Firm-Level Shocks and Labor Flows*

Mikael Carlsson[†], Julián Messina[‡] and Oskar Nordström Skans[§]

November 22, 2017

Abstract

We analyze how labor flows respond to idiosyncratic shifts in firm-level production functions and demand curves using very detailed Swedish micro data. Shocks to firms' physical productivity have only modest effects on firm-level employment decisions. In contrast, we document rapid and substantial employment adjustments through both hires and separations in response to firm-level demand shocks. The choice of adjustment margin depends on the sign of the shock: Firms adjust through increased hires if these shocks are positive and through increased separations if the shocks are negative.
Keywords: Technology, Demand, Job Flows, Worker Flows
JEL classifications: D22, J63, J23

*We would like to thank Marcela Eslava, Richard Freeman, Nils Gottfries, Erica Groshen, John Haltiwanger, Francis Kramarz, Per Krusell, Frédèric Warzynski and Nikolaj Angelov for helpful comments. In addition, we thank seminar participants at ELE, 2012; IAE-MOVE, 2012; Greater Stockholm Macro Group, 2013; CAP, 2013; EALE 2014; SOLE 2014; U. of Alicante; U. Carlos III; Bank of Spain; U. of Warwick; World Bank; IFAU; UCLS; Sveriges Riksbank; Linnaeus U. and Gothenburg U. We are also grateful to Karl Harmenberg, Mats Levander, Julia Tanndal, Matias Quiroz, and Juan Pablo Uribe for excellent research assistance. The data are confidential, but the authors' access is not exclusive. Carlsson and Skans are grateful for financial support from the Ragnar Söderberg Foundation. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank or the Inter-American Developement Bank.
[†]Uppsala University, UCLS and Sveriges Riksbank. e-mail: mikael.carlsson@nek.uu.se
[‡]Inter-American Development Bank and IZA. e-mail: julianm@iadb.org
[§]Uppsala University, IFAU, UCLS and IZA. e-mail: oskar.nordstrom_skans@nek.uu.se

1 Introduction

About one in every five jobs are either created or destroyed every year (Davis, Faberman and Haltiwanger, 2006). The bulk of this firm-level labor adjustment is truly idiosyncratic as firms operating in the same sector and area shrink and grow side-by-side. Hence, jobs are rapidly created and destroyed, even in sectors with stable net employment. Following the seminal work of Davis, Haltiwanger and Schuh (1996), the importance and magnitude of these labor flows has been documented for a large number of countries.[1] However, while the empirical regularities of job and worker flows have been abundantly documented, little is known about how job and worker flows respond to structural firm-level shocks.[2]

This paper presents novel evidence on how employment adjust through hires and separations when firms are hit by shocks that alter their positions in the performance distribution. We identify the two types of fundamental shocks highlighted by Foster, Haltiwanger, and Syverson (2008); *Technology shocks* shifting the firm-level physical production function (i.e., the ability to produce at a given level of inputs) and *Demand shocks* shifting the firm-level demand curve (i.e., the ability to sell at a given price).[3] We focus on the effects of idiosyncratic shocks, and thus concentrate on employment adjustments in a stable market environment, effectively abstracting from feedback effects through changes in market wages or aggregate unemployment. Moreover, we focus on permanent shocks, and show evidence consistent the notion that permanent shocks are the key drivers of employment adjustments as suggested by e.g. Franco and Philippon (2007) and Roys (2016). Our empirical analysis relies on a unique data base that links measures of firm-level input, output, and prices to individual worker-flow data for Swedish manufacturing firms.

The analysis adds to a vibrant empirical literature, surveyed in Syverson (2011), that documents the distinct impacts of firm-level technology and demand shocks on productivity and other firm-level outcomes. Most notably, Foster et al. (2008) shows

[1] See Davis, Faberman, and Haltiwanger (2012) for an overview. For evidence from Sweden, which is the empirical subject of this paper, see Andersson (2003).

[2] A small macro-oriented literature aims to identify the employment responses to technology-driven changes in firm-level productivity, see e.g., Carlsson and Smedsaas (2007) and Marchetti and Nucci (2005). The macro literature also contains a number of related studies, e.g., Galí (1999) and Michelacci and Lopez-Salido (2007), the latter of which distinguished between neutral technology shocks and investment-specific technology shocks and derived the consequences for job reallocation.

[3] Importantly, these shocks are defined according to their effects on firm-level optimization, not according to their origins.

that firm closures are driven primarily by changes in idiosyncratic demand and only to a lesser extent by changes in idiosyncratic physical productivity. Recent evidence in Foster, Haltiwanger, and Syverson (2016) suggests that the growth of young firms in the US is due to a shrinking product-demand gap relative to incumbents. Pozzi and Schivardi (2016), who uses Italian data to analyze how technology and demand affect firm output, show that firm-level technology shocks have a surprisingly low impact on firm growth and that demand shocks are at least as important. In addition, Carlsson, Messina, and Nordström-Skans (2016) shows that firm-level technology shocks affect workers' wages, using Swedish data.

This paper is, however, the first to focus on how firm-level technology and demand shocks affect firms' labor adjustments through hires and separations in response to shocks of different nature, signs and magnitudes, a question that speaks to a huge body of theoretical research regarding the relationship between firm-level revenue productivity and labor adjustments (Bentolila and Bertola, 1990; Davis and Haltiwanger, 1992; Hopenhayn and Rogerson, 1993; Mortensen and Pissarides, 1994; and more recently Cahuc, Postel-Vinay, and Robin, 2006; and Lise, Meghir, and Robin, 2016). The focus of this paper is to disentangle the separate roles in labor adjustments of two fundamental drivers of firm-level revenue productivity fluctuations: demand and technology shocks.

Our analysis departs from a stylized model (of monopolistically competitive firms) that motivates a set of restrictions on the long-run relationship between firm-level fundamentals and shocks. In the spirit of Franco and Philippon (2007), we then impose these long-run restrictions in a structural vector autoregression (SVAR) setting to filter out our empirical measures of permanent idiosyncratic demand and technology shocks.[4] This allows us to derive the shocks without imposing any restrictions on the firms' short-run behavior.

The most important restriction we rely on is instead the notion that the physical gross Solow residual is independent of all shocks except the technology shock *in the long run.* To take the analysis to the data, we benefit from a firm-specific price index. Using a strategy similar to Eslava et al. (2004), Carlsson, Messina, and Nordström-

[4]Demand shocks appear to have a non-trivial transitory component which we remove in the main analysis and then study separately. In contrast, the bulk of the (physical) technology shocks are persistent enough to emerge as permanent shocks from our SVAR filter. This is consistent with Carlsson, Messina, and Nordström-Skans (2016) who, when estimating an AR(1) process for the level of technology using Swedish data similar to ours, find a persistence estimate as high as 0.88. Eslava et al. (2004) find an even higher persistence of 0.92 for Colombia.

Skans (2016) and Smeets and Warzynski (2013), we deflate the (nominal) firm-level output series with firm-level price indices to derive measures of firm-level real output volumes. Importantly, the fact that we filter out the technology shocks using long-run restrictions implies that other shocks, or changes in factor utilization, or inventories, are allowed to have a transitory impact on the physical Solow residual without affecting the measured technology shocks. We further derive sufficient restrictions to identify permanent demand shocks without imposing any restrictions on the nature of short-run shocks or dynamics.[5]

Although our modelling assumptions are very similar to the single-equation approach taken by previous research, where the demand shocks are calculated as residuals from output-price relationships using an estimated demand elasticity, the SVAR has the advantage of not imposing any structure or restrictions on the firms' adjustments in the short run. It also provides us with results that are robust to transitory measurement errors and missmeasurement in the demand elasticity. A specific benefit we get from applying standard time-series econometrics tools is that they produce measures of shocks with known time-series properties in terms of persistence; properties which our results suggest are important when studying labor adjustments.

When implementing the SVAR, we deviate from standard time-series applications such as Blanchard and Quah (1989) and Franco and Philippon (2007). Because we have access to a wide panel of firms, we estimate the reduced-form equations using dynamic panel data methods building on Arellano and Bond (1991). This allows us to estimate both the parameters and the covariance matrix of the error terms with considerable precision, thereby avoiding standard macro-data concerns regarding the practical implementation of SVARs.

We start by showing that, despite being crucial for both firm-level prices and output, firm-level technology shocks have a relatively limited effect on labor inputs.[6] In contrast, we find that product demand is a key driving force behind firm-level labor adjustments. An idiosyncratic demand shock of 1 standard deviation increases em-

[5] Since we use data from a small, open economy, our system also explicitly allows for shocks to factor prices. In addition, the system allows for a (transitory) residual shock component to soak up any remaining short-run dynamics, including mean-reverting shocks to purely idiosyncratic factor prices.

[6] The employment elasticity to technology shocks is 0.05 (or 0.14 if assuming decreasing returns to scale) which is very close to what Pozzi and Schivardi (2016) found for Italy (0.08) using a static single-equation approach. The demand shock responses are, however, much larger in our case (0.39 on elasticity form).

ployment by 6 percent, whereas the corresponding number for technology shocks is 0.5 percent. These results which are robust to a wide range of variations in measures and specifications appear to ask for a model where price responses to technology shocks are muted by non-constant demand elasticities, which we can allow for without violating the identifying assumptions.[7] The employment adjustments induced by permanent demand shocks are both rapid and symmetric across hires and separations. Most of the adjustment takes place within a year. On average, firms adjust employment almost as much through changes in the separation rate as through changes in the hiring rate. One third the initial separation response is due to additional separations of short-tenured workers. However, separating from recent hires as well as adjustments in the use of marginal workers appear to primarily be important as short-run adjustment margins.

We further analyze non-linearities in the responses and show the choice of adjustment margin crucially depends on the sign of the shock. Negative shocks mostly induce additional separations, positive shocks mostly induce additional hires. Finally, we use the product demand shocks as instruments to uncover the causal component of the relationship between job flows and worker flows. This analysis builds on the seminal work of Abowd, Corbel, and Kramarz (1999) and Davis, Faberman, and Haltiwanger (2012) who provide descriptive studies of job flows and worker flows, decomposed into positive and negative changes using data from France and the United States, respectively. In contrast to these decomposition exercises, we analyze hires and separations in response to net employment changes driven by a well-identified and empirically relevant shock: permanent shifts in firms' product demand schedules. This allows us, for the first time, to provide an analysis that removes the potential impact of exogenous variations in worker flows (e.g. any type of worker-induced separation) which can create endogenous responses in the number of jobs, particularly in small firms. Our findings show that, although the average firm reduces hires in response to negative shocks somewhat, it also continues to recruit workers even when forced to reduce employment substantially. Thus, the firms are far from exploiting the full potential of downsizing through reduced hires.[8] The results thus imply that

[7] We further show that we can match estimated responses of output, prices and employment much better if we allow for endogenous demand elasticities.

[8] These results thus concur with the descriptive picture provided by Davis, Faberman, and Haltiwanger (2012) for the US, but differ from Abowd, Corbel, and Kramarz (1999) who documents that employment reductions in French firms primarily are associated with reduced hiring rates.

both hirings and separations should be treated as endogenous when modelling labor adjustments at the firm level.

Overall, the speed of adjustment, the symmetry between hires and separations as adjustment margins, the rapid separation of long-tenured workers, and the continued recruitment of workers in the face of negative shocks jointly suggest that firms facing permanent idiosyncratic demand shocks do adjust their labor input flexibly. In contrast, we show evidence suggesting that the firms' responses to transitory demand shocks may be much more muted, a which may be welfare-enhancing in the presence of uninsurable labor market risk (Bertola, 2004).[9]

The paper is organized as follows: Section 2 outlines a simple model that motivates the long-run restrictions needed to extract our permanent demand and technology shocks. Section 3 introduces the main characteristics of the firm-level data used in the analysis and discusses the empirical implementation of the SVAR and the validation of the shocks. Section 4 reports the main results, distinguishing employment, hiring, and separation margins in response to technology and demand shocks. Section 5 presents extensions of the basic analysis, focusing on non-linearities in the responses. Finally, section 6 concludes.

2 Model and Empirical Strategy

2.1 Shocks, Idiosyncratic Production, and Demand

In this section, we derive a stylized model of monopolistically competitive firms. The model focuses on two key exogenous idiosyncratic driving forces to the firm's relative performance: *technology shocks* affecting the firm's physical productivity and *demand shocks* affecting the firm's ability to sell to its products at a given price. The purpose of the paper is to analyze how these two disturbances affect firms' hiring and separation policies. The model, functional form assumptions, and statistical properties of the shocks are deliberately stylized, imposing the minimum amount of structure needed to identify these two structural shocks using VAR techniques.[10]

[9] See also Guiso, Schivardi, and Pistaferri (2005), who show that firms insure workers' wages relative to transitory shocks to value added, but not to permanent shocks.

[10] As we show, different representations of the demand structure give different mappings between empirical results and key structural parameters, while delivering unchanged long-run restrictions and thus unchanged findings in terms of labor-adjustment responses to the shocks.

The key distinction between our definitions of technology shocks and demand shocks lies in how the shock affects the producing firm, not in the origin of the shock. This approach, which is consistent with the existing (micro) literature (such as Foster et al., 2008, and Syverson, 2011), implies that we do not distinguish between shifts in the firm-specific demand curve that arise from changing preferences among final consumers, those that arise from increased demand among downstream firms, and those that arise from quality changes that increase product demand at a given price.[11]

To identify firm-level structural shocks, we need to make assumptions about the technology and market conditions faced by the firm. Our setup follows Eslava et al. (2004) and Foster et al. (2008, 2016) closely, by using a first-order approximation of both production technologies and product market demand and by modeling the key technology and demand shocks as neutral shifters of the production function and the demand curve, respectively. Following these papers, the firm-level production function is approximated by:

$$Y_{jt} = A_{jt} N_{jt}^{\alpha} K_{jt}^{\beta} M_{jt}^{1-\alpha-\beta} \text{ and } \alpha, \beta \in (0,1), \tag{1}$$

where physical gross output Y_{jt} in firm j at time t is produced using technology indexed by A_{jt} and combining labor input N_{jt}, capital input K_{jt}, and intermediate production factors (including energy) M_{jt}. Importantly, our data allow us to account for idiosyncratic price differences across firms, so that our measure of technology (the Solow residual, A_{jt}) refers to *physical* total factor productivity (TFPQ), rather than to *revenue* total factor productivity (TFPR) in the terminology of Foster et al. (2008). Equation (1) presupposes a constant-returns technology, which is an assumption we maintain in our main specification, but we also present robustness exercises where we relax this assumption. Note, however, that only the long-run returns are relevant for our empirical implementation, making decreasing returns to scale a less likely scenario.

We (for now) represent the firm-level demand curve by a constant-elastic function

[11] Franco and Philippon (2007) label these shocks as shocks to market shares, and model them formally as preference shocks. Note that the firm-level price index we use is based on unit prices for very detailed product codes (8/9-digit Harmonized System/Combined Nomenclature codes), which limits the scope for quality changes to be the key component in our demand shock. However, it is straightforward to show that if we added a quality shock to the system developed below (through a wedge between the measured firm-level price, based on unit values, and the quality-adjusted price), it would enter the system symmetrically to the demand shock.

according to

$$Y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\sigma} Y_t \Omega_{jt} \text{ and } \sigma > 1, \tag{2}$$

where P_{jt}/P_t is the firm's relative price, Y_t denotes aggregate market demand, and Ω_{jt} is a firm-specific demand shifter. The parameter σ denotes the elasticity of substitution between different competing goods and hence captures the demand elasticity for each firm in the economy. Here, we let σ represent a constant demand elasticity, but below we show that our identification remains consistent if we treat σ as a function of the shocks, i.e. $\sigma = \sigma(A_{jt}, \Omega_{jt})$, allowing for Kimball (1995)-style strategic complementarity in price setting.

Following Guiso et al. (2008) and Franco and Philippon (2007) we model the key shocks as permanent shifters. Formally:

$$A_{jt} = A_{jt-1} e^{\mu_j^a + \Phi^a(L)\eta_{jt}^a}, \tag{3}$$

$$\Omega_{jt} = \Omega_{jt-1} e^{\mu_j^\omega + \Phi^\omega(L)\eta_{jt}^\omega}, \tag{4}$$

where μ_j^a and μ_j^ω are constant drifts, and $\Phi^a(L)$ and $\Phi^\omega(L)$ are polynomials in the lag operator, L. The white-noise idiosyncratic technology and demand shocks are denoted by η_{jt}^a and η_{jt}^ω. The assumed functional form implies that the shocks' lag polynomials are linearly related to the log differences of A_{jt} and Ω_{jt}, respectively.[12] As is evident from the formulation, our focus is on permanent shocks, but in a variation of the model we also explicitly analyze the role of transitory disturbances (see section 4.3).

Our model also allows for sectoral shocks to factor prices other than labor. This is potentially important in the Swedish setting of a small open economy where factor prices are likely to vary across sectors and time (due to exchange-rate volatility, for example).[13] To simplify the notation, we next define a price index (consistent with cost minimization) for input factors other than labor, $P_{jt}^F = \left(P_{jt}^K/\beta\right)^\beta \left(P_{jt}^M/(1-\alpha-\beta)\right)^{1-\alpha-\beta}$, where P_{jt}^K is the capital price and P_{jt}^M is the price of intermediate materials at time t. Similarly to technology and demand, P_{jt}^F evolves according to $P_{jt}^F = P_{jt-1}^F e^{\mu_j^f + \Phi^f(L)\eta_{jt}^f}$, where μ_j^f is a firm-specific drift, $\Phi^f(L)$ is a polynomial in the lag operator, L, and η_{jt}^f is a white-noise factor-price shock.

[12] This, in turn, provides a convenient moving average (MA) representation of the VAR specified below (see Appendix C for details).

[13] Allowing for a factor price shock and, as discussed below, including a residual variable to soak up remaining transitory variation helps our VAR to pass standard diagnostic tests.

The specified shocks (together with aggregate conditions) are taken as state variables during firm-level wage determination. In addition, we assume cost minimization and that the firms have the right-to-manage so that factor choices are made taking wages as given.

2.2 Identifying Long-Run Restrictions

We rely on the stylized model presented above to derive a set of long-run restrictions that allow us to filter out the structural shocks of interest (η^a and η^ω). Table 1 summarizes the set of equations that motivate our restrictions, and Appendix A presents details and derivations. The second column of the table denotes variables that can all be constructed from our firm-level data as long as we have an estimate of the demand elasticity σ (as detailed in the next section).

The third column summarizes the three key predictions that we rely on for identification:

1. The measured physical Solow residual ($TFPQ$ in the terminology of Foster et al. 2008) is equal to A and hence independent of both demand (Ω) and factor prices (P^F).

2. The "wage-neutral" unit labor costs ($WNULC$), as defined in the second row, is a function of both A and P^F.

3. The "wage-neutral" demand (WND), as defined in the third row, is a function of A, Ω, and P^F.

We use the modifier "wage-neutral" to highlight that the measures are defined to neutralize the impact of potential wage shocks. The variables are constructed in order to deliver a set of recursive long-run restrictions that we can use for identification.

The recursive sequence of restrictions are highlighted in the fourth column: The Solow residual is independent of the innovations η^ω and η^f, and $WNULC$ is independent of η^ω. If invoked in the long run, these restrictions are sufficient to identify a VAR model in these variables using standard structural VAR (SVAR) techniques. In practice, we will also include a fourth residual variable in the system to soak up all remaining transitory dynamics in the system. We impose that this fourth shock has no long-run impact on the three variables within our core system. We return to this issue below.

Table 1: The Core Structural VAR Equations

Column	(1)		(2)	(3)
Variable:	Measured in data as:		Model expression:	Long-run restrictions:
$Solow$	$Y_{jt}\left(N_{jt}^{\alpha}K_{jt}^{\beta}M_{jt}^{1-\alpha-\beta}\right)^{-1}$	$=$	A_{jt}	Independent of η^{ω} and η^{f}
$WNULC$	$\left(W_{jt}N_{jt}/Y_{jt}\right)W_{jt}^{-\alpha}$	$=$	$\alpha^{1-\alpha}A_{jt}^{-1}P_{jt}^{F},$	Independent of η^{ω}
WND	$Y_{jt}W_{jt}^{\sigma\alpha}$	$=$	$\psi Y_{t}P_{t}^{\sigma}A_{jt}^{\sigma}\left(P_{jt}^{F}\right)^{-\sigma}\Omega_{jt}$	–

Note: $Solow$ is the physical Solow residual (TFPQ), $WNULC$ is wage-neutral unit labor cost and WND is wage-neutral demand. ψ is a constant such that $\psi\equiv\left(\frac{1}{\alpha}\right)^{-\sigma\alpha}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}$.

2.3 Benefits of the Empirical Approach

First, we only need to impose the zero-impact restrictions of the last column in the *long run*. Hence, we do not make any assumptions regarding the short-run dynamics or about transitory measurement errors. Notably, our identification of the technology shocks (η^{a}) are therefore consistent with changes in inventories, factor utilization, markups, or idiosyncratic input prices altering the Solow residual as long as these changes are mean reverting, i.e., as long as they do not affect the Solow residual in the long run.

Second, we do not require that all aspects of the motivating model are true, even in the long run. We only require that the impact of the shocks on the three variables ($Solow$, $WNULC$, WND) measured in column (1) of Table 1 does not violate the restrictions listed in column (3) of the same table. These restrictions are in fact consistent with a wider class of models than the one we used to derive the restrictions.[14] A particular possible extension, that for reasons presented below will turn out to be useful, is inspired by the literature on strategic complementarity in price setting (Kimball, 1995). To allow for such effects, we can let the elasticity of demand (and thereby the markup) be affected by technology and demand shocks. This replaces

[14] The essential assumptions are non-restrictive relative to a broad class of possible models. The key assumptions are cost minimization, the relevance of the first order approximation of the production function, the assumption of monopolistic competition and an assumption that firms has the "right to manage", i.e. that firms make employment decisions taking wages as given (regardless of whether they are set in bargaining or not).

equation (2) by

$$Y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\sigma(A_{jt},\Omega_{jt})} Y_t\Omega_{jt},\ \ \sigma(A_{jt},\Omega_{jt}) > 1 \text{ and } \sigma(\bar{A}_{jt},\bar{\Omega}_{jt}) = \sigma, \tag{5}$$

where a bar denotes an average across firms. Importantly, the only change relative to the measurement equations outlined in Table 1 is that the measure of wage-neutral demand (WND) acknowledges that σ is no longer constant, i.e. $Y_{jt}W_{jt}^{\sigma(A_{jt},\Omega_{jt})\alpha}$. The modified model-expression (for column 2, in Table 1) of WND is thus $WND = \psi(A_{jt},\Omega_{jt})Y_tP_t^{\sigma(A_{jt},\Omega_{jt})}A_{jt}^{\sigma(A_{jt},\Omega_{jt})}\left(P_{jt}^F\right)^{-\sigma(A_{jt},\Omega_{jt})}\Omega_{jt}$ and the long-run zero restrictions of column 3 of Table 1 remain unchanged.

Third, as we show in Section 4, the approach is completely robust to potential missmeasurement of σ since this parameter only enters on the third row of Table 1 with (as we show) a very low weight.

Fourth, the key assumption for distinguishing technology shocks from demand shocks is that technology shocks alter the physical Solow residual in the long run, whereas other shocks do not. This assumption implies that changes in the scale of operation are not allowed to permanently alter the efficiency of production as measured by TFP. The most straightforward reason why this assumption may prove invalid is that firms might use a production technology with non-constant returns to scale. It is, however, straightforward to incorporate non-constant returns to scale into the model.[15] In Section 4, we provide versions of the model where we vary the returns to scale across the full reasonable range.

3 Data and Estimation of the Shocks

3.1 Firm-Level Data and Measurement

Our primary data source is the Swedish Industry Statistics Survey (IS). It contains annual information on inputs, outputs, and firm-specific producer prices for all Swedish manufacturing plants with 10 employees or more from 1990 through 2002. We perform our analysis at the plant level, but because about 72 percent of the observations in our sample pertain to plants that are also firms, we refer to the plants as firms.

[15] Details regarding the necessary modifications for non-constant returns to scale cases, are found in Appendix A

In our model, the technology shock η^a is the only shock that affects the Solow residual in the long run. This assumption is only credible if the Solow residual is calculated from a measure of real output where nominal output has been deflated by firm-specific prices. This is important because gross output deflated by sector-level price deflators (a measure often used in empirical analyses) will be a function of firm-specific idiosyncratic prices, which themselves are likely to depend on shocks other than technology (see Carlsson and Nordström Skans, 2012, for direct evidence). As our data-set contains a firm-specific price index built from plant-specific unit price changes,[16] we can derive a measure of gross output that is robust to changes in relative prices across firms. See Eslava et al. (2004) and Smeets and Warzynski (2013) for a similar strategy.

To take our model to the data, we rely on gross output throughout. We first compute a measure of firm-level changes in the physical Solow residual for firm j at time t. Letting lowercase letters denote logs, we use

$$\Delta a_{jt} = \Delta y_{jt} - \Delta z_{jt}, \tag{6}$$

where Δy_{jt} is the growth rate of real gross output, and Δz_{jt} is a cost-share-weighted input index defined as $C_K \Delta k_{jt} + C_N \Delta n_{jt} + C_M \Delta m_{jt}$ where Δk_{jt} is the growth rate of the capital stock (see details in Appendix B), Δn_{jt} is the growth rate of labor input, and Δm_{jt} is the growth rate of intermediate materials and energy. C_J terms are the cost shares of factor J in total costs. To calculate the cost shares, we use industry-level averages over time and take total costs as approximately equal to total revenues. The cost share of capital is then given by one minus the sum of the cost shares for all other factors.[17]

[16] The index uses Paasche-type links. In cases where a plant-specific unit-value price is missing (e.g., when the firm introduces a new good), Statistics Sweden uses a price index for similar goods defined at the minimal level of aggregation (starting at the four-digit goods-code level). The disaggregated sectoral producer-price indices are only used when a plausible goods-price index is not available. Our identification is fully resilient to transitory errors in measured prices.

[17] Our monopolistic-competition model outlined above implies pure economic profits. However, similar to U.S. evidence discussed in e.g. Basu, Fernald, and Shapiro (2001), we find a very small time average $(1968-1993)$ for the share of economic profits (-0.001) when relying on the aggregate Swedish manufacturing data from Carlsson (2003) and a no-arbitrage condition from neoclassical investment theory (taking the tax system into account) to calculate the user cost of capital. This finding thus support the commonly used approximation in the literature of measuring (average) cost shares by (average) revenue shares, which is also used here. For simplicity, however, we do not complicate the cost structure in our model in order to explicitly accommodate the absence of economic profits in the data.

Using data on factor compensations, changes in output, and changes in inputs, we can thus calculate the residual Δa_{jt}, which provides an estimate of changes in the physical Solow residual. As argued above, this might not accurately measure technology shocks (η^a) due to varying factor utilization, inventories, or truly idiosyncratic factor prices, but the SVAR will filter out true technology shocks from equation (6) as long as η^a is the only factor that permanently shifts A_{jt}. Material inputs are deflated using three-digit sectoral price indices, which implies that we allow, not only for an arbitrary set of transitory factor price shocks, but also for permanent input price shocks within the manufacturing sector as long as these are shared with other similar (at the 3-digit level) firms.

We next compute $\Delta wnulc_{jt}$ and Δwnd_{jt}. Relying on cost minimization, we use C_N as the estimate of α and thus let it vary by two-digit industry. The rest of the components of $\Delta wnulc_{jt}$ are directly observed in the firm-level data. However, to compute wage-neutral demand (Δwnd_{jt}) we also need an estimate of the demand elasticity σ. We obtain this by estimating the demand equation (2) while instrumenting the firm idiosyncratic price using the Solow residual, as in Foster et al. (2008). The instrument is consistent with our initial assumptions, because the Solow residual is expected to affect firm-level sales only through firm-level prices. The results of this procedure suggest an elasticity of substitution equal to 3.306 (s.e. 0.075), which we use when computing Δwnd_{jt}. The σ estimate is well in line with standard calibration exercises (see e.g., Erceg, Henderson, and Levin, 2000) as well as recent Swedish micro-evidence provided by Heyman, Svaleryd, and Vlachos (2008). As robustness checks, we also show that the main results are robust to using sector-specific estimates of σ and to using a very wide span of assumed values of σ.

We extract our baseline shocks by estimating a VAR on a sample of $6,137$ firms and $53,379$ firm/year observations (see Appendix B for additional details on the data and for details on the construction of the final sample). Since the VAR model uses lags, we can extract structural shocks for $41,105$ firm/years.

To analyze the impact of the shocks on the use of labor and the flows into and out of the firms, we link a longitudinal employer-employee data base (Statistics Sweden's register-based labor market statistics, or RAMS) to the firm-level data. These data are based on tax records and include the identity of all employees within the plants at the end of the year (November). We restrict the analysis to full-time employees within their main jobs. In the end, we are able to match shocks and labor flows for

$40,451$ firm/year observations in $6,125$ firms. The final sample covers nearly two-thirds of all manufacturing employees.[18] For completeness, we further study how the use of marginal manpower (i.e. employees that do not satisfy these criteria) changes in response to the shocks.

3.2 Estimation and Validation

To derive the shocks of interest, we estimate a SVAR on the three variables defined in Table 1: Δa_{jt}, $\Delta wnulc_{jt}, \Delta wnd_{jt}$, which are constructed in order to provide us with the recursive set of long-run restrictions we need to identify the structural shocks, and a fourth residual variable (which will be output, Δy_{jt}, unless otherwise noted) which will soak up any remaining residual transitory dynamics. The details are outlined in Appendix C. The standard deviation of the demand shock is about 60 percent larger than the technology shock (16.2 and 10.1, respectively), see the appendix for distributions and impulse responses. Notably, we find a fairly limited amount of dynamics, in particular in the Solow residual since the Solow residual is defined in physical gross terms and much of the dynamics in standard measures of Solow residuals appear to be due to the dynamics of idiosyncratic prices (see Carlsson and Nordström-Skans 2012, for direct evidence on relative-price dynamics).

In Appendix C.4, we validate the interpretation of the derived shocks by showing that they have the expected impact on firm-specific prices and output. The idea is that both technology and demand shocks should affect output, whereas prices should fall if the output increase is due to a technology shock but not if it is due to a demand shock. The reason is that technology shocks only affect the cost of production, so firms need to lower their prices in order to increase their sales along a fixed demand curve. In contrast, demand shocks, shifts the firm-specific demand curve, allowing the firm to sell more at a given price. The Appendix validates these predictions: A 1 standard deviation (sd.) technology (demand) shock increases output by 6 (10) percent in the long run. Moreover, as expected, prices go down due to technology shocks but increase slightly due to demand shocks. Note that these results are not

[18] Note that the employment data used to construct the variables in the VAR are obtained from a different source (IS) than the employment, hiring and separation data used in the final analysis (which is obtained from RAMS). This insulates the analysis from the threat of joint measurement errors in the calculation of the shocks and the employment adjustment analysis. Estimates of the impact of the shocks on overall employment are, however, very similar using the two data sources, suggesting that the issue is of minor importance.

imposed from the construction of our variables: in particular, prices could well (from a pure measurement standpoint) respond in either direction to structural innovations in both technology and demand.[19]

4 Shocks, Employment and Labor Flows

4.1 Labor Flow Data

Our main outcomes are based on firm-level employment and labor flows. We either measure employment in logs or, when decomposing the results into flows, measured as net employment changes defined as the difference between hires and separations. We compute these flow based measures using annual individual-level employment data on end-of-the-year employment that are matched to our firm-level data, using the metrics proposed by Davis et al. (1996).

Table 2: Summary Statistics. Worker's data

	Category	(1) Mean	(2) sd	(3) p(25)	(4) p(75)	(5) Firms	(6) Observations
Net Emp. Rate	overall	0.012	0.208	-0.062	0.089	6,125	40,451
	within		0.195				
Hiring Rate	overall	0.150	0.151	0.063	0.200	6,125	40,451
	within		0.127				
Separation Rate	overall	0.138	0.152	0.061	0.174	6,125	40,451
	within		0.131				
ST Separation Rate	overall	0.061	0.082	0	0.083	6,125	40,451
	within		0.065				
Marginal Net Emp. Rate	overall	0.009	0.353	-0.069	0.082	6,125	40,451
	within		0.334				

Note: The"within"rows show the dispersion within establishments. p(N) denotes the Nth percentile of the data.

Following Davis et al. (1996), net employment growth is defined as the change in employment relative to the preceding year, divided by the average employment during the two years. Similarly, we define the hiring (separation) rate as the number of new (separated) employees between t and $t-1$, divided by the average number of employees during the two years. With these definitions, net employment growth will

[19] The Appendix also provide additional support for the interpretation based on theory-consistent signs for the three unrestricted responses within the VAR system.

be the difference between the hiring rate and the separation rate, and the timing of the flows are defined such that the flow equation of employment holds i.e. $Employment_t = Employment_{t-1} + Hires_t - Separations_t$.

We do not observe the contract type in the data, but in order to explore the role played by the (potential) flexibility provided by marginal workers, we use two additional flow margins. We (i) separate out the number of separations of short-tenured (less than three years) workers divided by average employment across the two years and (ii) measure the change in the number of marginal workers defined as the number of individuals who are employed during the year, but who are not included in the stock of end-of-the year employees.

Descriptive statistics are presented in Table 2, The average hiring rate during the observation period is 15 percent, and the average separation rate is 14 percent, whereof slightly less than half (6 percent) are separations of short-tenured workers .

4.2 Baseline Results

The objective of our analysis is to illustrate how employment flows at the firm level respond to permanent shifts in idiosyncratic production functions and demand curves. We have derived results for the employment responses within the VAR framework and present these results in the appendix. But in order to allow us to also explore non-linear response margins, we instead extract the measures of structural shocks from the SVAR and relate them to different outcomes in most of the analyses we present in the main paper. This gives us additional flexibility in the specifications and allows us to present the results in a more compact table format. None of the results differ between strategies, however. Empirically, we estimate the following equation in the (linear version of the) baseline specification:

$$Outcome_{jt} = \eta_{jt}^{a}\delta_1 + \eta_{jt}^{\omega}\delta_2 + \rho_t\beta_\rho + \mu_j + \xi_{jt,} \tag{7}$$

where *Outcome* denotes employment (or some measure of labor flows) for firm j at time t. The coefficients δ_1 and δ_2 capture the impact of the two structural shocks.[20] Moreover, we include time, ρ_t, and firm-fixed effects, μ_j, in line with the SVAR formu-

[20]Formally, the inference is exposed to a potential generated regressor bias, but we show that all key results hold when either estimating them internally in the VAR or when relying on an IV-strategy (see below), both of which are insensitive to generated regressor biases.

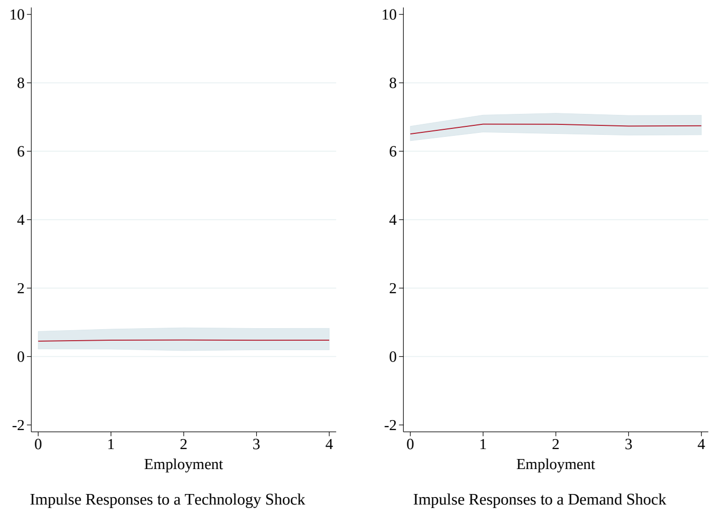
lation above. This ensures that identification is driven by idiosyncratic, rather than aggregate, shocks.[21] Equation (7) shows the short-run impact of the shocks. We also present the long-run impact, measured as the sum of the contemporary effect and the impact of the first lag in the shock series.

Our baseline specification, following equation (7) is presented in the first column of Table 3, measured as net employment changes whereas Figure 1 shows impulse responses of log employment with bootstraped confidence bands directly from within the VAR. The results between the table and the figure concur, as expected. They both show that idiosyncratic demand shocks have substantially more impact than the corresponding technology shocks on firm-level labor adjustments. Focusing on the Figure, we see that a positive demand shock of 1 sd. increases employment by slightly more than 6 percentage points, whereas the impact of an equivalent technology shock has a very limited impact on employment. It is also evident from Figure 1 that the dynamics of labor adjustments are fairly limited. More than 90 percent of the long-run adjustments in response to the permanent shocks occur within the first year. We show in Section 4.4 below that the finding that demand is more important than technology and that dynamics are limited is robust to varying assumptions about returns to scale, the demand elasticity, sectoral heterogeneity, sample selection, alternative treatments of dynamics, and models that account for firm exit. We return to a deeper discussion regarding the magnitudes in Section 4.5 below.

The results presented in the table where we focus on net employment changes (defined as the difference between hires and separations following the metric proposed by Davis, Haltiwanger, and Schuh (1996)) are, as expected, quite similar to the results presented in the figure. Estimates imply that the coefficient of a 1 sd. technology shock is 0.15 (not statistically different from 0). If we add one lag of the shocks to the regression and calculate the long-run employment responses (Column 4), the technology shock becomes somewhat larger and statistically significant. However, the table corroborates our findings of limited dynamics in the labor adjustment. As before, firms' demand shocks continue to be the main driver of employment adjustments: A positive 1 sd. shock to the demand curve increases employment in the long run in 6.4 percentage points, while the equivalent technology shock increases employment by 0.5 percentage points. Panel B of the table shows estimates in the form of elasticities

[21]Since the shocks are identified as structural orthogonal innovations, they are uncorrelated with each other conditional on the year and firm-fixed effects of the SVAR.

Figure 1: Employment Responses



Note: Impulse responses to a 1 sd. shock expressed in percentage points. Lines depict the mean of the bootstrap distributions. Shaded areas depict the bootstrapped 95 percent confidence intervals calculated from 1,000 replications.

(see Appendix C regarding the computation of the elasticities).

We proceed by estimating equation (7) for hires and separations. A normal demand shock is estimated to increase the hiring rate by 2.9 percentage points and reduce the separation rate by 2.7 percentage points in the short run (slightly more in the long run).These numbers should be compared with average hiring and separation rates of about 14 to 15 percent each, as shown in Table 2 above. The estimates imply that, on average, 52 percent of the net employment adjustment is obtained using the hiring margin, and 48 percent using the separation margin. Firms thus, on average, rely as much on variations in separations as on variations in hires when responding to the shocks. This result is also interesting in the light of the literature on labor flows and the business cycles (see Barnichon, 2012; Fujita and Ramey, 2009; and Shimer, 2012). It suggests that any quantitatively important asymmetries between hiring and separations over the business cycles should be explained by asymmetries in the market responses, and not as asymmetries in firm-level labor adjustment behavior.

The results further imply that the low response of net employment to technology shocks does not mask any substantive counteracting responses in terms of gross flows. Rather, idiosyncratic technology shocks appear to have a limited impact on both hiring and separation rates in both the short run and the long run.

Table 3: Contemporaneous and Long-Run Effect on Labor Flows

	SHORT RUN			LONG RUN		
	(1)	(2)	(3)	(4)	(5)	(6)
	Net Emp. Rt.	Hiring Rt.	Sep. Rt.	Net Emp. Rt	Hiring Rt.	Sep. Rt.
A) 1 sd. shock:						
Technology (η^a)	0.115	-0.050	-0.165*	0.412*	-0.093	-0.504**
	(0.119)	(0.075)	(0.078)	(0.163)	(0.116)	(0.128)
Demand (η^ω)	5.609**	2.906**	-2.703**	6.009**	3.125**	-2.884**
	(0.173)	(0.096)	(0.120)	(0.228)	(0.156)	(0.186)
B) Elasticities:						
Technology (η^a)	0.011	-0.005	-0.016*	0.0409*	-0.009	-0.050**
	(0.012)	(0.007)	(0.008)	(0.016)	(0.011)	(0.013)
Demand (η^ω)	0.347**	0.180**	-0.167**	0.371**	0.193**	-0.178**
	(0.011)	(0.006)	(0.007)	(0.014)	(0.010)	(0.012)
Observations	40,451	40,451	40,451	34,414	34,414	34,414
Firms	6,125	6,125	6,125	6,116	6,116	6,116

Note: Robust standard errors in parenthesis. Net Emp. Rt.: Net employment rate; Hiring Rt: Hiring rate; Sep. Rt: Separation rate. Hiring and separation rates are measured as the flow between the end points of two years divided by the average employment across these two points in time. The net employment rate is the difference between the hiring rate and the separation rate. Regressions include time dummies and firm fixed effects. The long-run impact is based on the sum of the contemporary effect and the effect of the first lag. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

Next, we isolate the analysis of separations of short-tenured (< 3 years) workers. The results in Column (2) of Table 4 show that these make up slightly more than one-third of the total separation response to demand shocks in the short run (Column 1) and even less in the longer-run (Column 4 vs. 5). Note that the lower relative contribution of short-tenured separations in the long run is consistent with a reduction in contemporary hires, which reduces the number of short-tenured workers who can be released in the next period.

As a final analysis, we document the responses in terms of "marginal workers", defined as the number of remunerated employees that do not at any time satisfy our criteria as being regular end-of-the year employees.[22] The results, presented in Table 4, show that the adjustments in terms of marginal workers is very similar to the adjustments of regular employees in the sense that most of the adjustment is due to demand shocks. We also see some evidence of initial overshooting in the sense

[22]Note that we measure the *number* of marginal employees and do not address the intensity by which these are used.

that the initial response is larger than the long-run adjustment in the use of marginal workers. The short-run response indicate an increase in the number of marginal workers corresponding to 3.8 percent of the number of full time employees in response to a 1 sd. positive demand shock. The corresponding long-run estimate is 3.0 percent.

Table 4: Contemporaneous and Long-Run Effect on Short Tenured Separations and Marginal Workers

	SHORT RUN			LONG RUN		
	(1)	(2)	(3)	(4)	(5)	(6)
	Sep. Rt.	ST Sep Rt.	Marg. Net. Emp Rt	Sep. Rt.	ST Sep Rt.	Marg. Net. Emp Rt.
Technology (η^a)	-0.165*	-0.117**	0.110	-0.504**	-0.177**	0.482
	(0.078)	(0.038)	(0.162)	(0.128)	(0.066)	(0.248)
Demand (η^ω)	-2.703**	-1.010**	3.796**	-2.884**	-0.416**	3.019**
	(0.120)	(0.052)	(0.213)	(0.186)	(0.076)	(0.278)
Observations	40,451	40,451	40,451	34,414	34,414	34,414
Firms	6,125	6,125	6,125	6,116	6,116	6,116

Note: Effect of one s.d. shock. Sep. Rt: Separation rate; ST Sep. Rt.: Short-tenured separation rate measured as the number of separations of short-tenured (< 3 years) workers; Marg. Net Emp Rt.: Adjustment of workers not fullfilling the criteria for a full-time primary employment. All rates are measured as the flow between the end points of two years divided by the average (full-time primary) employment across these two points in time. Regressions include time dummies and firm fixed effects. The long-run impact is based on the sum of the contemporary effect and the effect of the first lag. Robust standard errors in parenthesis. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

Overall our main results show that (i) employment and labor flows respond more heavily to permanent demand shocks than to permanent technology shocks, (ii) most labor adjustments happen within the year, (iii) hires and separations are equally important as adjustment margins, and (iv) short-tenured separations and adjustments of marginal workers follow similar patterns as adjustments of regular employees, but with a somewhat larger initial response.

4.3 Alternative Identification

The focus of the analysis so far has been on how firms adjust employment, hires, and separations when hit by *permanent* idiosyncratic shocks. Here we instead derive an alternative measure of demand and technology shocks that also include transitory shocks. As technology shocks, we now use the raw (physical) Solow residuals. We then use these residuals as an instrument for prices in an estimation of a log-linearized version of the demand equation (2), where time dummies control for aggregate shocks and firm fixed effects eliminate between-firm permanent heterogeneity. Since the ensuing residuals of the estimated equation represent changes in sales without price adjustments, they serve as a measure of demand shocks. This strategy is similar to Foster, Haltiwanger, and Syverson (2008). Thus, we label these shocks "FHS".

In contrast to the SVAR filter the FHS procedure does not differentiate between permanent and transitory shocks and the processes do not account for factor price shocks. The correlation between FHS demand and our SVAR demand shocks is 0.538. The standard deviation is considerably higher for the FHS demand shocks than the baseline demand shocks (0.24 versus 0.16). Thus, the two demand-shock series appear to contain a substantial common component without being identical. The correlation between FHS demand and the factor price component of the SVAR is considerably smaller (-0.25), but statistically significant. As expected, the FHS-demand series is uncorrelated with the SVAR technology shocks. Also as expected, given the limited dynamics observed in the physical Solow residual series, the physical Solow residual is highly correlated with the SVAR-technology shocks (0.98), and only marginally related to SVAR demand (correlation of 0.02) and factor price shocks (correlation of 0.06).

Table 5 shows how these measures relate to labor flows. Estimates for our SVAR shocks are reproduced in Columns 1 and 4. Clearly, the main message still holds for results in Column 2 and 5, using the FHS-series: the short-run impact of demand shocks is 10 times that of the technology shock in the short run, and 4 times in the long run. But it is also noticeable that the estimated impact of demand shocks is about half as large when using FHS demand as when using the SVAR demand shock.

To see what drives the difference, we proceed by purging the FHS series of our permanent structural shocks. To this end, we run a regression with FHS demand as the dependent variable and use our SVAR shocks (demand, technology and factor prices) as regressors and then repeat this for technology. We then label the residuals of this exercise *transitory* demand and technology shocks. Because these residuals are measured in the same units as the composite demand shock, we can directly compare its impact on employment adjustments with the impact of that series.[23]

[23] The decomposition resembles Guiso et al. (2005), which extracts the permanent component of

Table 5: Baseline Estimates vs. Solow Residuals and FHS Demand Shocks

	SHORT RUN			LONG RUN		
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	FHS	Transitory	Baseline	FHS	Transitory
Technology (η^a)	0.153	0.333*	-0.157	0.504*	0.993**	0.0754
	(0.159)	(0.168)	(0.164)	(0.214)	(0.250)	(0.216)
Demand (η^ω)	5.986**	3.406**	0.674**	6.357**	4.061**	0.863**
	(0.233)	(0.183)	(0.136)	(0.310)	(0.252)	(0.217)
Observations	40,451	40,451	40,451	34,414	34,414	34,414
Firms	6,125	6,125	6,125	6,116	6,116	6,116

Note: Effect of one s.d. shock. In the FHS column the technology shock is the Solow residual, and the demand shock is FHS demand, as defined in the main text. The transitory shocks are calculated as the residual component of the FHS series. Robust standard errors in parentheses. Regression includes time dummies and firm fixed effects. Long-run impact is based on the sum of the contemporary effect and the effect of the first lag. Regression sample limited to observations where the absolute value of both the technology and the demand shock is less than or equal to two sd. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

The results in Columns 3 and 6 of Table 5 show that the ensuing transitory demand shocks have a much more muted impact on employment than the SVAR shocks and the composite FHS shocks. This reinforces the picture that our SVAR strategy capture the most relevant parts of the shock process. The result hold for both the short and the long run responses. The fact that the long-run response to transitory demand shocks does not revert back when the lag is introduced suggests that the transitory series may still contain a persistent component.[24] With this caveat in mind, the fact that the part of the demand series which is certified to be permanent has a much larger effect suggest that firms' employment adjustment depends on the time-series properties of the shocks as in Franco and Philippon (2007) and Roys (2016). This is important because the welfare consequences of firms' lack of adjustment are likely to crucially depend on these properties. Labor hoarding in the face of negative transitory shocks may be welfare-enhancing in the presence of uninsurable labor market risk (Bertola, 2004), whereas the ability of firms to adjust to permanent shocks is likely to be crucial for long-run allocative efficiency.

firm-level value added using high-order polynomials of lags as instruments. Although the mechanics of the methods differ, the underlying logic is similar. Note, however, that an additional value added from our strategy is that we are able to remove the factor price component.

[24]If we purge the series of the fourth, residual, shock of the SVAR, the estimates are about half the size and the long run estimate is insignificant.

4.4 Further Robustness

In this subsection, we present a wide set of further robustness checks. We focus on the impact on employment measured in logs.

Table 6: Contemporaneous and Long-Run Effect on Log Employment under Different Returns to Scale Assumptions

	SHORT RUN			LONG RUN		
	(1)	(2)	(3)	(4)	(5)	(6)
	RTS=1	RTS=0.9	RTS = 1.1	RTS=1	RTS=0.9	RTS = 1.1
Technology (η^a)	0.153	0.955**	-0.492**	0.504*	1.378**	-0.244
	(0.159)	(0.161)	(0.149)	(0.214)	(0.211)	(0.232)
Demand (η^ω)	5.986**	6.149**	5.541**	6.357**	6.313**	5.978**
	(0.233)	(0.233)	(0.223)	(0.310)	(0.310)	(0.301)
Observations	40,451	41,132	39,788	34,414	35,031	33,811
Firms	6,125	6,193	6,065	6,116	6,184	6,055
Sd. η^a	10.06	10.04	10.37	10.06	10.04	10.37
Sd. η^ω	16.18	18.74	13.45	16.18	18.74	13.45

Note: Effect of one s.d. shock. Robust standard errors in parenthesis. Regression includes firm fixed effects and time dummies. Long-run estimates are obtained by adding the contemporaneous impact and one lag. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

The constant returns to scale (RTS) assumption used in the construction of the Solow residual is potentially controversial. In Carlsson, Messina, and Nordström-Skans (2016), we estimate RTS separately for the durables and non-durables sectors among Swedish manufacturing firms, obtaining 1 for durables and 0.9 for non-durables. In both cases we cannot reject the null of constant RTS. These results are very similar to what Basu, Fernald, and Kimball (2006) report for the U.S. Note also that what matters is the long-run returns to scale which implies that the theoretical case for assuming constant returns to scale becomes stronger.

The model can be altered to accommodate increasing or decreasing RTS. Changing the assumed RTS affects the measures that are fed into the SVAR (for details, see Appendix A) and hence also the estimated magnitudes of employment adjustments. However, the main message remains robust throughout. Column 2 in Table 6 reports results from imposing RTS of 0.9 in the construction of the Solow residual. A positive technology shock of 1 sd. raises employment now by 1 percentage point in the short run (1.4 in the long run, see column 5). But this estimate still remains far below the estimated impact of a demand shock: an increase of 6.1 percentage points in the short run and 6.3 in the long run. If instead we impose an RTS coefficient of 1.1, the results change in the other direction (the impact of technology turns negative), but the main message regarding the strong relative importance of demand remains unaltered.

We proceed by carrying out a battery of checks to assess the robustness of our first set of findings—namely, that (i) firm-level demand shocks are more important in the determination of labor adjustments than firm-level technology shocks, and (ii) employment adjustment to the permanent shocks is very rapid, exhibiting limited short-term dynamics. In all cases we use the specification presented in equation (7). We discuss the main findings here, but present the regression tables in Appendix D to conserve space.

Demand elasticity. The baseline specification uses an estimated demand elasticity of 3.3. As a robustness check we have verified that our key results are robust to demand elasticities that vary within what we believe to be the full range of plausible values (from 1.1 to 10); the results in Table D1 (columns 2 and 3) show that the estimated coefficients of technology and demand shocks are remarkably stable despite this large interval of measured demand elasticities. Additional tests in Table D1 allow for industry-specific estimates of the demand elasticity, and as shown in column 4 of the table, this does not alter the results. The reason for this robustness is that measured σ only enters our system in order to handle idiosyncratic wage movements, and these are much smaller than than the movements in output which it is weighted against.[25] The main results also stay unaffected if we instead replace the year dummies by industry-by-year dummies, which controls for different employment trends across sectors (column 5).

Sectoral heterogeneity. The dynamic panel approach used for estimation took advantage of our large-N small-T panel setting to estimate the VAR system with considerable precision. This is a key advantage relative to standard SVAR estimations in the macro literature. A potential cost, however, is that the underlying dynamic processes are assumed to be equal across different firm types. To address this concern, we have allowed for separate dynamics for each two-digit industry, and the employment

[25] To recap, $\Delta wnd_{jt} = \Delta y_{jt} + \sigma(\alpha * \Delta w_{jt})$. In the data, the within-firm standard deviation in Δy_{jt} (0.326) is seven times larger than the within-firm standard deviation in $\alpha \Delta w_{jt}$ (0.046). Furthermore, the two elements are positively correlated (0.27). As a consequence, the within-firm correlation between Δwnd_{jt} as measured with $\sigma = 1.1$, and Δwnd_{jt} as measured with $\sigma = 10$, respectively, is 0.81.

adjustment results remain unchanged (see column 6 in appendix Table D1).

Sample selection. The data appendix (Appendix B) explains that the output allocation across plants within (the relatively few) multi-plant firms after 1996 is imputed in the IS data set. We have therefore redone the analysis for the single-plant firms in the sample (column 2 of appendix Table D2), as well as for a mixed sample including multi-plant firms until 1996, but not thereafter (column 3 of Table D2). The results are robust in these alternative samples which is not surprising since the bulk of the original sample is unaffected.[26] The results are also unchanged when the shock distribution is truncated into the Lester range of -2 to 2 sd. (see column 4 of Table D2).

Alternative fourth variable in the SVAR. To ensure that the limited dynamics in the employment adjustments we find is not due to the specific way we handle the residual dynamics in the system, we have used sales per worker, output, and employment from our two data sources (RAMS and IS) as alternative fourth variables. Table D3 shows that these variations only have minor impacts on both the estimated dynamics and the long-run adjustments.[27]

Firm exit. Finally, a possible concern with the analysis is that we disregard the firm-exit process. Firms are likely to exit in response to severe negative demand or technology shocks, and this process may impact labor dynamics. To address this concern,we have analyzed the employment impact of the shocks using a two-periods specification instead of the one-period baseline (see Appendix Table D4). In practice, this implies that we relate the shock to the net employment growth across two years, defined as the change in employment divided by the average employment in the two years as in Davis et al. (1996). Since the labor flows are defined even if all workers exit the year after the shock, we can calculate the impact of the shocks while excluding or including the firms that exit. Reassuringly, the results are insensitive to whether we include or exclude exiting firms.[28]

Overall, our findings strongly suggest that (i) permanent shifts in firms' idiosyncratic demand curves are a key determinant of firms' idiosyncratic net employment

[26] As explained in Section 3, 72 percent of plants are in single-plant firms and the imputation only affect the later half of the sample.

[27] That the fourth variable plays a negligible role in employment adjustment is also suggested in the variance decomposition shown in Appendix C.

[28] We have also analyzed the explicit relationship between the shocks and the probability of firm exit from the sample. The main driver of firm exits are large negative demand shocks which is well in line with results for the United States in Foster, Haltiwanger, and Syverson (2008).

adjustments, and (ii) the pace of labor adjustment is relatively fast. In contrast, permanent shifts in firms' physical production functions (i.e., technology shocks) appear to play a much more limited role in firms' labor adjustment, despite being crucial to the evolution of both output and relative prices.

4.5 Discussion of Magnitudes

The impact of technology shocks on employment is small, but fully in line with the finding of Pozzi and Schivardi (2016) for the Italian manufacturing industry. In particular, if assuming constant (decreasing) returns to scale our implied elasticity of the technology shock is 0.05 (0.14), whereas Pozzi and Schivardi (2016) find 0.08.[29]

The key novel finding relative to the previous studies is instead the strong employment effect we find from the demand shocks.[30] Here, it is worth noting that our demand shocks are permanent, and these are likely to have a larger impact than transitory shocks as indicated by our results presented above.

It is also notable that the full set of responses we observe are difficult to reconcile with the constant-σ assumption on which most of the literature so far has relied, including Foster, Haltiwanger, and Syverson (2008), Foster, Haltiwanger, and Syverson (2016) and Pozzi and Schivardi (2016) (when operating below full capital utilization), and which we also used to derive the restrictions in Section 2. The constant-σ model predicts that employment responses to technology and demand shocks are related by a factor of $\frac{1}{\sigma-1}$ (in our case; only in the long-run). The empirical employment responses would thus suggest that σ should be smaller than the value of 3.3 which we use in the measurement of WND. Although we could, as shown above, in principle choose any reasonable number for σ without affecting the results, the single parameter σ pins down all responses of prices, output and employment according to the standard constant-σ model (see Appendix C.6 for the full Jacobian). Unsurprisingly, we are unable to simultaneously match all of these responses regardless of which value we choose for σ. However, due to the flexible nature of our identifying restrictions,

[29] Pozzi and Schivardi (2016) find strongly decreasing returns to scale (0.8) for Italian firms in the textile, leather, metals and machinery sectors. This is lower than the average overall manufacturing returns to scale that has been found for Sweden, see e.g. Carlsson, Messina, and Nordström-Skans (2016) and the U.S., see e.g. Basu, Fernald, and Kimball (2006). The qualitative conclusions hold even if we impose a returns to scale of 0.8 however (demand is still three times as important as technology).

[30] The implied employment elasticity is 0.39, compared to, e.g., 0.08 found by Pozzi and Schivardi (2016)

it is possible to use an extended versions motivating model which can be reconciled with our identifying assumptions and which allows us to match the empirical results much better. Given the fact that our observed price response to the technology shock is substantially smaller than the unit response implied by the standard constant-σ model, the data seem to ask for a model that is richer in its description of product market responses to the shocks. A straightforward generalization in this direction is to assume that the elasticity of demand (and thereby the markup) can be affected by technology and demand shocks, i.e. a model where we replace σ by the function $\sigma(A_{jt}, \Omega_{jt})$ as already hinted at in Section 2. Such an extension is discussed in more detail in Appendix C.6.

5 Asymmetry and Non-Linearities

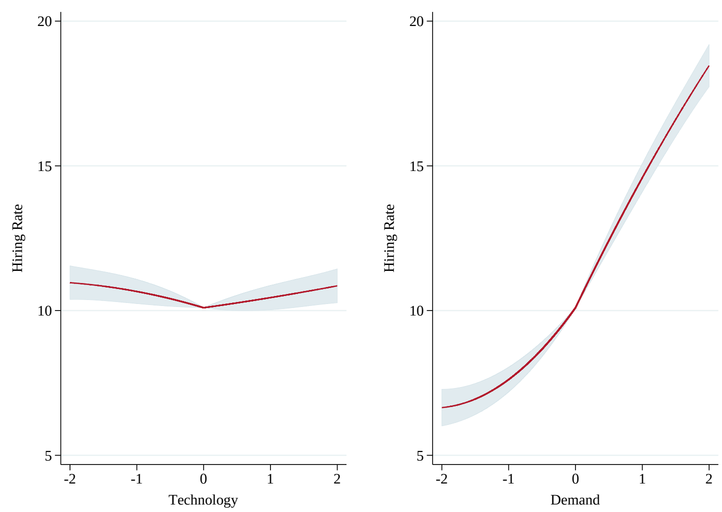
To examine if hiring and separation responses depend on the signs and magnitudes of the shocks, we extend equation (7) by allowing for separate second-order polynomials in the shocks, separately above and below zero. Because we showed above that the dynamics add few insights, we focus on the short-run impact.

Figure 2 shows how firms adjust their hiring rates in response to positive and negative shocks of different magnitudes. To facilitate the interpretation, the graphs show the sum of the average hiring rate among firms that do not adjust employment (about 10 percent) and the predicted estimates for various deviations from a zero-shock state. For completeness, we show the responses to both technology and demand shocks, but we focus our attention toward the demand-shock responses. (Throughout, we find limited adjustments in response to technology shocks, as expected from the results presented above.)

Two patterns are particularly noteworthy: First, the hiring response is considerably smaller if the shocks are negative. Second, the impact of a 2 sd. positive shock is exactly twice that of a 1 sd. positive shock, suggesting that the costs of increasing hires are a linear function of the magnitude of the adjustment.

Figure 3 shows the corresponding patterns for separations. The shapes and magnitudes (again focusing on the demand shocks) are not far from mirror images of the impact on the hiring rate. Thus, separations primarily respond to negative shocks. Although separations do go down somewhat when shocks are positive, this impact is even smaller than the hiring cuts in response to negative demand shocks. Symmetri-

Figure 2: Shocks and the Hiring Rate



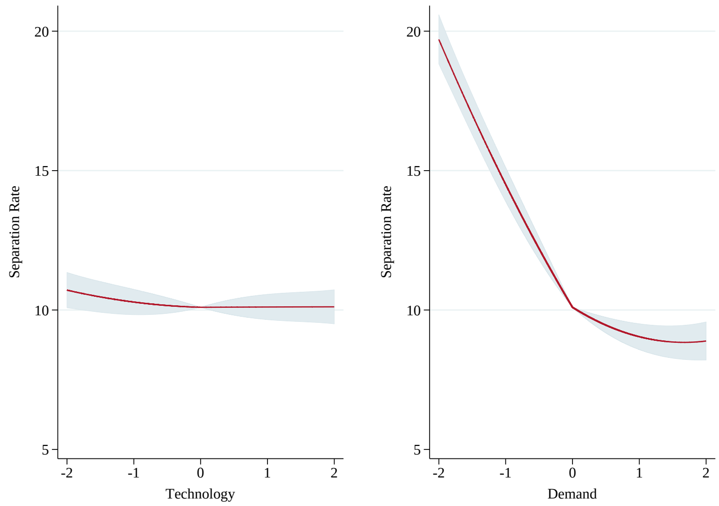
Note: Each line represents the sum of the average hiring rate among firms that do not adjust employment (10 percent) and the response of the hiring rate in percentage units as a (non-linear) function of an x sd. technology or demand shock. Shaded areas depict 95 percent confidence intervals.

cally to the hiring response, the estimates imply that a 2 sd. negative shock causes a separation response that is exactly twice as large as the response to a 1 sd. negative shock, which suggests that the costs of increasing separations are approximately linear on average. Notably, the results of Figures 2 and 3 imply that firms primarily use separations when responding to permanent negative demand shocks, an issue to which we return to below.

Finally, Figure 4 shows the impact on net employment and, as could be imagined from the combination of Figure 2 and Figure 3, these effects add up to a fairly linear relationship. The somewhat more curved pattern on the positive side arises because the kink at zero is more pronounced for hires than for separations. This difference in curvature is statistically significant, but the magnitude is fairly small: The net employment changes in response to a 2 sd. positive demand shock (9 percentage points) is reasonably close to the response to a 2 sd. negative shock (-13 percentage points) in absolute values.

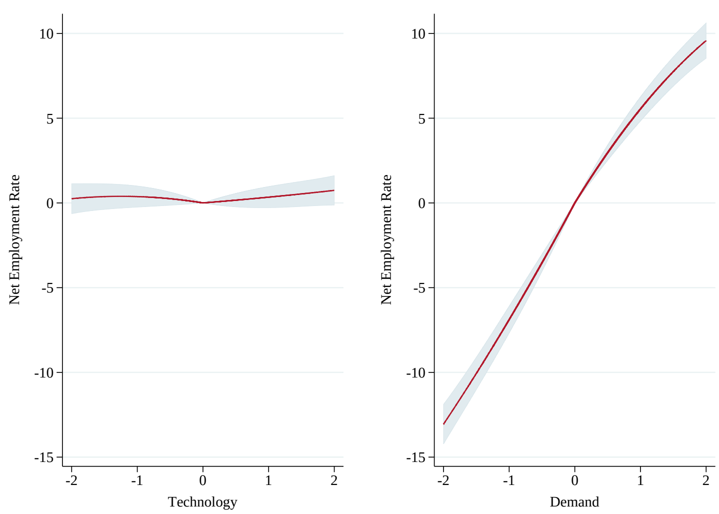
In Figure 5, we analyze the impact of transitory shifts in product demand on net employment changes using the data explained in Section 4.3. For comparison, the figure also reproduces the baseline response to a permanent demand shock (as in the

Figure 3: Shocks and the Separation Rate



Note: Each line represents the sum of the average separation rate among firms that do not adjust employment (10 percent) and the response of the separation rate in percentage units as a (non-linear) function of an x sd. technology or demand shock. Shaded areas depict 95 percent confidence intervals.

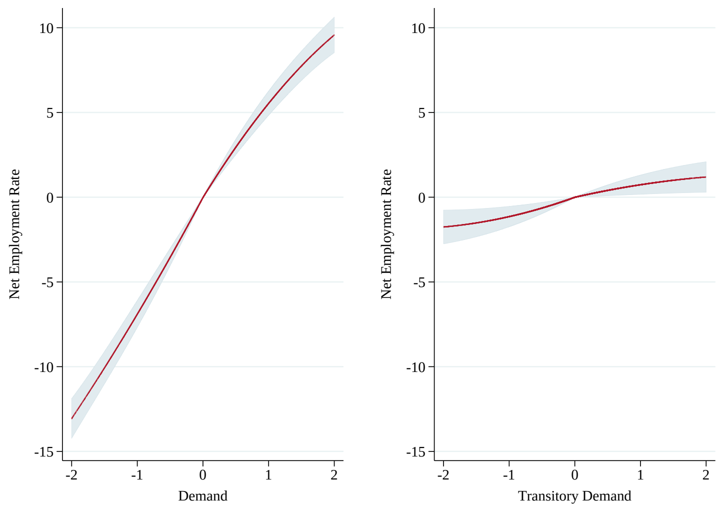
Figure 4: Shocks and the Net Employment Rate



Note: Each line represents the response of the net employment rate in percentage units as a (non-linear) function of an x sd. technology or demand shock. Shaded areas depict 95 percent confidence intervals.

right-hand side panel of Figure 4). As before, we allow for second-order polynomials of negative and positive shocks, respectively. The results show that the impact of the transitory shocks is substantially lower than the impact of the permanent shocks regardless of the sign or magnitude of the shock.[31]

Figure 5: Net Employment, Permanent and Transitory Demand Shocks



Note: Contemporaneous net employment rate in percentage units as a (non-linear) function of an x sd. of the baseline permanent demand shock and of the transitory demand shock (calculated as the residual component of FHS demand). Shaded areas depict 95 percent confidence intervals.

5.1 Decomposing Employment Responses

This subsection provides an analysis of how firm-level employment adjustments in response to permanent demand shocks translate into worker flows.[32] This analysis is similar in spirit to Abowd et al. (1999), and Davis et al. (2012), which provide decomposition exercises of the relative contribution of various worker flows to the observed employment changes in French and U.S. firms, respectively. In contrast to these previous studies, however, we analyze changes in hires and separations induced by employment adjustments due to a demand shock. This allows us to obtain a causal correspondent to the decompositions in the earlier literature. In our case,

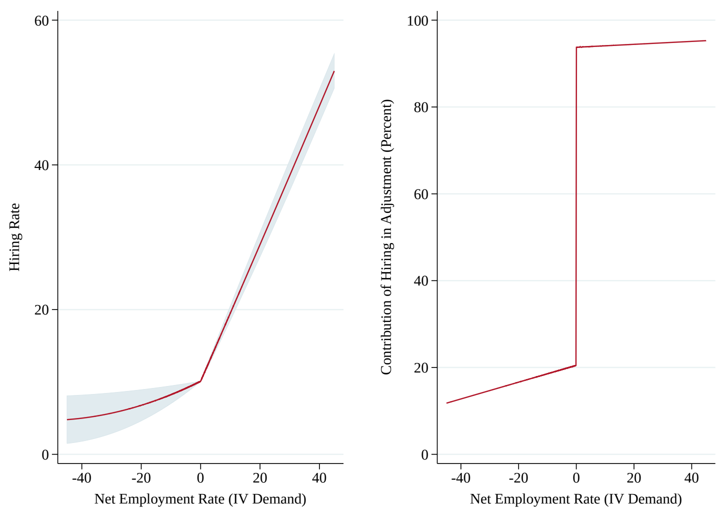
[31]Notably, the effect is even smaller, and insignificant, if we do not remove the factor price shock component.

[32]We focus on permanent demand shocks because technology shocks are found to have negligible impacts on net employment.

demand shocks drive the changes in employment, and we can therefore abstract from, for example, possible exogenous separations which may affect firms' employment levels in the short run.

In practice, we characterize labor adjustments by two second-order polynomials, one for positive values and one for negative values. We then instrument this adjustment by a similarly constructed set of polynomials in the demand shock. We use the hiring rate as our outcome, but since net employment adjustment is identical to the difference between hires and separations, the impact on separations is easily deduced.[33]

Figure 6: The Hiring Rate and Net Employment Changes. IV Results



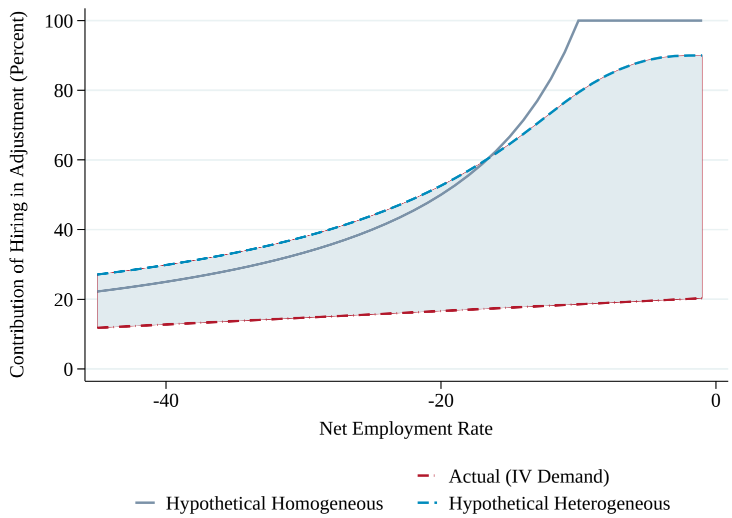
Note: Left-side panel: Contemporaneous hiring rate in percentage units as a (non-linear) function of employment adjustment in percentage units. Employment adjustments are instrumented by demand shocks. Shaded areas depicts 95 percent confidence intervals. Right-side panel: Implied fraction of employment adjustment achieved through changes in hirings as a function of the size and magnitude of the employment adjustment.

The results are presented in the left-hand panel of Figure 6. They imply a strong and linear relationship between net employment adjustments and hires when the employment adjustments are positive, but a very modest relationship when the employment adjustments are negative. The right-hand panel of Figure 6 shows the share of employment adjustment through hires as a function of demand-induced net employ-

[33]The instrumental variable (IV) strategy essentially implies that we scale the shock impact on the hiring rate presented in Figure 2 above with the first stage, which corresponds to Figure 4.

ment changes. This share jumps from 20 percent to 95 percent when employment adjustments become positive instead of negative.[34]

Figure 7: Actual (IV) and Simulated Hiring Responses to Employment Changes



Note: Actual (estimated from data) and hypothetical maximum (simulated) fraction of negtive employment adjustments (in percentage units) achieved through changes in hirings. Employment adjustments are instrumented by demand shocks. "Hypothetical homogenous" assumes that the same fraction of workers always leaves the firm. "Hypothetical heterogenous" imposes a random individual quit rate on the actual firm-size distribution.

Figure 6 also suggests that firms are relatively unconstrained in their use of separations, since they rely on increased separations even when they could have adjusted through reduced hires. To make this point precise, Figure 7 repeats the patterns shown in the right-hand panel of Figure 6 but focuses on negative values. As benchmarks illustrating what the firms *could have done*, the figure also depicts two hypothetical adjustment curves. The first, denoted "hypothetical homogeneous," assumes homogenous firms and imposes the empirical steady-state (i.e., without employment changes) separation rate of 10 percent on all the firms. In this case, as long as the need for adjustment is 10 percent or less, reduced hires could fully accommodate the necessary adjustments. If the shock is 20 (30) percent instead, the firm could instead accommodate half (one-third) of the adjustment through reduced hires. Notably, this curve

[34] Note that, in contrast to Figures 2 and 3 (where the zeros refer to the absence of an idiosyncratic shock), zero here refers to the state when net employment adjustment is predicted to be zero based on the full first stage (i.e., based on the combination of the shock polynomials, the year dummies, and the firm-fixed effects).

assumes that 10 percent of employees leave each firm every year, which clearly cannot be the case.

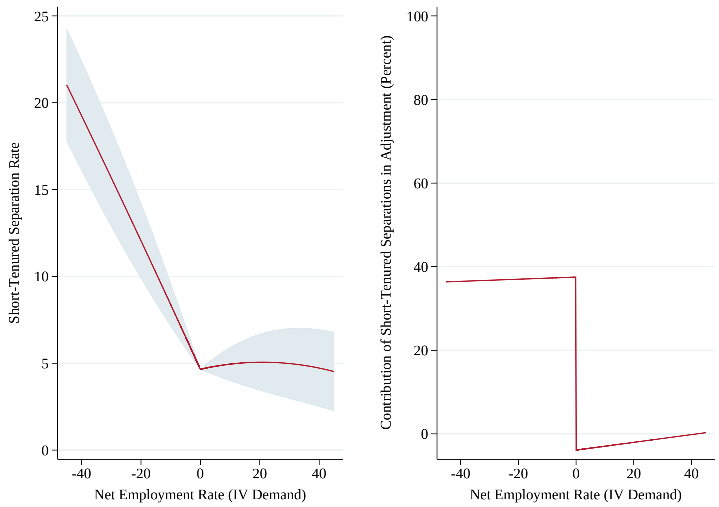
We therefore also provide a second benchmark, assuming instead that the *individual* probability of leaving a firm is 10 percent. By randomly allocating quits across the workers in our full sample and then aggregating to the firm level, we get the firm-level distribution of quit rates. With this distribution, which naturally widens if firms are small, some firms will not experience any quits at all, which means that they cannot accommodate even the smallest employment adjustment through reduced hires, whereas other firms will experience many random separations, allowing them to accommodate large employment reductions through reduced hires. The curve denoted "hypothetical heterogeneous" displays the simulated frontier of adjustments with random individual quits, within our actual distribution of firm sizes.

The logic behind the hypothetical curves is that they provide a baseline indicating how firms would behave in a completely rigid world where firing is prohibitively costly as long as firms are hiring someone. In this case, firms would always adjust according to the hypothetical heterogeneous curve in Figure 7. As is evident, the observed employment adjustments are far from this rigidity benchmark. The actual share of adjustment through reduced hires is much lower than the hypothetical reliance on separations would allow for. The shaded area between the heterogeneous hypothetical curve and the actual behavior of the firm could be interpreted as a region of flexibility because it depicts the amount of negative labor adjustments through induced separations (i.e., separations above the random rate) which could have been accomplished through reduced hires instead.

One reason for the observed patterns may be that firms adjust by releasing marginal, short-tenured workers who are more likely to be on temporary contracts. Sweden is a country with slightly above-average levels of employment protection (OECD, 2014), but the use of temporary contracts is flexible, whereas protection for workers with open-ended contracts is more restrictive. It is thus possible that the labor market responses studied here may hide important heterogeneity across workers, depending on their contract type and tenure with the firm.

We do not observe the contract type in the data, but in order to explore the role played by the (potential) flexibility provided by marginal workers, we have estimated the IV specification using the separation of short-tenured (less than three years) workers divided by average employment across the two years. The results, shown in Figure

Figure 8: The Separation Rate of Short-Tenured Workers and Net Employment Changes. IV Results



Note: Left-side panel: Contemporaneous separation rate of short tenured workers in percentage units as a (non-linear) function of employment adjustment in percentage units. Employment adjustments are instrumented by demand shocks. Shaded areas depicts 95 percent confidence intervals. Right-side panel: Implied fraction of employment adjustment achieved through changes in separation rate of short tenured workers as a function of the size and magnitude of the employment adjustment.

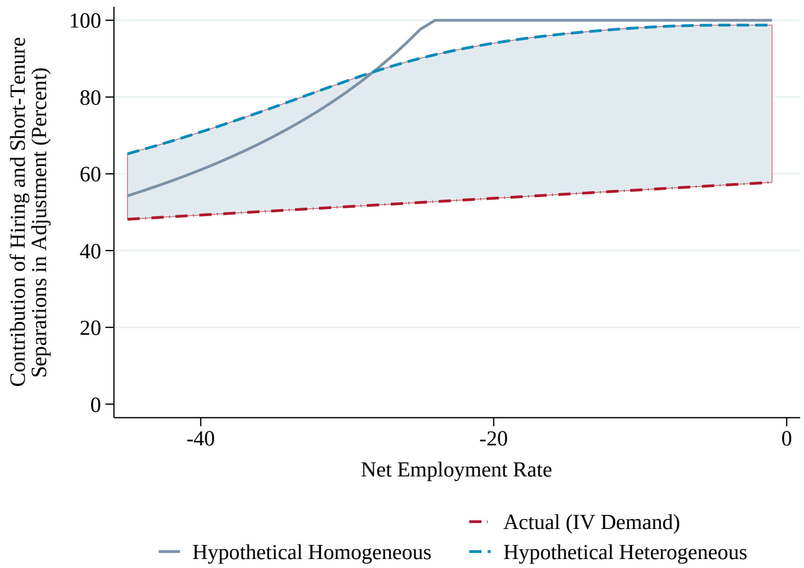
8, suggest that about half of the response to negative shocks come through reductions of short-tenured workers.

We have also repeated the simulation exercise presented in Figure 7 above, but instead contrasting the actual combined adjustment of reduced hires and increased separations of short-tenured workers with the maximum possible adjustment levels The results, presented in Figure 9, show that firms are far from using the flexibility provided by these two margins. The substantial shaded area in the figure implies that firms rely much more on separations of *long-tenured workers* than they would have needed to in order to achieve the same level of net employment reduction.

5.2 Firm-Level Heterogeneity

Taken at face value, our results imply that firms either bear few costs to separate long-tenured workers, or rely heavily on a well-defined mix of worker types that is hard to change when demand changes. If the latter is true, it is more than likely that

Figure 9: Actual (IV) and Simulated Hiring plus Short-Tenured Separation Responses



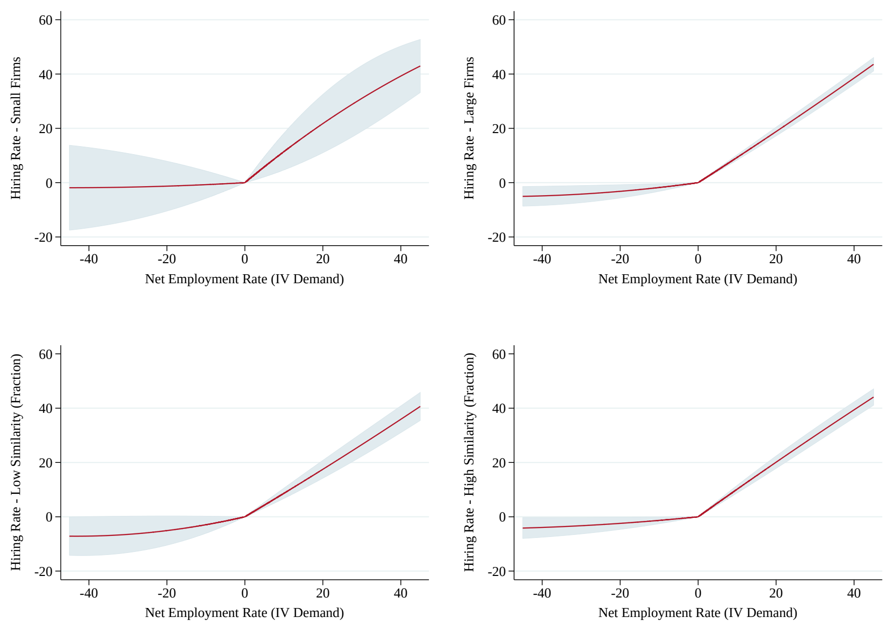
Note: Actual (estimated from data) and hypothetical maximum (simulated) fraction of negtive employment adjustments (in percentage units) achieved through changes in hirings and short-tenured separations. Employment adjustments are instrumented by demand shocks. "Hypothetical homogenous" assumes that the same fraction of workers always leaves the firm and are employed on short tenure (less than three years). "Hypothetical heterogenous" imposes a random individual quit rate and short-tenure rate on the actual firm-size distribution.

the workers who leave, or who are on temporary contracts, differ from the types of workers that the firms would like to separate from. Hence, it is not possible for the firm to fully exploit worker attrition or their pool of short-tenured workers to adjust to the shock.

To explore this further, we provide estimates separately for firms with a homogenous workforce in terms of field and education level and for firms with a heterogeneous workforce in the same dimensions. The idea is that firms with a more homogenous set of employees should care less about whom they separate from and thus rely more on attrition and the separation of short-tenured workers when adjusting their net employment.

In practice, we calculate the fraction of coworkers (to each worker in the data) that has the exact same type of education (three-digit field and two-digit level) and take the average of this share for each firm. This gives an index of the average worker's exposure to similarly trained workers within the firm (in spirit, similar to measures of

Figure 10: The Hiring Rate and Net Employment Changes: Firm Size and Worker Heterogeneity. IV Results



Note: Contemporaneous hiring rates in percentage units as a (non-linear) function of employment adjustments (in percentage units) in subsamples defined by employee heterogeneity (lower graphs) and firm size (higher graphs). Employment adjustments are instrumented by demand shocks. Low (high) similarity firms are those with a similarity index (described in the text) below (above) the median. Small (large) firms are those with fewer (more) than 20 employees. Shaded areas depicts 95 percent confidence intervals.

workforce diversity). In a second step, we split our firm-level data across the median of this index and analyze the two samples separately.

Figure 10 presents the results for the two samples, i.e., for firms with high versus low degrees of educational similarity among workers. As before, we characterize labor adjustments by two second-order polynomials, one for positive values and one for negative values. We then instrument this adjustment by a similarly constructed set of polynomials in the demand shock.

Quite surprisingly, we find little support for the notion that within-firm heterogeneity is an important explanation for the low reliance on separations when firms are hit by negative demand shocks. We would, however, like to acknowledge that our measures of staff heterogeneity may well be too crude to capture the role of firm-level heterogeneity in the adjustment patterns.

Figure 10 also shows results separately by firm size (more than 20 employees or fewer than 20 employees). The idea is again that if worker heterogeneity is important for the results, smaller firms are more likely to have difficulties using attrition and separation of short-tenured workers to adjust their staffs. The results, however, are very similar across the two size classes, displaying as before little signs of systematic heterogeneity.

6 Conclusions

This paper has analyzed how firms adjust their labor inputs in response to permanent idiosyncratic firm-level shocks to technology and demand. We identify the shocks by imposing a set of long-run restrictions in an SVAR estimated on firm-level data. The restrictions are derived from a stylized model of a monopolistically competitive firm. The SVAR is estimated using dynamic panel-data methods, allowing us to identify the parameters of the reduced form with considerable precision. To estimate the model, we rely on a unique data-set that merges information about inputs, outputs, and prices of Swedish manufacturing firms with a linked employer-employee data-set.

The shocks derived from the SVAR affect output and prices in a theory-consistent way, which lends support to their interpretation as demand and technology disturbances. Firm-level output responds vigorously to both technology and demand shocks. In contrast, firm-level prices fall in response to positive technology shocks, but they remain independent of product demand innovations.

Our labor-adjustment results show that both the nature and the time-series properties of the shocks matter. Permanent demand shocks, which affect output but not relative prices, have a pronounced impact on employment. In line with other recent studies, technology shocks have relatively limited employment effects despite affecting both output and relative prices.

A possible limitation of our study is the focus on the manufacturing sector, the sector for which technology shocks can be reasonably approximated. However, it seems likely that the overwhelming force of idiosyncratic demand shocks as a source of employment adjustments in manufacturing firms should provide a lower bound for the importance of demand within other sectors. Demand is likely to play an even more important role for reallocation in service sectors, where product differentiation (and hence demand shocks) is likely to be even more important than in manufacturing.

We further provide the first analysis of the causal impact of job flows on the composition of worker flows, using our permanent demand shocks as an instrument for adjustments in the number of jobs. The results suggest that employment adjustments in response to permanent shifts in the product demand curve are fast and symmetric. By far the largest part of employment adjustment takes place within a year. Almost as much of the employment adjustments are through changes in the separation rates as through changes in the hiring rates, suggesting that both margins should be considered endogenous at the firm-level. Moreover, there are no signs of non-linear responses in hires or separations. Finally, the sign of the shock determines the primary margin of adjustment: firms primarily adjust through separations if shocks are negative and primarily though hires if shocks are positive.

The speed of adjustment, the symmetry between hires and separations as adjustment margins, and the continued recruitment of workers in the face of negative shocks jointly suggest that labor market rigidities play a very limited role in hampering firm-level labor adjustments in the face of permanent idiosyncratic demand shocks. However, the adjustments with respect to transitory shocks appear to be muted. Thus, firms accommodate the impact of permanent shocks, but may hoard labor and refrain from hiring when hit by transitory shocks.

Overall, our results imply that cross-country comparisons of labor flows need to be careful in accounting for the types of the shocks that hit these economies, because responses depend not only on the nature of the shocks (technology versus demand) but also on the time-series properties of these shocks: Labor market adjustments will differ depending on the prevalence of permanent versus transitory components within the shock distribution.

Building on this notion, our empirical approach also suggests a route forward in trying to understand the forces behind the declining rates of labor adjustments observed in many countries. Essentially, our empirical approach provides a tool for assessing whether this development is due to a changing nature of firm-level shocks or due to a reduced impact of these shocks on labor reallocation. Although this question is beyond the scope of this paper, it serves as a good example of the questions that future research can answer by combining data on labor flows and well-identified firm-level shocks.

References

ABOWD, J. M., P. CORBEL, AND F. KRAMARZ (1999): "The Entry and Exit of Workers and the Growth of Employment: An Analysis of French Establishments," *Review of Economics and Statistics*, 81, 170–187.

ANDERSSON, F. (2003): *Causes and Labor Market Consequences of Producer Heterogeniety,* chap. Job Flows in Swedish Manufacturing, pp. 13–52, Economic Studies 73. Uppsala University.

ARELLANO, M., AND S. BOND (1991): "Some Tests of Specifications for Panel Data: Monte Carlo Evidence and Applicationto Employment Equations," *Review of Economic Studies*, 58, 277–297.

BARNICHON, R. (2012): "Vacancy Posting, Job Separation and Unemployment Fluctuations," *Journal of Economic Dynamics and Control*, 36, 315–330.

BASU, S., J. FERNALD, AND M. KIMBALL (2006): "Are Technology Improvements Contractionary?," *American Economic Review*, 95, 1418–1448.

BASU, S., J. FERNALD, AND M. SHAPIRO (2001): "Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?," *Carnegie-Rochester Conference Series on Public Policy*, 55, 117–165.

BENTOLILA, S., AND G. BERTOLA (1990): "Firing Costs and Labour Demand: How Bad Is Eurosclerosis?," *Review of Economic Studies*, 57, 381–402.

BERTOLA, G. (2004): "A Pure Theory of Job Security and Labour Income Risk," *The Review of Economic Studies*, 71(1), 43–61.

BLANCHARD, O., AND D. QUAH (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79, 655–673.

CAHUC, P., F. POSTEL-VINAY, AND J.-M. ROBIN (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 74(2), 323–64.

CARLSSON, M. (2003): "Measures of Technology and the Short-Run Response to Technology Shocks," *Scandinavian Journal of Economics*, 105, 555–579.

Carlsson, M., J. Messina, and O. Nordström-Skans (2016): "Wage Adjustment and Productivity Shocks," *Economic Journal*, 126, 1739–1773.

Carlsson, M., and O. 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, M., and J. Smedsaas (2007): "Technology Shocks and the Labor Input Response: Evidence from Firm-Level Data," *Journal of Money, Credit and Banking*, 39, 1509–1520.

D'Agostino, R. B., A. J. Belanger, and R. B. D'Agostino Jr. (1990): "A Suggestion for Using Powerful and Informative Tests of Normality," *American Statistician*, 44, 316–321.

Davis, S., and J. Haltiwanger (1992): "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, 107, 819–863.

Davis, S. J., R. J. Faberman, and J. Haltiwanger (2012): "Labor Market Flows in the Cross Section and Over Time," *Journal of Monetary Economics*, 59, 1–18.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1996): *Job Creation and Job Destruction*. MIT Press, MA.

Erceg, C., D. Henderson, and A. Levin (2000): "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics*, 46, 281–313.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler (2004): "The Effects of Structural Reforms on Productivity and Profitability Enhancing Reallocation: Evidence From Colombia," *Journal of Development Economics*, 75, 333–371.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Foster, L., J. Haltiwanger, and C. Syverson (2016): "The Slow Growth of New Plants: Learning and Demand," *Economica*, 83(329), 91–129.

Franco, F., and T. Philippon (2007): "Firms and Aggregate Dynamics," *Review of Economics and Statistics*, 89, 587–600.

Fujita, S., and G. Ramey (2009): "The Cyclicality Of Separation And Job Finding Rates," *International Economic Review*, 50, 415–430.

Galí, J. (1999): "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 89, 249–271.

Guiso, L., F. Schivardi, and L. Pistaferri (2005): "Insurance within the Firm," *Journal of Political Economy*, 113, 1054–1087.

Heyman, F., H. Svaleryd, and J. Vlachos (2008): "Competition, Takeovers and Gender Discrimination," *forthcoming Industrial and Labor Relations Review*.

Hopenhayn, H., and R. Rogerson (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101, 915–938.

Kimball, M. (1995): "The Quantitative Analytics of the Basic Neomonetarist Model," *Journal of Money, Credit and Banking*, 27, 1241–1277.

Lise, J., C. Meghir, and J.-M. Robin (2016): "Matching, Sorting and Wages," *Review of Economic Dynamics*, 19, 63–87.

Marchetti, D., and F. Nucci (2005): "Price Stickiness and the Contractionary Effect of Technology Shocks," *European Economic Review*, 49, 1137–1163.

Melander, O. (2009): "The Effect of Cash Flow on Investment: An Empirical Test of the Balance Sheet Channel," Sveriges Riksbank Working Paper Series No. 228.

Michelacci, C., and D. Lopez-Salido (2007): "Technology Shocks and Job Flows," *Review of Economic Studies*, 74, 1195–1227.

Mortensen, D., and C. Pissarides (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61(3), 397–415.

OECD (2014): "Indicators of Employment Protection," *Database available at: http://www.oecd.org/employment/emp/oecdindicatorsofemploymentprotection.htm.*

Pozzi, A., and F. Schivardi (2016): "Demand or Productivity: What Determines Firm Growth?," *Rand Journal of Economics*, 47, 608–630.

Ramey, V. (2011): "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126, 1–50.

Roodman, D. (2006): "How to Do xtabond2: An introduction to "Difference" and "System" GMM in Stata," Working Paper Number 103, Center for Global Development.

Roys, N. (2016): "Persistence of Shocks and the Reallocation of Labor," *Review of Economic Dynamics*, 22, 109–130.

Shimer, R. (2012): "Reassessing the Ins and Outs of Unemployment," *Review of Economic Dynamics*, 15, 127–148.

Smeets, V., and F. Warzynski (2013): "Estimating Productivity with Multi-Product Firms, Pricing Heterogeneity and the Role of International Trade," *Journal of International Economics*, 90(2), 237–244.

Syverson, C. (2011): "What Determines Productivity?," *Journal of Economic Litterature*, 49, 326–365.

Appendices - For Online Publication

A Derivation of Long-Run Restrictions

We use the stylized model presented in the paper to filter out shocks that permanently shift the firms' production functions and demand curves. To filter out the shocks of interest, we first note that the assumptions of the model ensure that the only shock that can affect the physical gross output Solow residual (A) is the technology shock. Since we only impose this restriction in the long run, we can allow for temporary variations in factor utilization and inventories.

Further, we use the standard result that a firm's optimal pricing rule under these conditions is to set the price, P_{jt}, as a constant markup $\sigma/(\sigma-1)$ over marginal cost, MC_{jt}. Marginal cost is, in optimum, equal to

$$MC_{jt} = A_{jt}^{-1}\left(\frac{W_{jt}}{\alpha}\right)^{\alpha} P_{jt}^{F}. \tag{A1}$$

Using (A1) and that $MC_{jt} = (W_{jt}N_{jt})/(\alpha Y_{jt})$ in optimum to get

$$(W_{jt}N_{jt}/Y_{jt})W_{jt}^{-\alpha} = \alpha^{1-\alpha}A_{jt}^{-1}P_{jt}^{F}. \tag{A2}$$

Thus, expression (A2) will be affected by technology and factor-price shocks but not demand shocks. It is also worth noting that any direct shocks to the firm-level wage-setting relationship (such as changes in the degree of competition over similar types of labor) will not drive this expression. Essentially, expression (A2) is a measure of unit labor cost ($W_{jt}N_{jt}/Y_{jt}$) net of wage-setting disturbances.[35] We therefore refer to the variable as wage-neutral labor cost ($WNULC_{jt}$).

Using the demand equation (2) and expression (A1), we arrive at

$$Y_{jt}W_{jt}^{\sigma\alpha} = \psi Y_t P_t^{\sigma} A_{jt}^{\sigma}\left(P_{jt}^{F}\right)^{-\sigma}\Omega_{jt}, \tag{A3}$$

where $\psi = \left(\frac{1}{\alpha}\right)^{-\sigma\alpha}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}$. Thus, expression (A3) will be driven by shocks to technology, factor prices other than labor, and demand (apart from aggregate factors that will be captured by time dummies in the empirical implementation of the model). In effect, expression (A3) is demand adjusted for wage-setting disturbances. Thus, we

[35] Note also that unit labor cost is proportional to marginal cost.

refer to it as wage-neutral demand (WND_{jt}) in the text.

B Data

The firm data-set we use is primarily drawn from Sweden's Industry Statistics Survey (IS) and contains annual information for the years 1990–2002 on inputs and output for all Swedish manufacturing plants with 10 employees or more and a sample of smaller plants. Here we focus on firms that have at least 10 employees and that we observe in a spell with at least five observations (the minimum panel dimension required for the SVAR to pass diagnostic tests).

Our measure of real output, Y_{jt}, is the value of total sales taken from the IS deflated by a firm-specific producer-price index. The firm-specific price index is a chained index with Paasche-type links that combines plant-specific unit values and detailed disaggregated producer-price indices (either at the goods level, when available, or at the most disaggregated sectoral level available). Note that when a plant-specific unit-value price is missing (e.g., when the firm introduces a new good), Statistics Sweden tries to find a price index for similar goods defined at the minimal level of aggregation (starting at four-digit goods-code level). The disaggregated sectoral producer-price indices are only used when a plausible goods-price index is unavailable.

To compute the input index (Δz_{jt}), which is necessary for the computation of the Solow residual (Δa_{jt}), real intermediate inputs (M_{jt}) are measured as the sum of costs for intermediate goods and services (including energy) collected from the IS deflated by a three-digit (SNI92/NACE) producer-price index collected by Statistics Sweden. The real capital stock (K_{jt}) is computed using a variation of the perpetual inventory method. In the first step, we calculate the forward recursion

$$K_{jt} = \max((1-\delta)K_{jt-1} + I_{jt}, BookValue_{jt}), \tag{B1}$$

where δ is sector-specific depreciation rate (two-digit SNI92/NACE) and is computed as an asset-share-weighted average between the machinery and buildings depreciation rates (collected from Melander (2009), Table 2); I_{jt} is real net investments in fixed tangible assets (computed using a two-digit SNI92/NACE sector-specific investment deflator collected from Statistics Sweden); and $BookValue_{jt}$ is the book value of fixed tangible assets taken from the Firm Statistics data base maintained by Statistics

Sweden, deflated using the same deflator as for investment. Moreover, K_{j0} is set to zero if the initial book value is missing in the data. Since, for tax reasons. the firms want to keep the book values low, we use the book values as a lower bound of the capital stock. In a second step, we then calculate the backward recursion

$$K_{jt-1} = \frac{K_{jt} - I_{jt}}{(1-\delta)}, \tag{B2}$$

where the ending point of the first recursion, K_{jT}, is used as the starting point for the second backward recursion. This is done to maximize the quality of the capital-stock series given that we lack a perfectly reliable starting point and the time dimension is small. The labor input (i.e., number of employees) is taken from the IS. To compute the cost shares, we also need a measure of the firms' labor cost, which is defined as total labor cost (including payroll taxes) in the IS.

When computing Δa_{jt}, we take an approach akin to the strategy outlined by Basu, Fernald, and Shapiro (2001). Thus, the C_J (i.e., the output elasticities) are treated as constants. Second, the cost shares are estimated as the time average of the cost shares for the two-digit industry to which the firm belongs (SNI92/NACE).[36] Third, to calculate the cost shares, we take total costs as approximately equal to total revenues.[37] The cost share of capital is then given by one minus the sum of the cost shares for all other factors.

Since 1996, Statistics Sweden has imputed the allocation of production across different plants within multi-plant firms. For this reason, we have explored various cuts of the data either focusing on single-plant firms throughout or use multi-plant firms before 1996 but only single-plant firms thereafter. The results are shown in Table D2 in Appendix D and discussed in the robustness section of the paper.

When computing $\Delta wnulc_{jt}$ and Δwnd_{jt}, we use C_N as the estimate of α and the measure of the firms' labor costs together with the measure of real output and labor input (all discussed above). Also, when computing Δwnd_{jt}, we set σ equal to our

[36] In the calculation we drop firm/year observations in which the (residual) capital share is below -25 percent of sales. This procedure generates reasonable aggregate cost shares, and ensures that the cost shares in all industries are positive.

[37] Using the data underlying Carlsson (2003), and relying on a no-arbitrage condition from neoclassical investment theory (also taking the tax system into account) to calculate the user cost of capital, we find that the time average $(1968 - 1993)$ for the share of economic profits in aggregate Swedish manufacturing revenues is about -0.001, thus supporting the the approximation of cost shares by revenue shares. The result of approximately zero economic profits on average is similar to findings in U.S. data; See e.g. Basu, Fernald, and Shapiro (2001) for a discussion.

estimate of 3.306. Finally, we remove 2 percent of the observations in each tail for each of the distributions of Δa_{jt}, $\Delta wnulc_{jt}$, Δwnd_{jt}, and Δy_{jt}. This has little effect on estimated coefficients, but it ensures that the SVAR passes diagnostic tests. We finally require the firm to be observed in spells of at least five years (because we are interested in the within-firm dynamics when estimating the SVAR).

In the end, we construct series for Δa_{jt}, $\Delta wnulc_{jt}$, Δwnd_{jt}, and Δy_{jt} for $7,940$ ongoing firms (observed at least during five consecutive years), over the $1991-2002$ period. All in all, this amounts to $70,077$ firm/year observations. Removing extreme tail events reduces the sample to $6,137$ firms and $53,379$ firm/year observations (in the specification with output growth as the fourth variable). For these firms we can compute the structural shocks for $41,105$ firm/years (due to lags in the model). Finally, we can match on labor flows from RAMS for $6,125$ firms and $40,451$ firm/year observations. Note that the procedure outlined above implies that changing the fourth variable in the VAR introduces small changes in the sample size.

C The SVAR

C.1 Identification

The model outlined in the paper and presented in detail in Appendix A provides a set of three equations that depend on the three structural shocks (i.e., demand, technology, and intermediate inputs). The left-hand-side variables in these equations can all be constructed from our firm-level data, and the model motivates a recursive sequence of long-run restrictions regarding the impact of the structural shocks on these variables. To extract the shocks of interest from the system, we estimate a VAR and proceed along the lines of Blanchard and Quah (1989).

Since we are interested in how other variables (such as output, prices, and employment) respond to structural shocks, we start by including these other variables as fourth variables in the system, allowing each to have a long-run effect on itself but not on the other variables in the system. These variables will thus also soak up all remaining transitory dynamics. In practice, we rotate across these variables while keeping the core system of the first three equations intact as in Ramey (2011). Parts of our analysis rely on extracting the technology and demand shocks from the system. In these exercises we use output as the fourth variable, but we also present several robustness checks showing that the results are insensitive to this choice. The VAR system, a fully interacted dynamic system of the variables, can, under standard regularity conditions, be written in a vector moving average (MA) form. Using lowercase letters for logarithms and denoting the fourth variable by θ, the MA representation of the system follows:[38]

$$\begin{bmatrix} \Delta a_{jt} \\ \Delta wnulc_{jt} \\ \Delta wnd_{jt} \\ \Delta \theta_{jt} \end{bmatrix} = \begin{bmatrix} C_{11}(L) & C_{12}(L) & C_{13}(L) & C_{14}(L) \\ C_{21}(L) & C_{22}(L) & C_{23}(L) & C_{24}(L) \\ C_{31}(L) & C_{32}(L) & C_{33}(L) & C_{34}(L) \\ C_{41}(L) & C_{42}(L) & C_{43}(L) & C_{44}(L) \end{bmatrix} \begin{bmatrix} \eta^a_{jt} \\ \eta^f_{jt} \\ \eta^\omega_{jt} \\ \eta^\theta_{jt} \end{bmatrix}. \tag{C1}$$

We assume that the shocks ($[\eta^a_{jt}, \eta^f_{jt}, \eta^\omega_{jt}, \eta^\theta_{jt}]$) are structural innovations and hence mutually orthogonal and serially uncorrelated. Because the shock associated with the fourth variable lacks a theoretical interpretation, we refer to it as the "residual" shock in what follows. The terms $C_{rc}(L)$ are polynomials in the lag operator, L, with coefficients $c_{rc}(k)L^k$ at each lag k. The shocks are orthogonal, and using a standard normalization we get $E\boldsymbol{\eta}_t'\boldsymbol{\eta}_t = \mathbf{I}_t$, where $\boldsymbol{\eta}_t = [\eta^a_{jt}, \eta^f_{jt}, \eta^\omega_{jt}, \eta^\theta_{jt}]'$.

Following standard practice, we denote the elements of the matrix of *long-run* multipliers corresponding to (C1) as $C_{rc}(1)$. Relying on the model outlined above, we know that the technology shock, η^a_{jt}, is the only shock with a long-run impact on a_{jt}, so $C_{12}(1) = C_{13}(1) = C_{14}(1) = 0$ in the matrix of long-run multipliers.[39] Similarly, only the technology and the factor-price shocks have a long-run effect on $wnulc_{jt}$, so $C_{23}(1) = C_{24}(1) = 0$. Finally, since the residual shock has no long-run effects on wage-neutral demand, it follows that $C_{34}(1) = 0$.

Given these assumptions, we can recover the time series of the firm's structural shocks $\boldsymbol{\eta}_{jt}$ from an estimate of the VAR(p) formulation of the system (C1), i.e., from

$$\Delta \mathbf{x}_t = \textstyle\sum_1^P \mathbf{A}_p \Delta \mathbf{x}_{t-p} + \mathbf{e}_t, \tag{C2}$$

where $\mathbf{A}_p$ denotes the matrices with coefficients, $\Delta \mathbf{x}_t = [\Delta a_{jt}, \Delta wnulc_{jt}, \Delta wnd_{jt}, \Delta \theta_{jt}]'$, $\mathbf{e}_t$ is a vector of reduced-form disturbances, and we have suppressed constants to save on notation.

Under standard regularity conditions, there exists a VAR representation of the

[38] Note that the assumed functional form of the processes for demand and technology shifters specified in equations (3) and (4) directly leads to equation (C1).

[39] That is, the coefficients $c_{12}(k)$ are such that $\sum_{k=0}^{\infty} c_{12}(k) = 0$, and similarly for the coefficients $c_{13}(k)$ and $c_{14}(k)$.

MA representation (C1) of the form

$$\mathbf{x}_t = \mathbf{A}(L)L\mathbf{x}_t + \mathbf{e}_t, \tag{C3}$$

where $\mathbf{x}_t = [\Delta a_{it}, \Delta wnulc_{jt}, \Delta wnd_{jt}, \Delta\theta_{jt}]$, $A_{rc}(L) = \sum_{k=0}^{\infty} a_{rc}(k)L^k$ and $\mathbf{e}_t$ is a vector of reduced-form errors. Since the errors in the VAR, $\mathbf{e}_t$, are one-step-ahead forecast errors, we will have that

$$\mathbf{e}_t = \mathbf{c}(0)\boldsymbol{\eta}_t, \tag{C4}$$

where $\mathbf{c}(0)$ is the matrix of $c_{rc}(0)$ coefficients from the MA representation and $\boldsymbol{\eta}_t = [\eta^a_{jt}, \eta^f_{jt}, \eta^\omega_{jt}, \eta^\theta_{jt}]'$. Thus, if the 16 coefficients in $\mathbf{c}(0)$ were known, we could recover $\boldsymbol{\eta}_t$.

In practice, we first use that $E\boldsymbol{\eta}_t\boldsymbol{\eta}_t' = \mathbf{I}_t$ together with an estimate of $\Omega = E\mathbf{e}_t'\mathbf{e}_t$ from our estimates of equation (C3) to obtain 10 restrictions. In addition, we impose the 6 long-run restrictions. Finally, rewriting equation (C3), we can obtain the MA form by using equation (C4) in terms of coefficients from equation (C3) and the $\mathbf{c}(0)$ coefficients as

$$\mathbf{x}_t = [I - \mathbf{A}(L)L]^{-1}\mathbf{c}(0)\boldsymbol{\eta}_t. \tag{C5}$$

Then, our 6 long-run restrictions imply an equal number of restrictions on the matrix $[I - \mathbf{A}(L)L]^{-1}\mathbf{c}(0)$, that together with an estimate of (C3) yields 6 additional restrictions on $\mathbf{c}(0)$. Jointly, these 16 restrictions provide an estimate of the $\mathbf{c}(0)$ matrix, $\hat{\mathbf{c}}(0)$, and using these we can solve for the structural shocks using equation (C4):

$$\hat{\mathbf{c}}(0)^{-1}\hat{\mathbf{e}}_t = \hat{\boldsymbol{\eta}}_t. \tag{C6}$$

When deriving results in term of elasticities, and to obtain an estimate of the standard deviation of the structural shocks, we use a re-normalized $\hat{\mathbf{c}}(0)$ where each element is divided by its column diagonal element.

C.2 Estimation

To derive the shocks of interest, we estimate a SVAR on the three variables defined in Table 1: Δa_{jt}, $\Delta wnulc_{jt}, \Delta wnd_{jt}$, which are constructed in order to provide us with the recursive set of long-run restrictions we need to identify the structural shocks, and a fourth residual variable (which will be output, Δy_{jt}, unless otherwise noted) which will soak up any remaining residual transitory dynamics. In practice, we first estimate four reduced-form equations where Δa_{jt}, $\Delta wnulc_{jt}$, Δwnd_{jt}, and the residual variable are explained by two lags of all four variables. We then invoke the long-run restrictions (including the long-run independence of the core system to the fourth residual shock) to derive the impulse responses of the structural shocks. Details regarding identification and estimation are found in Appendix C.

The specification includes firm-specific fixed effects to capture the drift terms of equations (3) and (4) as well as year dummies to capture aggregate shocks shared by different firms within the manufacturing sector, hence allowing us to concentrate on idiosyncratic disturbances. As a robustness check, we also use specifications accounting for sector-specific year dummies.

We use dynamic panel data methods building on Arellano and Bond (1991) for estimation because the asymptotic properties of the estimator rely on the cross-sectional dimension. This is a very useful feature in the current context of a large N ($6,137$ firms), but short T (12 years) panel because the identification of structural shocks with long-run restrictions crucially hinges on the quality of the estimated reduced-form coefficients and covariance matrix.

Table C1 shows descriptive statistics of the structural shocks derived for our baseline sample and specification. The standard deviation of the demand shock is about 60 percent larger than the technology shock (16.2 and 10.1, respectively). Appendix C depicts the shock distributions in graphs and also shows impulse responses and variance decompositions related to the main SVAR model. In addition, the appendix discusses specification tests.

Two particular results are relevant for the analysis ahead. First, we find a fairly limited amount of dynamics, in particular in the Solow residual. The main reason for this finding is that the Solow residual is defined in physical gross terms and much of the dynamics in standard measures of Solow residuals appear to be due to the dynamics of idiosyncratic prices (see Carlsson and Nordström-Skans 2012, for direct evidence on relative-price dynamics). Second, shocks to the residual fourth variable explain little of the variance in our key variables at all horizons. Since the VAR model is estimated conditional on time dummies, this finding is in line with the result of Franco and Philippon (2007), which shows that transitory shocks, although highly correlated across firms (and therefore of macroeconomic importance), matter only marginally at the firm level.

Table C1: Demand and Technology Shocks

	(1) Mean	(2) Sd.	(3) p(25)	(4) p(75)	(5) Firms	(6) Observations
Technology (η^a)	–	0.101	-0.056	0.058	6,137	41,105
Demand (η^ω)	–	0.162	-0.086	0.085	6,137	41,105

Note: p(N) denotes the Nth percentile of the data.

C.3 Impulse Responses, Variance Decompositions, and Tests

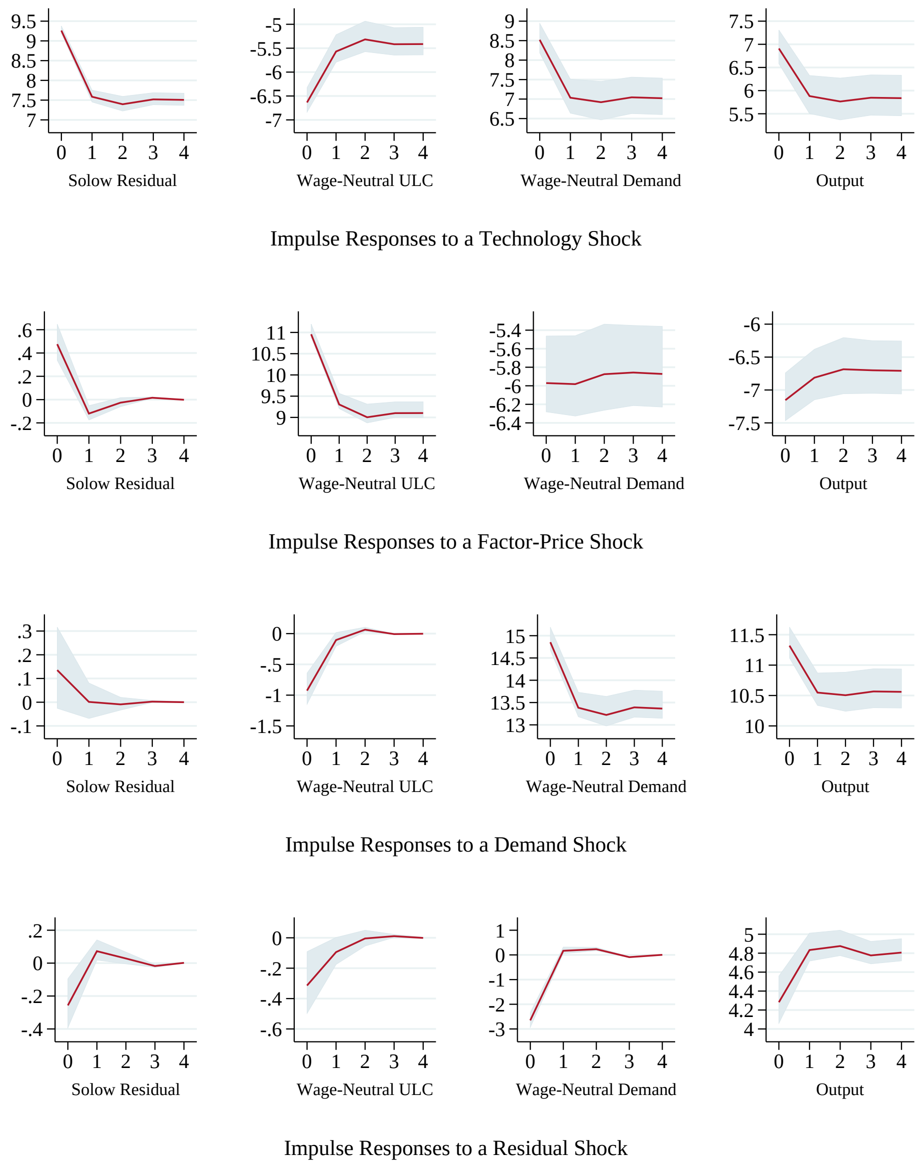
Relying on the Arellano and Bond (1991) autocorrelation test of the differenced residual, two lags in the VAR are enough to remove any autocorrelation in the residuals in all four equations. Here we rely on the two-step Arellano and Bond (1991) difference estimator, using the second to the fourth lag levels as instruments. It is worth noting, though, that the parameter estimates are not sensitive to the actual choice of where to cut the instrument set. The results are also insensitive to the inclusion of more lags as instruments. As an additional precaution, we collapse the instrument set to avoid overfitting. That is, we impose the restriction that the relationships in the "first stage" are the same across all time periods (see Roodman, 2006, for a discussion). For all specifications, the Hansen test of the overidentifying restrictions cannot reject the null of a correct specification and valid instruments.

Figure C1 shows the impulse responses of each of the variables in the baseline VAR in levels to each of the structural shocks. Since the estimated system converges fairly rapidly, we only plot the initial five periods. All impulse responses are precisely estimated as indicated by the tight (95 percent) confidence bands based on $1,000$ bootstrap replications. The high level of precision is not surprising, given that we estimate the impulse responses on a much larger sample than is common in macroeconomic applications.

Unfortunately, we have not been able to find any statistical tests of stationarity that are suitable for a setting with a short but wide panel. However, it should be clear from Figure C1 that this issue is of little importance in the current setting. Importantly, the figure is expressed in log-levels, and the flat, non-zero-end-segments in the responses imply that shocks do have permanent effects on the levels of the series (i.e., the levels are I(1)) and that the differenced series are stationary (I(0)).

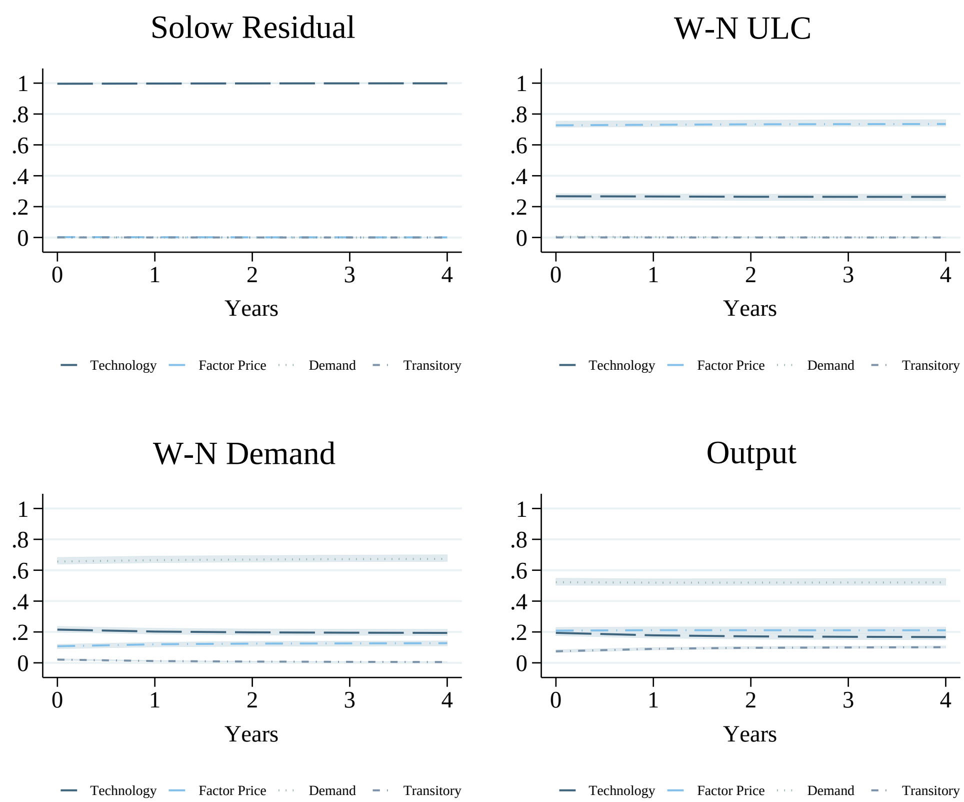
The first row of Figure C1 traces out the impulse responses of the Solow residual, the $wnulc$, the wnd, and output to a 1 sd. technology shock, η^a_{jt}. Technology shocks

Figure C1: Impulse Responses



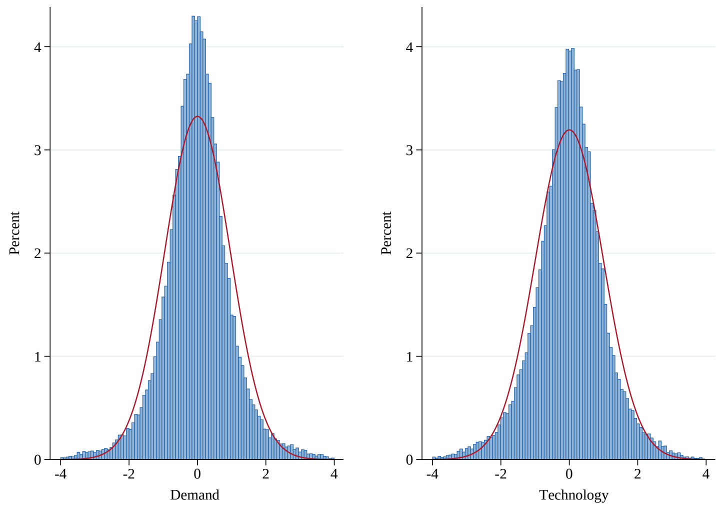
Note: Impulse responses of the Solow residual, wage neutral unit labor costs (wage neutral ULC), wage-neutral demand and output in the baseline VAR to a 1 sd. shock in percentage points. Each line depicts the mean of the bootstrap distributions. Shaded areas depict the bootstrapped 95 percent confidence intervals calculated from 1,000 replications.

Figure C2: Variance Decompositions



Note: Forecast-error variance decompositions of the VAR in levels. W-N Demand denotes wage-neutral demand. W-N ULC denotes wage-neutral unit labor costs.The left-most panel shows the percentage of the forecast-error variance in the Solow residual that can be explained by each structural shock at different horizons. Each line depicts the mean of the bootstrap distributions. Shaded areas depict the bootstrapped 95 percent confidence intervals calculated from 1,000 replications.

Figure C3: The Distribution of Demand and Technology Shocks



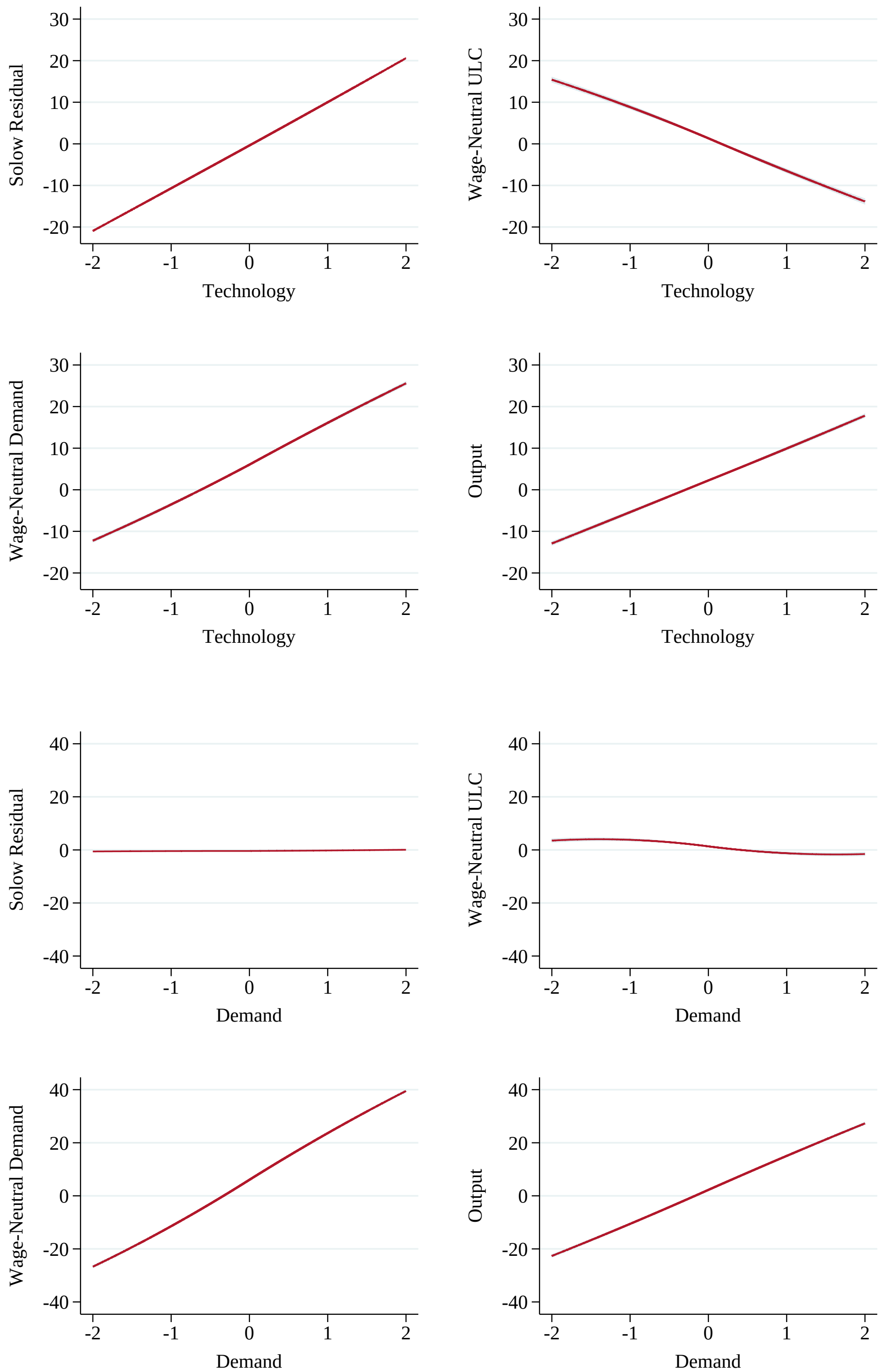
Note: Histograms of demand and technology shocks. Distributions normalized to a have unit standard deviation. Dashed lines depict a normal distribution.

have a positive permanent effect on the Solow residual: a "normal" (i.e., 1 sd) shock increases the Solow residual slightly less than 10 percent in the long run. The estimated VAR model does not impose any restrictions on how technology shocks affect *wnulc* and *wnd*. However, the results do concur with predictions from expression (A2) in the sense that *wnulc* falls permanently in response to the (permanent) technology shock. Similarly, we find that a permanent technology shock raises *wnd*, as predicted from expression (A3).

The second row in Figure C1 reports the impulse responses to a 1 sd. permanent factor-price shock. A "normal" factor-price shock increases *wnulc* and lowers *wnd* permanently (theoretically working through marginal cost, price setting, and demand). The latter result is, again, an unconstrained result in line with predictions from expression (A2). By the same logic, output also falls permanently in response to a factor-price shock. The Solow residual is affected in the very short run by factor-price shocks but converges to the long-run restriction fairly rapidly.

The impulse responses to a permanent demand shock are shown in the third row of Figure C1. In this case, *wnd* is permanently increased in response to a permanent demand shock. In the short run, demand shocks increase the Solow residual and reduce *wnulc*. As expected, a demand shock also has permanently positive effects on

Figure C4: Non-Linear Responses to a Technology Shock



Note: Contemporaneous response of variables included in the baseline VAR in percentage units as a (non-linear) function of an x sd. technology or demand shock. Shaded areas depict the 95 percent confidence intervals.

output. A "normal" demand shock increases it by about 10 percent in the long run. For completeness, Figure C1 also reports the responses to the residual shock in the last row. A "normal" residual shock raises output permanently by slightly more than 5 percent.

Figure C2 presents forecast error variance decompositions for each of the variables in the VAR in levels, decomposing the movements of the three variables. Again, bootstrapped confidence bands are extremely tight. Quantitatively, the Solow residual is solely driven by technology shocks on all horizons. The *wnulc* is mostly driven by factor-price shocks (75 percent of the variation) and partly by technology shocks (25 percent). Demand shocks explain about 65 percent of the movements in *wnd*, whereas factor-price shocks explain about 20 percent. We also see in Figure C2 that there is a role for technology shocks in explaining wage-neutral demand movements, accounting for about 15 percent. For output, we see that about 55 percent of the variation is driven by demand shocks, the rest being explained by factor-price shocks (about 20 percent), technology (about 15 percent), and the residual shock (about 10 percent).

Overall though, we find the residual shock to be of little importance. Given that we include time dummies in the VAR, this finding is in line with the results of Franco and Philippon (2007), which finds that transitory shocks are not very important on the firm level but account for most of the volatility of aggregates because they are correlated across firms.

Figure C3 shows the distributions for extracted innovations to technology and demand. As the two panels of the figure show, neither the demand nor the technology shock distributions are particularly skewed (skewness coefficients of -0.02 and -0.14, respectively), whereas both are leptokurtic (kurtosis coefficients of 5.85 and 4.25). This is also clearly visible in the graphs where the dashed line depicts a normal distribution, and a standard skewness/kurtosis test (D'Agostino, Belanger, and D'Agostino, 1990) rejects the null of normality for both distributions (p-value of 0.00 in both cases). The shock distributions depicted in Figure C3 are normalized to have a unit standard deviation. When re-normalizing the system (see Appendix A), we find that the standard deviation of the demand shock is about 35 percent larger than the technology shock (standard deviations of 16.02 and 11.86 percentage points, respectively).

A maintained assumption in the analysis is that the baseline VAR is linear in the structural shocks. In Figure C4 we plot the predicted contemporaneous responses of the variables included in the VAR as (possibly non-linear) functions of structural shocks (allowing for a separate second-order polynomial above and below zero). As the graphs show, the results do support the maintained linearity assumption.

C.4 Validation

Because the shocks we are analyzing are idiosyncratic, we cannot use correlations with known aggregate shocks such as oil-price or exchange-rate movements to cross-validate their interpretation, at least not without strong priors regarding differences between firms in the sensitivity to these aggregate shocks. Instead, we perform two alternative corroboration exercises.

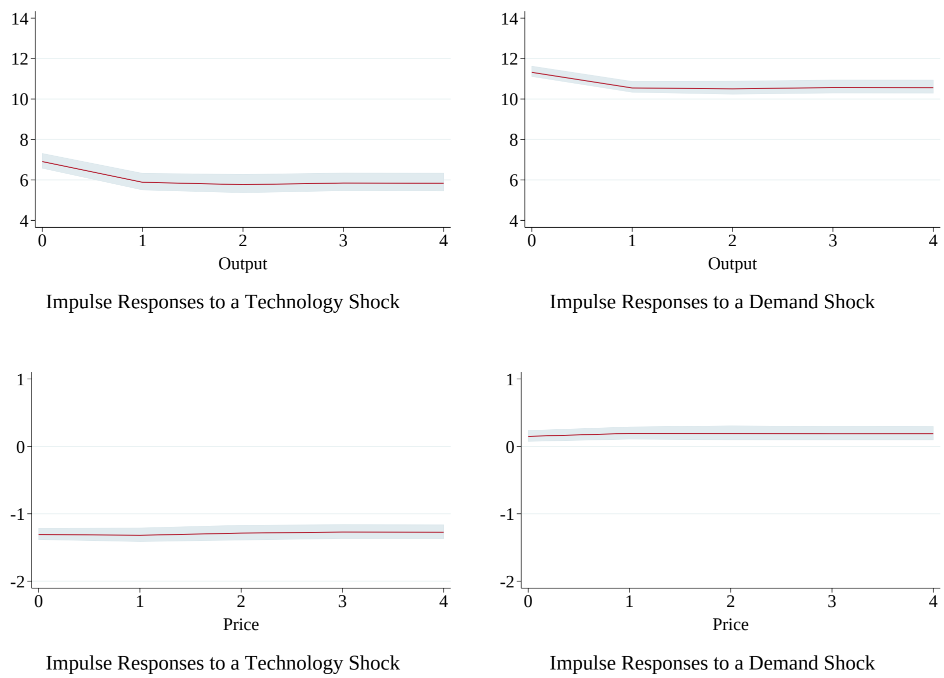
A first piece of evidence supporting our interpretation of the shocks is presented in Appendix C, which shows theory-consistent signed impulse responses for the three unrestricted responses within the VAR system: The estimated response of $\Delta wnulc_{jt}$ to a technology shock is negative, as predicted from the theoretical model. Similarly, the estimated responses from both technology shocks and factor prices on Δwnd_{jt} are negative.

A second piece of evidence comes from relating the permanent shocks to the firm-specific price index and to output. If technology shocks only affect the cost of production, we should expect technology shocks to reduce prices since firms would need to set lower prices in order to increase their sales along a fixed demand curve. In contrast, demand shocks, defined as shifts in the firm-specific demand curve, allow the firm to sell more at a given price. This suggests that prices should remain unchanged or increase under reasonably pricing strategies.

Hence, economic theory suggests that both technology and demand shocks should affect output, whereas prices should fall if the output increase is due to a technology shock (but not if it is due to a demand shock). To assess these general predictions, we reestimate the SVAR and compare responses of output and prices to the two shocks (using output and prices, in turn, as the fourth variable in the SVAR system).

Figure C5 shows the impulse responses of output and idiosyncratic prices to technology and demand shocks—indicating that both types of shocks are important for firm-level aggregates. The figure also clearly validates the general predictions discussed above: A 1 standard deviation (sd.) technology shock increases output by 6 percent in the long run. In the case of a 1 sd. demand shock, output rises by 10 percent. Moreover, as expected, prices go down in the case of a technology shock. In contrast, prices *increase* slightly when the demand curve shifts. In our view, the finding that the demand shock permanently changes output without lowering relative prices strongly supports the interpretation of the demand shock as an idiosyncratic shift in the demand curve. Note that these results are not imposed from the construction of our variables: in particular, prices could well (from a pure measurement standpoint) respond in either direction to structural innovations in both technology and demand.

Figure C5: Output and Price Responses



Note: Impulse responses to a 1 sd. shock expressed in percentage points. Lines depict the mean of the bootstrap distributions. Shaded areas depict the bootstrapped 95-percent confidence intervals calculated from 1000 replications.

C.5 Non-Constant Returns to Scale

The model can be easily extended to accommodate non-constant returns to scale. Define the overall returns to scale as $\lambda = \alpha + \beta + \gamma$. Notice that under non-constant returns to scale, it is straightforward to show that the measurement of the variables in the system of equations needs to be changed to those of Table C2 to retain the recursive form of the long-run impact of the structural shocks. Also note that the cost share of a factor will equal the output elasticity divided by the overall returns to scale in optimum, which we use in the empirical implementation provided in columns 2 and 3 of Table 6 in the main text.

Table C2: Summary of Structural System (Non-Constant Returns)

Variables:	Measured as:
$Solow:$	$Y_{jt}\left(N_{jt}^{\alpha}K_{jt}^{\beta}M_{jt}^{\gamma}\right)^{-1}$
$WNULC:$	$\left(W_{jt}N_{jt}/Y_{jt}\right)W_{jt}^{-\frac{\alpha}{\lambda}}Y_{jt}^{\left(1-\frac{1}{\lambda}\right)}$
$WND:$	$Y_{jt}^{\left(1+\sigma\left(\frac{1}{\lambda}-1\right)\right)}W_{jt}^{\sigma\frac{\alpha}{\lambda}}$

C.6 Non-constant demand elasticities

In this appendix section we show that the three identifying long-run restrictions can be reconciled with our observed impulse responses. As shown in the paper, the theoretical long-run predictions for the response of prices, output and employment to technology and demand shocks under the constant-σ assumption are given by

$$\mathbf{J}^{T}=\begin{bmatrix}\frac{\partial\ln P_{jt}}{\partial\ln A_{jt}} & \frac{\partial\ln P_{jt}}{\partial\ln\Omega_{jt}}\\ \frac{\partial\ln Y_{jt}}{\partial\ln A_{jt}} & \frac{\partial\ln Y_{jt}}{\partial\ln\Omega_{jt}}\\ \frac{\partial\ln N_{jt}}{\partial\ln A_{jt}} & \frac{\partial\ln N_{jt}}{\partial\ln\Omega_{jt}}\end{bmatrix}=\begin{bmatrix}-1 & 0\\ \sigma & 1\\ (\sigma-1) & 1\end{bmatrix}, \tag{C7}$$

which implies a proportionality factor in the employment responses to technology and demand shocks of $\frac{1}{\sigma-1}$. The corresponding empirical Jacobian, derived from the implied elasticities associated with the impulse responses presented in Figures C5 and 1 in the paper, is

$$\mathbf{J}^{E}=\begin{bmatrix}\underset{(0.008)}{-0.215} & \underset{(0.003)}{0.015}\\ \underset{(0.010)}{0.637} & \underset{(0.007)}{0.711}\\ \underset{(0.021)}{0.050} & \underset{(0.019)}{0.393}\end{bmatrix}, \tag{C8}$$

with robust standard errors presented in parenthesis. Thus, one may be tempted to use the Jacobians J^T and J^E to derive an estimate of the structural parameter σ by either looking directly at the response of output to the technology shock, or by evaluating the relative impact of technology and demand shocks. However, as we note in the text, the constant-σ model is deliberately stylized to facilitate a reduced form identification, and not designed to provide the basis for structural estimation of the parameters of the model. Specifically, since small departures from this original model that retain the same identification restrictions would lead to very different interpretations of the structural parameters. Assuming, as in Section C.6 of the paper, that the elasticity of demand (and thereby the markup) may be affected by technology and demand shocks, i.e.

$$Y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\sigma(A_{jt},\Omega_{jt})} Y_t\Omega_{jt},\ \ \sigma(A_{jt},\Omega_{jt}) > 1 \text{ and } \sigma(\bar{A}_{jt},\bar{\Omega}_{jt}) = \sigma, \tag{C9}$$

where a bar denotes an average across firms, long-run restrictions remain the same as those imposed in the baseline model.[40] However, moving beyond the constant-σ assumption dramatically changes the structural interpretation of the results, as shown in Section C.6 of the paper we get the following Jacobian

$$\mathbf{J}^{TExtended} = \begin{bmatrix} \left(\frac{-\sigma'_{A_{jt}}}{\sigma(\sigma-1)} - 1\right) & \frac{-\sigma'_{\Omega_{jt}}}{\sigma(\sigma-1)} \\ \left(\frac{\sigma'_{A_{jt}}}{(\sigma-1)} + \sigma\right) & \frac{\sigma'_{\Omega_{jt}}}{(\sigma-1)} + 1 \\ \left(\frac{\sigma'_{A_{jt}}}{(\sigma-1)} + \sigma - 1\right) & \frac{\sigma'_{\Omega_{jt}}}{(\sigma-1)} + 1 \end{bmatrix}, \tag{C10}$$

where $\sigma'_{A_{jt}}$ and $\sigma'_{\Omega_{jt}}$ denote the derivatives of $\sigma(A_{jt},\Omega_{jt})$ with respect to A_{jt} and Ω_{jt}, respectively. Interestingly, imposing $\sigma = 3.306$ (as is the baseline in the paper) and minimizing a loss function (akin to how overidentification is handled in Generalized Method of Moments estimation) in terms of the sum of the squares of the six elements of $\left[\mathbf{J}^{TExtended} - \mathbf{J}^E\right]$, weighted by the inverse of the standard deviation of the respective element in J^E, with respect to $\sigma'_{A_{jt}}$ and $\sigma'_{\Omega_{jt}}$, yields $\sigma'_{A_{jt}} = -5.857$ and $\sigma'_{\Omega_{jt}} = -0.763$. The implied reduction in σ from a 1 standard deviation technology shock is moderate (calculated as the derivative times the standard deviation, i.e. $-5.857 * 0.101 = -0.592$) and even smaller in the case of a demand shock ($-0.763 * 0.162 = -0.124$). In fact, both are tiny compared to the variations of σ (from 1.1 to 10) considered in the main text as robustness exercises. Interestingly,

[40] Moreover, as discussed in the main text, our estimation strategy provide employment responses to technology and demand shocks that are insensitive to large variations in estimated values of σ. Thus, treating σ as constant or not will be irrelevant for the main results for all reasonable variations in σ.

the effect on the markup from a technology (demand) shock equals $-1/(1-\sigma)^2$ times the derivative $\sigma'_{A_{jt}}$ ($\sigma'_{\Omega_{jt}}$), which implies that the firm increases the markup slightly in response to a 1 standard deviation technology shock, 0.111 whereas the markup response to a 1 standard deviation demand shock is very small (0.023). These results are thus in line with the standard "smoothed-off kinked" demand-curve interpretation suggested by Kimball (1995). More importantly, computing the elements in $J^{TExtended}$ using $\sigma = 3.306$, $\sigma'_{A_{jt}} = -5.857$ and $\sigma'_{\Omega_{jt}} = -0.763$ gives

$$\begin{bmatrix} -0.232 & 0.100 \\ 0.766 & 0.669 \\ -0.234 & 0.669 \end{bmatrix}, \tag{C11}$$

which is well in line with the estimated responses of prices, output and employment to the two shocks $\left(\mathbf{J}^E\right)$. It should be noted that alternative permutations of the original model may be consistent with the results. The main point of this exercise is to show that the identifying assumptions we rely on are consistent with the responses we observe. Obviously, it would be straightforward to decrease the distance for any particular (set of) element(s) within the matrix by giving it a higher relative weight when minimizing the loss function.

D Appendix Tables

Table D1: Contemporaneous and Long-Run Effect on Log Employment - Different Values of σ and Sectoral Dynamics

	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline ($\sigma = 3.3$)	$\sigma = 1.1$	$\sigma = 10$	σ by sector	σ by sector	Sectoral Dynamics
			SHORT RUN			
η^a	0.153	0.207	0.259	0.192	0.147	0.126
	(0.159)	(0.158)	(0.152)	(0.161)	(0.162)	(0.162)
η^ω	5.986**	6.591**	4.060**	5.693**	5.520**	5.506**
	(0.233)	(0.240)	(0.198)	(0.221)	(0.222)	(0.225)
Observations	40,451	41,046	39,207	40,214	39,580	39,580
Firms	6,125	6,189	5,998	6,102	5,997	5,997
			LONG RUN			
η^a	0.504*	0.512*	0.643**	0.599**	0.510*	0.490*
	(0.214)	(0.208)	(0.220)	(0.214)	(0.217)	(0.221)
η^ω	6.357**	7.009**	4.267**	5.996**	5.811**	5.737**
	(0.310)	(0.312)	0.291	(0.302)	(0.306)	(0.302)
Observations	34,414	34,612	33,291	34,198	33,667	33,667
Firms	6,116	6,181	5,991	6,094	5,989	5,989
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Sectoral Sigma	No	No	No	Yes	Yes	Yes
Pooled Dynamics	Yes	Yes	Yes	Yes	Yes	No
Sector by Time FE	No	No	No	No	Yes	Yes
Sd. η^a	10.06	10.16	9.98	10.03	9.94	9.88
Sd. η^ω	16.18	13.87	27.23	17.09	16.98	16.80

Note: Columns (2) and (3) impose large variation in values of σ. Column (4), (5) and (6) allow for a sectoral σ (for sufficiently large two-digit indstries). Column (4) retains joint time dummies. Column (5) lets the time dummies be sector specific. Column (6) reestimates the entire SVAR for each two-digit industry. All estimates are the effect of a 1 sd. shock. Robust standard errors in parentheses. Long-run estimates are the sum of the contemporaneous impact and one lag. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

Table D2: Contemporaneous and Long-Run Effect on Log Employment - Sample Variations

Sample	(1) Baseline	(2) Single Plant Always	(3) Single Plant After 1996	(4) $\leq \pm 2$ Sd. Shocks
		SHORT RUN		
η^a	0.153	0.421**	0.312*	0.040
	(0.159)	(0.158)	(0.151)	(0.164)
η^ω	5.986**	5.500**	6.244**	6.317**
	(0.233)	(0.236)	(0.238)	(0.205)
Observations	40,451	20,877	30,234	36,072
Firms	6,125	3,246	5,259	6,111
		LONG RUN		
η^a	0.504*	0.534*	0.669**	0.336
	(0.214)	(0.233)	(0.215)	(0.234)
η^ω	6.357**	5.715**	6.657**	6.397**
	(0.310)	(0.309)	(0.326)	(0.294)
Observations	34,414	17,638	25,040	30,693
Firms	6,116	3,246	5,250	6,066
sd. η^a	10.06	9.13	9.41	10.06
sd. η^ω	16.18	15.07	14.79	16.18

Note: Column (2) restricts the sample to single-plant firms; column (3) includes a mixed sample with multi-plant firms until 1996, but not thereafter; column (4) shows results for a trimmed sample where we focus on shocks in the Lester range of $\pm$ 2 standard deviations. Estimates are the effects of a 1 sd. shock. Robust standard errors in parentheses. Regression includes firm fixed effects and time dummies. Long-run estimates are the sum of the contemporaneous impact and one lag. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

Table D3: Contemporaneous and Long-Run Effect on Log Employment - Varying the Fourth Variable in the VAR

	(1)	(2)	(3)	(4)
Fourth Variable of VAR:	Output	Sales per Worker	Employment (IS)	Employment (RAMS)
			SHORT RUN	
η^a	0.153	0.524**	0.263	0.499**
	(0.159)	(0.154)	(0.143)	(0.061)
η^ω	5.986**	5.840**	5.261**	6.986**
	(0.233)	(0.234)	(0.212)	(0.086)
Observations	$41,105$	$40,284$	$38,213$	$37,234$
Firms	$6,125$	$6,113$	$5,879$	$5,703$
			LONG RUN	
η^a	0.504*	0.812**	0.644**	0.643**
	(0.214)	(0.218)	(0.209)	(0.097)
η^ω	6.357**	6.134**	5.477**	7.514**
	(0.310)	(0.317)	(0.266)	(0.121)
Observations	$34,414$	$34,260$	$32,407$	$31,531$
Firms	$6,116$	$6,102$	$5,871$	$5,703$
sd. η^a	10.06	9.980	9.971	9.964
sd. η^ω	16.18	16.39	15.35	15.13

Note: Column (2) derives shocks from a VAR in which the fourth variable is sales per worker; in column (3) the fourth variable is annual employment measured in the IS data set; in column (4) the fourth variable is end-of-the-year employment measured from the RAMS data-set.All estimates are the effect of a 1 sd. shock. Robust standard errors in parentheses. All regressions include time dummies and firm fixed effects. Long-run estimates are the sum of the contemporaneous impact and one lag. ** and * denote statistical significance at the 1 and 5 percent levels, respectively.

Table D4: Contemporaneous and Long-Run Effect on Net Employment Growth - Two Period Specifications.

	(1) One Period – Baseline	(2) Two Period – Without Exits	(3) Two Period – With Exits
		SHORT RUN	
η^a	0.115	0.328*	0.278
	(0.119)	(0.153)	(0.164)
η^ω	5.609**	5.431**	5.749**
	(0.173)	(0.375)	(0.376)
Observations	40,451	39,822	40,238
Firms	6,125	6,114	6,121
		LONG RUN	
η^a	0.412*	0.420	0.380
	(0.163)	(0.350)	(0.368)
η^ω	6.009**	4.112**	4.696**
	(0.228)	(0.391)	(0.422)
Observations	34,414	33,830	34,243
Firms	6,116	6,099	6,110

Note: The dependent variable in column (1) is the employment change between t and $t-1$ divided by the average employment in the two years. In columns (2) and (3) the dependent variable is defined as the employment change between $t+1$ and $t-1$ divided by the average employment in the two years. Columns (1) and (2) exclude frims that exit the sample in the calculation of the flows, and column (3) includes them. The reported coefficients are the effect of 1 sd. shock. Robust standard errors in parentheses. Regression includes firm fixed effects and time dummies. Long-run estimates are the sum of the contemporaneous impact and one lag. ** and * denote statistical significance at the 1 and 5 levels, respectively.