

Gottfries, Nils; Mickelsson, Glenn; Stadin, Karolina

Working Paper

Deep Dynamics

CESifo Working Paper, No. 8873

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Gottfries, Nils; Mickelsson, Glenn; Stadin, Karolina (2021) : Deep Dynamics, CESifo Working Paper, No. 8873, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/232470>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Deep Dynamics

Nils Gottfries, Glenn Mickelsson, Karolina Stadin

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Deep Dynamics

Abstract

How do firms adjust their output, inventories, employment and capital in response to demand-side shocks? To understand this, we estimate a reduced-form model using firm-level panel data and we construct a theoretical model that can match the estimated impulse-response functions. A combination of convex adjustment costs and implementation lags explains input adjustment very well. Although inputs adjust slowly, production responds quickly to the demand shock and this adjustment is explained by a combination of increasing returns and increased utilization of the production factors. To avoid stock-outs, firms increase their inventories when demand increases.

JEL-Codes: E220, E230, E240, E320.

Keywords: production function, productivity, Solow residual, labor hoarding, effort, organizational capital, capacity, returns to scale, markup, inventory investment.

Nils Gottfries
Department of Economics, Uppsala University
Uppsala / Sweden
nils.gottfries@nek.uu.se

Glenn Mickelsson
Handelsbanken
Stockholm / Sweden
glenn.mickelsson@gmail.com

Karolina Stadin
Department of Economics, Uppsala University
Uppsala / Sweden
karolina.stadin@nek.uu.se

This is a substantially revised version of WP 2018:10 from the Department of Economics, Uppsala university. We have received helpful comments from Martin M. Andreasen, Susanto Basu, Mikael Carlsson, Daria Finocchiaro, Peter Fredriksson, Georg Graetz, Bertil Holmlund, Per Krusell, James Kahn, Jesper Lindé, Florin Maican, Fabio Schiantarelli, Ingvar Strid, Karl Wallentin, Erik Öberg, and seminar participants at the CESifo Conference of Maroecconomics and Survey Data, NASMES, the Dynare Conference, Greater Stockholm Macro Group, Ratio Institute, Stockholm School of Economics, Boston College, and the Universities of Lund, Penn State, Stockholm, Uppsala, and Örebro. We received excellent research assistance from Irina Andone and financial support from the Wallander-Hedelius-Browaldh Foundation and the Marianne and Marcus Wallenberg Foundation.

1. Introduction

Understanding how firms in general react to shocks is important for understanding business cycles and the role of stabilization policy. Alternative theories provide potential explanations of key stylized facts, such as the pro-cyclicality of investment, labor input, factor productivity, and inventory holdings and these theories have been incorporated into DSGE models that have been estimated on macro data. In this paper, we investigate the relevance of some of these theories using panel data for manufacturing firms. We study how firms react to specific demand-side shocks and we obtain estimates of key structural parameters that can serve as benchmarks in the construction of macroeconomic models

Many studies have documented pro-cyclical factor productivity and the Solow residual has been used to measure technology shocks (Prescott, 1986) but many researchers have questioned the interpretation of the Solow residual as a measure of technology shocks.¹ Hall (1988) considered variations in labor input and production that arise due to shocks that should be uncorrelated with technology shocks (military spending, oil prices, and the political party of the president). He showed that variations in labor input that are associated with these shocks lead to more than proportional changes in production and he interpreted this as evidence of strongly increasing returns to scale. With increasing returns, firms will make losses if the price is equal to the marginal cost, and since firms typically do not make losses, Hall concluded that firms must have very substantial market power. Methods similar to Hall's method have recently been applied by e.g. de Loecker and Warzynski (2012), Akerberg, Caves and Frazer (2015) and de Loecker, Eeckhout and Unger (2020).

An alternative explanation is that demand shocks lead to variations in factor utilization that shows up as pro-cyclical factor productivity as we measure it. To explain why firms do not always make full use of their production factors, there must be some costs that limit factor utilization. This cost may be a direct cost of extra effort as in the models of Burnside, Eichenbaum and Rebelo (1993), Sbordone (1996, 1997), and Basu, Fernald and Kimball (2006). Alternatively, working time that is directed to current production may come at the expense of activities that increase *future* production such as maintenance and training (Fay and Medoff, 1985, Bean, 1990, Kim and Lee, 2007). Variations in effort spent on current production may generate variations in output per worker that look like technology shocks or increasing returns to scale.

¹ Hart and Malley (1999), Baily, Bartelsman and Haltiwanger (2001), and Field (2010) document pro-cyclical productivity for different countries and time periods. The literature on the “paradox of short run increasing returns to labor” goes back many years; see Fay and Medoff (1985) and Biddle (2014) for reviews.

Another stylized fact is that inventory holdings are pro-cyclical. If the main role of inventories were to smooth production, we would expect inventories to decrease in periods of high demand. However, inventory investment is pro-cyclical and some researchers have viewed this as an indication that the cost of producing must be low in boom periods due to technology shocks or positive externalities (Christiano, 1988, Blinder and Maccini, 1991, Khan and Thomas, 2007). Alternatively, the stock-out avoidance motive can explain pro-cyclical inventory investments in the face of demand-side shocks (Kahn 1987, 1992; Bils and Kahn, 2000). The basic idea is that firms need to have stocks of finished goods on the shelves in order to sell, and to satisfy higher demand, they need to have more goods on the shelves. More recently, Galeotti, Maccini and Schiantarelli (2005), Wen (2005) and Kryvtsov and Midrigan (2013) have found support for the stock-out avoidance theory.

In this paper, we use firm-level panel data to investigate how firms respond to specific demand-side shocks and we build a theoretical model that can match those responses. Our study does not address the question of the relative importance of supply and demand shocks for business cycle fluctuations, but we obtain estimates of key structural parameters that can serve as benchmarks in the construction of macroeconomic models. We find slow adjustment of labor and capital and this adjustment can be well explained by a combination of convex adjustment costs and implementation lags (time to build). Still, production responds immediately to the demand shock and a large fraction of this adjustment is achieved by increasing utilization of the production factors for current production. Increasing returns to scale also play a role. Inventories respond strongly to demand-side shocks because intermediate goods are necessary for production and because firms increase the stock of finished goods to avoid stock-outs when demand increases.

Our analysis proceeds in four steps. First, we use input-output tables and aggregate data to construct a product market *demand index* that varies across firms because the shares of production that are used for consumption and investment vary across industries and because the share of the firm's production that is sold in the export market varies across industries and firms. This approach is similar to Hall (1988) and Kryvtsov and Midrigan (2013) in that we try to construct demand-side shocks that should be uncorrelated with technology and cost shocks that affect individual firms and industries.

Second, we try to capture the responses of firms to these shocks by estimating an empirical *reduced-form model* using panel data for manufacturing firms. The empirical model includes production, the inventory stock, the number of employees, the capital stock, and the firm-specific demand index. The endogenous variables depend on lags of all the endogenous

variables and on the demand index, which we take to be exogenous for the individual firm and industry. We include firm and time fixed effects and firm-specific trends in the estimation. We find that employment and especially the capital stock respond with substantial lags to the demand shocks that we have constructed but production and inventory holdings respond quickly. This implies positive responses of factor and labor productivity – as they are normally measured – to demand-side shocks.

Third, we construct a *theoretical model* that incorporates many of the explanations of demand-driven fluctuations in factor productivity and inventory holdings that the literature suggests. We assume that firms have market power and that there may be increasing returns to scale in production. Hiring and investment are subject to adjustment costs and implementation lags (time to build). Workers spend a fraction of their time on activities which increase “organizational capital” and future production, but do not contribute to production in the current period.² Firms hold inventories of two kinds: inputs and finished goods. Inventories of inputs are necessary for production and inventories of finished goods increase sales because a higher inventory stock reduces stock-outs.

In the fourth step, we investigate the relevance of these different theoretical mechanisms by *estimating the structural parameters* of our model. We follow the approach of Christiano, Eichenbaum and Evans (2005) by choosing the structural parameters to match the estimated impulse-response functions from our empirical reduced-form model. Distributions for the estimates are obtained by bootstrapping, i.e., resampling from the population of firms with replacement and re-estimating the parameters. We find that our theoretical model can explain the estimated responses very well.

We can learn five important lessons from this exercise. The first is that a combination of convex adjustment costs and implementation lags can explain the average response of capital and labor to demand-side shocks. These two components suffice well to match the average responses of manufacturing firms. Convex adjustment costs and implementation lags both play a role: without implementation lags, we cannot match the delayed response of the capital stock.³

Second, output responds immediately to the demand shock and the first-year response of production is three times as strong as that of employment while capital responds

² We also consider an alternative version where firms pay a convex cost for varying the utilization of both labor and capital.

³ Christiano, Eichenbaum and Evans (2005) and Eberly, Rebelo and Vincent (2012) use adjustment costs for investments to generate sluggish capital adjustment, but we find implementation lags and convex costs for adjusting the capital stock to be more plausible, and sufficient to match the average response of the firms.

very sluggishly. This is hard to explain without a very strong effect of the demand shock on factor utilization. Thus, we should not interpret all fluctuations in the Solow residual as technology shocks and studies of production dynamics may produce biased results if variations in factor utilization are disregarded.

Third, and equally important, if there are variations in factor utilization, there must also be some *cost* of utilization and this matters for how we estimate the markup. If we disregard factor utilization, we will interpret the large change in production relative to inputs as evidence of strongly increasing returns to scale and we will overestimate the markup because we omit the marginal cost of increased utilization.

Fourth, we find evidence of returns to scale, but increasing returns to scale are not sufficient to explain the dynamic relation between inputs and output. A demand shock generates an increase in labor productivity as we normally measure it, but like Sbordone (1996, 1997) we find that this increase is only temporary, which contradicts explanations based on internal or external economies of scale.

Finally, there is a strong “accelerator effect” of a demand shock on inventory investments and this response is well explained in a model where inputs are needed in order to produce and stocks of finished goods are needed in order to sell the goods. Thus, inventory investment amplifies the effects of demand shocks and contributes to the volatility of production.⁴

As far as we know, this is the first paper to estimate a structural model of the joint dynamics of production, inventory stock, employment and investment using panel data for individual firms. These decisions are intimately linked, so it makes sense to model them jointly.⁵ Our approach to identification follows Hall (1988), Shea (1992), Perotti (2008) and Nekarda and Ramey (2011) in that we try to isolate movements in the endogenous variables that are the result of demand-side shocks which should be orthogonal to productivity and cost shocks. In terms of the estimation, we follow Christiano, Eichenbaum and Evans (2005), estimating the structural parameters by matching empirical impulse-response functions. But contrary to these studies, we use micro data instead of macro or industry data.

⁴ Firm-level data do not allow us to distinguish different types of inventories, but the theoretical model has two types of inventories: inputs and finished goods.

⁵ Adjustment costs and implementation lags interact with utilization to determine the dynamic responses of inputs, production and the marginal cost and variations of the marginal cost are intimately linked with the evolution of the inventory stock (Bils and Kahn 2000). To increase the stock of finished goods, the firm needs to produce more for given sales, which leads to a higher marginal cost of production.

Copeland and Hall (2011) estimate structural parameters by matching the coefficients of a reduced-form model of production, sales, prices, and inventories which has been estimated on data for specific car models. This approach is similar but there are two main differences: we have an exogenous demand shock variable, so the reduced form is a recursive system and we match the impulse-response functions rather than the coefficients in the reduced form model. As we explain in the Appendix, we do not view the parameters of the reduced-form model as interpretable by themselves.

An alternative method would be to estimate the structural equations directly on the data. An example of this approach is Galeotti, Maccini and Schiantarelli (2005), who estimated a model of inventories, employment and hours worked on industry data. Aside from the fact that we use micro data, a main difference is that we do the estimation in two steps. In our view, there are two advantages of this approach compared to direct estimation of a structural model where one has to make specific assumptions about the unobserved shocks. First, we can remain agnostic about whether the unobserved shocks are shocks to technology, factor prices, competitors' prices, or something else. As explained in detail in the *Appendix*, the basic idea is that a reduced form model with lags of observable variables can mop up the dynamic effects of the unobserved shocks. Second, by estimating a reduced-form model, we let the data speak more freely, and by comparing the impulse-response functions in the theoretical model to their empirical counterparts, we can see which features of the model help to explain the dynamics. For example, we find that a model with adjustment costs but no implementation lags is unable to explain the delayed response of investments and that a model with increasing returns but without variations in utilization can explain the responses for some time horizons but not the whole profile of the impulse-response functions.

We present the data and the estimated reduced-form model in Section 2. The theoretical model is presented in Section 3 and Section 4 explains how we estimate the structural parameters. The results are presented in Section 5 and Section 6 concludes.

2. Data and empirical model

In this section, we present the firm-level data and the construction of the demand index followed by the presentation of the empirical model and the estimated impulse-response functions.

2.1 Firm-level data

The firm-level panel consists of yearly data from Statistics Sweden for all firms in Sweden. Our main sample consists of firms with at least ten employees in the manufacturing sector 1996-2016.⁶ Real production (Yr) is the value of the firm's total output deflated by the producer price for the industry. As a robustness check, we instead use real value added deflated by the value-added deflator for the industry (VAr) to measure production. The real inventory stock (Zr) is the value of the firm's inventories at the end of the year deflated by the producer price for the industry.

N is the number of persons employed by the firm, measured in terms of "full-year equivalents". This measure takes account of people who are working part-time or employed for only part of the year, but it does not take account of overtime. This means that variations in factor utilization in our theoretical model may reflect variations in official and unofficial hours worked and leisure on the job as well as other variations in utilization. However, employment and official hours worked are very closely correlated for manufacturing as a whole; the correlation between yearly changes in registered hours worked and the number of people employed is 0.91 (see *Figure A1* in the Appendix).

The real capital stock (Kr) consists of machines and buildings. Generous depreciation allowances imply that the book value is much lower than the economic value. Therefore, we constructed economic capital stocks using the perpetual inventory method as described in the *Appendix*.

Firms may use workers who are employed by staffing companies or sub-contractors and they may rent some of the capital that they use. We cannot take account of this because of lack of data. This means that we treat these inputs as material inputs, which are assumed to be proportional to production.

2.2 Sample selection

In this study, we are interested in profit-maximizing firms that produce goods using labor and capital and that have inventory stocks consisting of inputs and goods they have produced, rather than firms, which are primarily engaged in trading goods. For this reason, we chose to study firms in the manufacturing sector (industries 15-36 according to SNI92) with at least ten employees and a production value of at least five yearly wages of workers with high-school in

⁶ We use the FAD units from Statistics Sweden to identify firms. FAD units are based on legal organizational numbers, but the FAD number changes when there are mergers or splits affecting more than 50 percent of the workforce, even if the legal organizational number remains the same. The estimated impulse-response functions are similar if we simply use organizational numbers.

all their years of existence. We exclude very small firms for two reasons. First, export data are missing for many small firms, and second, the dynamic responses of very small firms may be different from those of medium-sized and large firms.⁷ Publicly owned firms are dropped because they may have different objectives than privately owned firms. Also, we exclude firms that in some period had a non-positive value for the capital or inventory stock.

Descriptive statistics for the resulting sample are shown in *Table 1*.

To limit the influence of measurement errors, we exclude firms that in some period had “extreme” levels of production per worker, inventory stock relative to production or capital stock relative to production. “Extreme” is defined as being in the top or bottom 1 percent of the sample when the variable is measured relative to the median value for the relevant year. In the baseline estimation, we include only firms for which we have no missing observations 1997-2016 and since we have two lags, the estimation period is 1999-2016.⁸ This baseline sample is a balanced panel with 1150 firms and 20700 observations over 18 years. By considering a balanced panel, we minimize the “Nickell bias” that arises when we estimate with fixed effects for firms and a small number of periods.⁹ We do robustness checks considering more or less strict delimitations of the sample, finding similar results.

2.3 The firm-specific demand index

We construct a firm-specific demand index, $D_{i,t}$, as a weighted average of domestic and international demand. The demand index is constructed to be as exogenous as possible to the firm and the industry by using only data for components of aggregate demand in Sweden and foreign countries and firm- and industry-specific weights that are kept fixed over time. The weights are measures of the firm’s *exposure to demand fluctuations* in different segments of the product market, at home and abroad. We provide a formal motivation in the *Appendix*.¹⁰

⁷ Log employment changes may become very large if a firm has very few employees. By including only firms which have at least ten employees in *all* years we reduce attrition bias.

⁸ Output data is not available for 1996 but value added is available for that year.

⁹ The estimation method is OLS. Nickell bias means that the estimated coefficient for the first lag of the dependent variable tends to be underestimated because some of its variation is instead picked up by the firm fixed effect when there are few observations in the time dimension. We tried to deal with the Nickell bias by conducting diff-GMM estimation (Arellano-Bond), but we were unable to find an instrument set that is both valid and sufficiently relevant to provide good identification.

¹⁰ Similar approaches have been used by Lundin et al. (2009), Carlsson, Eriksson and Gottfries (2013), Eriksson and Stadin (2017) and Stadin (2015). Nekarda and Ramey (2011) used a similar approach but they considered only one component of demand: changes of government expenditure multiplied by an industry-specific measure of exposure.

Table 1. Descriptive statistics for firm-year observations, manufacturing

Total sample	Yr	N	Zr	Kr	Yr/N	Zr/Yr	Kr/Yr	VAr/N	Zr/VAr	Kr/VAr	VA/Y	Kb/Y
Mean	242000	108	32200	128000	1540	0.1558	0.5143	554	0.6411	1.476	0.3837	0.2109
Std. d.	1590000	466	171000	879000	1257	0.2505	1.1728	473	44.8414	30.4123	0.2158	0.5241
1%	5500	10	61	261	387	0.0034	0.0133	62	0.0026	0.0126	0.0384	0.0028
5%	8836	12	2614	1076	546	0.0153	0.0466	231	0.0272	0.1076	0.1500	0.0110
25%	19400	18	1849	5927	850	0.0686	0.1918	361	0.1572	0.4671	0.2864	0.0581
50%	42300	32	5542	16500	1203	0.1257	0.3813	463	0.3200	0.9663	0.3834	0.1449
75%	114000	71	16700	48700	1809	0.1970	0.6581	619	0.5573	1.6900	0.4841	0.2734
95%	745000	356	98700	346000	3594	0.3776	1.3367	1147	1.2948	3.6827	0.6406	0.6017
99%	3300000	1164	480000	1980000	6035	0.6570	2.3490	2261	2.77654	7.8255	0.7611	1.0532
Observations	78844	83738	83738	72395	78844	78844	67501	83738	83738	72395	78844	78844

Note: Panel with all private firms in manufacturing which for all years of their existence have at least 10 employees, production value of at least 5 yearly wages, and positive stocks of capital and inventories. The industries included are SNI 15-36, and the years included are 1996-2016. X% denotes the Xth percentile, Yr is real production (output), N is full-time equivalent employees, Zr is real inventory stock, Kr is the real capital stock, and VAr is real value added. Real values are in thousands of SEK in prices as of year 2000. PPI for the two-digit industry is used to deflate Y and Z, and the value-added deflator is used to calculate the real value added. The calculation of the real capital stock is described in the appendix. Kb/Y is the nominal book value of the capital stock relative to the nominal value of production (output).

The domestic component of demand is defined as $\phi_j^C \ln C_t^s + \phi_j^G \ln G_t^s + \phi_j^I \ln I_t^s$, where C_t^s, G_t^s and I_t^s are aggregate real private consumption, public consumption and gross fixed investment in Sweden in year t . The subscript j denotes the industry (SNI92, two-digit level) and the weights ϕ_j^C, ϕ_j^G and ϕ_j^I , which sum to unity, represent shares of domestic final use of production from industry j . We calculate the shares using input-output tables for Sweden 2005.

When it comes to foreign demand, we first construct an *industry-specific* measure of demand in country m as $\phi_j^C \ln C_t^m + \phi_j^G \ln G_t^m + \phi_j^I \ln I_t^m$ where C_t^m, G_t^m and I_t^m are real private consumption, public consumption and gross fixed investment in country m . Thus, we assume that the allocation across final uses is approximately the same in the foreign country as in Sweden; a good that is used for private consumption in Sweden is used for private consumption if it is exported to another country. We then construct the firm-specific demand index as

$$\ln D_{i,t} = (1 - \delta_i) (\phi_j^C \ln C_t^s + \phi_j^G \ln G_t^s + \phi_j^I \ln I_t^s) + \delta_i \sum_m \omega_{j,m} (\phi_j^C \ln C_t^m + \phi_j^G \ln G_t^m + \phi_j^I \ln I_t^m). \quad (1)$$

The weight δ_i is a measure of the firm's exposure to demand fluctuations in the export market and the weight $\omega_{j,m}$ is the share of industry j 's direct exports that went to country m in the period 2000-2005. The countries included are 26 of Sweden's main trading partners.¹¹

We have access to data on the share of output that was exported by each firm and this allows us we calculate the firm's average *direct* export share over the sample period, δ_i^d . But a large part of the output from manufacturing is used as input by other firms, which may then export their products. To take account of this, we use input-output data for 2005 to calculate the share of the *industry* output that was ultimately used for exports, directly and indirectly. Denoting this share $\bar{\delta}_j$, we then we calculate the *indirect* export share for a typical firm in industry j as $\bar{\delta}_j^i = \bar{\delta}_j - \bar{\delta}_j^d$ where $\bar{\delta}_j^d$ is the average direct export share for the firms in industry j , calculated from firm-level data. Assuming that all firms in the industry have the same indirect export share we can then calculate the total (direct and indirect) export market exposure for a particular firm as $\delta_i = \delta_i^d + \bar{\delta}_j^i = \bar{\delta}_j + \delta_i^d - \bar{\delta}_j^d$. Thus, a firm that has a

¹¹ In the theoretical model below, *potential sales* by firm i are determined by $\hat{D}_{i,t} = \Phi D_{i,t}^\Sigma P_{i,t}^{-\eta}$, where $D_{i,t}$ is the firm-specific demand shifter and $P_{i,t}$ is the price set by the firm.

relatively high share of *direct* exports is more exposed to demand fluctuations in the export market than the average firm in the industry.¹²

Note that we do not use time series data measuring industry production or prices to construct the demand index. We avoid using such variation because this would induce spurious correlations arising from unobserved industry-specific shocks. If foreign steel production increases, we do not know if this is due to a demand shock or a supply shock to the steel industry abroad.

It should be noted, however, that even if our demand variable is constructed so as to be exogenous to the *industry*, the effect on the *individual* firm goes via different channels. Most likely, industry prices respond to an industry-specific demand shock, so we should view the effect on an individual firm as the combined effect of an exogenous shift in industry demand *and* the industry price response. As explained in Section 3, we take account of this in the estimation of the structural model.

What type of variation does the demand variable represent? We include time dummies in our estimated model and these will mop up fluctuations in demand that are common to all firms. Still, there will be a business cycle element in our demand variable because investment varies more than consumption over the cycle. Additionally, our demand variable will reflect structural changes in the composition of domestic demand and differences in economic developments between Sweden and foreign markets, which will affect firms differently depending on their exposure to demand shocks in the different markets.

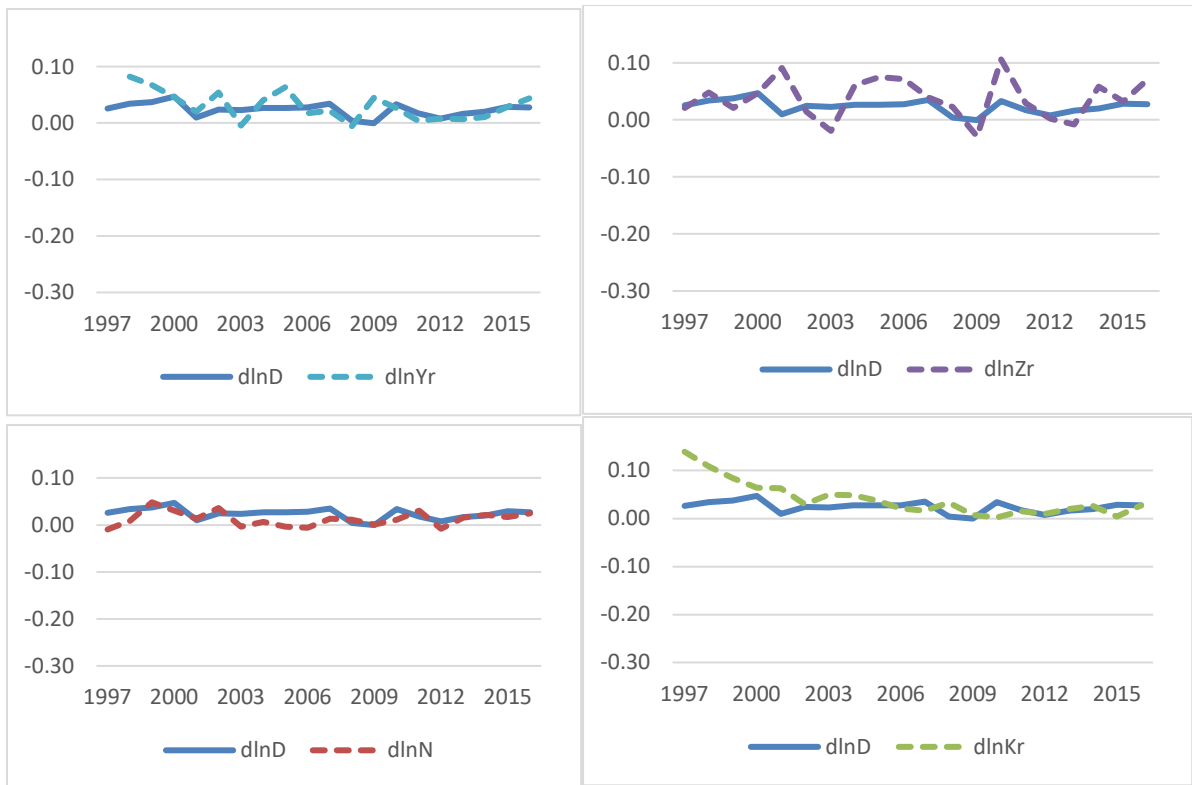
Figure 1 illustrates the data for the five industries with the largest number of firms in the baseline panel.¹³ The figure shows yearly industry averages of the log changes of the firm-level variables. We see that production and inventories co-vary strongly with the demand index without any lag while employment and the capital stock appear to respond with a lag to demand shocks. The lag is particularly clear for capital. Demand for metals and machinery is much more volatile than it is for food and the endogenous variables reflect this difference. It is primarily this difference in the volatility and timing of demand between different industries that allows us to identify how firms on average respond to shocks.

¹² The export share over the sample period does depend on firm-specific shocks during the sample period but we have firm fixed effects in the estimation that pick up this effect. The estimation relies on variation over time *within* the sample period; see Nekarda and Ramey (2011) for a discussion. The estimated IRFs look very similar if we use the initial export share of the firm to measure exposure; see *Figure A3* in the Appendix.

¹³ For a detailed list of the number of observations per industry, see the *Table A1* in the Appendix. More than 80% of the firms are in the same industry throughout the sample period. A firm that changes its industry was assigned to the industry to which it belonged for the longest period of time. Typically, a firm does not change its production entirely but simply passes a threshold in the composition of goods that leads to a change of industry classification.

Figure 1. Log changes of firm-level variables, industry averages for the five industries with the largest number of firms, baseline panel

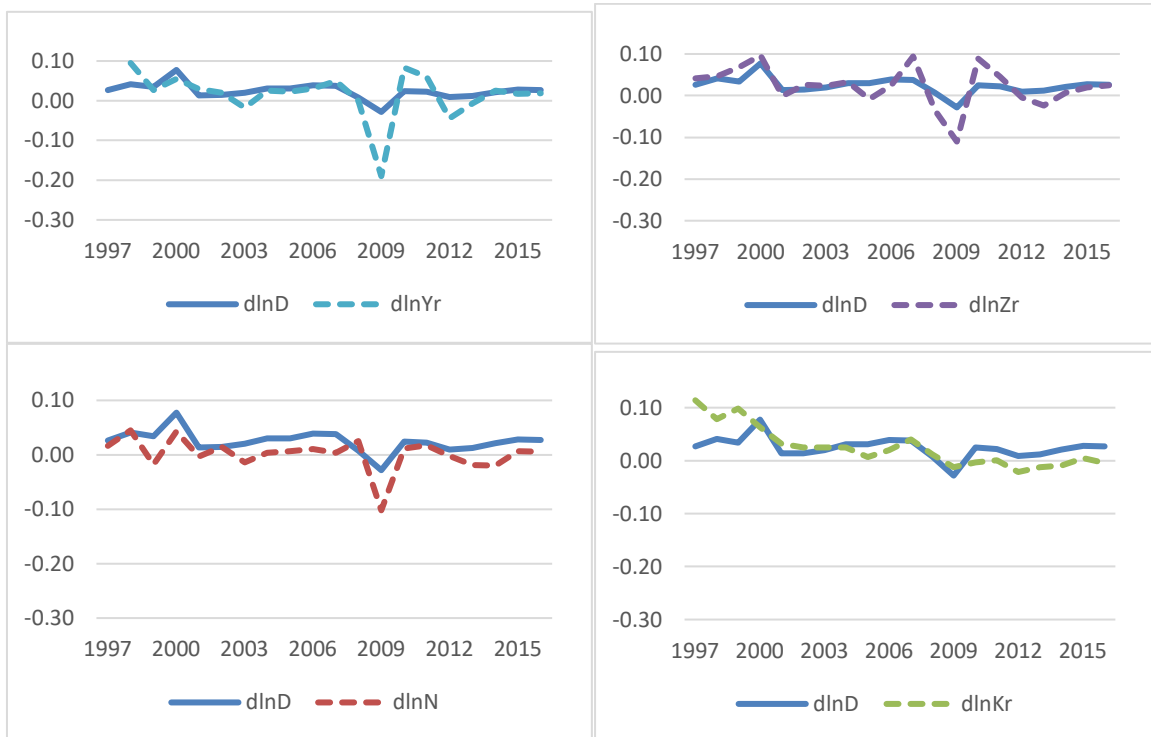
15 Food products and beverages



20 Wood products and wood



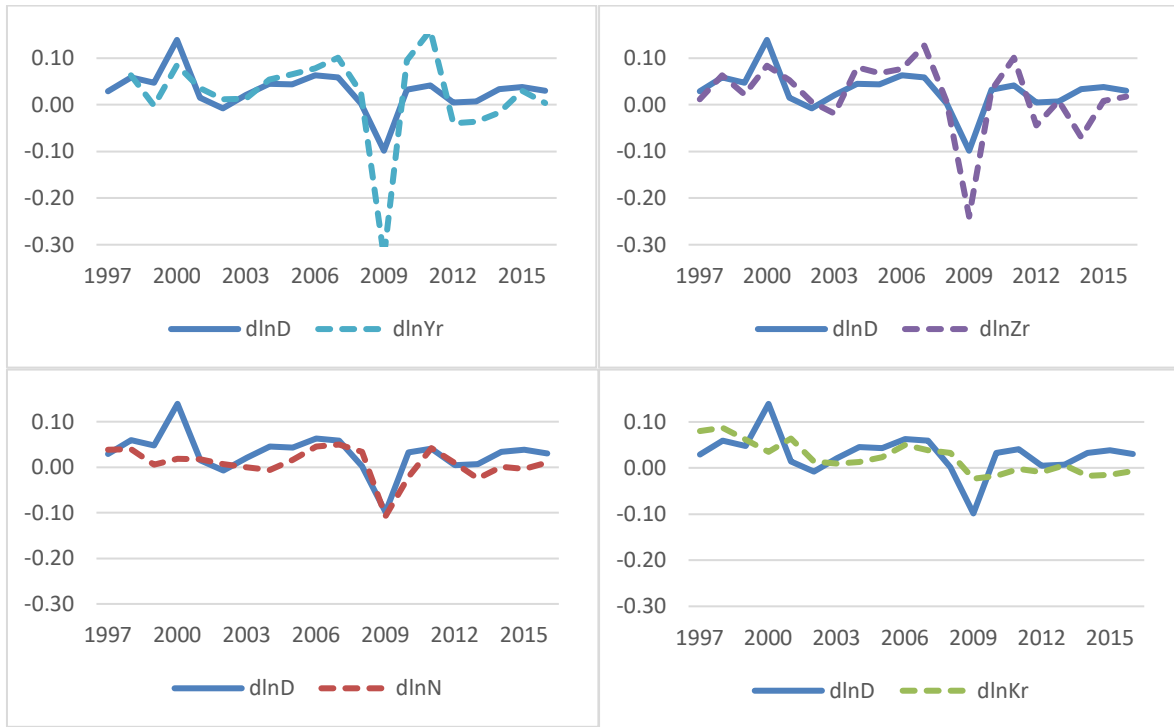
25 Rubber and plastics



28 Fabricated metal products except machinery and equipment



29 Machinery and equipment (not electronics and computers)



Note: The figure shows log changes on the firm level, aggregated to the two-digit level (SNI92). *Yr* is real production, *Zr* is real inventory holdings, *N* is full-time equivalent employees, and *Kr* is the real capital stock. Baseline panel: private firms with at least 10 employees and no extreme observations or missing values.

2.4 Empirical model

To capture the effects of demand shocks on real production (*Yr*), the real capital stock (*Kr*), employment (*N*) and real inventory holdings (*Zr*), we estimate a *reduced-form model* of the firm with two lags of the endogenous variables and the firm-specific product demand index as an exogenous variable:

$$\begin{aligned} \ln Y_{r,t} = & \beta_1^Y \ln Y_{r,t-1} + \beta_2^Y \ln Y_{r,t-2} + \beta_3^Y \ln N_{i,t-1} + \beta_4^Y \ln N_{i,t-2} + \beta_5^Y \ln Kr_{i,t-1} + \beta_6^Y \ln Kr_{i,t-2} \\ & + \beta_7^Y \ln Z_{r,t-1} + \beta_8^Y \ln Z_{r,t-2} + \beta_9^Y \ln D_{i,t} + \beta_{10}^Y \ln D_{i,t-1} + \varepsilon_{i,t}^Y \end{aligned} \quad (2)$$

$$\begin{aligned} \ln N_{i,t} = & \beta_1^N \ln Y_{r,t-1} + \beta_2^N \ln Y_{r,t-2} + \beta_3^N \ln N_{i,t-1} + \beta_4^N \ln N_{i,t-2} + \beta_5^N \ln Kr_{i,t-1} + \beta_6^N \ln Kr_{i,t-2} \\ & + \beta_7^N \ln Z_{r,t-1} + \beta_8^N \ln Z_{r,t-2} + \beta_9^N \ln D_{i,t} + \beta_{10}^N \ln D_{i,t-1} + \varepsilon_{i,t}^N \end{aligned} \quad (3)$$

$$\begin{aligned} \ln Kr_{i,t} = & \beta_1^K \ln Y_{r,t-1} + \beta_2^K \ln Y_{r,t-2} + \beta_3^K \ln N_{i,t-1} + \beta_4^K \ln N_{i,t-2} + \beta_5^K \ln Kr_{i,t-1} + \beta_6^K \ln Kr_{i,t-2} \\ & + \beta_7^K \ln Z_{r,t-1} + \beta_8^K \ln Z_{r,t-2} + \beta_9^K \ln D_{i,t} + \beta_{10}^K \ln D_{i,t-1} + \varepsilon_{i,t}^K \end{aligned} \quad (4)$$

$$\begin{aligned} \ln Zr_{i,t} = & \beta_1^Z \ln Yr_{i,t-1} + \beta_2^Z \ln Yr_{i,t-2} + \beta_3^Z \ln N_{i,t-1} + \beta_4^Z \ln N_{i,t-2} + \beta_5^Z \ln Kr_{i,t-1} + \beta_6^Z \ln Kr_{i,t-2} \\ & + \beta_7^Z \ln Zr_{i,t-1} + \beta_8^Z \ln Zr_{i,t-2} + \beta_9^Z \ln D_{i,t} + \beta_{10}^Z \ln D_{i,t-1} + \varepsilon_{i,t}^Z \end{aligned} \quad (5)$$

$$\ln D_{i,t} = \rho_1 \ln D_{i,t-1} + \rho_2 \ln D_{i,t-2} + \varepsilon_{i,t}^D. \quad (6)$$

We estimate these equations using OLS with firm fixed effects and we include time dummies to control for common unobserved macro shocks and trends. Also, we include firm-specific trends in all equations to take account of firm-specific productivity growth and other long-term structural changes.

This empirical model is a reduced-form model. All variables on the right-hand side are assumed to be either predetermined or strictly exogenous relative to the firm's decision variables, but these equations should *not* be interpreted as the closed-form solution to the firm's decision problem. Instead they represent a mix of the decision rules and the stochastic processes for the unobserved exogenous shocks. As explained in the *Appendix*, the basic idea is that the firm's decision problem contains firm-specific technology and cost shocks and other state variables, which we cannot observe, but the effects of these omitted state variables are "mopped up" by lags of the variables that we can observe. Thus, the first four equations should *not* be considered as structural and the shocks in these equations are a mix of shocks that we cannot measure. The shock to the last equation is considered to be an exogenous demand shock, however, and the model can be used to trace out how firms respond to such a shock.

2.5 Identification

We include time dummies and firm-specific trends in all equations, so the impulse-response functions are identified because of medium-term variations of demand that differ between firms. This "difference-in-difference" variation has two primary sources. One is that some firms produce goods that are mainly used for investment while others produce consumption goods. The other is that some firms export much of what they produce while others sell mainly in the domestic market. To interpret these shocks as demand shocks rather than supply shocks, we need to assume that this cross-firm variation is unrelated to technology and cost shocks that differ systematically between Swedish firms.

To see when this can be problematic, consider a simple case where we have two industries that both cater exclusively to the domestic market. Industry A produces investment

goods while industry B produces consumption goods. In this case, we identify effects of demand shocks from the fact that an increase in investment raises demand only for firms in industry A. But suppose that there is a positive technology shock that affects only the firms that produce investment goods. Such a shock will reduce the price of investment goods, which will lead to an increase in aggregate investment.¹⁴ This shock will affect firms in industry A directly as well as our measure of demand. This means that the demand variable will not only capture demand shocks but also supply-side shocks.

To provide some idea whether shocks of this type were important in the relevant period, we plot the ratio of investment to consumption together with the ratio of the corresponding deflators in *Figure A2* in the Appendix. Investment fluctuates more than consumption over the cycle but the relative price varies much less. Both ratios have trends but in the medium term, there is little relation between these two variables. Medium-term variations in investment relative to consumption do not appear to be driven by technology or cost shocks because such shocks should show up in the relative price.

A similar argument could be made with respect to domestic and foreign demand. A productivity or cost shock that affects firms, which produce for the domestic market, but leaves exporters unaffected, may lead to a change in domestic demand relative to foreign demand. However, it is hard to imagine important shocks that would affect firms or industries differentially in this way. Thus, we feel confident that our firm-specific demand variable captures fluctuations in demand rather than supply.

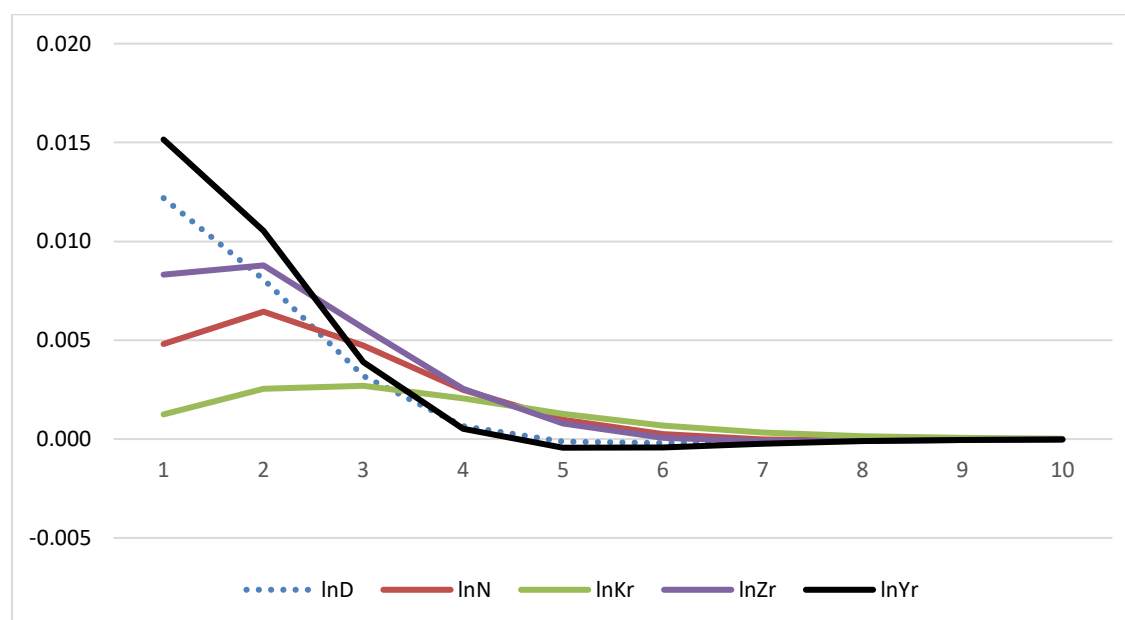
2.6 Impulse-response functions

Figure 2 shows how firms on average respond to the demand shock.¹⁵ Production responds immediately to a change in demand and the effect on demand and production lasts for about 4 years. For inventories, employment and capital, we see hump-shaped responses. The inventory stock and employment peak one year after the peak in production while the capital stock peaks 2 years after the peak in production. The first-year effect on the number of workers is one third of the effect on output, and the capital stock responds very little in the first year. Firms are able to cater to increased demand with a relatively small increase in registered inputs, which implies very strong effects of demand shocks on labor and factor productivity as they are commonly measured.

¹⁴ For a macroeconomic analysis of the effects of investment-specific technology shocks, see Greenwood, Hercowitz and Krusell (2000).

¹⁵ For a detailed presentation of the regression results, see *Table A2* in the Appendix. Confidence intervals for the impulse-response functions are shown in *Figure 5* below.

Figure 2. Effects of a firm-specific demand shock in empirical model



Note: Impulse-response functions from reduced-form model estimated on a balanced panel consisting of private manufacturing firms with at least ten employees, no extreme observations, and no missing values 1997-2016 (1150 firms). Extreme observations are defined as having production/employee, inventory/output or capital/output in top or bottom 1 percentile. Firm and time fixed effects and firm-specific trends are included. The variables are in logs and the time units on the horizontal axis are years. With two lags, the number of observations included in the estimations is 20700.

To check robustness, we present results for alternative samples and model specifications in *Figure A3* in the *Appendix*. We consider an unbalanced panel, more stringent elimination of outliers, estimation with only foreign or only domestic demand shocks, and measuring export market exposure using the firm's initial export share. We divide the firms into groups based on size or capital intensity. Our conclusion is that the results are reasonably robust. Production, inventories, labor and capital respond positively to the demand shock and, in most cases, the order of the response is the same: production responds quickly, followed by inventories and labor while capital stock responds least and with the longest lag.¹⁶

However, when we include time dummies interacted with industry fixed effects, so as to identify the model from the variation between firms *within* industries, we found very small and uncertain effects of the demand shock. Our interpretation is that export shares do not vary sufficiently between firms within industries to provide good identification. Thus, our identification relies primarily on between-industry variation over time. Having firm-level data

¹⁶ To check whether demand is exogenous, we ran a regression with the demand variable as dependent variable and two lags of all variables on the right-hand side. The coefficients for lagged production turned out to be barely statistically significant, but all coefficients (except those for lags of demand itself) were found to be close to zero.

is still a great advantage because we avoid problems of aggregation and reclassification of firms, which affect data that have been aggregated to the industry level.

Our results are consistent with Nekarda and Ramey (2011), who also found that demand shocks have large and immediate effects of on production. However, they did not find any significant effects on labor productivity. One reason may be that they used production worker hours rather than the total number of workers to measure labor input. In the theoretical model, we allow for variations in effective labor input per worker. Another reason may be that there are more flexible rules for hiring and firing in the U.S. than in Sweden.

3. Theory

There appear to be some adjustment lags and/or costs that slow down the adjustment of capital and labor input, but production and inventories respond quickly and strongly to the demand shock. Below, we present the elements of a theoretical model that can potentially explain these empirical responses. First, we discuss adjustment costs, implementation lags, increasing returns to scale and factor utilization. Then, we specify the relation between output and value added, price rigidity and how sales are related to inventories because of the possibility of stock-out. Finally, we present the firm's maximization problem in Section 3.7.

3.1 Adjustment costs and implementation lags

We assume that firms can hire workers at a given wage and buy capital at a given price, but subject to adjustment costs. We include *quadratic adjustment costs* for labor and capital as a simple representation of various types of adjustment costs that affect the average response of firms. The adjustment costs are equal to $c_n (H_t - \delta_n \bar{N})^2 / 2 + c_k (I_t - \delta_k \bar{K})^2 / 2$, where δ_k is the rate at which capital depreciates, and δ_n is an exogenous separation rate for labor.

\bar{N} and \bar{K} denote the steady-state levels of N_t and K_t , so there are quadratic costs for hiring and investing more than the steady-state levels $\delta_n \bar{N}$ and $\delta_k \bar{K}$. These costs take the form of reduced production due to disruptions in the production process.¹⁷

¹⁷ With this specification, adjustment costs are zero in the steady state. This helps to solve analytically for the steady state, which is necessary for the estimation.

We also include an *implementation lag* in investment by assuming that the investment that is decided in year t , I_t , is implemented in years t , $t+1$ and $t+2$. The capital stock at the end of period t is determined as follows:

$$K_t = (1 - \delta_k) K_{t-1} + \lambda_{k1} I_t + \lambda_{k2} I_{t-1} + (1 - \lambda_{k1} - \lambda_{k2}) I_{t-2}. \quad (7)$$

Thus, a fraction λ_{k1} of the investment that is decided in year t is implemented in the same year. This approach is similar to “time to build” (Kydland and Prescott 1982) and it is more flexible than assuming either no lag or a one-period implementation lag, as in Burnside-Eichenbaum-Rebelo (1993). Similarly, we assume that hiring is implemented in the current and the coming year: $N_t = (1 - \delta_n) N_{t-1} + \lambda_n H_t + (1 - \lambda_n) H_{t-1}$ where N_t is employment during period t , and H_t is the hiring *decided* in period t .

3.2 Increasing returns to scale in production

Our estimated impulse-response functions show that firms can increase production in the short run with much smaller percentage increases in the registered inputs of capital and labor. One possible explanation is that there are *increasing returns to scale* so that changes in inputs lead to proportionally larger changes in production (Hall 1988). To model this, we assume that the capital stock consists of two components. First, there is a flexible part \hat{K}_t that enters a CES production function with constant returns to scale, and second, there is a fixed amount of capital F_k that the firm must have in order to produce at all. Thus, the total observed capital is given by $K_t = F_k + \hat{K}_t$. Similarly, we distinguish between fixed employment (overhead) and flexible employment that enters the production function: $N_t = F_n + \hat{N}_t$.

3.3 Factor utilization

Looking at the dynamic response in *Figure 2*, we see that production increases much more than registered inputs in the first year, but after two years, employment has caught up with production. This makes it unlikely that increasing returns can explain the whole picture. Firms appear to have some form of excess capacity that they can use to meet demand and a natural interpretation is that firms can vary the *utilization* of the production factors that they have. The key question, then, is why the firm did not make full use of its resources for production before the shock occurred. As Bean (1990) noted, hiring and firing costs cannot explain why firms operate *within* their production frontier during recessions. There must be some cost of

increasing resource utilization, or else the firm would always make full use of its resources in production.

Several authors have noted that workers spend substantial amounts of time on activities, which do not contribute to *current* production, but which increase *future* production (Fay and Medoff, 1985; Bean, 1990; Kim and Lee, 2007). There are many such activities we can think of, e. g. cleaning, maintenance, reorganizing and training. To capture this, we assume that the firm has a stock of *organizational capital* Ω , and the larger this stock is, the more it can produce with given inputs. Workers spend a share x_t of their time on activities that increase current production and a fraction $1-x_t$ of their time accumulating organizational capital that increases future production. We write the production function for value added

$$F(u_t, \Omega_{t-1}, \hat{K}_{t-1}, x_t, \hat{N}_t, H_t, I_t) = Au_t \left(a - \frac{a-1}{\Omega_{t-1}^\xi} \right) \left(\alpha \hat{K}_{t-1}^{\frac{\sigma-1}{\sigma}} + (1-\alpha) (x_t \hat{N}_t)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} - \frac{c_n}{2} (H_t - \delta_n \bar{N})^2 - \frac{c_k}{2} (I_t - \delta_k \bar{K})^2 \quad (8)$$

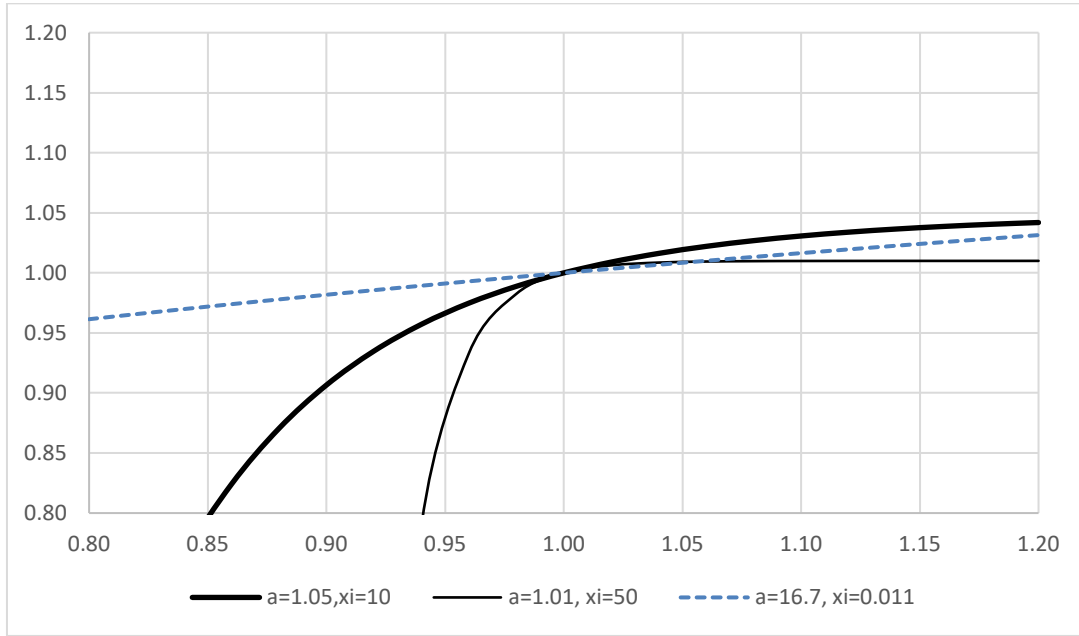
where $a > 1$ and where \hat{K}_t and Ω_t denote the stocks of flexible capital and organizational capital at the end of period t .

Figure 3 illustrates the function $a - (a-1)/\Omega_{t-1}^\xi$ for different values of the parameters a and ξ . We normalize so that $\Omega = 1$ in the steady state, and thus the function value is one in the steady state. As organizational capital increases, the value of this function increases asymptotically toward a , and if organizational capital falls to $(a-1)/a$, the function value falls to zero. Roughly speaking, the parameter a determines the slope of the function, while ξ determines its concavity. We assume that the accumulation of organizational capital is governed by

$$\Omega_t = (1 - \delta_\omega) \Omega_{t-1} + \chi (1 - x_t) \quad (9)$$

where δ_ω is the depreciation rate of organizational capital, and $1 - x_t$ is the fraction of time spent accumulating organizational capital. The parameter χ is set to be consistent with the normalization $\Omega = 1$ in the steady state. The basic idea behind this specification is that when there is an unexpected increase in demand, the firm can tell its workers to cut down maintenance and training and to increase the time spent on current production.

Figure 3. Productivity contribution of organizational capital



Note: The function shows the contribution of organizational capital to factor productivity $\left(a - (a-1)/\Omega_{t-1}^\xi\right)$ as a function of Ω for different values of the parameters a and ξ .

An alternative way to model utilization is to allow for costly variations in *effort* as in Burnside, Eichenbaum, Rebelo (1993) and Sbordone (1996, 1997). We can allow for variations in effort (effective working time) that affects the utilization of both factors of production by including a variable u_t , which enters multiplicatively in the production function, and a cost of utilization given by $\Phi_u \left(u_t - 1 + (c_u / 2)(u_t - 1)^2\right) \hat{N}_t$. This can represent effort or overtime, which increase the use of both labor and capital. In the latter interpretation, the convex cost may reflect an overtime premium that may be part of an explicit or implicit contract.¹⁸ As it turns out, these two ways of modeling utilization have similar implications and we allow only the first version in the baseline specification that we estimate.

4.4 Output and value added

We assume that value added and materials inputs are combined in a Leontief production function

$$Y_t = \min \left\{ F \left(u_t, \Omega_{t-1}, \hat{K}_{t-1}, x_t, \hat{N}_t, H_t, I_t \right), M_t / m \right\} \quad (10)$$

¹⁸ As mentioned in Section 3, employment is measured as “full-time equivalent workers” and this measure does not take account of variations in overtime.

where Y_t is the quantity produced, $F(\cdot)$ is as defined above, and M_t is the quantity of intermediate inputs used. Cost minimization then implies that $Y_t = F(\cdot) = M_t / m$; a fixed amount of cloth is used to make a shirt. Normalizing the price of intermediate inputs to one, the cost of inputs is mY_t .

If there is no substitutability between value added and materials, materials inputs and total output will always be proportional to value added, and it should not matter whether we measure production by output or value added. In fact, the impulse-response functions are very similar when we use value added to measure production. Note that we allow for increasing returns to scale by including fixed costs in terms of capital and labor, but not in terms of materials.¹⁹

3.5 Price rigidity

Another factor that can prevent firms from always optimally utilizing their resources is *price rigidity*. If demand falls and the firm cannot (or does not want to) reduce its price, quasi-fixed resources will become less utilized (Rotemberg and Summers, 1990). We incorporate price rigidity in a simple way by including a quadratic adjustment cost for prices:

$\theta(P_t / P_{t-1} - 1)^2 / 2$. We assume that firms always satisfy demand, which makes sense if firms have sufficient market power so that the markup is always positive.

3.6 Inventories, sales and stock-outs

The estimated responses show that firms increase their inventory stocks when demand increases. This result is the opposite of the production-smoothing idea that by drawing down inventories in periods of high demand, firms can stabilize production. To explain the observed pattern, we follow the ideas of Kahn (1987, 1992) and Bils and Kahn (2000) and assume that inventories of finished goods are needed in order to sell the goods. Below, we present a very stylized model that provides a reasonable functional form that we can include in our estimated model.

Consider a firm that sells goods, e.g., steel bars, that come in M different varieties, which we will call sizes. We follow Kahn (1987, 1992) and assume that a customer will buy a steel bar only if he/she finds the right size.²⁰ The firm has a sales department and a production department, and the sales department sends an order to the production department T times per

¹⁹ See section VI in Basu (1996) and Basu-Fernald (1997) page 255 for discussions of these issues.

²⁰ Bils (2011) refers to empirical evidence that supports this assumption.

year to replenish the inventory stock. For concreteness, we can think of the case when $T=12$, so inventories are replenished every month.²¹

Let \hat{D}_t be the *potential sales* of all varieties during year t ; \hat{D}_t is what the firm would sell if it would never stock out. We assume that $\hat{D}_t = \Phi D_t^\Sigma (P_t / \Upsilon_t)^{-\eta}$ where Φ is a constant, D_t is a demand shifter, P_t is the price set by the firm, and Υ_t is the market price in the relevant market. To make the model as simple as possible, we assume that D_t is observed at the beginning of the year and remains constant over the year and the price is set after observing D_t . Demand in a particular month for a particular size is assumed to be $\lambda \hat{D}_t$, where λ is a stochastic variable that is uniformly distributed between λ_1 and λ_2 . The supports of the distribution are given by $\lambda_1 = (1 - \Psi) / (TM)$ and $\lambda_2 = (1 + \Psi) / (TM)$ where T is the number of inventory periods (months), and M is the number of sizes, so without stock-outs, expected monthly sales of a variety would be $\hat{D}_t / (TM)$ and the density is $TM / (2\Psi)$. The parameter Ψ , which has a value between zero and unity; reflects the degree of uncertainty about the demand for individual varieties.

Since demand is assumed to be symmetrically and independently distributed across sizes, the sales department will ensure that they stock up with the same quantity of each size whenever they replenish inventories. Let z_t be the inventory stock of a specific size held at the beginning of each month of year t . It follows immediately that $\lambda_1 \hat{D}_t < z_t < \lambda_2 \hat{D}_t$. With a smaller inventory stock, the firm would always stock out, and there is no reason to hold a larger inventory stock than the maximum possible sales of a particular size. If the realization of λ is such that $\lambda \hat{D}_t \leq z_t$, sales of that specific size will be $\lambda \hat{D}_t$, and if $\lambda \hat{D}_t > z_t$, sales of that specific size will be z_t . Letting $\hat{\lambda}_t$ be the critical value of λ where the firm runs out of stock ($\hat{\lambda}_t \hat{D}_t = z_t$) we obtain the expected sales of a particular size in a given month of year t :

$$\begin{aligned}
 s_t &= \int_{\lambda_1}^{\hat{\lambda}_t} \lambda \hat{D}_t \frac{TM}{2\Psi} d\lambda + \int_{\hat{\lambda}_t}^{\lambda_2} z_t \frac{TM}{2\Psi} d\lambda = \frac{TM}{2\Psi} \left(\frac{\hat{\lambda}_t^2 \hat{D}_t}{2} - \frac{\lambda_1^2 \hat{D}_t}{2} + \lambda_2 z_t - \hat{\lambda}_t z_t \right) \\
 &= \frac{1 + \Psi}{2\Psi} z_t - \frac{(1 - \Psi)^2}{4\Psi TM} \hat{D}_t - \frac{TM}{4\Psi} \frac{z_t^2}{\hat{D}_t}
 \end{aligned} \tag{11}$$

²¹ Contrary to the Ss model, T is taken as exogenous, so our model differs from the Ss model in the same way as the Taylor model of wage/price setting differs from state-contingent pricing. This type of model is called “periodic review” in the inventory literature; see Urban (2004).

We assume that \hat{D}_t is observed at the beginning of the year and constant throughout the year, so the total expected sales of storable finished goods during the year are:

$$TMs_t = \frac{1+\Psi}{2\Psi} T\hat{Z}_t - \frac{(1-\Psi)^2}{4\Psi} \hat{D}_t - \frac{T^2\hat{Z}_t^2}{4\Psi\hat{D}_t} \quad (12)$$

where \hat{Z}_t is the total stock of finished goods at the beginning of each month in year t :

$\hat{Z}_t = Mz_t$. To keep the model simple, we assume that the firm sells a large number of varieties, so we can view this function as a deterministic function that determines sales. This function has intuitive properties:

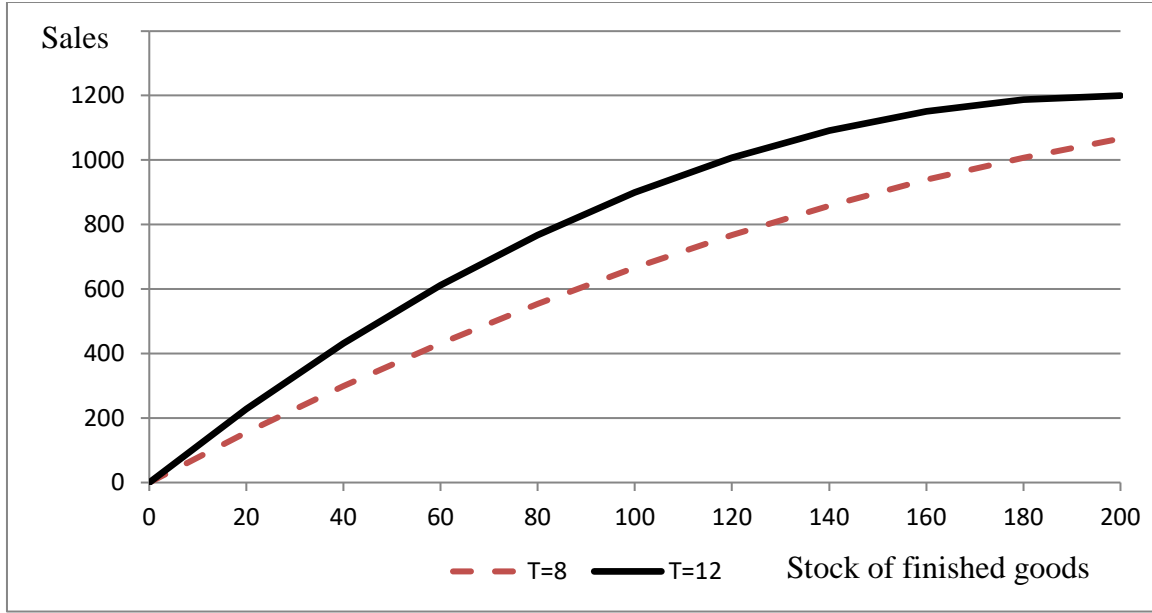
- i) For a given inventory stock, the maximum sales are $T\hat{Z}_t$, and sales approach that limit as \hat{D}_t goes to $\hat{Z}_t T / (1-\Psi)$. For a lower level of demand, some varieties will not sell out.
- ii) For a given level of demand, the inventory stock that maximizes sales is $\hat{Z}_t = (1+\Phi)\hat{D}_t / T$. To ensure that the required sizes are always available, the firm needs finished goods inventories of each variety that correspond to the maximum possible demand during an inventory period. In the model, there are costs of financing, depreciation and storage, so the optimal stock will be smaller than this amount.
- iii) Starting from a low T , more frequent review (higher T) will increase sales.

Figure 4 shows the sales of finished storable goods (TMs_t) as a function of the inventory stock, when potential sales per year are 1200, $T=12$ so the stock is replenished every month, and $\Psi = 1$ so that the demand for a particular size is uniformly distributed between zero and twice the expected demand for that size. In this case, sales are equal to $12 \cdot \hat{Z}_t - 0.03 \cdot \hat{Z}_t^2$ in the relevant interval. To never stock out, the firm needs to have an inventory stock that is 200, twice as large as the potential monthly sales. With a smaller stock, it will sell less because some sizes will run out. If the stock is replenished more seldom, this will reduce the sales for a given stock.²²

²² We do not explicitly address adjustments during the year, but the following very stylized timing assumptions can serve to motivate the specification: i) Inventory stocks are replenished at the beginning of the month, and customers buy the good at the end of the month. ii) Firms learn the level of demand for the coming year at the start of the year. iii) Workers dislike putting in more effort during a year, but they are indifferent to how effort is allocated during the year. This means that firms adjust the level of inventories at the beginning of the year.

Figure 4. Sales of storable goods as a function of the inventory stock

$$(\Psi = 1, \hat{D} = 1200)$$



Note: T denotes the number of times the firm replenishes inventories per year.

To this we add yet another modification by assuming that there are some goods that are sold without holding any inventory stock. These may be goods produced on order. Sales of these goods are simply assumed to be equal to \hat{D}_t . Letting the fraction of storable finished goods be Λ , we obtain total sales as

$$S_t = \Lambda T M S_t + (1 - \Lambda) \hat{D}_t = \kappa_1 \hat{Z}_t + \kappa_2 D_t^\Sigma (P_t / \Upsilon_t)^{-\eta} - \kappa_3 \hat{Z}_t^2 D_t^{-\Sigma} (P_t / \Upsilon_t)^\eta \quad (13)$$

where $\kappa_1 = \Lambda(1 + \Psi)T / (2\Psi)$, $\kappa_2 = 1 - \Lambda - \Lambda(1 - \Psi)^2 \Phi / (4\Psi)$ and $\kappa_3 = \Lambda T^2 / (4\Psi\Phi)$. For given demand, sales are an increasing and concave function of the inventory stock. Higher demand increases the marginal effect of the inventory stock on sales and conversely. The accumulation of finished goods inventories is governed by the function

$$\hat{Z}_t = (1 - \delta_z) \hat{Z}_{t-1} + Y_t - S_t \quad (14)$$

where \hat{Z}_t is the finished goods inventory stock at the end of the year, and δ_z is the rate at which inventories depreciate during the year. We also include a cost $c_z \cdot \hat{Z}_t$ that reflects other costs of holding inventories, such as the cost of providing storage space and managing the inventories.

Finally, inventories consist of not only finished goods but also of inputs and goods in process and our data do not allow us to distinguish between different types of inventories. To take this into account, we simply assume that the firm holds a stock of intermediate inputs that is proportional to current production: $h_z Y_t$.²³ We assume that these inputs can be bought without delay; hence, the total observed inventory stock Z_t is given by $Z_t = \hat{Z}_t + h_z Y_t$.

3.7 Profit maximization

The firm's profit-maximization problem is to choose $S_t, Y_t, P_t, \hat{K}_t, I_t, \hat{N}_t, H_t, u_t, x_t, \Omega_t, \hat{Z}_t$ to maximize

$$E_t \left\{ \sum_{\tau=t}^{\infty} \beta^{\tau-t} \left[\begin{array}{l} S_{\tau} P_{\tau} - mY - W \hat{N}_{\tau} - \Phi_u \left(u_{\tau} - 1 + \frac{c_u}{2} (u_{\tau} - 1)^2 \right) \hat{N}_{\tau} \\ - P^K \left(\hat{K}_{\tau} - (1 - \delta_k) \hat{K}_{\tau-1} \right) - c_z \hat{Z}_{\tau} - \frac{\theta}{2} \left(\frac{P_{\tau}}{P_{\tau-1}} - 1 \right)^2 \end{array} \right] \right\} \quad (15)$$

subject to the following constraints (with associated shadow prices) that hold for all t :

$$Y_t = A \left(a - \frac{a-1}{\Omega_{t-1}^{\xi}} \right) u_t \left(\alpha \hat{K}_{t-1}^{\frac{\sigma-1}{\sigma}} + (1-\alpha) (x_t \hat{N}_t)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} - \frac{c_n}{2} (H_t - \delta_n \bar{N})^2 - \frac{c_k}{2} (I_t - \delta_k \bar{K})^2 \quad v_t \quad (16)$$

$$S_t = \kappa_1 \hat{Z}_t + \kappa_2 D_t^{\Sigma} \Upsilon_t^{\eta} P_t^{-\eta} - \kappa_3 \hat{Z}_t^2 D_t^{-\Sigma} \Upsilon_t^{-\eta} P_t^{\eta} \quad \mu_t \quad (17)$$

$$\hat{Z}_t = (1 - \delta_z) \hat{Z}_{t-1} + Y_t - S_t \quad g_t \quad (18)$$

$$F_k + \hat{K}_t = (1 - \delta_k) (F_k + \hat{K}_{t-1}) + \lambda_{k1} I_t + \lambda_{k2} I_{t-1} + (1 - \lambda_{k1} - \lambda_{k2}) I_{t-2} \quad q_t \quad (19)$$

$$F_n + \hat{N}_t = (1 - \delta_n) (F_n + \hat{N}_{t-1}) + \lambda_n H_t + (1 - \lambda_n) H_{t-1} \quad \gamma_t \quad (20)$$

$$\Omega_t = (1 - \delta_{\omega}) \Omega_{t-1} + \chi (1 - x_t). \quad \phi_t \quad (21)$$

The shadow price v_t is the marginal value (and cost) of a unit of value added, g_t is the marginal value (and cost) of a unit of the final good, μ_t is the value of an additional unit of sales (price minus marginal cost), and q_t , γ_t and ϕ_t are the shadow price prices associated

²³ We disregard the stock-flow aspect of input inventories; see Humphreys, Maccini and Schuh (2001) and Iacoviello, Schiantarelli and Schuh (2011) for more elaborate models.

with investments, hiring and organizational capital. Defining

$$\hat{Y}_t = \left(\alpha \hat{K}_{t-1}^{\frac{\sigma-1}{\sigma}} + (1-\alpha) \left(x_t \hat{N}_t \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \text{ we can write the first-order conditions:}$$

$$S_t: \mu_t = P_t - g_t \quad (22)$$

$$Y_t: g_t = v_t + m \quad (23)$$

$$P_t: S_t - \mu_t \eta \left(\frac{\kappa_2 D_t^\Sigma \Upsilon_t^\eta}{P_t^{1+\eta}} + \frac{\kappa_3 \hat{Z}_t^2}{D_t^\Sigma \Upsilon_t^\eta P_t^{1-\eta}} \right) - \theta \left(\frac{P_t}{P_{t-1}} - 1 \right) \frac{1}{P_{t-1}} + \beta \theta E_t \left\{ \left(\frac{P_{t+1}}{P_t} - 1 \right) \frac{P_{t+1}}{P_t^2} \right\} = 0 \quad (24)$$

$$K_t: q_t = -P^k + \beta E_t \left\{ v_{t+1} A u_{t+1} \left(a - \frac{a-1}{\Omega_t^\xi} \right) \alpha \left(\frac{\hat{Y}_{t+1}}{\hat{K}_t} \right)^{\frac{1}{\sigma}} + (1-\delta_k) (P^k + q_{t+1}) \right\} \quad (25)$$

$$I_t: v_t c_k (I_t - \delta_k \bar{K}) = \lambda_{k1} q_t + \beta \lambda_{k2} E_t (q_{t+1}) + \beta^2 (1-\lambda_{k1} - \lambda_{k2}) E_t (q_{t+2}) \quad (26)$$

$$\hat{N}_t: \gamma_t = v_t A u_t \left(a - \frac{a-1}{\Omega_{t-1}^\xi} \right) (1-\alpha) x_t^{\frac{\sigma-1}{\sigma}} \left(\frac{\hat{Y}_t}{\hat{N}_t} \right)^{\frac{1}{\sigma}} - W_t - \Phi_u \left(u_t - 1 + \frac{c_u}{2} (u_t - 1)^2 \right) + \beta (1-\delta_n) E_t (\gamma_{t+1}) \quad (27)$$

$$H_t: v_t c_n (H_t - \delta_n \bar{N}) = \lambda_n \gamma_t + \beta (1-\lambda_n) E_t (\gamma_{t+1}) \quad (28)$$

$$u_t: \Phi_u (1-c_u + c_u u_t) \hat{N}_t = v_t A \left(a - \frac{a-1}{\Omega_{t-1}^\xi} \right) \hat{Y}_t \quad (29)$$

$$x_t: \phi_t \chi = v_t A u_t \left(a - \frac{a-1}{\Omega_{t-1}^\xi} \right) (1-\alpha) \hat{N}_t^{\frac{\sigma-1}{\sigma}} \left(\frac{\hat{Y}_t}{x_t} \right)^{\frac{1}{\sigma}} \quad (30)$$

$$Z_t: g_t = \mu_t \left(\kappa_1 - 2\kappa_3 \frac{\hat{Z}_t}{D_t^\Sigma \Upsilon_t^\eta P_t^{-\eta}} \right) - c_z + \beta (1-\delta_z) E_t (g_{t+1}) \quad (31)$$

$$\Omega_t: \phi_t = \beta E_t \left\{ v_{t+1} A u_{t+1} \frac{a-1}{\Omega_t^{\xi+1}} \hat{Y}_{t+1} + (1-\delta_\omega) \phi_{t+1} \right\}. \quad (32)$$

The total amounts of capital, labor and inventory stock are $K_t = F_k + \hat{K}_t$, $N_t = F_n + \hat{N}_t$, and

$Z_t = h_z Y_t + \hat{Z}_t$. We also need to specify the stochastic process for the demand shock. In line

with our empirical model, we assume that the demand shock follows an AR(2) process:

$$D_t = 1 + \rho_1 (D_{t-1} - 1) + \rho_2 (D_{t-2} - 1) + \varepsilon_t. \quad (33)$$

In order to simulate responses to demand shocks, we need to take account of induced changes in the market price Υ_t . Since we cannot observe competitors' prices, we need to make an assumption about the market price. Since all firms in the same market are subject to the same demand shock, we assume that, *on average*, other firms in the same market respond in the same way as the firms we are considering. Thus, we set $\Upsilon_t = P_t$ in the simulation. Note that we do not assume that all firms in the same market set the same price but that, on average, they *respond* in the same way to the demand shock. A more elaborate model should take account of market prices responding differently in different markets, pricing to market, and customer relations but such an elaboration is beyond the scope of the present paper.

To sum up, adjustment costs and implementation lags may help to explain the sluggish adjustment of labor and capital, while increasing returns to scale in production and variable utilization could potentially explain the observed increase in factor productivity (as it is normally measured) in response to a demand shock. A positive inventory response may arise because of the stock-out motive and because a large fraction of inventories consists of inputs and goods in process. We now turn to the estimation, which will help us to evaluate these alternative explanations.

4. Estimation method

We follow Christiano, Eichenbaum and Evans (2005) and estimate the structural parameters in the theoretical model by finding the set of parameter values that make the impulse responses in the theoretical model match the impulse responses of its empirical counterpart.

4.1 Matching impulse-response functions

The target function is constructed as in Christiano, Eichenbaum and Evans (2005):

$$J = \min \left[\hat{\Psi} - \Psi(\vec{\gamma}) \right] V^{-1} \left[\hat{\Psi} - \Psi(\vec{\gamma}) \right]. \quad (34)$$

$\Psi(\vec{\gamma})$ contains the impulse response coefficients calculated with the theoretical model for different horizons as a function of the model parameter vector $\vec{\gamma}$, and $\hat{\Psi}$ is the empirical counterpart. V is a diagonal matrix with the variances from the empirical estimation. These variances are related to the 95% confidence intervals, which are generated by bootstrapping and shown in *Figure 5* below. We include IRFs for 20 years in the estimation.

4.2 Prior constraints

We constrain the estimated parameters to be in an economically meaningful range, indicated in the columns denoted min and max in *Table 2* below. Also, parameter estimates that are grossly inconsistent with what we know about the levels of the variables would be uninteresting and for this reason, we impose restrictions on some steady-state ratios. In the baseline sample, the median cost of personnel relative to value added is 0.78, and we constrain this ratio to be between 0.73 and 0.83 in the steady state. The median real capital (calculated as described in the Appendix) relative to output is 0.42 and we constrain this ratio to be between 0.37 and 0.47 in the steady state.

The median ratio of inventories to production is 0.16, but we should note that Z_t in the theoretical model is the inventory stock when the firm has just replenished inventories of finished goods. This means that Z_t in the model should be higher than the number observed in the data. We therefore constrain total inventories relative to production, Z_t / Y_t , to be between 0.10 and 0.60 in the steady state. We cannot distinguish between different types of inventories in our data, but we know the proportions of finished goods, inputs and goods in process for manufacturing as a whole. If we count 50 percent of goods in process as finished goods, then roughly half the inventory stock consists of finished goods. Therefore, we constrain the ratio of finished goods to stored inputs, $\hat{Z}_t / (h_z Y_t)$, to be between 2/3 and 3/2 in the steady state. Finally, we constrain the return on total capital to be between 3 and 30 percent in the steady state and the ratio of inputs to total output (m) is set to 0.6 based on aggregate data for manufacturing.²⁴

4.3 Search algorithm

We use the search algorithm from Mickelsson (2016), which is based on the local algorithm of Nelder and Mead (1965). The basic idea behind this algorithm is to start with a large number of starting vectors that are spread out across the parameter space and then combine these vectors in a smart manner to approach the global maximum without getting stuck at local maxima or iterating too long on flat surfaces. Mickelsson (2016) shows that this algorithm does better than most commonly used search algorithms when the objective function has many local minima and flat surfaces.

²⁴ The return on capital is calculated as $(S - W * N - m * Y - P_k * \delta_k * K - cz * \hat{Z}) / (P_k * K + Z)$.

4.4 Confidence intervals

To obtain confidence intervals, we generate distributions of the estimates in the following way:

1. First, we create a new sample of firms of the same size as the original sample by drawing firms randomly from the original sample with replacement.
2. This sample is used to obtain a new estimate of the empirical IRFs.
3. The impulse responses from the empirical model are then used to estimate the parameters of the theoretical model, as described above.
4. The vector of parameter estimates is saved, and steps 1-3 are repeated 1000 times to obtain a distribution of estimates.²⁵

4.5 Fixing of poorly identified parameters

There are many parameters in our theoretical model, so it is not surprising that some are poorly identified. Attempts to estimate the discount factor and some depreciation rates indicated that these parameters are not well identified. For this reason, we set the discount factor (β) to 0.96 and the parameters δ_n and δ_z were set to 0.10, consistent with numbers in other studies.²⁶ The depreciation rate for capital (δ_k) was set to 0.074 based on a weighted average of the depreciation rates for machines and buildings used by Statistics Sweden. Reasonable changes of these parameters have small effects on the dynamic responses, which confirms that they are poorly identified.

We included two ways of modelling factor utilization in the theoretical model. One is that increased effort that is directed to current production (x_t) comes at the cost of less accumulation of organizational capital and lower future productivity. The other is that increased utilization (u_t) is associated with a direct cost in the form of overtime payments or some other form of compensation. As it turns out, these two ways of modeling utilization are very similar so we decided to include only the first one in our baseline model.

²⁵ We tested bootstrapping with the restriction that the number of firms from each industry should be the same in each sample, but this made very little difference to the results.

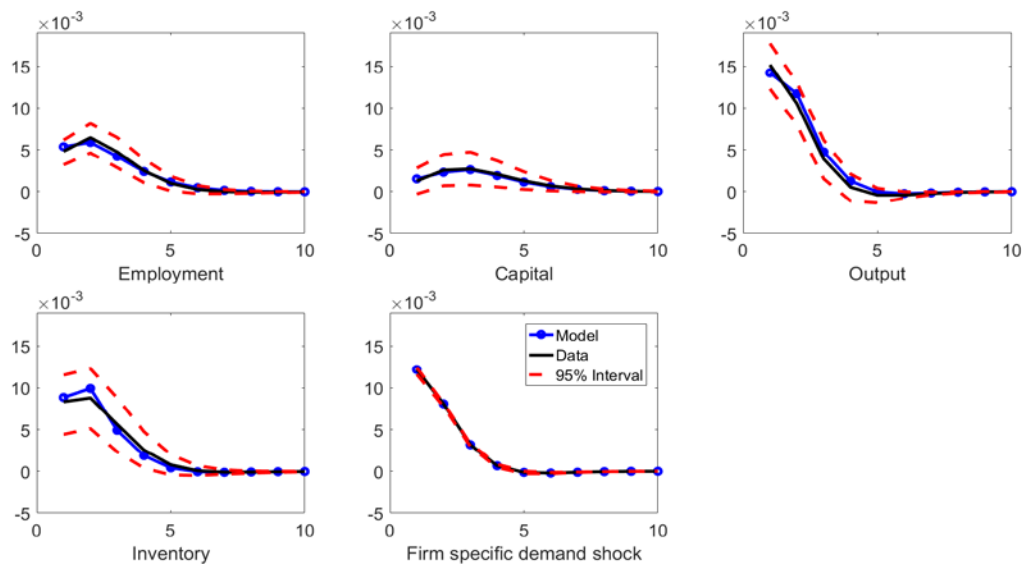
²⁶ Bartelsman, Haltiwanger, and Scarpetta (2013) and Justiniano, Primiceri, and Tambalotti (2010) set the depreciation of capital to 0.10. Kryvtsov and Midrigan (2013) use monthly depreciation rates that imply yearly depreciation rates of 0.12 for capital and 0.13 for inventories. According to Statistics Sweden (AM 63 SM 1201), 12 percent of the permanently employed workers in Sweden left their jobs each year 1990-2011.

5. Results

5.1 Replication of empirical responses

As we see in *Figure 5*, the model can replicate the empirical responses quite well. A combination of adjustment costs and implementation lags explains sluggish adjustment of labor and capital very well. Although capital and labor adjust slowly, the model explains a 1.4 percent increase in production in the first year. This increase in production is achieved by increasing employment by 0.53 percent, which translates to a 1.11 percent increase in flexible labor input (production labor) and by increasing each worker's time spent in current production by 0.84 percent (*Figure 6*). Together, this sums to a 2 percent increase in effective labor input into production.²⁷ Thus, output increases almost three times as much as employment in the first year and this is due to a combination of increasing returns to scale and increased utilization. By assumption, investment does not contribute to production in the first year and investment responds very little in the year when the shock occurs, so capital contributes only marginally to the increase in production in the first two years.

Figure 5. Effects of a firm-specific demand shock in the theoretical and empirical model

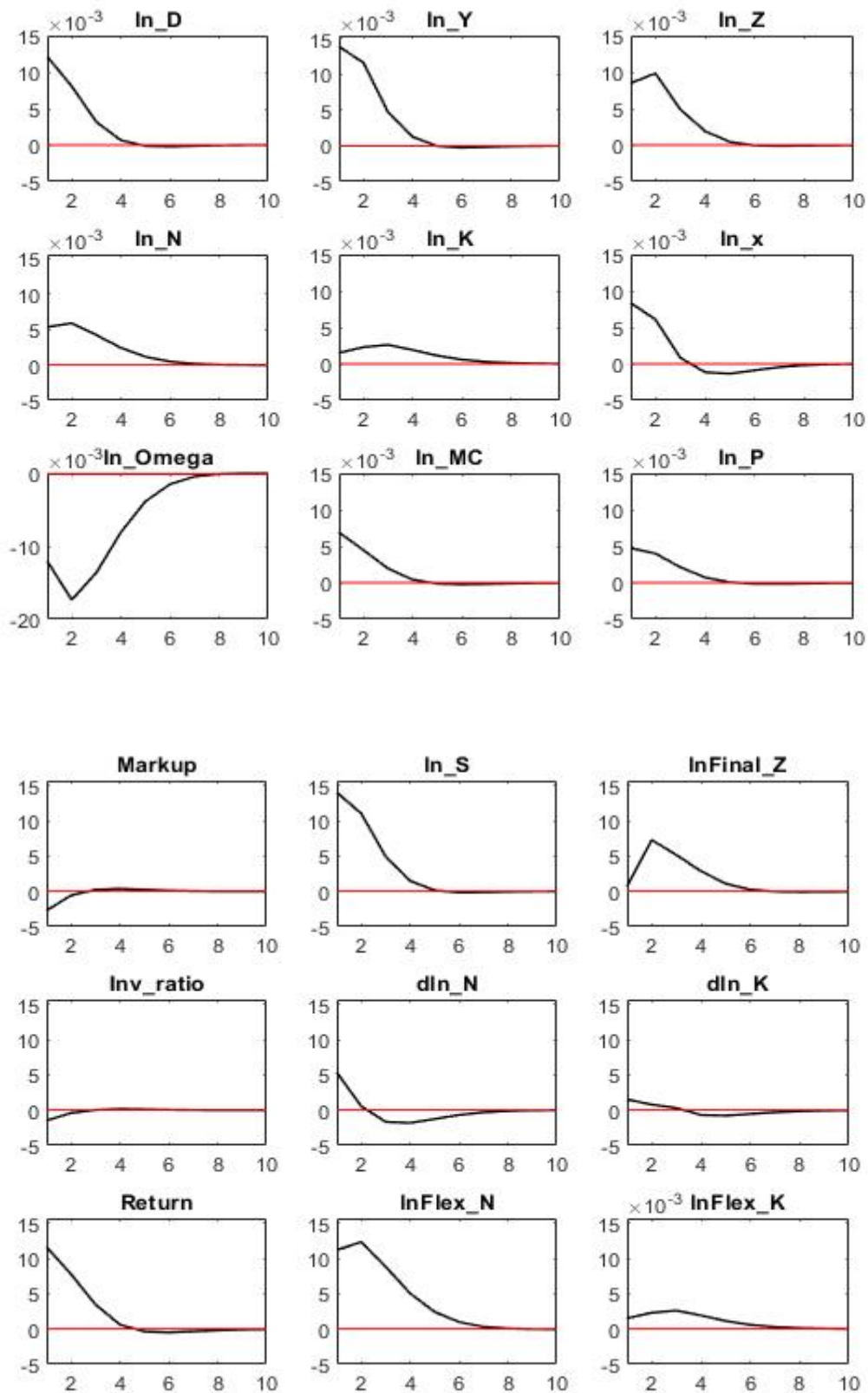


Note: Confidence intervals for the empirical IRFs have been calculated by bootstrapping, i.e. resampling the firms 1000 times with replacement and re-estimating the parameters.

²⁷ An approximate calculation goes as follows:

$$\frac{\Delta Y}{Y} = \left(\frac{\Delta \hat{N}}{\hat{N}} + \frac{\Delta x}{x} \right) (1 - \alpha) = \left(\frac{\Delta \hat{N}}{N} \frac{N}{\hat{N}} + \frac{\Delta x}{x} \right) (1 - \alpha) = \left(0.0053 \cdot \frac{1 + 1.093}{1} + 0.0084 \right) (1 - 0.34) = 0.019 \cdot 0.66 = 0.013$$

Figure 6. Effects of a firm-specific demand shock on variables in the theoretical model



The model explains the strong inventory response to the demand shock. Note, however, that inventories of final goods increase very little in the first year. According to the model, the initial increase in inventories consists mainly of an increase in the stock of inputs.

5.2 Parameter estimates

Parameter estimates and confidence intervals are shown in *Table 2* and *Figure 7* shows the distributions obtained by bootstrapping. In the following, we discuss how adjustment costs, implementation lags, increasing returns and variable utilization help to explain the dynamic adjustment and we try to relate our estimates to previous research.

Distribution of factor returns: $\alpha = 0.342$

The parameter α and the markup determine factor returns. The cost of labor is 73% of value added, which is the lower bound that we have set for the labor share.

Elasticity of substitution: $\sigma = 0.391$

The estimate of the elasticity of substitution is well below unity as is commonly found in the literature; see Knoblach, Roessler and Zwerschke (2020). It is close to the estimates of Chirinko, Fazzari and Meyer (2011) and Chirinko and Mallick (2017) which were based on long run changes of user costs.

Adjustment costs, implementation lags: $c_k = 2.632, c_n = 1.706, \lambda_{k1} = 0.491, \lambda_{k2} = 0.254, \lambda_n = 1$

We match the adjustment of employment and capital very well with a combination of convex adjustment costs and implementation lags. Many authors have found evidence of asymmetric and lumpy adjustment costs and we are not questioning these results, but we see quadratic adjustment costs as a useful stand-in for various types of adjustment costs.²⁸ To interpret the convexity of the adjustment costs, note that units have been chosen so that $K = 1 + F_k$ in steady state, so if investments are ten percent higher than the steady state level (δK) the marginal adjustment cost is $c_k (I - \delta K) = c_k \cdot 0.10 \cdot \delta_k \cdot (1 + F_k) = 0.0195$.²⁹ Since $P_k = 0.47$

²⁸ Alternative models of adjustment costs can be found in Hamermesh and Pfann (1996), Adda and Cooper (2003), Cooper and Haltiwanger (2006), and Bond and Van Reenen (2007). Adjustment costs associated with employment could represent search frictions, but the results in Carlsson, Eriksson and Gottfries (2013) and Stadin (2015) contradict this interpretation.

²⁹ The parameter c_k reflects the *convexity* of the adjustment cost. As discussed by Erickson and Whited (2000) c_k is not informative about the total adjustment cost. Adjustment costs that are linear in investment are indistinguishable from the price of capital P_k and cannot be identified from the IRFs.

this makes the marginal cost of investment 4 percent higher than it would be if investments were at the steady state level. This degree of convexity is much smaller than what is typically found when researchers estimate investment models based on Tobin's Q or forecasts of future marginal returns as in Abel and Blanchard (1986).³⁰ However, most of these studies do not allow for implementation lags. Gilchrist and Himmelberg (1998) assume that there is a one year information that lag and that the time to build is one year and they find an adjustment cost for capital, which is similar in magnitude to our estimate ($c_k \approx 1.7$).

The estimate of λ_{k1} is significantly different from unity and indicates that about half of the investment in capital that is planned in a certain year is implemented in the same year. If we omit implementation lags in investment by setting $\lambda_{k1} = 1$ we are unable to match the delayed and hump-shaped response of the capital stock and we get a much higher adjustment cost for capital: $c_k = 9.367$. In yearly data, we see no evidence of implementation lags in hiring.

Organizational capital: $\xi = 0.011, a = 16.7, \delta_o = 0.287$

The parameters ξ and a determine the slope and concavity of the function that maps the effect of organizational capital on production. As we can see in *Figure 3*, the estimated function has very little concavity, making it desirable for the firm to vary investments in organizational capital when an unexpected shock occurs and inputs adjust sluggishly. As demand increases, time is reallocated from investment in organizational capital to current production and this leads to a 1.7 percent decrease in organization capital in the year following the shock (*Figure 6*).

This finding is consistent with evidence that workers spend a substantial fraction of their time on tasks that increase future rather than current production; see Fay and Medoff (1985) and Kim and Lee (2007). Kim and Lee showed theoretically that even without adjustment costs, skill accumulation will be countercyclical in a real business cycle model because the opportunity cost of skill accumulation is higher when productivity is high, and

³⁰ The estimates reported by Chirinko (1993), Blundell, Bond, Devereux and Schiantarelli (1992) and Erickson and Whited (2000) imply values of c_k equal to 25 or higher. Hall (2004) combines Euler equations for labor, capital and materials to estimate convex adjustment costs and finds a wide dispersion of adjustment costs across industries. He does not take account of implementation lags and estimation is done on first differences, which should be sensitive to timing issues.

similar ideas have emerged in the growth literature.³¹ Bean (1990) constructed a real business cycle model where production factors are used either in production or to accumulate human capital. In his model, there will be less human capital accumulation in periods of high government expenditures. Bean argued that this is consistent with UK data showing high measured productivity during wars. Consistent with his model, he found that a shock to government expenditures has a negative long-run effect on growth. We have incorporated these ideas by assuming that workers spend some of their time building “organizational capital” that increases future productivity, but organizational capital depreciates at a substantial rate ($\delta_\omega = 0.287$) so we do not have endogenous growth.

Model with variations in effort instead of organizational capital: $\Omega = x_t = 1, c_u = 52.9$

An alternative way to model labor hoarding and utilization is to allow for costly variations in effort. Aggregate data for manufacturing shows that official hours worked co-vary very closely with the number of workers employed (see *Figure A1* in the *Appendix*) but there may still be variations in hours per worker or effort on the job. These can be included by allowing effort to affect utilization via the variable u_t , which enters multiplicatively in the production function. Variations in effort and variations in time spent in current production (x_t) have very similar effects on production today. The difference between these two aspects of factor utilization is that variations in effort are associated with a direct cost because workers must be compensated for effort in some way, while disinvestment in organizational capital shows up as lower productivity in the future.³²

To investigate the differences between the two ways of modeling utilization and labor hoarding, we re-estimated the model, allowing effort (u_t) to vary but omitting variations in time spent building organizational capital, i.e. setting $x_t = \Omega_t = 1$. As we can see in *Figure 8*, these two models produce similar impulse-response functions and the overall fit is very similar as measured by the value of the target function in *Table 3*. With the data we have, it is hard to distinguish between these two models, but on a general level, they are consistent with the findings of Basu (1996), Sbordone (1996, 1997), and Basu, Fernald and Kimball (2006).³³ Factor utilization is an important margin of adjustment.

³¹ See Aghion and Saint-Paul (1998) and DeJong and Ingram (2001). Cooper and Johri (2002) assume instead that there is learning-by-doing, so the accumulation of “organizational capital” is positively related to the level of production.

³² If variations in effort were not associated with some cost, the firm would always ask for maximal effort.

³³ One may argue that the model with effort variation produces less reasonable results, with no substitutability of production factors and an implausibly large inventory stock; see *Table 3*

Table 2. Parameter estimates and confidence intervals

Description		Estimate	95% min	95% max	90% min	90% max	min	max
Std. of the shock	ε	0.012	0.012	0.012	0.012	0.012	0.000	0.020
Capital intensity	α	0.342	0.236	0.569	0.259	0.519	0.010	0.990
Elasticity of subst.	σ	0.391	0.133	0.810	0.179	0.652	0.010	2.000
Price stickiness	θ	1.264	0.257	6.387	0.315	5.433	0.000	100.000
Demand sensitivity	Σ	1.528	1.213	2.078	1.267	1.988	0.100	3.000
Demand elasticity	η	6.509	5.121	8.344	5.437	7.735	2.000	500.000
Fixed capital	F_k	0.000	0.000	0.242	0.000	0.138	0.000	3.000
Fixed employment	F_n	1.093	0.669	1.573	0.758	1.424	0.000	3.000
Adjustment cost N	c_n	1.706	0.876	2.495	0.998	2.307	0.000	300.000
Adjustment cost K	c_k	2.632	1.094	5.710	1.325	4.272	0.000	300.000
Investment in t	λ_{k1}	0.491	0.164	0.700	0.227	0.658	0.000	1.000
Investment in t+1	λ_{k2}	0.254	0.160	0.494	0.178	0.450	0.000	1.000
Hiring in t	λ_n	1.000	0.561	1.000	0.594	1.000	0.000	1.000
Utilization cost	c_u	100000	100000	100000	100000	100000	100000	100000
Chi	χ	1.712	0.702	4.031	0.793	3.076	0.000	1000.000
Deprec. org. cap.	δ_ω	0.287	0.013	0.464	0.066	0.443	0.000	0.990
Xi	ξ	0.011	0.010	8.047	0.010	4.503	0.010	100.000
Kappa1	κ_1	3.409	2.377	8.111	2.523	6.626	0.100	1000.000
Finished goods in ss	\bar{Z}	0.111	0.060	0.155	0.069	0.144	0.000	0.500
Input stock/prod.	h_z	0.158	0.040	0.207	0.049	0.199	0.000	0.500
Cost of inventory	c_z	0.360	0.200	0.637	0.249	0.557	0.000	2.000
Demand lag 1	ρ	0.663	0.653	0.674	0.655	0.672	-2.000	2.000
Demand lag 2	ρ_2	-0.179	-0.197	-0.163	-0.193	-0.165	-2.000	2.000
Restrictions in ss.		Estimate					min	max
Return on assets		0.030					0.030	0.300
$W*N/((P-m)*Y)$		0.730					0.730	0.830
P_k*K/Y		0.470					0.370	0.470
Z/Y		0.269					0.100	0.600
h_z/Z		1.426					0.667	1.500
Implied Values								
a		16.715						
$\mu = (P - MC) / P$ in ss		0.206						
κ_2		0.674						
κ_3		5.132						
target function		25.256						

Note to Table 2: The following parameters have been fixed in the estimation:

$c_u = 10^5$, $\beta = 0.96$, $\delta_k = 0.074$, $\delta_n = 0.1$, $\delta_z = 0.1$. Return on assets in steady state is defined as

$$\left(S * P - W * N - m * Y - P_K * \delta_k * K - c_z * \hat{Z} \right) / (P_K * K + Z)$$

where all variables are steady state values.

Figure 7. Distributions for deep parameters

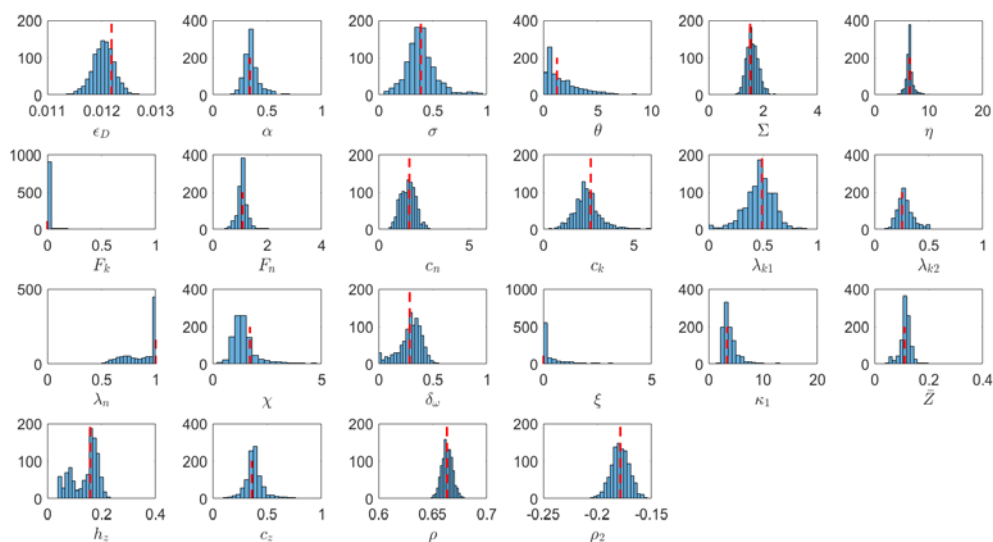
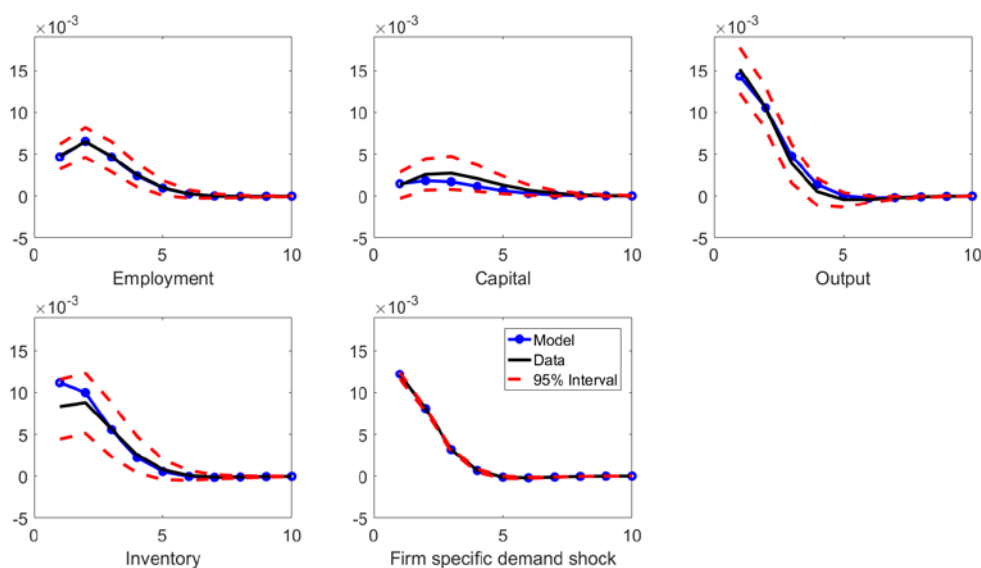


Figure 8. Effects of a firm-specific demand shock in model with costly variation in effort u_t but no variation in organizational capital: $\Omega_t = 1$.



Straight time and overtime

Lucas (1970), Sargent (1978) and Hansen and Sargent (1988) showed that imperfect substitution between straight time and overtime can help to explain the pro-cyclical pattern of the standard Solow residual. Hansen and Sargent (1988) wrote the production function

$Y_t = h_1 A_t K_t^\alpha n_{1t}^{1-\alpha} + h_2 A_t K_t^\alpha n_{2t}^{1-\alpha}$, where n_{1t} is straight time and n_{2t} is overtime. This production function is equivalent to our production function if we fix organizational capital, allow effort to vary and set the elasticity of substitution to unity. To see this, rewrite it as

$Y_t = A_t K_t^\alpha n_{1t}^{1-\alpha} \left(h_1 + h_2 (n_{2t} / n_{1t})^{1-\alpha} \right)$. Changing notation so that $n_{1t} = N_t$ and

$u_t = h_1 + h_2 (n_{2t} / n_{1t})^\alpha$, we obtain our production function with variable effort and substitution elasticity equal to unity.

Returns to scale: $F_k = 0, F_n = 1.093$

Increasing returns to scale were proposed by Hall (1988) as an explanation for why production varies more (in percent) than labor input in response to demand-side shocks. Hall (1988) showed that shocks that should be uncorrelated with technology lead to variations in output that are several times larger (in percent) than the corresponding variation in inputs and he interprets this as evidence of increasing returns to scale. With increasing returns, firms make losses if the price is equal to the marginal cost, and since firms typically do not make losses, not even in periods of low demand, Hall concluded that firms must have very substantial market power. He estimated markups of more than 100 percent for most industries.

Burnside, Eichenbaum and Rebelo (1993) offered an alternative explanation of these observations. They showed that a model with constant returns to scale, perfect competition, implementation lags in employment and variations in effort can account for a positive correlation between the growth rates of the Solow residual and government expenditures. The markup is zero in their model once we take account of the marginal cost of increasing utilization. Models with utilization have been analyzed by e.g. Fairise and Langot (1994), Basu (1996), Braun and Evans (1998), Imbs (1999) and Basu, Fernald and Kimball (2006).³⁴

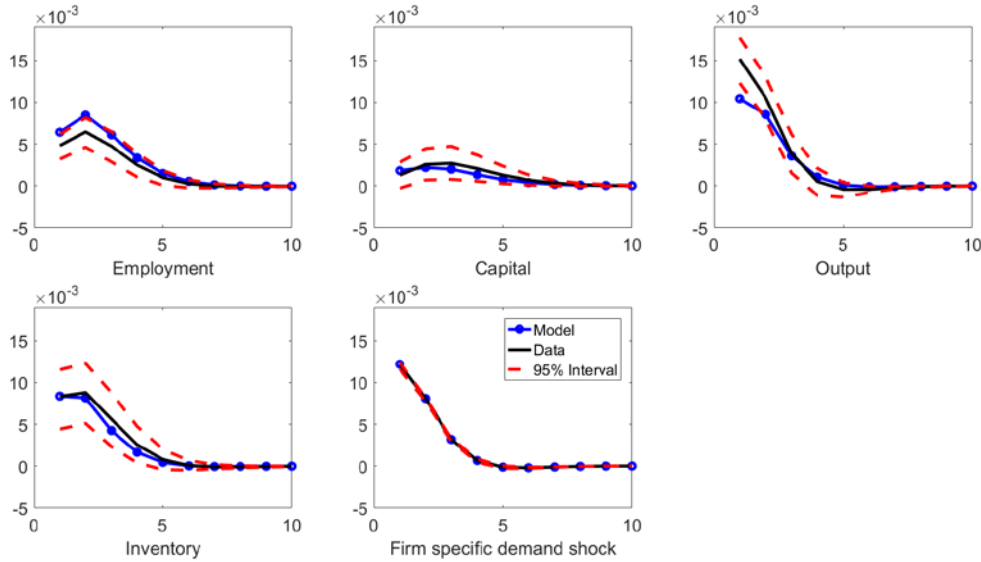
³⁴ Imbs (1999) adjusts Solow residuals for variations in utilization of capital and labor, and he finds that the adjusted residuals are substantially less pro-cyclical than standard series. In his model, utilization can be backed out due to specific functional forms for the utilization costs. Our utilization cost function is more general, so our results are more data-driven and less dependent on theoretical assumptions, but the conclusions are similar.

Table 3. Parameter estimates for alternative models

Description		Baseline model	95% min	95% max	Variable effort	CRS	No utilization	Flex price
Std. of the shock	ε	0.012	0.012	0.012	0.012	0.012	0.012	0.012
Capital intensity	α	0.342	0.236	0.569	0.128	0.180	0.194	0.278
Elasticity of subst.	σ	0.391	0.133	0.810	0.010	0.132	0.053	2.000
Price stickiness	θ	1.264	0.257	6.387	100	100	4.839	0
Demand sensitivity	Σ	1.528	1.213	2.078	1.596	1.286	1.418	1.521
Demand elasticity	η	6.509	5.121	8.344	64.43	129.8	5.542	6.617
Fixed capital	F_k	0.000	0.000	0.242	0.000	0.000	0.440	0.000
Fixed employment	F_n	1.093	0.669	1.573	0.000	0.000	1.371	1.067
Adjustment cost N	c_n	1.706	0.876	2.495	53.103	1.698	1.074	0.624
Adjustment cost K	c_k	2.632	1.094	5.710	32.737	2.784	12.610	0.400
Investment in t	λ_{k1}	0.491	0.164	0.700	0.683	0.710	0.677	0.444
Investment in t+1	λ_{k2}	0.254	0.160	0.494	0.158	0.145	0.161	0.278
Hiring in t	λ_n	1.000	0.561	1.000	0.696	0.743	0.835	0.597
Utilization Cost	c_u	100000			52.887	100000	100000	100000
Chi	χ	1.712	0.702	4.031		3.410		16.619
Deprec. org. cap.	δ_w	0.287	0.013	0.464		0.217		0.015
Xi	ξ	0.011	0.010	8.047		0.020		0.010
Kappa1	κ_1	3.409	2.377	8.111	480.4	25.98	18.97	3.340
Finished goods in ss	\bar{Z}	0.111	0.060	0.155	0.347	0.360	0.043	0.095
Input stock/prod.	h_z	0.158	0.040	0.207	0.231	0.240	0.064	0.143
Cost of inventory	c_z	0.360	0.200	0.637	0.011	0.001	1.456	0.445
Demand lag 1	ρ	0.663	0.653	0.674	0.663	0.663	0.663	0.663
Demand lag 2	ρ_2	-0.179	-0.197	-0.163	-0.179	-0.179	-0.178	-0.179
Restrictions in ss			min	max				
Return on assets		0.030	0.030	0.300	0.044	0.036	0.030	0.030
$W*N/((P-m)*Y)$		0.730	0.730	0.830	0.730	0.753	0.730	0.730
$PK*K/Y$		0.470	0.370	0.470	0.370	0.370	0.370	0.470
Z/Y		0.269	0.100	0.600	0.578	0.600	0.107	0.238
h_z/Z		1.426	0.667	1.500	0.667	0.667	1.500	1.500
Implied Values								
a		16.715				4.614		1.247
$\mu = (P - MC) / P$ in ss		0.206			0.065	0.056	0.247	0.204
κ_2		0.674			-82.757	-4.128	0.457	0.702
κ_3		5.132			689.067	32.875	148.549	3.325
target function		25.256			24.945	59.868	80.063	24.857

Note: See note to *Table 1*. The baseline model has variable organizational capital but no variation in effort. The variable effort model has no variation in organizational capital and c_u is estimated. In the constant returns model there are no fixed costs: $F_n = F_k = 0$. The no utilization model has no variation in organizational capital or effort: $\Omega = u = x = 1$. In the flex-price model we set $\theta = 0$.

Figure 9. Effects of a firm-specific demand shock in model with constant returns to scale: $F_n = F_k = 0$.



Our identification strategy is conceptually similar to that of Hall (1988) and in line with his results, we find that increasing returns to scale are needed to match the empirical responses. If we impose constant returns to scale ($F_k = F_n = 0$) the model produces a too large employment response but still a too small output response to the demand shock; see *Figure 9*. The point estimate $F_n = 1.093$ suggests that roughly half the workers constitute overhead labor that is independent of how much the firm produces. This estimate should be interpreted with some caution, however, because it may reflect the medium-term time horizon that we are considering. The firm may be organized in a certain way in the medium term, but some part of F_n may still be variable in the very long run.³⁵

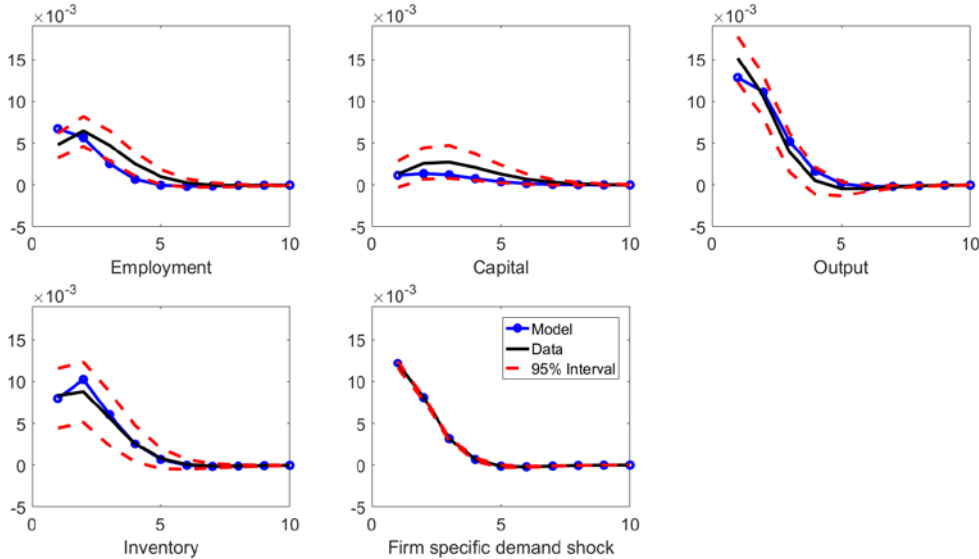
However, increasing returns are not sufficient to match the impulse-response functions without variations in utilization. *Figure 10* shows estimates where we have eliminated all variations in utilization by setting $\Omega = u = 1$. In this case, we find much stronger increasing returns ($F_k = 0.440, F_n = 1.371$) but the model matches the impulse-response

³⁵ Basu (1996) considers a similar specification in Section VI. He also finds evidence of returns to scale (but smaller) in the value added production function.

functions less well. The employment response is too large in the first period but still the increase in output is too small. The model can match the output response in periods 3 and 4 but then the employment response is too small.

To understand why increasing returns cannot be the whole story, note that the empirically estimated increase in employment is only one third of the increase in output in the first year. We would need very strong returns to scale to match this response without variations in utilization. But at the same time, the employment response is equal to that of output two years after the shock and this observation is inconsistent with extreme returns to scale. Put differently, an increase in demand generates an *immediate* increase in factor productivity, as we normally measure it, but most of this effect is *temporary* and disappears after some time. This should not be the case if economies of scale were the main explanation of pro-cyclical factor productivity. Sbordone (1996, 1997) makes a similar observation.

Figure 10. Effects of a firm-specific demand shock in model without variations in utilization: $u_t = \Omega_t = 1$.



External economies of scale

Caballero and Lyons (1990, 1992) showed that when aggregate output is included in industry-level production-function regressions, the estimated coefficient for this variable becomes positive and significant for many industries. Their interpretation is that there are *external* economies of scale and such “Marshallian” externalities have been studied by Cooper and Haltiwanger (1996), Braun and Evans (1998) and others.

Sbordone (1996, 1997) argued that labor hoarding is a better explanation of the results of Caballero and Lyons (1990, 1992). Her point is that when there are demand-side shocks and it takes time to adjust labor and capital, aggregate output may act as a proxy for factor utilization. Sbordone shows empirically that a shock to aggregate output has a *persistent* effect on industry-level output but only a *transitory* effect on industry-level productivity, which supports the labor hoarding interpretation. Our results point strongly in the same direction. Aggregate externalities cannot explain our results because any such effects would be picked up by the time dummies.

Quasi-fixed factors

Increasing returns may take the form of increasing returns in the long-run production function, or there may be some form of *short-run increasing returns* because some factors of production are *quasi-fixed*. Hall (1988) discussed a case where some predetermined amount of overhead labor determines the firm's maximum production capacity, and a certain amount of production labor is needed per unit of output actually produced. When the firm operates below capacity, the marginal cost is the cost of the required production work. Importantly, Hall assumed that there is no cost of increasing the *utilization* of overhead labor.

The problem with this argument is that firms should not choose to always have excess capacity. Sometimes, they will hit their capacity constraints and then the marginal cost will be very high. If the marginal cost curve has one flat part and one part that is very steep, and a positive demand shock increases the fraction of firms that are on the steep part, the *average* of firms' marginal costs will increase with production – as it does in the present model.

Inventory model: $c_z = 0.360$, $\bar{Z} = 0.111$, $h_z = 0.158$, $\kappa_1 = 3.409$, $\kappa_2 = 0.674$, $\kappa_3 = 5.132$

In this model, there are three costs of holding finished goods inventories: the financing cost, depreciation of the inventory stock ($\delta_z \cdot \hat{Z}$) and the storage cost ($c_z \cdot \hat{Z}$). Having set

$\beta = 0.96$ and $\delta_z = 0.10$, we find that the yearly storage cost is large (36 percent of the value), inventories of finished goods correspond to 11 percent of production in steady state and the stock of inputs corresponds to 16 percent of production. There are four structural parameters in the inventory model (Λ, Ψ, T, Φ) but only three estimated parameters ($\kappa_1, \kappa_2, \kappa_3$) so we cannot identify the structural parameters. However, if we exogenously set $\Psi = 1$, so demand

for a certain variety is uniformly distributed between zero and twice expected demand, we get $T = 10$ and $\Lambda = 0.33$; the stock is refilled ten times per year and about one third of the goods need to be in stock in order to be sold and 2/3 of the goods are produced on order.

We find that a demand shock has a strong positive effect on inventory holdings and this response is consistent with a model where firms stock inputs and where stocks of finished goods are held to avoid stock-outs, as suggested by Kahn (1987, 1992). Other studies finding support for inventory models with a stock-out motive are Bils and Kahn (2000), Wen (2005), Galeotti, Maccini and Schiantarelli (2005) and Kryvtsov and Midrigan (2013).³⁶

The inventory stock responds positively to the demand shock but it fails to keep pace with shipments. *Figure 6* shows that in the period when the shock occurs, the stock of finished goods increases much less than production, so the *ratio* of finished-goods inventories to production (and sales) decreases. To understand this, note that the first-order condition for finished-goods inventory holdings is

$$MC_t + c_z - \beta(1 - \delta_z) E_t(MC_{t+1}) = \left(\kappa_1 - 2\kappa_3 \frac{\hat{Z}_t}{D_t^\Sigma P_t^{-\eta}} \right) (P_t - MC_t). \quad (35)$$

where $MC_t = v_t + m$. The left-hand side is the cost of holding an additional unit of finished-goods inventories: the marginal cost today plus the storage cost minus the expected discounted marginal cost next year. The right-hand side is the effect of inventories on sales (due to reduced stock-outs) times the markup (the value of selling one more unit). Without stock-outs, inventories would have no effect on sales, so the right-hand side would be zero. This would imply a pure production-smoothing role for inventories, which would adjust until the sum of the marginal cost today and the storage cost is equal to the discounted marginal cost next year. When inventories contribute to sales, the desired ratio of inventories to demand depends on the markup. As noted by Bils and Kahn (2000), at least one of the following conditions must hold for the *ratio* of inventories to sales to *decrease* when there is an increase in demand:

- i) marginal costs increase relative to discounted future marginal costs, making it more expensive to hold inventories, or
- ii) the markup declines so the gain from holding inventories is reduced.

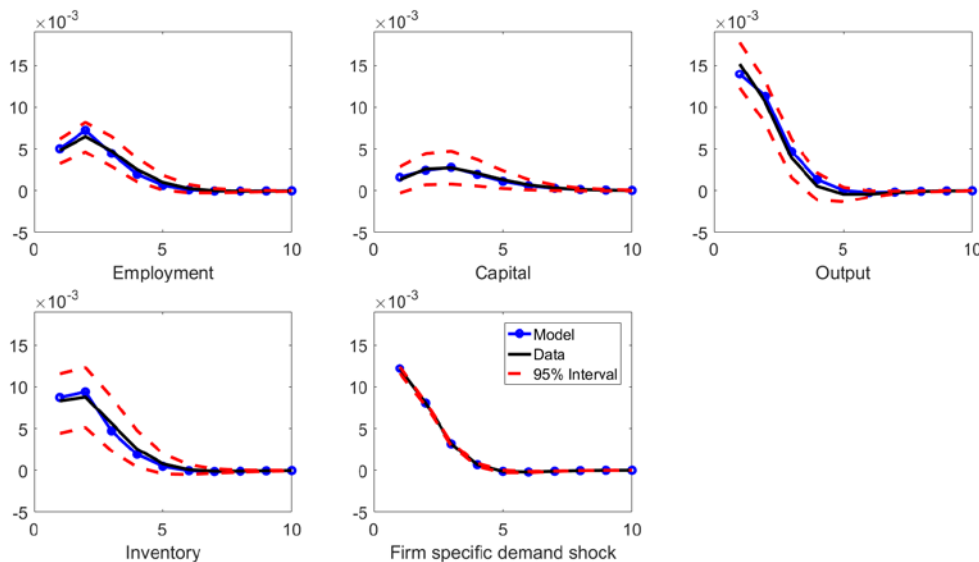
³⁶ An alternative but similar approach is to treat inventories as a factor of production; see Kydland and Prescott (1982), Christiano (1988) and Ramey (1989).

Bils and Kahn argue that there is little evidence of predictable changes in marginal costs but that the markup is lower in booms.³⁷ Looking closely at *Figure 6* we can see that both mechanisms are at work in our model: marginal cost increases relative to discounted future marginal cost and the markup declines.

Price rigidity ($\theta = 1.264$)

Price data are available for some firms but we do not use price data because we do not have a sufficient amount of price data to estimate a model, which includes prices, with robust results. If we eliminate price rigidity by setting $\theta = 0$ we get an equally good fit of the impulse-response functions; see *Figure 11*. However, the substitution elasticity hits the upper bound of 2; see *Table 3*. We find this less plausible and inconsistent with other estimates suggesting that the substitution elasticity is smaller than unity; see e. g. Chirinko, Fazzari and Meyer (2011).

Figure 11. Effects of a firm-specific demand shock in model with flexible prices: $\theta = 0$.



Rotemberg and Summers (1990) argued that price rigidity could explain the results found by Hall (1988). They suggested a model with perfect competition where firms must fix prices before demand is known. With free entry, the price must be equal to the *average cost*, so the price will be above the marginal cost in a recession. In booms, firms produce at the point where the marginal cost equals the predetermined price, so there is rationing in periods of

³⁷ Bils (2016) uses a similar first-order condition to analyze stockouts.

high demand. Thus, the price will be higher than the marginal cost on average. In our model, the markup is sufficiently large and utilization is sufficiently flexible so that the price exceeds the marginal cost throughout the adjustment to a one standard deviation shock and this means that firms always want to satisfy demand.³⁸ However, there will be increased rationing after a positive demand shock as a larger fraction of customers will be unable to find their desired variety.

Price elasticity and markup: $\eta = 6.509$, $\mu_{ss} = (P_{ss} - MC_{ss}) / P_{ss} = 0.206$, *Markup* = 0.26

In a model without inventories, a price elasticity of 6.509 would imply a markup equal to 18 percent of marginal cost, but the possibility of stock-out makes the effective demand curve less elastic for a given inventory stock, so the markup is higher, 26 percent

$(\mu_{ss} / (1 - \mu_{ss}) = 0.26)$. This markup is within the range of estimates reported by De Loecker and Warzynski (2012) and close to the median markup found by de Loecker, Eeckhout and Unger (2020). Carlsson and Smedsaas (2007) estimated the markup for Swedish manufacturing firms to be 17%.

If we close down factor utilization by setting $\Omega = u = 1$, we get stronger returns to scale and a higher markup over marginal cost: $\mu_{ss} / (1 - \mu_{ss}) = 0.33$ (see *Table 3*). Our interpretation is that this markup estimate is biased upwards because we omit the marginal cost of increased utilization.

Demand sensitivity: $\Sigma = 1.528$

A demand shock of one percent shifts the demand curve of the typical firm 1.5 percent for given prices. An elasticity above unity is plausible because many manufacturing firms produce investment goods and durable goods and the demand for such goods is relatively sensitive to shocks. To construct the firm-specific demand variable, we have used data on aggregate consumption and investment, which contain large portions of services, and the demand for services is much more stable than the demand for manufactured goods. Furthermore, investment in stocks of intermediate inputs by *other* firms in the same industry will respond positively to the demand shock and contribute to the volatility of demand for the goods produced by an individual firm. When we constructed the firm-specific demand variable, we did not include inventory investments.

³⁸ The finding of a relatively flat marginal cost curve is consistent with Carlsson and Nordström-Skans (2012), but not with Galeotti, Maccini and Schiantarelli (2005).

Permanent and transitory shocks

Our autoregressive model allows for a persistent demand shock, but we do not allow for unit roots and we do not distinguish permanent and transitory shocks to demand. An alternative would be to allow for permanent and transitory shocks, as in Franco and Philippon (2007), but we leave this for future research. Carlsson, Messina and Nordström Skans (2017) used long-run restrictions to identify demand shocks. Consistent with the present study, they found positive effects of demand shocks on production, employment, and production per worker.

6. Conclusion

Investment, hours worked, labor productivity and inventory holdings are all pro-cyclical, but there is no consensus on how to interpret these correlations. A positive correlation between production and output per worker may arise because productivity shocks drive both variables or because demand-side shocks lead to variations in factor utilization. A positive correlation between production and inventory holdings may arise because firms build up inventories when it is cheap to produce or because firms need more inventory holdings in order to not stock out when demand is high. Since all variables are endogenous, it is difficult to establish causality using macroeconomic data.

The same problem occurs if we use panel data for individual firms: a positive correlation between production and output per worker can be interpreted in different ways, and without some exogenous source of variation, it is difficult to establish causality.

In this paper, we have tried to study the *causal effects* of demand-side shocks on firms' dynamic adjustments. We used macro data, input-output tables, and micro-level export shares to construct measures of demand-side shocks that we take to be exogenous for individual firms and industries. We used firm-level panel data to study how firms in general react to such shocks. We found that production and inventory holdings respond quickly to an increase in demand while the registered inputs of labor and capital respond slowly. We found that these estimated responses can be explained by a theoretical model with adjustment costs, implementation lags, increasing returns, variations in utilization, and where inventories affect sales because they reduce stock-outs.

We study only the effects of specific demand-side shocks, so we cannot draw any conclusions about the relative importance of supply and demand shocks for business cycle fluctuations, but our estimates may be useful as benchmarks for researchers who estimate or calibrate macro models of the business cycle. Most importantly, we have found evidence of

important variations in factor utilization. If we ignore utilization, variations in utilization may erroneously be interpreted as technology shocks, marginal cost will be underestimated and the markup will be overestimated.

This analysis could be extended in many ways. Financial constraints could be included in the theoretical model. Other sectors than manufacturing could be examined and heterogeneity between firms could be analyzed. Effects of other shocks could be analyzed using the same method and other outcome variables could be included in the empirical analysis.

An outcome variable of particular interest is the price that the firm sets. We do have access to producer price data for some of the firms, but we did not include prices in the empirical model. The primary reason is that the sample size would be drastically reduced if we would include only firms for which we have price data and we need a large number of observations to estimate the reduced form model. A second reason is that, to properly model price setting, we would need to take account of customer relations in the product market. We omitted customer relations in the present study because the theoretical model is already quite complicated.³⁹

References

- Abel, Andrew B. and Olivier J. Blanchard, The Present Value of Profits and Cyclical Movements in Investment, *Econometrica*, 54, 249-273.
- Akerberg, Daniel A., Kevin Caves, and Garth Frazer, 2015, Identification Properties of Recent Production Function Estimators, *Econometrica*, 83, 2411-2451.
- Adda, Jérôme, and Russel Cooper, 2003, *Dynamic Economics – Quantitative Methods and Applications*, MIT Press, Cambridge, Massachusetts.
- Aghion, Philippe, and Gilles Saint-Paul, 1998, Virtues of Bad Times, *Macroeconomic Dynamics*, 2, 322-344.
- Arellano, Manuel, and Stephen Bond, 1991, Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *The Review of Economic Studies*, 58, 277–297.

³⁹ Lundin et al. (2009) and Gilchrist et.al (2017) study the implications of customer relations and financial constraints for price setting.

Bagliano, Fabio C. and Carlo A. Favero, 1999, Information from Financial Markets and VAR Measures of Monetary Policy, *European Economic Review*, 43, 825-837.

Baily, Martin N., Eric J. Bartelsman and John Haltiwanger, 2001, Labor Productivity: Structural Change and Cyclical Dynamics, *Review of Economics and Statistics* 83, 420-433.

Bartelsman, Erik, John Haltiwanger, and Stefano Scarpetta, 2013, Cross-Country Differences in Productivity: The Role of Allocation and Selection, *American Economic Review*, 103, 205-334.

Basu, Susanto, 1996, Procyclical Productivity: Increasing Returns or Cyclical Utilization?, *Quarterly Journal of Economics*, 111, 719-751.

Basu, Susanto, and John G. Fernald, 1997, Returns to Scale in U.S. Production: Estimates and Implications, *Journal of Political Economy*, 105, 249-283.

Basu, Susanto, John G. Fernald and Miles S. Kimball, 2006, Are Technology Improvements Contractionary? *American Economic Review*, 96, 1418-1448.

Bean, Charles R., 1990, Endogenous Growth and the Procyclical Behavior of Productivity, *European Economic Review*, 34, 355-363.

Biddle, Jeff E., 2014, The Cyclical Behavior of Labor Productivity and the Emergence of the Labor Hoarding Concept, *Journal of Economic Perspectives*, 28, 197-212.

Bils, Mark, 2016, Deducing Markups from Stockout Behavior, *Research in Economics*, 70, 320-331.

Bils, Mark, and James A. Kahn, 2000, What Inventory Behavior Tells Us about Business Cycles, *American Economic Review*, 90, 458-481.

Blinder, Alan S., and Louis J. Maccini, 1991, Taking Stock: A Critical Assessment of Recent Research on Inventories, *Journal of Economic Perspectives*, 5, 73-96.

Blundell, Richard, Stephen Bond, Michael Devereux and Fabio Schiantarelli, 1992, Investment and Tobin's Q, *Journal of Econometrics*, 51, 233-257.

Bond, Stephen, and John Van Reenen, 2007, Microeconomic Models of Investment and Employment, in *Handbook of Econometrics*, 6A, Elsevier.

Braun, R. Anton, and Charles Evans, 1988, Seasonal Solow Residuals and Christmas: A Case for Labor Hoarding and Increasing Returns, *Journal of Money, Credit and Banking*, 30, 306-330.

- Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, 1993, Labor Hoarding and the Business Cycle, *Journal of Political Economy*, 101, 245–273.
- Carlsson, Mikael, Stefan Eriksson, and Nils Gottfries, 2013, Product Market Imperfections and Employment Dynamics, *Oxford Economic Papers*, 65, 447–470.
- Carlsson, Mikael, Julián Messina, and Oskar Nordström-Skans, 2017, Firm-Level Shocks and Labor Flows, mimeo, Uppsala University.
- Carlsson, Mikael, and Oskar Nordström-Skans, 2012, Evaluating Microfoundations for Aggregate Price Rigidities: Evidence from Matched Firm-Level Data on Product Prices and Unit Labor Cost, *American Economic Review*, 102, 1571-1595.
- Carlsson, Mikael, and Jon Smedsaas, 2007, Technology Shocks and the Labor-Input Response: Evidence from Firm-Level Data, *Journal of Money, Credit and Banking*, 39, 1509–1520.
- Chirinko, Robert S., 1993, Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications, *Journal of Economic Literature*, 31, 1875-1911.
- Chirinko, Robert S., Steven M. Fazzari, and Andrew P. Meyer, 2011, A New Approach to Estimating Production Function Parameters: The Elusive Capital–Labor Substitution Elasticity, *Journal of Business & Economic Statistics*, 29, 587–594.
- Chirinko, Robert S. and Debdulal Mallick, 2017, The Substitution Elasticity, Factor Shares, and the Low-Frequency Panel Model, *American Economic Journal: Macroeconomics*, 9, 225-253.
- Christiano, Lawrence J., 1988, Why Does Inventory Investment Fluctuate So Much? *Journal of Monetary Economics*, 21, 247-280.
- Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, 2005, Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy, *Journal of Political Economy*, 113, 1–45.
- Cooper, Russel W., and John C. Haltiwanger, 2006, On the Nature of Capital Adjustment Costs, *Review of Economic Studies*, 73, 611-633.
- Cooper, Russel W., and Alok Johri, 2002, Learning-by-doing and Aggregate Fluctuations, *Journal of Monetary Economics*, 49, 1539-1566.

Copeland, Adam and George Hall, 2011, The Response of Prices, Sales and Output to Temporary Changes in Demand, *Journal of Applied Econometrics*, 26, 232-269.

DeJong, David, and Beth F. Ingram, 2001, The Cyclical Behavior of Skill Acquisition, *Review of Economic Dynamics*, 4, 536-561.

De Loecker, Jan and Frederic Warzynski, 2012, Markups and Firm-level Export Status, *American Economic Review*, 102, 2437-2471.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger, 2020, The Rise of Market Power and the Macroeconomic Implications, *Quarterly Journal of Economics*, 135, 561-664

Eberly, Janice, Sergio Rebelo and Nicolas Vincent , 2012, What Explains the Lagged Investment Effect? *Journal of Monetary Economics*, 59, 370-380.

Erickson, Timothy and Toni M. Whited, 2000, Measurement Error and the Relationship between Investment and q, *Journal of Political Economy*, 108, 1027-1057.

Eriksson, Stefan, and Karolina Stadin, 2017, What are the Determinants of Hiring? The Importance of Product Market Demand and Search Frictions, *Applied Economics*, 49, 5144-5165.

Fairise, Xavier, and Francois Langot, 1994, Labor Productivity and the Business Cycle: Can R.B.C Models be Saved?, *European Economic Review*, 38, 1581-1594.

Fay, Jon A., and James L. Medoff, 1985, Labor and Output over the Business Cycle: Some Direct Evidence, *American Economic Review*, 75, 638-655.

Field, Alexander J., 2010, The Procyclical Behavior of Total Factor Productivity in the United States, 1890-2004, *Journal of Economic History*, 70, 326-350.

Franco, Francesco, and Thomas Philippon, 2007, Firms and Aggregate Dynamics, *Review of Economics and Statistics*, 89, 587-600.

Garber, Peter M. and Robert G. King, 1983, Deep Structural Excavation? A Critique of Euler Equation Methods, *NBER technical working paper* 31.

Gilchrist, Simon and Charles Himmelberg, 1998, Investment: Fundamentals and Finance, *NBER Macroeconomics Annual*, 13, 223-262.

Gilchrist, Simon, Raphael Schoenle, Jae Sim and Egon Zakrajšek, 2017, Inflation Dynamics during the Financial Crisis, *American Economic Review*, 107, 785-823.

Galeotti, Marzio, Louis J. Maccini and Fabio Schiantarelli, 2005, Inventories, Employment and Hours, *Journal of Monetary Economics*, 52, 575-600.

Gertler, Mark, and Peter Karadi, 2015, Monetary Policy Surprises, Credit Costs, and Economic Activity, *American Economic Journal: Macroeconomics*, 7, 44-76.

Greenwood, Jeremy, Zvi Hercowitz and Per Krusell, 2000, The Role of Investment-Specific Technological Change in the Business Cycle, *European Economic Review*, 44, 91-115.

Hall, Robert E., 1988, The Relation between Price and Marginal Cost in U.S. Industry, *Journal of Political Economy*, 96, 921-947.

Hall, Robert E., 2004, Measuring Factor Adjustment Costs, *Quarterly Journal of Economics*, 119, 899-927.

Hansen, Gary D., and Thomas J. Sargent, 1988, Straight Time and Overtime in Equilibrium, *Journal of Monetary Economics*, 21, 281-308.

Hamermesh, Daniel S., and Gerard Pfann, 1996, Adjustment Costs in Factor Demand, *Journal of Economic Literature*, 34, 1264-1292.

Hart, Robert A., and James R. Malley, 1999, Procyclical Labor Productivity: A Closer Look at a Stylized Fact, *Economica*, 66, 533-550.

Humphreys, Brad R., Louis J. Maccini and Scott Schuh, 2001, Input and Output Inventories, *Journal of Monetary Economics*, 47, 347-375.

Iacoviello, Matteo, Fabio Schiantarelli, and Scott Schuh, 2011, Input and Output Inventories in General Equilibrium, *International Economic Review*, 52, 1179-1213.

Imbs, Jean M., 1999, Technology, Growth, and the Business Cycle, *Journal of Monetary Economics*, 44, 65-80.

Ireland, Peter N., 2004, A Method for Taking Models to the Data, *Journal of Economic Dynamics and Control*, 28, 1205-1226.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti, 2010, Investment Shocks and Business Cycles, *Journal of Monetary Economics*, 57, 132-145.

Kahn, James A., 1987, Inventories and the Volatility of Production, *American Economic Review*, 77, 667-679.

Kahn, James A., 1992, Why is Production More Volatile than Sales? Theory and Evidence on the Stockout-Avoidance Motive for Inventory-Holding, *Quarterly Journal of Economics*, 107, 481-510.

Khan, Aubhik, and Julia K. Thomas, 2007, Inventories and the Business Cycle: An Equilibrium Analysis of (S,s) Policies, *American Economic Review*, 97, 1165-1188.

Kim, Daehaeng, and Chul-In Lee, 2007, On-the-job Human Capital Accumulation in a Real Business Cycle Model: Implications for Intertemporal Substitution Elasticity and Labor Hoarding, *Review of Economic Dynamics*, 10, 494-518.

Knoblauch, Michael, Martin Roessler and Patrick Zwerschke, 2020, The Elasticity of Substitution Between Capital and Labor in the US Economy: A Meta-Regression Analysis, *Oxford Bulletin of Economics and Statistics*, 82, 62-81.

Kryvtsov, Oleksiy, and Virgiliu Midrigan, 2013, Inventories, Markups, and Real Rigidities in Menu Cost Models, *Review of Economic Studies*, 80, 249-276.

Kydland, Finn, and Edward Prescott, 1982, Time to Build and Aggregate Fluctuations, *Econometrica*, 50, 1345-1370.

Levinsohn, James, and Amil Petrin, 2003, Estimating Production Functions Using Inputs to Control for Unobservables, *Review of Economic Studies*, 70, 317-341.

Lucas, Robert E. Jr., 1970, Capacity, Overtime, and Empirical Production Functions, *American Economic Review*, 60, Papers and Proceedings 23-27.

Lundin, Magnus, Nils Gottfries, Charlotte Bucht and Tomas Lindström, 2009, Price and Investment Dynamics: Theory and Plant-Level Data, *Journal of Money, Credit, and Banking*, 41, 907-934.

Mickelsson, Glenn, 2016, Estimation of DSGE Models with Uninformative Priors, in Mickelsson, Glenn, *DSGE Model Estimation and Labor Market Dynamics*, PhD Thesis, Economic Studies 163, Department of Economics, Uppsala University.

Nekarda, Christopher J. and Valerie A. Ramey, 2011, Industry Evidence on the Effects of Government Spending, *American Economic Journal: Macroeconomics*, 3, 36-59.

Nelder, J. A., and R. Mead, 1965, A Simplex Method for Function Minimization, *The Computer Journal*, 7, 308-313.

Olley, Steven, and Ariel Pakes, 1996, The Dynamics of Productivity in the Telecommunications Equipment Industry, *Econometrica*, 64, 1263-1295.

Paul, Pascal, 2020, The Time-varying Effect of Monetary Policy on Asset Prices, *Review of Economics and Statistics*, 102, 690-704.

Perotti, Roberto, 2008, In Search of the Transmission Mechanism of Fiscal Policy, *NBER Macroeconomics Annual*, 2007, 169-226, ed. Daron Acemoglu, Kenneth Rogoff and Michael Woodford, University of Chicago Press.

Prescott, Edward C., 1986, Theory Ahead of Business Cycle Measurement, *Carnegie-Rochester Conference Series on public Policy*, 25, 11-44.

Ramey, Valerie A., 1989, Inventories as Factors of Production and Economic Fluctuations, *The American Economic Review*, 79, 338-354.

Ramey, Valerie A., 2016, Macroeconomic Shocks and Their Propagation, *Handbook of Macroeconomics*, 2, 71-162.

Rotemberg, Julio J., and Lawrence H. Summers, 1990, Inflexible Prices and Procyclical Productivity, *Quarterly Journal of Economics*, 105, 851-873.

Rotemberg, Julio J., and Michael Woodford, 1999, The Cyclical Behavior of Prices and Costs, *Handbook of Macroeconomics*, edited by John Taylor and Michael Woodford, Amsterdam: North-Holland Elsevier.

Sargent, Thomas J., 1978, Estimation of Dynamic Labor Demand Schedules under Rational Expectations, *Journal of Political Economy*, 86, 1009-1044.

Sbordone, Argia M., 1996, Cyclical Productivity in a Model of Labor Hoarding, *Journal of Monetary Economics*, 38, 331-361.

Sbordone, Argia M., 1997, Interpreting the Procyclical Productivity of Manufacturing Sectors: External Effects or Labor Hoarding? *Journal of Money, Credit, and Banking*, 29, 26-45.

Shea, John, 1993, Do Supply Curves Slope Up? *Quarterly Journal of Economics*, 108, 1-32.

Stadin, Karolina, 2015, Firms' Employment Dynamics and the State of the Labor Market, Working Paper Series 2015:20, IFAU, forthcoming in *Macroeconomic Dynamics*.

Urban, Timothy L., 2005, Inventory models with inventory-level-dependent demand: A comprehensive review and unifying theory, *European journal of Operations Research*, 162, 792-804.

Wen, Yi, 2005, Understanding the Inventory Cycle, *Journal of Monetary Economics*, 52, 1533-1555.

APPENDIX

Using a reduced-form model to find firms' responses

Studying firm dynamics is difficult. To describe an individual firm's dynamic optimization problem, we would need to observe a large set of state variables that are relevant to the firm's decisions. The problem is that we cannot observe all the relevant state variables, and if we estimate the decision rules without some of the state variables, this will lead to biased estimates. If, for example, we find that past employment helps to explain current employment, we may interpret this as evidence of adjustment costs, but the result may equally well be due to some omitted state variable that affects employment. The same argument applies to Euler equation estimation, as was already pointed out by Garber and King (1983). In this section, we explain the underlying logic of our approach and we compare it to alternative estimation methods.

The idea in this study is to represent the relevant set of state variables using current and lagged values of the variables that we can observe. By estimating a reduced-form model, we determine how firms respond to specific demand shocks that we consider as exogenous relative to technology and cost shocks that affect the firm and the industry. Aggregate state variables are "mopped up" by time dummies, so the estimation is based on "diff-in-diff" variation. Then we estimate the structural parameters by matching the responses to the demand shocks that we have estimated.

To see how this might work (or not work), consider *as a simplified example* a standard model of a firm with convex adjustment costs related to changes in labor and capital and a convex cost of utilization. The firm faces a downward-sloping demand curve, and production is given by the production function: $y_t = (1 - \alpha)n_t + \alpha k_{t-1} + u_t + a_t$, where y_t is production, n_t is employment, k_t is the capital stock at the end of the period t and u_t is utilization. All variables are in logs and firm-specific. There are two exogenous state variables that matter for the firm: factor productivity a_t , which we cannot observe, and a demand shifter d_t , which we can observe. The exogenous state variables follow AR(1) processes:

$$d_t = \rho_d d_{t-1} + \varepsilon_{dt} \quad \text{and} \quad a_t = \rho_a a_{t-1} + \varepsilon_{at} \quad \text{where} \quad E(\varepsilon_{dt} \varepsilon_{at}) = 0.$$

An approximate solution to the firm's dynamic optimization problem consists of log-linear decision rules relating employment, the capital stock, and utilization to the initial levels of capital and employment as well as demand and factor productivity:

$$n_t = b_{11}n_{t-1} + b_{12}k_{t-1} + b_{13}d_t + b_{14}a_t + \varepsilon_{nt} \tag{36}$$

$$k_t = b_{21}n_{t-1} + b_{22}k_{t-1} + b_{23}d_t + b_{24}a_t + \varepsilon_{kt} \quad (37)$$

$$u_t = b_{31}n_{t-1} + b_{32}k_{t-1} + b_{33}d_t + b_{34}a_t. \quad (38)$$

We have added white noise shocks to the decision rules for labor and capital. These may represent temporary variation in factor prices, for example.

Now, we do not observe a_t so we cannot estimate these decision rules directly. However, we can use the equation for the productivity process to substitute for current productivity and then the production function in period t-1 combined with the decision rule for utilization to substitute for lagged productivity. By doing so, we obtain a reduced-form model with serially uncorrelated shocks and a recursive structure:⁴⁰

$$n_t = b_{11}n_{t-1} + b_{12}k_{t-1} + b_{13}d_t + b_{14} \left(\rho_a \frac{y_{t-1} - (1-\alpha)n_{t-1} - \alpha k_{t-2} - b_{31}n_{t-2} - b_{32}k_{t-2} - b_{33}d_{t-1} + \varepsilon_{a,t}}{1 + b_{34}} + \varepsilon_{n,t} \right) + \varepsilon_{n,t} \quad (39)$$

$$k_t = b_{21}n_{t-1} + b_{22}k_{t-1} + b_{23}d_t + b_{24} \left(\rho_a \frac{y_{t-1} - (1-\alpha)n_{t-1} - \alpha k_{t-2} - b_{31}n_{t-2} - b_{32}k_{t-2} - b_{33}d_{t-1} + \varepsilon_{a,t}}{1 + b_{34}} + \varepsilon_{k,t} \right) + \varepsilon_{k,t} \quad (40)$$

$$y_t = (1-\alpha)[b_{11}n_{t-1} + b_{12}k_{t-1} + b_{13}d_t + \varepsilon_{n,t}] + \alpha k_{t-1} + ((1-\alpha)b_{14} + b_{34} + 1) \left(\rho_a \frac{y_{t-1} - (1-\alpha)n_{t-1} - \alpha k_{t-2} - b_{31}n_{t-2} - b_{32}k_{t-2} - b_{33}d_{t-1} + \varepsilon_{a,t}}{1 + b_{34}} + \varepsilon_{a,t} \right) \quad (41)$$

$$d_t = \rho_d d_{t-1} + \varepsilon_{dt}. \quad (42)$$

Effectively, we have “mopped up” the effect of the initial firm-specific productivity level by including lagged values of production, labor, and capital on the right-hand side. If productivity follows an AR(2) process, we can account for this in the same way by including additional lags. Note that the shocks in the first three equations are mixtures of various structural shocks and that the coefficients are hard to interpret because they reflect both the decision rules and the dynamics of the unobserved productivity shock. However, the estimated *response* to a demand shock, ε_{dt} , has a clear causal interpretation, so this model can be estimated and used to trace out how firms respond to demand shocks.

In any practical application, there will be many unobserved state variables, so linear combinations of observed state variables will represent the unobserved state variables in an imperfect way – the “mopping up” will be less than perfect. Still, we can hope that our reduced-form model captures firms’ dynamic responses to demand shocks in a rough way.

⁴⁰ Here we use the fact that $y_{t-1} = (1-\alpha)n_{t-1} + \alpha k_{t-2} + b_{31}n_{t-2} + b_{32}k_{t-2} + b_{33}d_{t-1} + b_{34}a_{t-1} + a_{t-1}$. We are grateful to Fabio Schiantarelli for pointing out that this specific example could be generalized to allow for variable utilization.

This reduced form model is a VAR model with an exogenous variable. It is similar to the VAR-X models with shocks from financial markets, which have been estimated by Bagliano and Favero (1999) and Paul (2020), but we allow the exogenous variable to follow an autoregressive process rather than being serially uncorrelated and we include an equation for the exogenous variable in the system.

An alternative would be to estimate a standard vector-autoregressive model with the endogenous variables and then use the innovations to d_t as instruments for the shocks (see Gertler and Karadi, 2015, Ramey, 2017). By including the exogenous variable d_t explicitly in the system, we take explicit account of the serial correlation in this variable. Of course, the assumption that d_t follows an autoregressive process should be seen as an approximation.

Ireland (2004) considers a structural model with an economically interpretable (productivity) shock and to this structural model he adds a reduced-form system of unspecified shocks. This is conceptually similar to our approach but the key difference is that we have a measure of the interpretable (demand) shock.

Levinsohn and Petrin (2003) and Akerberg, Caves and Frazer (2015) invert the input demand function $m_t = f(n_t, k_{t-1}, a_t)$ and use the result to substitute for the unobserved productivity shock. Then, the production function is estimated in a two-step procedure as in Olley and Pakes (1996). This procedure requires input prices to be the same for all firms and it does not allow for variations in utilization.

Calculation of the real capital stock

The real capital stock (Kr) consists of machines and buildings. In the firm-level panel data, we have firms' book values of buildings and machinery, but generous depreciation allowances imply that the book values are much lower than the economic values of these stocks. With a too low value of the stock of capital, we would exaggerate the volatility of the capital stock measured as log changes. For this reason, we tried to construct a better measure of the real capital stock. We did this in three steps:

First, we obtained industry-level estimates of capital stocks and book values from Statistics Sweden. Using these data for the years 2000-2005, we calculated an *average* ratio of book value to economic value at the industry level (2-digit SNI92) for buildings and machines separately. This ratio was then used to scale up the book values of buildings and machines for the first year that a firm appears in the sample.

Adding buildings and machines together and dividing by a price index for investments, with base year 2000, we express the first-year capital stock in prices in the year 2000.

Finally, we calculated capital stocks for subsequent years by the perpetual inventory method, subtracting depreciation and adding investments in machines and buildings deflated by the investment price index. This was repeated for each year that the firm appears in the sample. We set the depreciation rate of capital to 7.4 percent based on a weighted average of the depreciation rates for buildings and machines used by the Statistics Sweden. The median ratio of (owned) real capital to total real production (output) is 0.38, more than twice as large as the median book value of capital relative to production which is 0.14 (see *Table 1*). This is because depreciation is much higher in the accounting than the estimated economic depreciation rate that we use.

The median ratio of value added of total production (output) is 0.38 in this sample and the median ratio of real capital to value added is 0.97. This may strike readers as a low value, but firms rent a substantial fraction of the capital that they use. Firms may lease cars, trucks and other types of equipment and the buildings in which they operate. We do not have data for rented capital and we effectively treat rented capital as a flexible input.

Motivation for the demand index

To motivate the demand index, let us consider an economy where goods produced in J different industries (indexed j) are used for consumption and investment. Let aggregate investment be a Cobb-Douglas aggregate of composite goods produced in different industries, where the latter are CES aggregates of goods produced for investment by different firms within the industry:

$$I = \prod_{j=1}^J I_j^{\theta_j^I} \quad \text{where} \quad I_j = \left(\sum_{i \in j} (Q_i^I)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \quad \text{and} \quad \sum_{j=1}^J \theta_j^I = 1. \quad (43)$$

Q_i^I denotes the amount of goods produced by firm i and used for investment. Let P_i be the price charged by firm i . Investors minimize the cost of a given investment I subject to these constraints. Maximizing

$$L = - \sum_{i=1}^N P_i Q_i^I + \Upsilon_I \left(\prod_{j=1}^J I_j^{\theta_j^I} - I \right) + \sum_{j=1}^J \Upsilon_j \left(\left(\sum_{i \in j} (Q_i^I)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} - I_j \right) \quad (44)$$

with respect to I_j and Q_i^I , we obtain first-order conditions $\Upsilon_j I_j = \theta_j^I \Upsilon_I I$ and

$Q_i^I = \left(P_i / \Upsilon_j \right)^{-\eta} I_j$, where Υ_I is the relevant price index of aggregate investment, $\Upsilon_I I$ is total investment expenditures, and Υ_j is the relevant price index for goods produced in industry j .

Applying the same reasoning to aggregate consumption, we obtain the total demand for goods produced by firm i in industry j :

$$\hat{D}_i = Q_i^I + Q_i^C = \frac{\theta_j^I \Upsilon_I I \left(\frac{P_i}{\Upsilon_j} \right)^{-\eta}}{\Upsilon_j} + \frac{\theta_j^C \Upsilon_C C \left(\frac{P_i}{\Upsilon_j} \right)^{-\eta}}{\Upsilon_j} = \left(\theta_j^I \Upsilon_I I + \theta_j^C \Upsilon_C C \right) P_i^{-\eta} \Upsilon_j^{\eta-1}. \quad (45)$$

Taking logs on both sides and linearizing with respect to $\ln C$, $\ln \Upsilon_I$, $\ln I$, and $\ln \Upsilon_j$ we obtain

$$\begin{aligned} \ln \hat{D}_i &= \ln \left(\theta_j^I e^{\Upsilon_I \ln I} + \theta_j^C e^{\Upsilon_C \ln C} \right) - \eta \ln P_i + (\eta - 1) \ln \Upsilon_j \\ &\approx \tilde{D} + \frac{\theta_j^I \bar{\Upsilon}_I \bar{I} (\ln \Upsilon_I + \ln I) + \theta_j^C \bar{\Upsilon}_C \bar{C} (\ln \Upsilon_C + \ln C)}{\theta_j^I \bar{\Upsilon}_I \bar{I} + \theta_j^C \bar{\Upsilon}_C \bar{C}} - \eta \ln P_i + (\eta - 1) \ln \lambda_j \\ &= \tilde{D} + \phi_j^I \ln I + (1 - \phi_j^I) \ln C - \eta (\ln P_i - \ln \Upsilon_j) + \left[\phi_j^I \ln \Upsilon_I + (1 - \phi_j^I) \ln \Upsilon_C - \ln \Upsilon_j \right] \end{aligned} \quad (46)$$

where \tilde{D} is a constant, bars denote steady-state values and $\phi_j^I = \bar{I}_j / (\bar{I}_j + \bar{C}_j)$. We see that the weights are the steady-state shares of production in industry j that are used for investment and consumption. Omitting the term in square brackets, we have the demand specification that we use in the theoretical model.⁴¹ The same logic can be applied to government expenditures and sales in different countries.

Construction of the steady state

The estimation involves a very large number of repeated simulations of the model. To save time in this process, we need to calculate the steady-state values analytically instead of searching for the steady state. Let variables with a bar denote steady-state values. Without loss of generality, we can choose units so that

$$\bar{Y} = \bar{K} = \bar{N} = \bar{\Omega} = \bar{P} = \bar{D} = \bar{u} = 1 \quad (47)$$

where \bar{K} and \bar{N} denote the steady-state values of flexible capital and employment. For given values of δ_ω and χ , we can calculate x in the steady state: $\bar{x} = 1 - \delta_\omega / \chi$, and then our normalizations imply a value for A :

⁴¹ In the special case when there is a set of symmetric sectors specializing in either investment or consumption goods, this term will be zero because $\phi_j^I = 1$ and $\ln \Upsilon_j = \ln \Upsilon_I$ for sectors specializing in investment goods and $\phi_j^I = 1$ and $\ln \Upsilon_j = \ln \Upsilon_C$ for sectors specializing in consumption goods.

$$A = \left(\alpha + (1-\alpha) \bar{x}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{1-\sigma}}. \quad (48)$$

It is convenient to view the steady-state inventory stock of finished goods as a parameter to be estimated. Denoting this value \bar{Z} , we obtain steady-state sales as $\bar{S} = 1 - \delta_z \bar{Z}$. In the steady state, we have

$$\kappa_1 \bar{Z} + \kappa_2 - \kappa_3 \bar{Z}^2 = \bar{S} \quad \frac{\bar{S} / \eta}{1 - \bar{v} - m} = \kappa_2 + \kappa_3 \bar{Z}^2 \quad (49)$$

$$\frac{r_z (\bar{v} + m) + c_z \bar{\mu}}{1 - \bar{v} - m} = \kappa_1 - 2\kappa_3 \bar{Z} \quad \text{where } r_z = 1 + \beta(1 - \delta_z).$$

Multiplying the last equation by \bar{Z} and summing both sides of these equations, we can solve for the marginal cost of real value added in the steady state:

$$\bar{v} = \frac{1 - m - 1/\eta - (r_z m + c_z) \bar{Z} / \bar{S}}{1 + r_z \bar{Z} / \bar{S}}. \quad (50)$$

For a given estimate of κ_1 , we can then use the equations above to solve for κ_3 and κ_2 .

We can also find the capital price and wage that are consistent with our normalizations:

$$P^k = \beta \bar{v} \alpha A^{1-1/\sigma} / (1 - \beta(1 - \delta_k)) \quad \text{and} \quad W = \bar{v} (1 - \alpha) (A \bar{x})^{1-1/\sigma}. \quad (51)$$

The first-order conditions for x and Ω yield

$$a = \frac{(1 - \beta(1 - \delta_\omega))(1 - \alpha)}{\beta \xi \chi \bar{x}^{1/\sigma}} + 1. \quad (52)$$

Furthermore, the normalizations imply that $\Phi_u = \bar{v}$ and $\bar{\phi} = \beta \bar{v} (a - 1) \xi / (1 - \beta(1 - \delta_\omega))$.

Table A1. Distribution of the observations across industries in the baseline panel (all in the manufacturing sector)

Industry (SNI 92)	Percent of observations
15 Food products and beverages	7.1
16 Tobacco products	0
17 Textiles	1.7
18 Wearing apparel; furs	0.3
19 Leather and leather products	0.4
20 Wood and products of wood and cork (except furniture)	8.4
21 Pulp, paper, and paper products	3.2
22 Printed matter and recorded media	3.2
23 Coke, refined petroleum products and nuclear fuels	0.2
24 Chemicals, chemical products, and man-made fibers	3.4
25 Rubber and plastic products	7.5
26 Other non-metallic mineral products	2.2
27 Basic metals	3.0
28 Fabricated metal products, except machinery and equipment	24.3
29 Machinery and equipment n.e.c.	16.4
30 Office machinery and computers	0.4
31 Electrical machinery and apparatus n.e.c.	3.0
32 Radio, television, and communication equipment & apparatus	0.5
33 Medical, precision and optical instruments, watches, & clocks	3.2
34 Motor vehicles, trailers, and semi-trailers	4.9
35 Other transport equipment	1.6
36 Furniture; other manufactured goods n.e.c.	5.3

Note: Industry definitions in SNI92 and SNI2002 are almost the same at the 2-digit level.

Table A2. Estimated empirical model

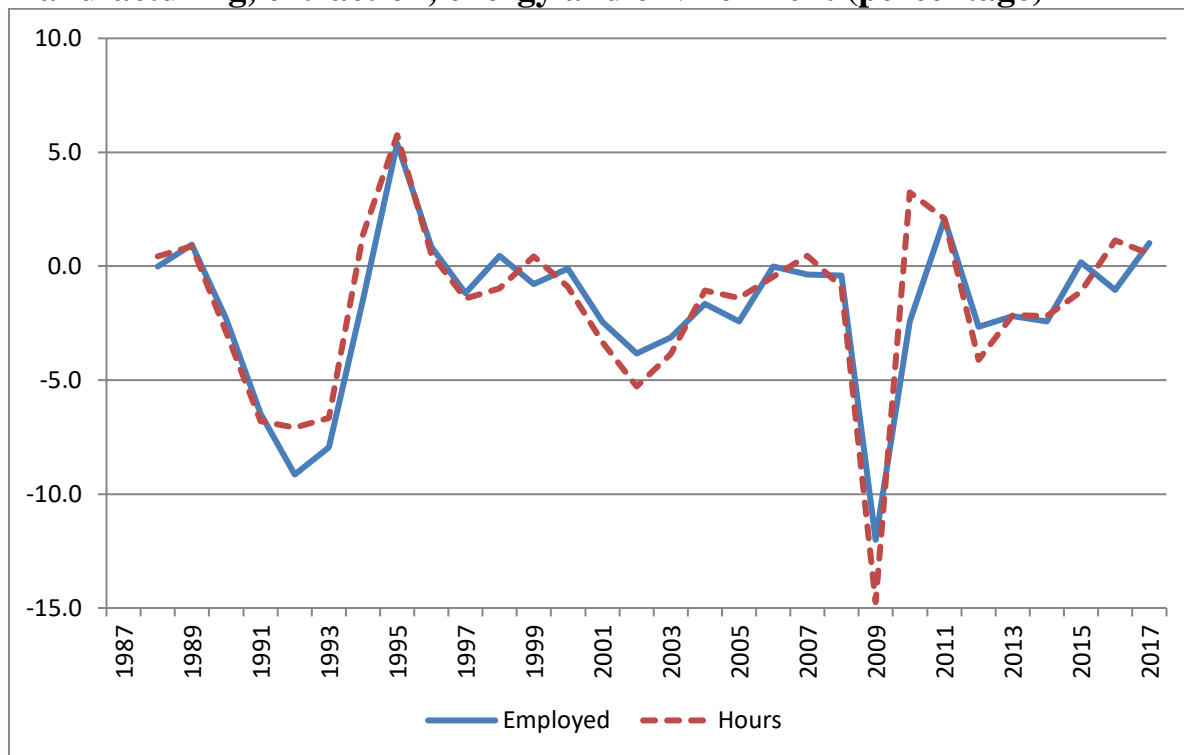
Equations for firm level production, employment, capital, and inventories				
	(1)	(2)	(3)	(4)
	lnYr	lnN	lnKr	lnZr
L.lnYr	0.438*** (0.015)	0.107*** (0.008)	0.046*** (0.007)	0.098*** (0.018)
L2.lnYr	-0.052*** (0.010)	-0.023*** (0.007)	-0.002 (0.007)	-0.007 (0.016)
L.lnN	0.081*** (0.017)	0.545*** (0.019)	0.025*** (0.010)	0.101*** (0.025)
L2.lnN	-0.020 (0.016)	-0.043*** (0.014)	0.011 (0.009)	-0.058** (0.023)
L.lnKr	0.024 (0.015)	0.050*** (0.009)	0.767*** (0.015)	0.052** (0.021)
L2.lnKr	-0.043*** (0.013)	-0.058*** (0.008)	-0.144*** (0.012)	-0.040** (0.020)
L.lnZr	0.071*** (0.007)	0.045*** (0.004)	0.008** (0.003)	0.400*** (0.020)
L2.lnZr	-0.023*** (0.006)	-0.013*** (0.004)	-0.007** (0.003)	-0.006 (0.013)
lnD	1.243*** (0.111)	0.394*** (0.066)	0.103 (0.065)	0.683*** (0.146)
L.lnD	-0.587*** (0.102)	-0.116* (0.060)	-0.010 (0.061)	-0.172 (0.141)
Observations	20,700	20,700	20,700	20,700
Number of firms (FAD id)	1,150	1,150	1,150	1,150
R-squared	0.706	0.796	0.872	0.629

AR2 process for product demand

lnD	
L.lnD	0.663*** (0.005)
L2.lnD	-0.179*** (0.009)
Observations	20,700
Number of firms (FAD id)	1,150
R-squared	0.989
St. d. of residual	0.012

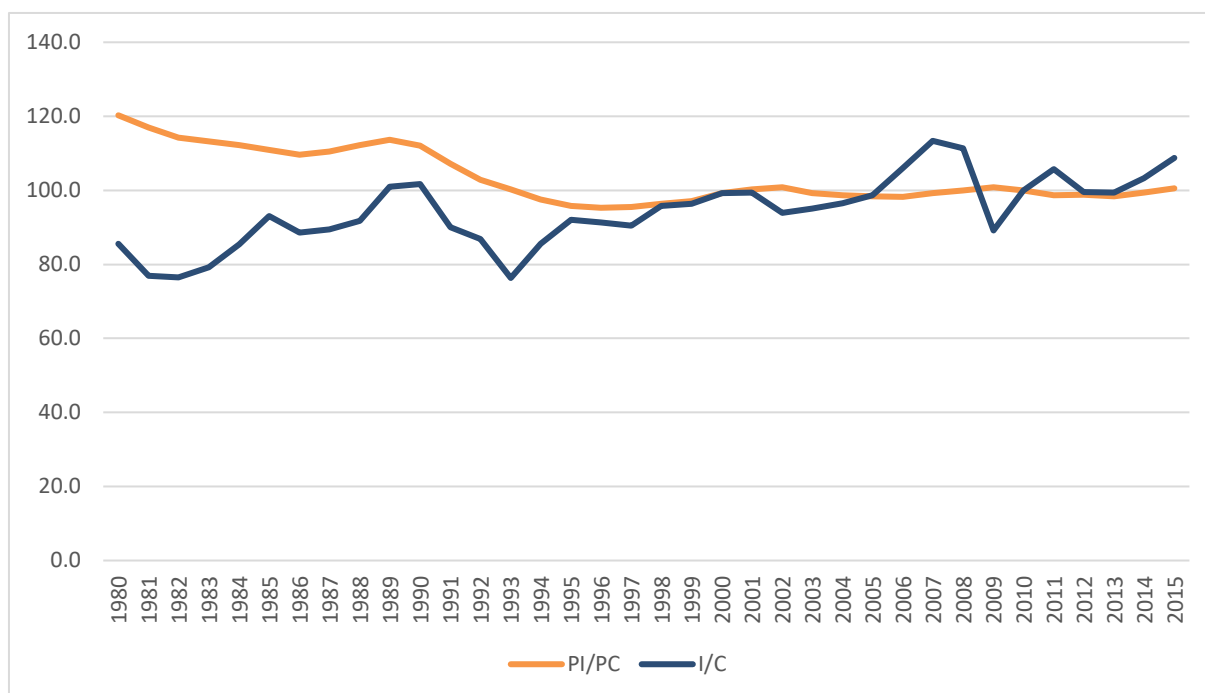
Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1
Time dummies, firm trends, and firm fixed effects are included in all regressions.

Figure A1. Growth rate of the number employed and hours worked in manufacturing, extraction, energy and environment (percentage)



Note: The series includes extraction, energy and environment because of data availability. The correlation between the two variables is 0.91. Source: Statistics Sweden homepage.

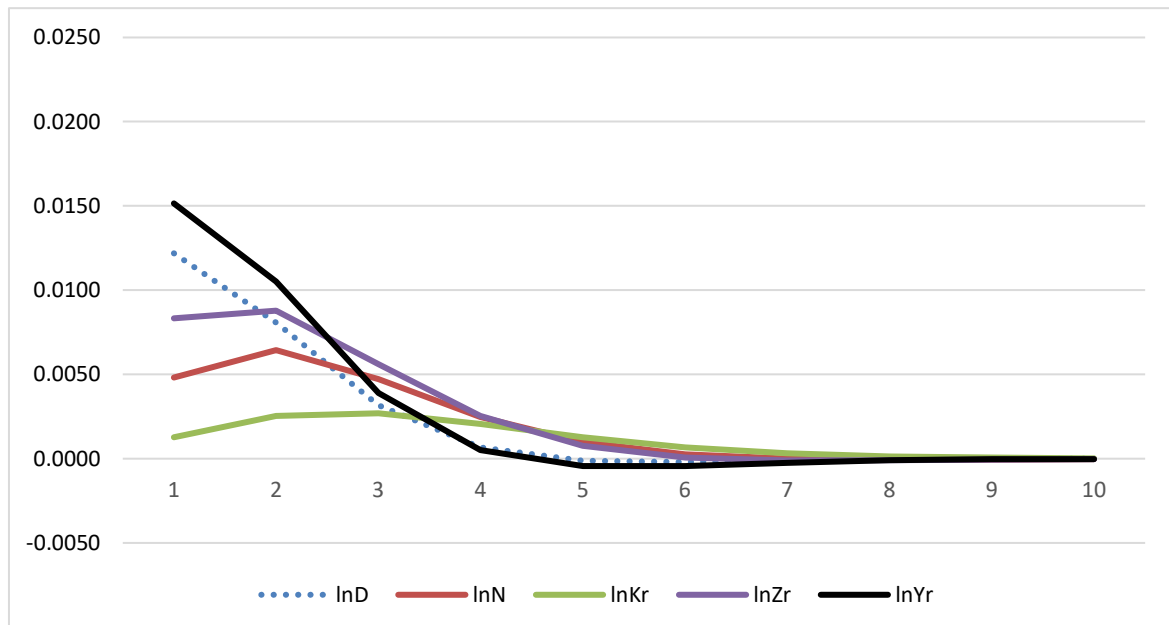
Figure A2. The relative price of investment goods



Note: I/C is the ratio of investment to consumption (volume indexes), and PI/PC is the ratio of the corresponding deflators. Source: Aggregate national account data from Statistics Sweden.

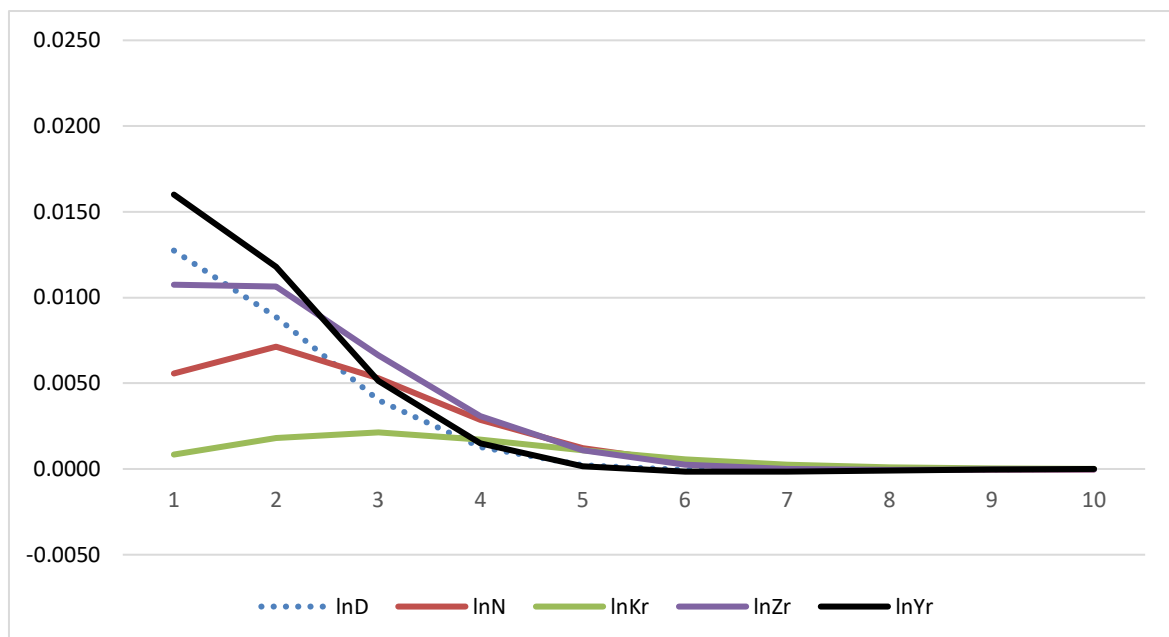
Figure A3. Robustness of impulse-response functions

a) Baseline estimation: balanced panel



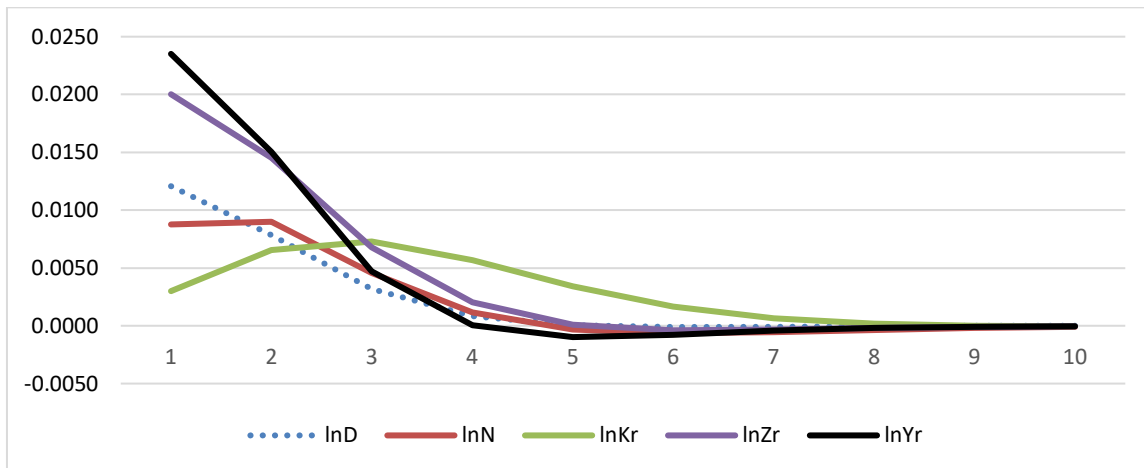
Note: Impulse-response functions from reduced-form model estimated on a balanced panel consisting of private manufacturing firms with at least ten employees, no extreme observations, and no missing values 1997-2016 (1150 firms). Extreme observations are defined as having production/employee, inventory/output or capital/output in top or bottom 1 percentile. Firm and time fixed effects and firm-specific trends are included. The variables are in logs and the time units on the horizontal axis are years. With two lags, the number of observations included in the estimations is 20700.

b) Unbalanced panel



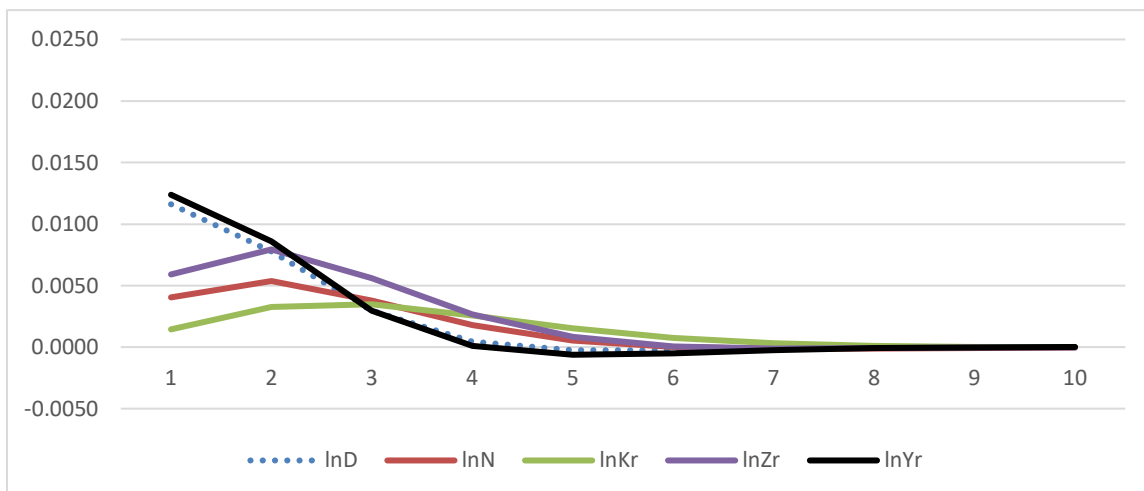
Note: Same as baseline but unbalanced panel. 42831 observations.

c) Excluding firms with top/bottom 10 percent “extreme” observations



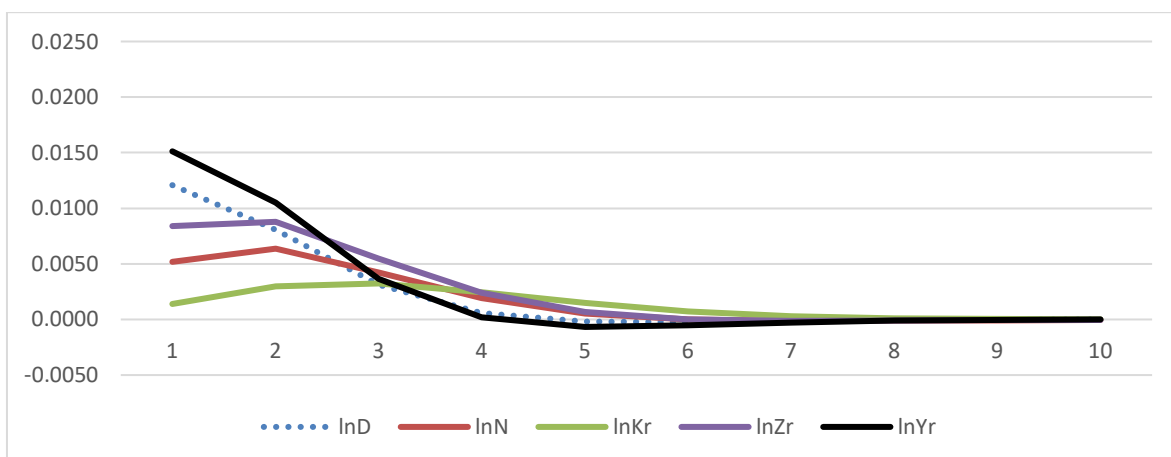
Note: Same as baseline but “extreme” observations defined as being in top or bottom 10 percent. 5148 observations.

d) One-plant firms only



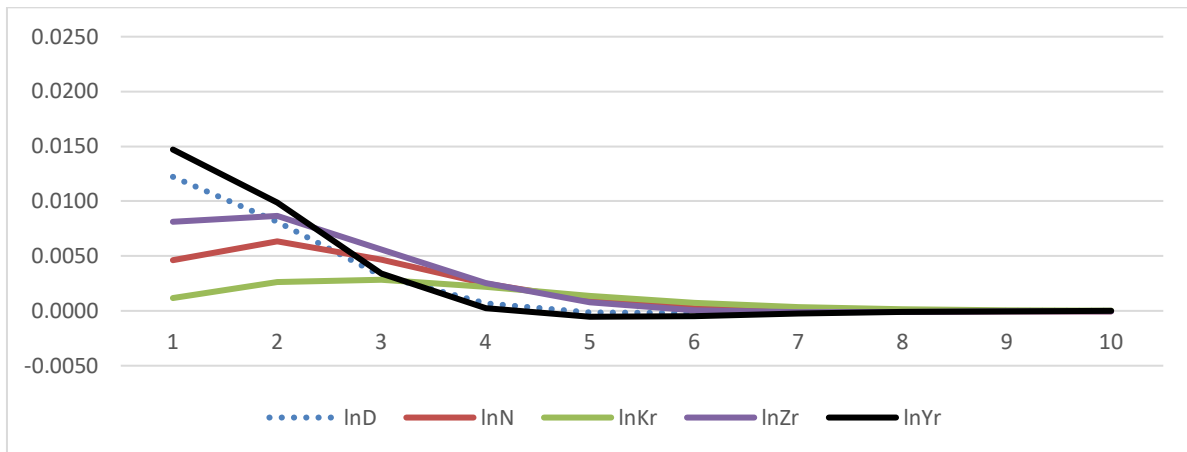
Note: Same as baseline but only firms with one plant. 13986 observations.

e) Using organizational number as identifier instead of FAD-code



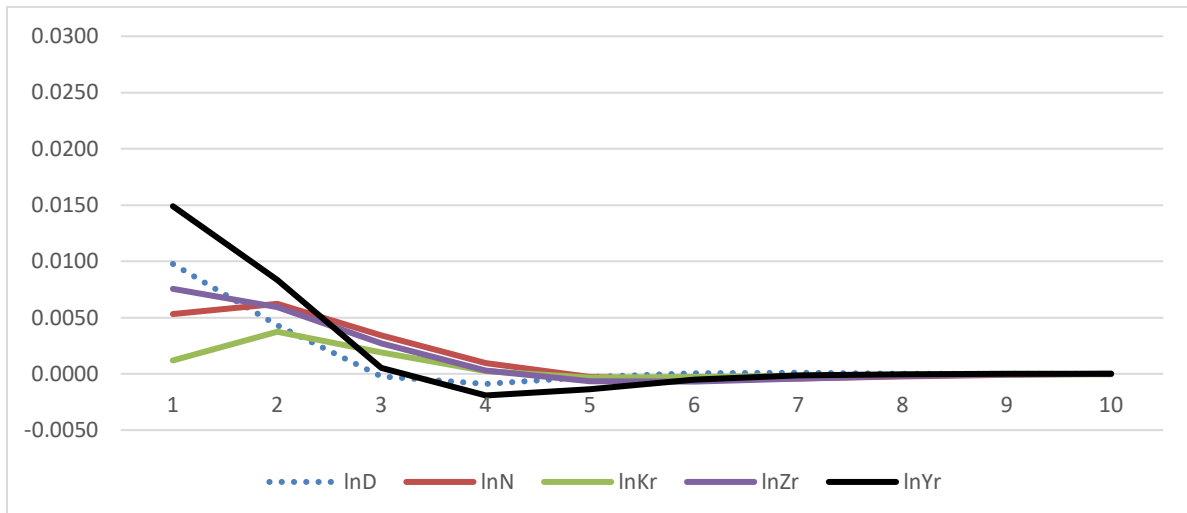
Note: Same as baseline using organizational number instead of FAD-code as identifier. 21586 observations.

f) Using initial export share of the firm to construct demand variable



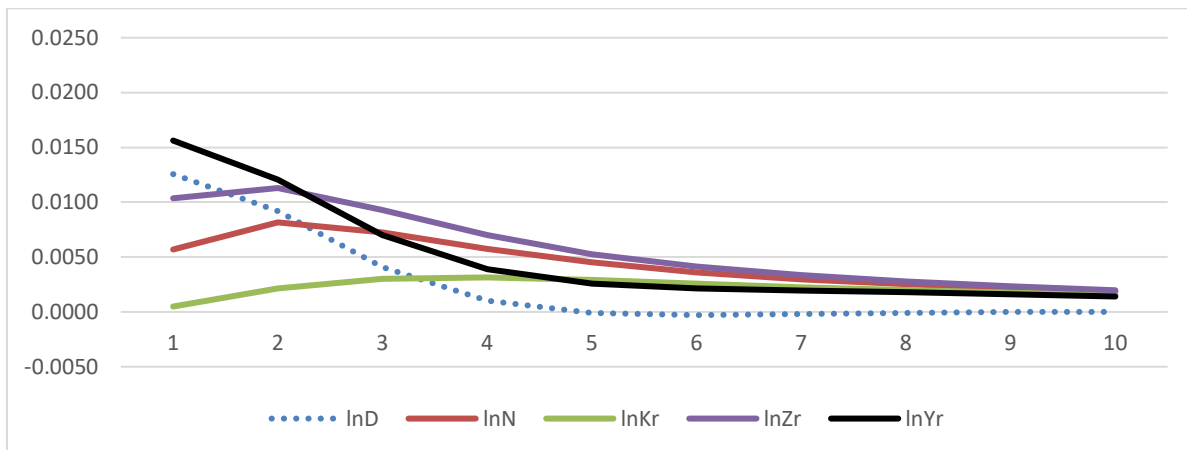
Note: Same as baseline but using initial export share to measure exposure to export market. 20700 observations.

g) Including three lags of the endogenous variables



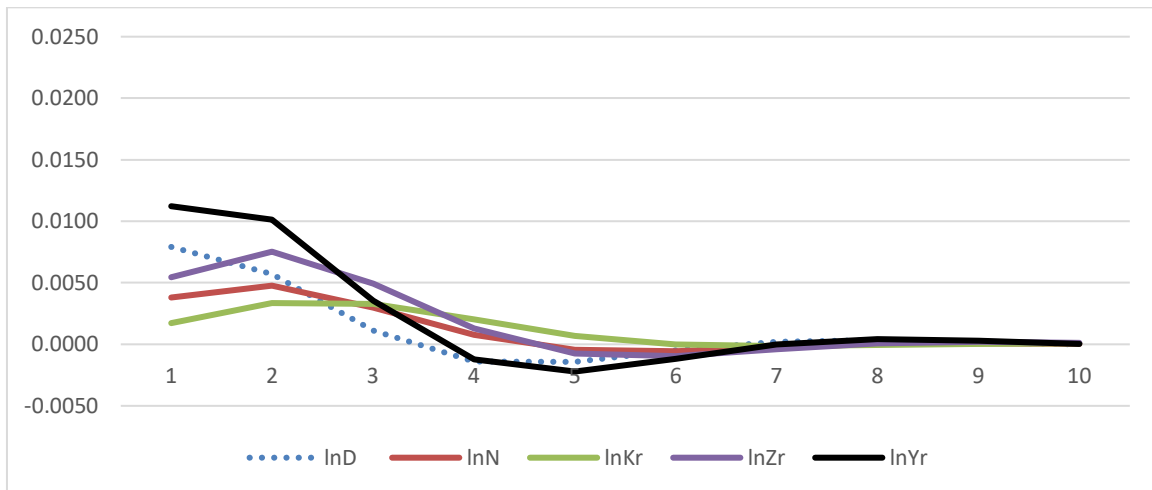
Note: Same as baseline but including three lags of endogenous variables and AR(3) process for demand variable. 19550 observations.

h) Including linear industry trends instead of firm-specific trends



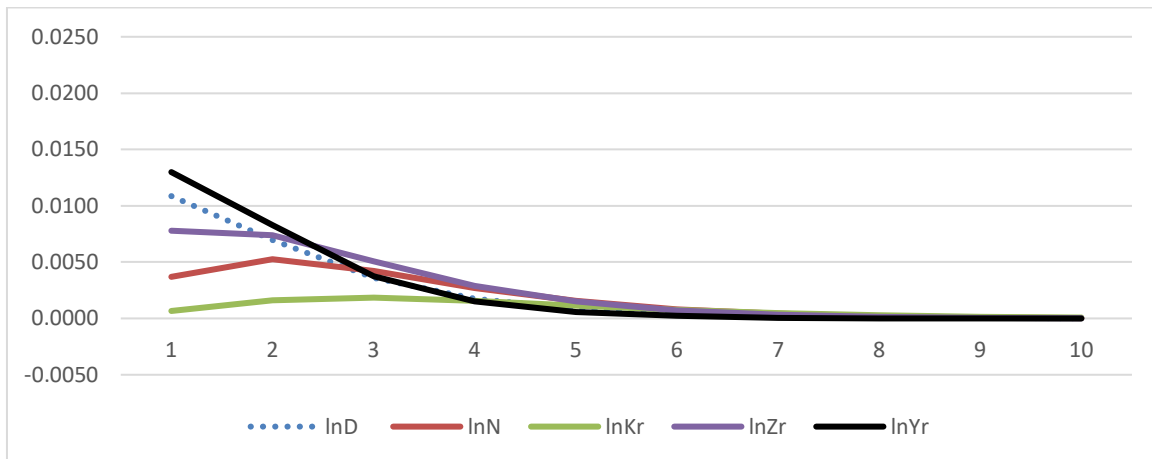
Note: Same as baseline but with industry-specific trends instead of firm-specific trends. 20700 observations.

i) Using only the domestic part of the demand variable



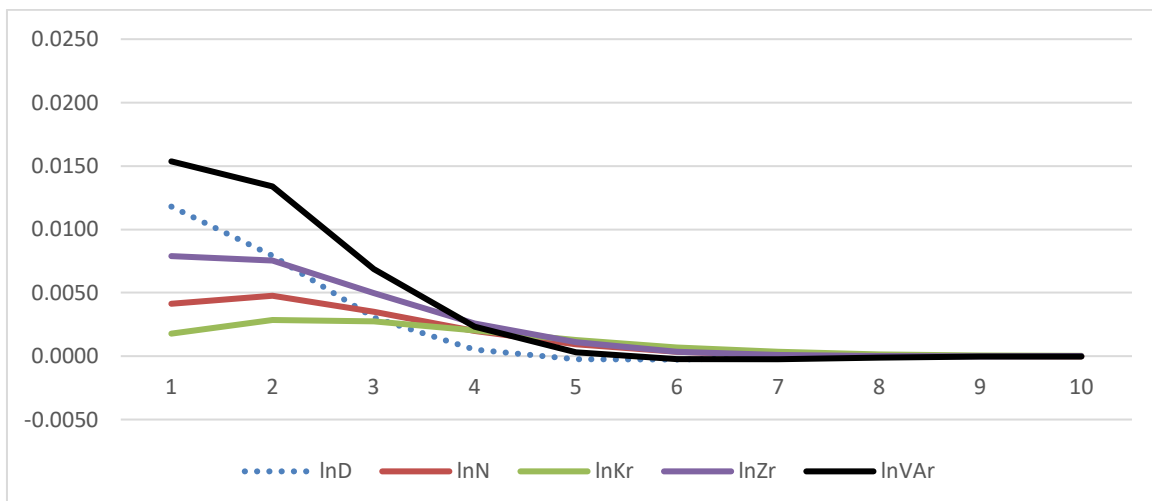
Note: Same as baseline but using only the domestic demand variable. 20700 observations.

j) Using only the foreign part of the demand variable



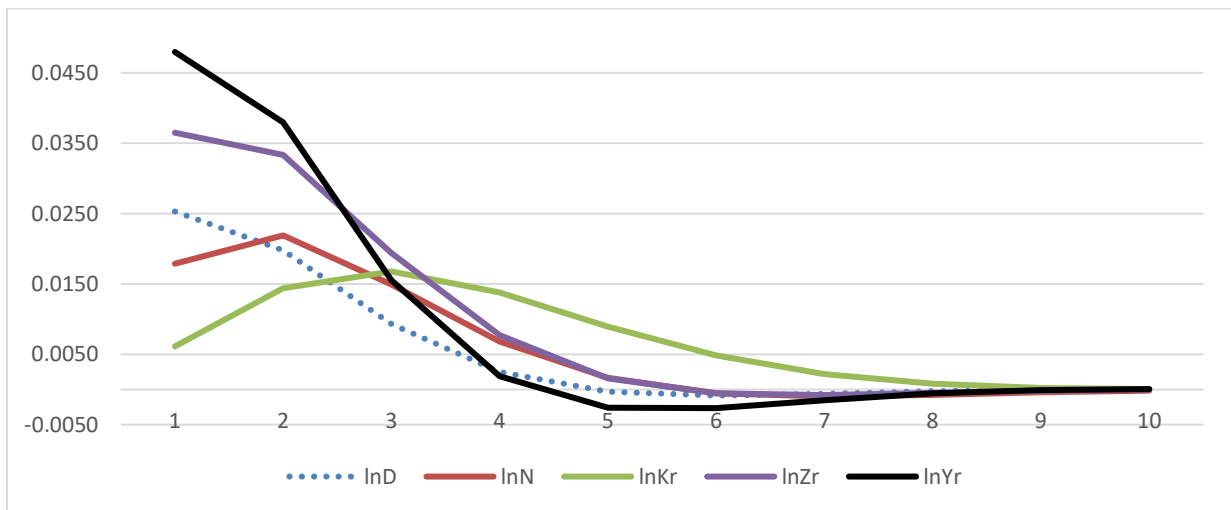
Note: Same as baseline but using only the foreign part of the demand variable. 20700 observations.

k) Using value added as the measure of production



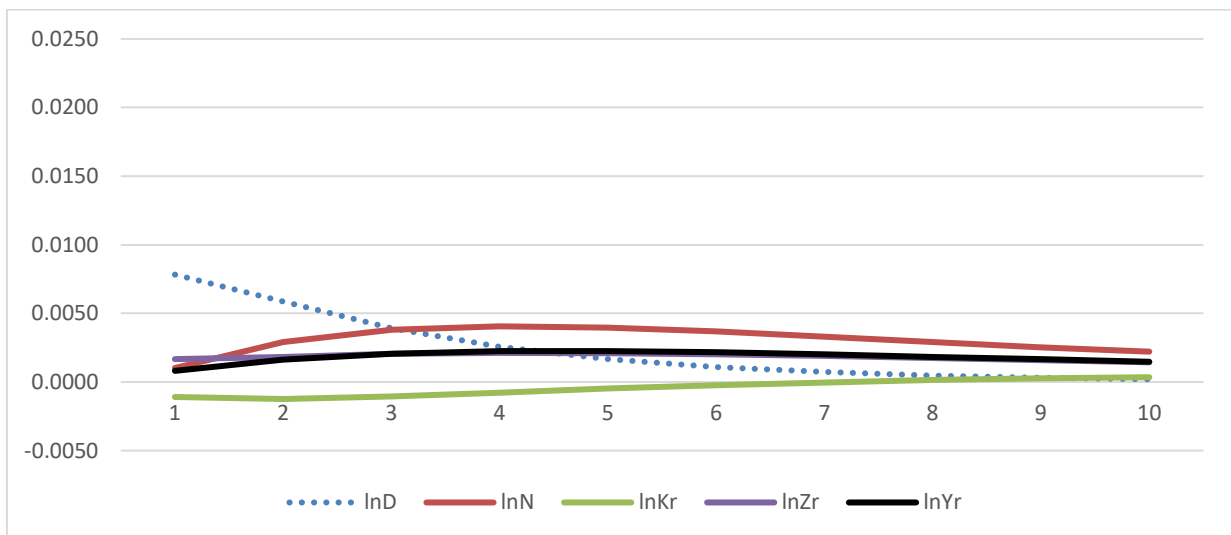
Note: Same as baseline but using value added to measure production. 19476 observations.

l) Excluding time dummies



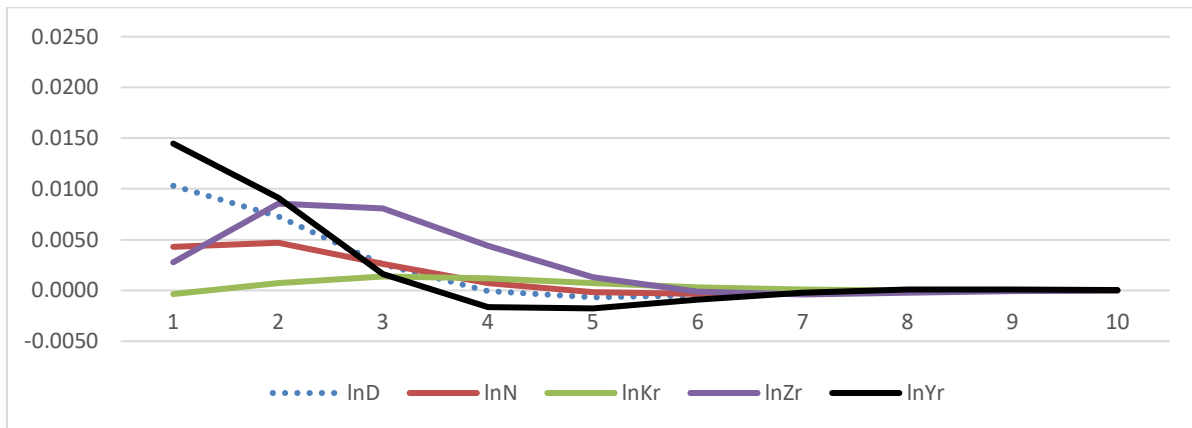
Note: Same as baseline but excluding time dummies. 20700 observations.

m) Including industry dummies interacted with time dummies



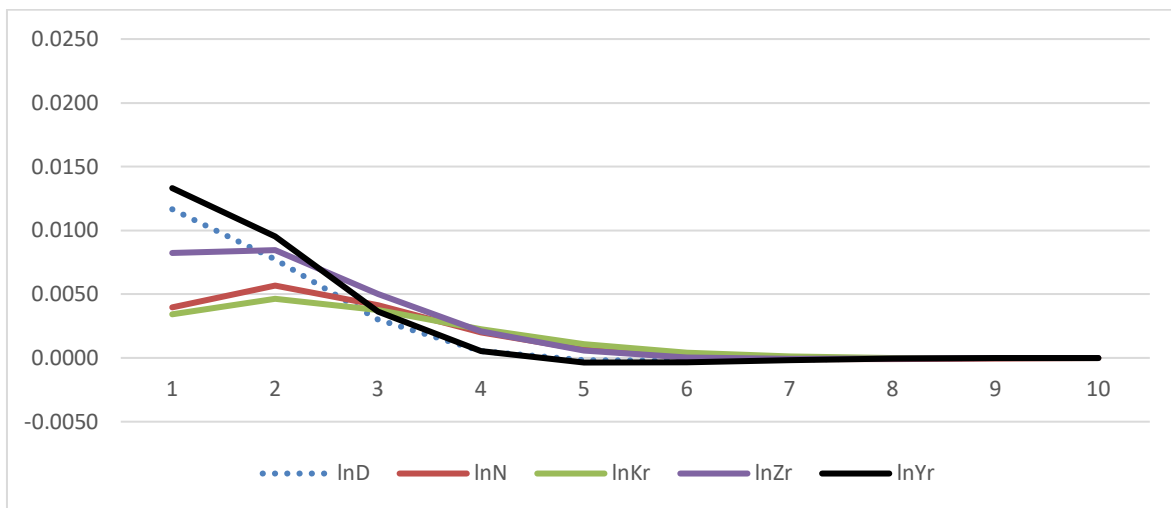
Note: Same as baseline but excluding firm-specific trends and including industry dummies interacted with time dummies. 20700 observations.

n) Small size firms



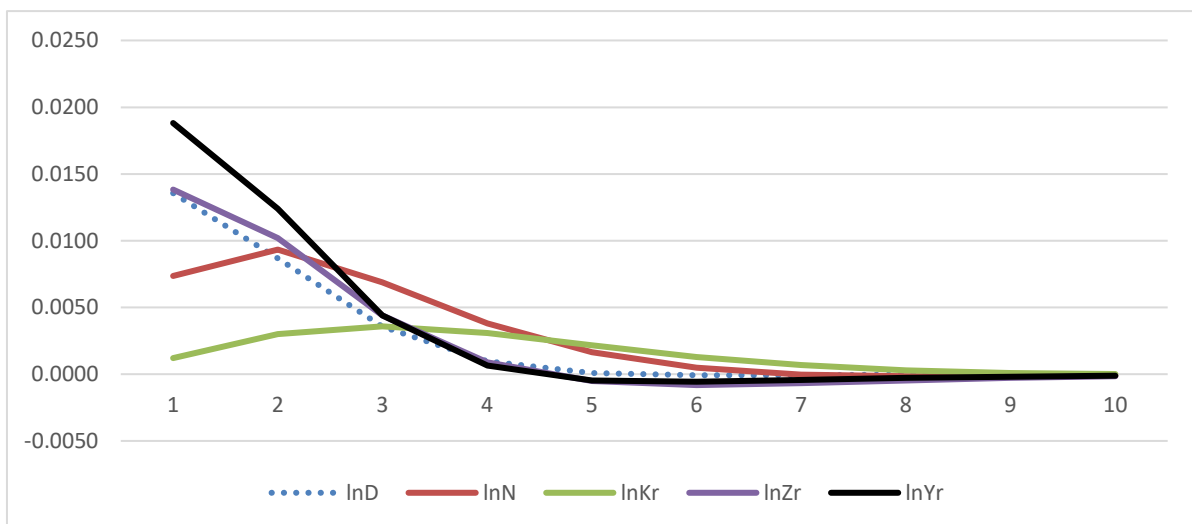
Note: Same as baseline but firms with a mean of 11 to 28 employees. 6894 observations.

o) Medium size firms



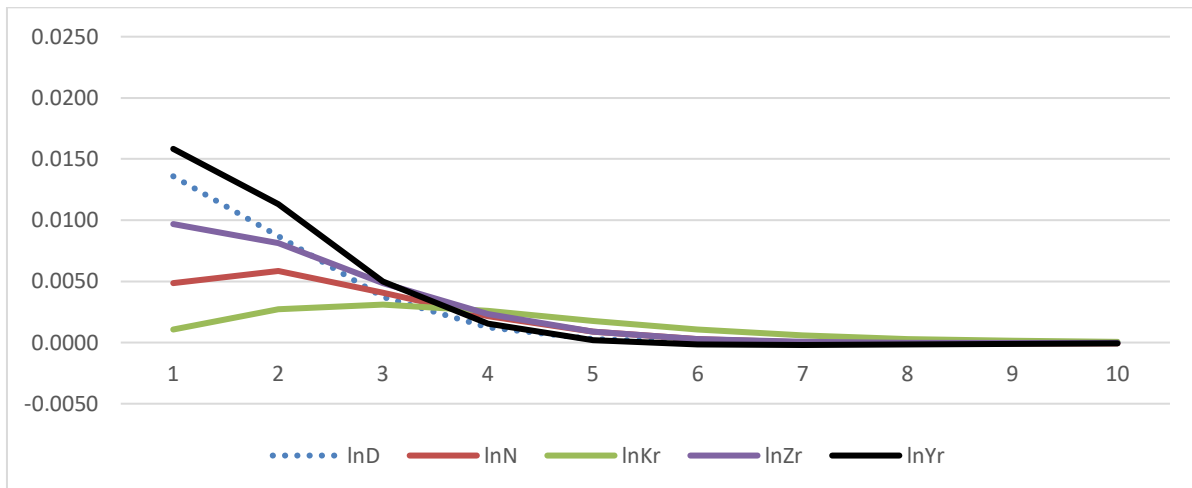
Note: Same as baseline but firms with a mean of 28 to 73 employees. 6912 observations.

p) Large size firms



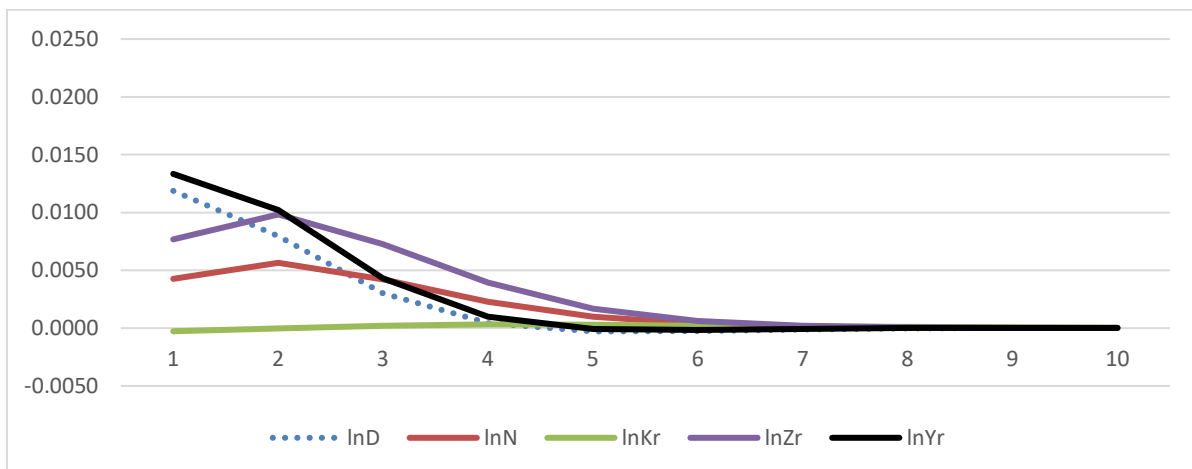
Note: Same as baseline but firms with a mean of 73 employees or more. 6894 observations.

q) Firms with low capital/output



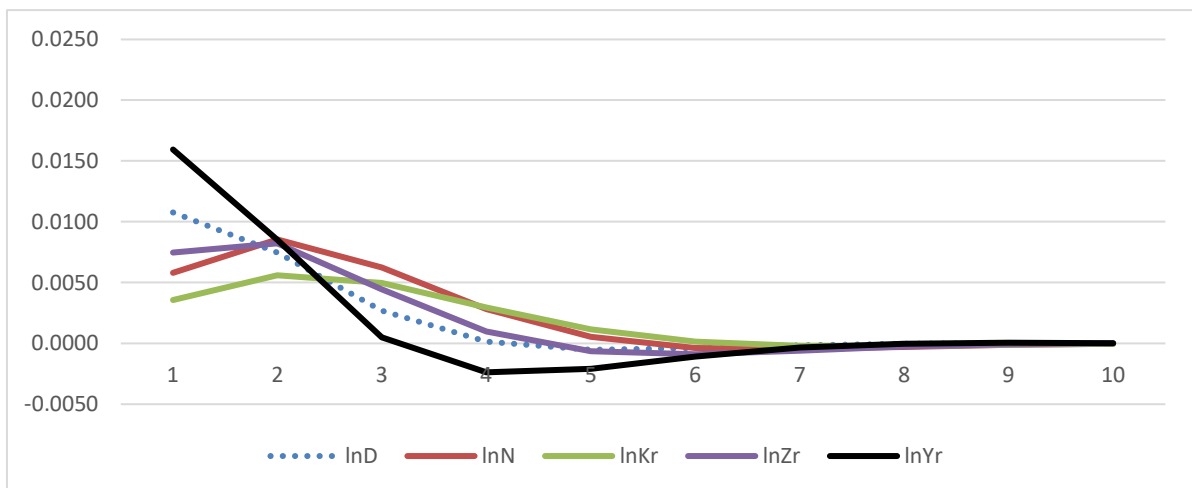
Note: Same as baseline but firms with low levels of capital/ output. 6894 observations.

r) Firms with medium capital/output



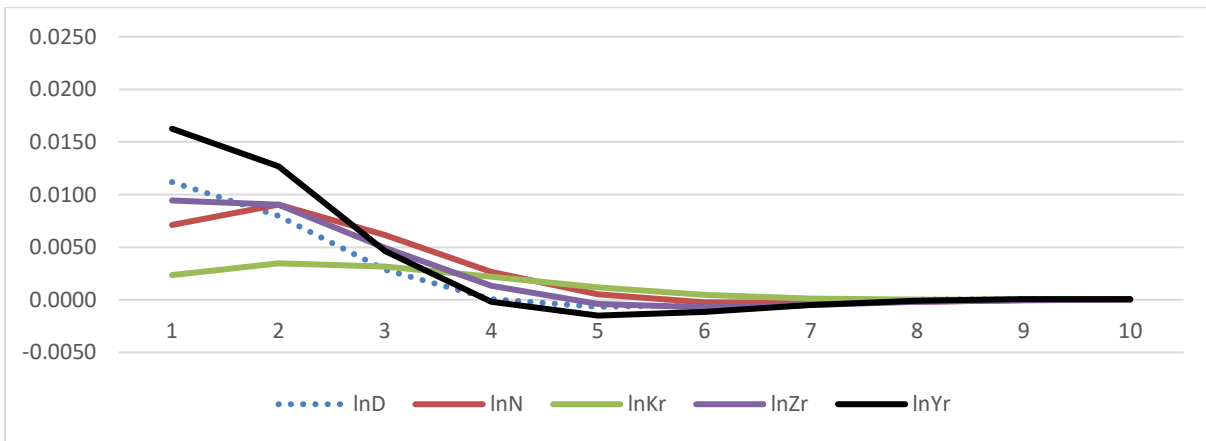
Note: Same as baseline but firms with medium levels of capital/ output. 6912 observations.

s) Firms with high capital/output



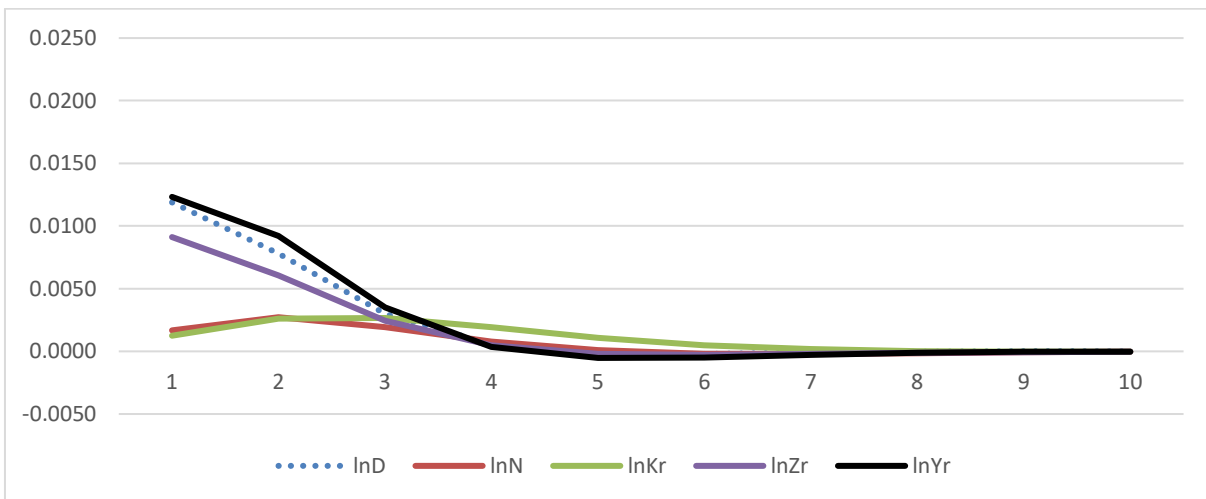
Note: Same as baseline but firms with high levels of capital/ output. 6894 observations.

t) Firms with low inventory/output



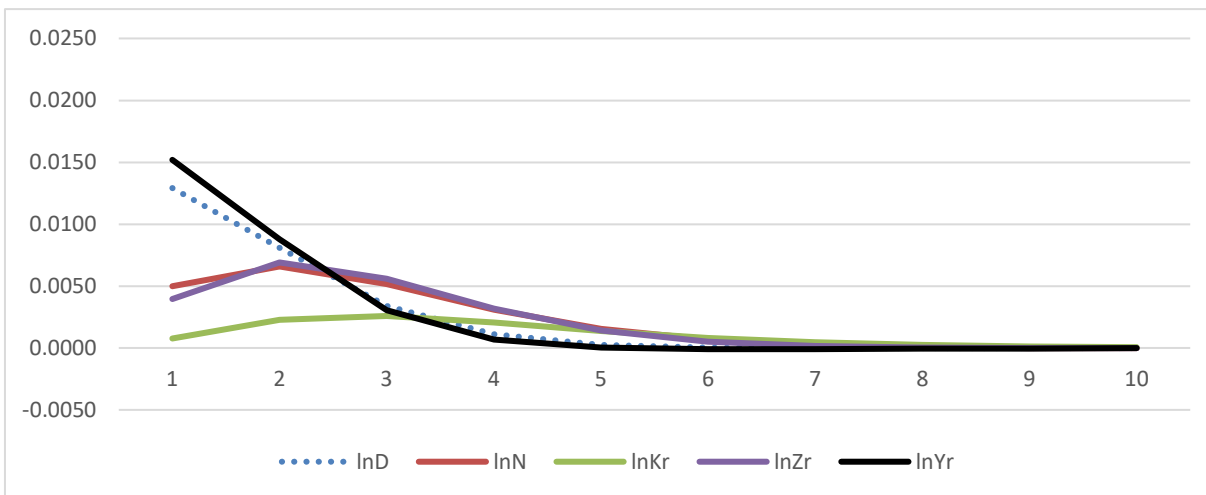
Note: Same as baseline but firms with low levels of inventory/ output. 6894 observations.

u) Firms with medium inventory/output



Note: Same as baseline but firms with medium levels of inventory/output. 6912 observations.

v) Firms with high inventory/output



Note: Same as baseline but firms with high levels of inventory/output. 6894 observations.