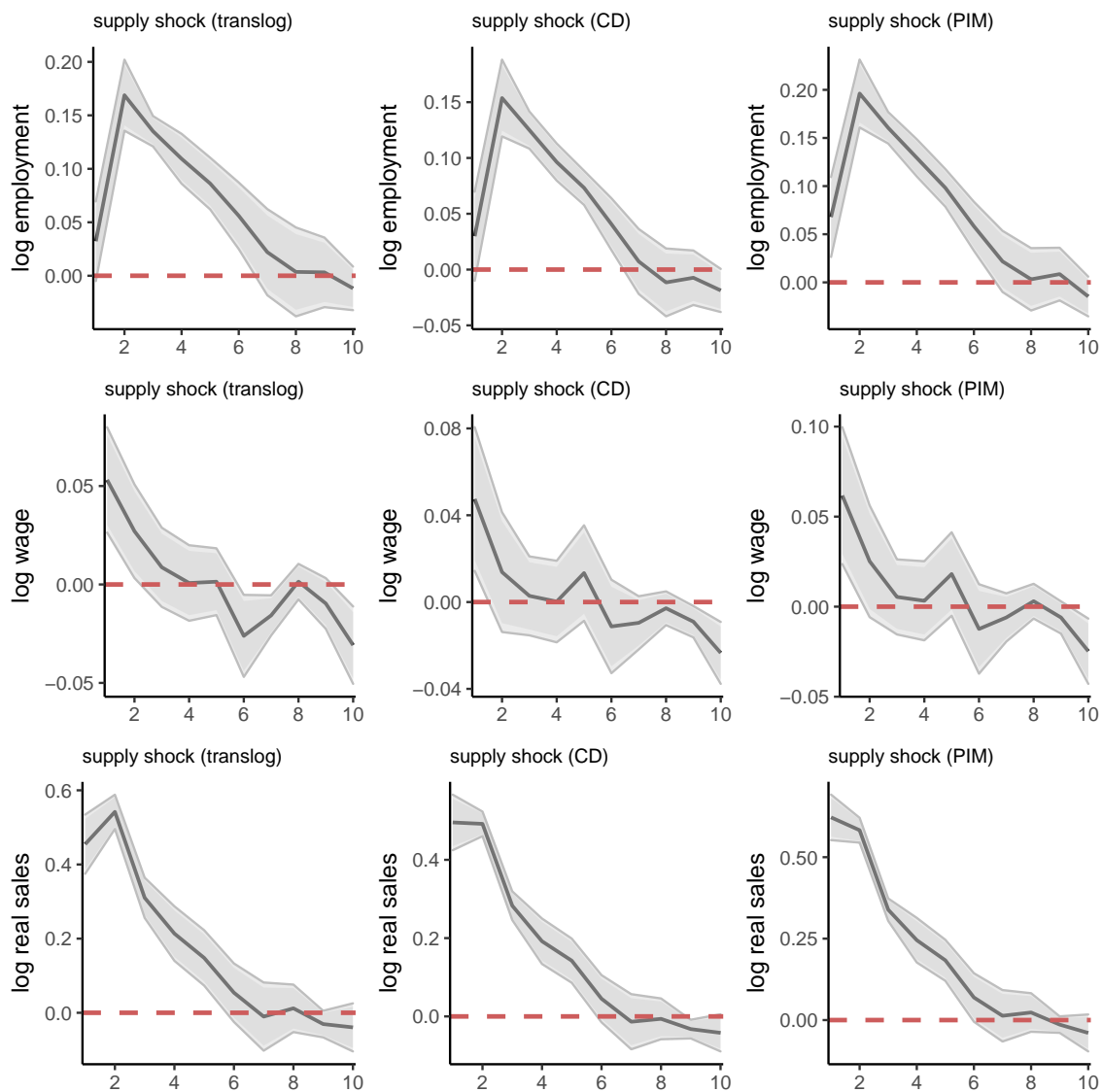


B Robustness of the main results

B.1 The identification of supply shocks

The identification scheme used to detect shocks in the data leaves more room to maneuver in choosing details of econometric procedure for productivity shock than for demand shocks. Here, we present our main results limited to the impulse responses of the most important variables (due to space considerations), with important changes in the identification scheme of the supply shock: 1/ using the translog production function instead of Cobb-Douglas, and 2/ using the Perpetual Inventory Method (PIM) to measure capital (in both cases, holding the other elements of the identification scheme constant and identical to the baseline case).

Figure B.1: Robustness: responses of log employment, log wage and log sales to supply shocks in the benchmark Cobb-Douglas case versus translog and PIM capital



Notes: Cobb-Douglas is used in the baseline; translog – unrestricted coefficients translog production function; PIM – Cobb Douglas production function with capital calculated using the perpetual inventory method.

Table B.1: Robustness - properties of supply shocks

supply shock	mean	sd	skewness	kurtosis	q10	q25	q75	q90
baseline	0.04	0.13	2.37	179.48	-0.07	-0.01	0.08	0.14
translog	0.03	0.11	3.90	342.82	-0.05	-0.01	0.07	0.12
PIM	0.03	0.12	0.53	68.05	-0.07	-0.01	0.08	0.13

Notes: Cobb-Douglas is used in the baseline; translog – unrestricted coefficients translog production function; PIM – Cobb Douglas production function with capital calculated using the perpetual inventory method.

In order to check the robustness of our results for the shape of the production function we assume an unrestricted translog with all squared and cross terms for all production factors. The parameters are estimated separately within 1-digit NACE groupings, using fixed effects (within transformation) in the identification. Most of the parameters are identified in the first step of the control function approach, with only linear and square capital terms left for the second stage. The first column of Figure B.1 shows the impulse responses of log employment, wages, and sales to supply shocks identified using the translog function, compared with the benchmark Cobb-Douglas case in the second column. The comparison shows that our main results are robust to the shape of the production function, which can be attributed to the high correlation of both measures of supply shocks (0.897). Table B.1 shows that most of the properties of the distributions of shocks in the translog case are similar to the baseline Cobb-Douglas case.

Using the perpetual method requires: 1/ a starting point for capital (we use the value of capital in the first period for the company in the data), 2/ depreciation rates (we use separate depreciation rates for buildings, machinery, and vehicles that follow Fraumeni, 1997), and 3/ a long and continuous time series of investments. We tried to minimize the loss of observations generated by short spells of missing data due to companies being out of the sample for one or two years. In these cases we imputed investments using the mean values from the adjacent periods. Despite these efforts, the effective number of observations of capital decreased by 17%, and the effective number of supply shocks identified declined by 13.2%. The correlation of supply shocks identified using the PIM definition of capital with the supply shocks in the baseline case proved to be high, at 0.854. Table B.1 indicates that most of the properties of the distributions of the two types of shocks were similar. The third row of Figure B.1 shows that the impulse responses hardly changed. Thus, our main results are robust to the change in the definition of capital.

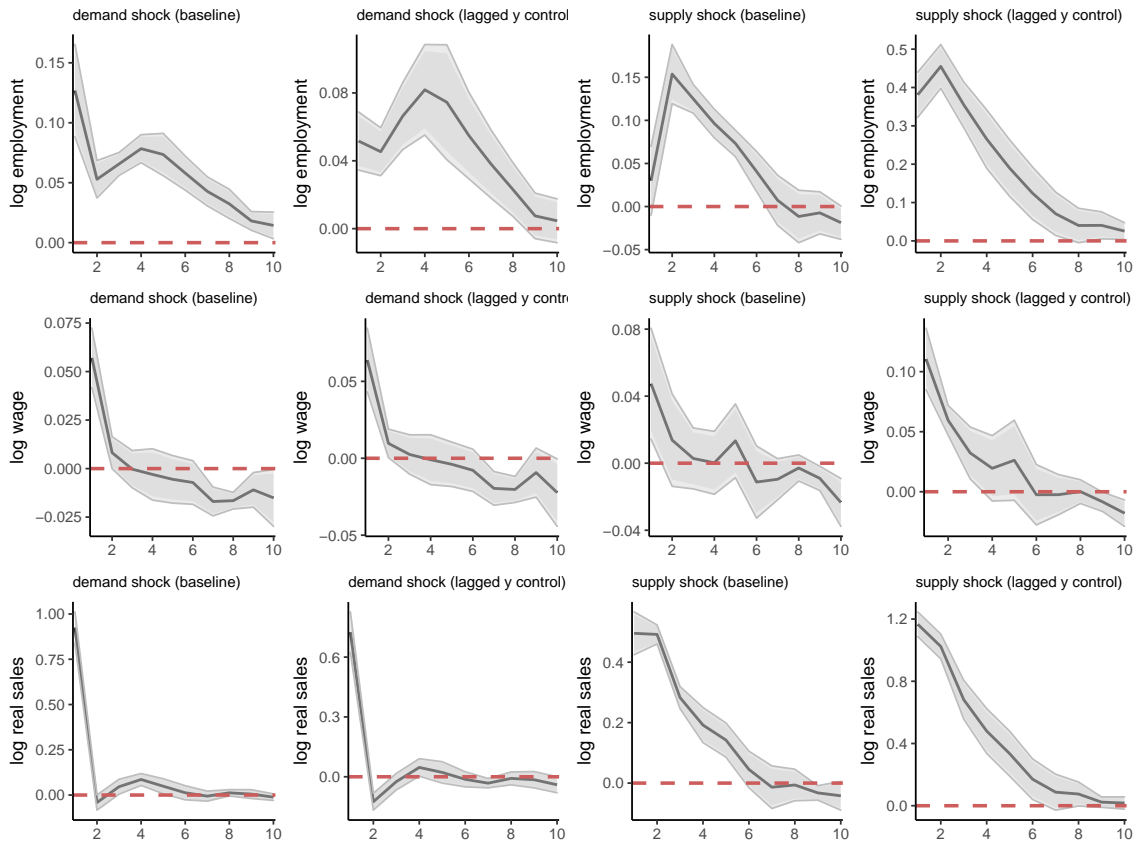
B.2 Controlling for the lagged dependent variable

A recent analysis by Olea and Plagborg-Møller (2021) showed that controlling for the lagged dependent variable leads to a more robust inference in applied local projections. This means that a set of equations defined in (13) becomes:

$$y_{it+h} = \alpha_{ih} + \gamma_t + \rho y_{it-1} + \beta_h shock_{it} + \epsilon_{it+h}, \text{ for } h \in 0, 1, \dots, H-1, \quad (17)$$

The augmentation is particularly relevant with highly persistent data and lag-augmented local projections with normal critical values that are asymptotically valid with both stationary and non-stationary data, and that also over a wide range of response horizons. Figure B.2 compares the impulse response functions in the baseline case with the impulse response functions derived with lag-augmented local projections, and shows that the general shape of responses is not affected by the assumptions of the estimation method.

Figure B.2: Robustness: responses of log employment, log wage and log sales to supply and demand shocks in the baseline specification and in the specification augmented for the lagged dependent variable



Notes: The first and third columns show the responses of selected variables in the baseline case, while the second and the fourth columns show the responses of selected variables in the specification augmented for the lagged dependent variable, as suggested in Olea and Plagborg-Møller (2021).