Digitalization, Factor-Biased Productivity, and Market Power*

Abiodun M. Musbaudeen † Subal C. Kumbhakar‡

October 28, 2026

Click here for the latest version

Abstract

Digitalization can improve labor efficiency, raise overall input efficiency, or increase firm markups. To disentangle these three channels, we develop a sequential production and markup estimator. Technology is first identified from quantity-side moments without imposing a specific model of competition. Markups are then recovered from revenue shares. We apply this framework to Spanish manufacturing firms from 2000 to 2019. Our central finding is that accounting for market power significantly changes the estimated composition of digital returns. On a common sample under perfect competition, a one-point increase in digitalization raises overall input efficiency by 6.5% and labor efficiency by 4.8%. Under imperfect competition, these responses shift to 4.3% and 5.5%, while markups rise by 0.7%. Paired bootstrap tests confirm a significant drop in the overall efficiency response and a 3.0 percentage-point shift toward labor efficiency. Two patterns further shape these gains. First, labor-efficiency gains are significantly larger for small firms, even though large firms adopt more digital technology. Second, among distinct digital technologies, physical automation is the sole component that significantly increases markups. Ignoring market power and technology differences thus distorts the sources of digital productivity growth.

Keywords: Digitalization, Factor-Biased Productivity, Labor-Augmenting Productivity, Markups, Market Power, Spanish Manufacturing

JEL Classification: D24, L13, L60, O33, J23, C33

*We thank the Department of Economics at Binghamton University for funding the dataset used in this paper. We are also grateful to Shunan Zhao, Ivan Korolev, Chris Hanes, Yingjun Su, and JoonHwan Cho for providing valuable feedback on an earlier version. We especially thank Jordi Jaumandreu for directing us to the dataset and for providing guidance on constructing the price indices from price variations.

†Department of Economics, State University of New York, Binghamton, NY, USA, amusbau1@binghamton.edu

‡Corresponding author, Department of Economics, State University of New York, Binghamton, NY, USA, kkar@binghamton.edu

1 Introduction

Firms invest in digital technologies expecting higher productivity. Managers increasingly describe digital adoption as a key route to operating efficiency and financial performance (PwC, 2026; Gartner, 2026). Yet a measured productivity gain can operate through three channels with different beneficiaries. Digital technology may make each hour of work more productive. This raises the marginal product of labor and alters labor demand. Alternatively, it may make all inputs more productive. This lowers unit costs and can reach buyers through lower prices. Finally, it may allow firms to raise prices relative to marginal cost, shifting surplus from buyers to firm owners. Identifying the dominant channel reveals whether digitalization is mainly a story of labor efficiency, cost reduction, or market power.

Existing studies have examined the relationship between digital adoption and firm productivity (Gal et al., 2019; Nucci et al., 2023) and documented higher markups in digitally intensive sectors (Calligaris et al., 2018). Related methodological work has developed production-based approaches to recover markups from output elasticities and revenue shares (De Loecker and Warzynski, 2012) and shown the importance of accounting for factor-biased technological change alongside Hicks-neutral efficiency (Doraszelski and Jaumandreu, 2019; Demirer, 2026). Modeling input-specific and overall efficiency jointly ensures that the output elasticities used in markup estimation reflect underlying technology rather than market power. In parallel, research on technology adoption has introduced measures covering multiple dimensions of firms' digital use, spanning physical automation, information technologies, digital human capital, and digital stakeholder interfaces (Añón Higón and Bonvin, 2024). Combining these strands allows us to examine which specific digital capabilities drive labor efficiency, overall productivity, or markups, and how these effects vary across operating environments. Two firm-level questions are particularly important. First, large firms adopt considerably more digital technology than small firms, raising the question of whether greater adoption yields larger efficiency gains. Second, given that digitally intensive industries display higher markups, we test whether further adoption raises markups within those industries.

In this study, we build on these contributions by developing a sequential estimator that identifies technology parameters before recovering markups under imperfect competition (IPC), with perfect competition (PC) as a restricted benchmark. We apply this framework to Spanish manufacturing firms in the Encuesta Sobre Estrategias Empresariales (ESEE) from 2000 to 2019 to separate labor-augmenting productivity, Hicks-neutral productivity, and markups. To unpack the mechanisms behind overall digital adoption, we construct an aggregate 13-variable digitalization index alongside four sub-indices: physical automation, information technologies, digital human capital, and stakeholder-facing technologies. This structure allows us to test how distinct digital capabilities separately affect each productivity and markup channel.

Our methodological contribution resolves a central identification challenge when revenue shares reflect both technology and market power. Under IPC, the revenue-share equation identifies only the ratio of the degree of homogeneity in labor and materials parameter, ν, to the firm's markup. Prior studies address this indeterminacy by using set identification or firm-level input price variation (Malikov et al., 2024), imposing constant returns to scale (Flynn et al., 2019), or relying on control functions (Demirer, 2026). Instead, we recover labor-augmenting productivity from the ratio of input first-order conditions. We then identify ν and the remaining technology parameters directly from the dynamic production equation combined with the Hicks-neutral law of motion, where the markup does not appear. With the technology parameters identified, revenue shares provide a clean firm-year markup signal. Assuming that productivity and markup innovations are unpredictable given period $t-1$ information, our estimates capture the effect of past digital adoption while still allowing adoption to respond endogenously to past productivity (Doraszelski and Jaumandreu, 2018).

Empirically, accounting for market power substantially changes the estimated composition of digital productivity gains. Under the restricted PC benchmark, a one-point increase in the digitalization index raises next-period Hicks-neutral productivity by 6.2% and labor-augmenting productivity by 3.7% in the overall sample. Under IPC, however, the Hicks-neutral response falls to 4.3% while the labor-augmenting response rises to 5.5%. Markups increase by a modest 0.7%, roughly one-eighth of the labor-augmenting effect. Comparing both models on an identical sample via paired bootstrap draws confirms that the gap between labor-augmenting and Hicks-neutral gains widens by a statistically significant 3.0 percentage points under IPC, driven primarily by the downward revision in Hicks-neutral productivity.

Our results also speak directly to the heterogeneity of digital returns across firms, industries, and technologies. First, across both specifications, the labor-augmenting response is significantly larger for small firms than for large firms, even though large firms are the most digitalized. Higher baseline adoption therefore, does not imply larger marginal gains in labor efficiency. Second, the markup response to digitalization is significantly smaller in high-digital-intensity industries than in low-digital-intensity industries. Finally, disaggregating the index reveals that these effects are highly technology-specific. Under IPC, digital human capital is the strongest driver of labor-augmenting productivity, raising it by 2.2% overall and 2.5% among small firms, with no effect on markups. Physical automation raises both labor-augmenting productivity by 1.1% and Hicks-neutral productivity by 1.7%, and it is the only component that significantly increases firm markups.

The rest of the paper is organized as follows. Section 2 reviews the literature. Section 3 presents the general IPC model and its PC restriction. Section 4 develops identification and estimation. Section 5 describes the data. Section 6 reports the results, and Section 7 examines robustness. Section 8 discusses implications. Section 9 concludes.

2 Related Literature

This paper relates to four literatures: the productivity effects of digitalization, factor-biased technological change, production-based markup estimation, and the relationship between digitalization and market power.

The first literature studies how information and communication technologies and broader digital adoption affect productivity. Early firm- and industry-level studies generally find positive effects of IT investment on productivity (Brynjolfsson and Hitt, 1996a, 2000; Stiroh, 2005; Draca et al., 2008). Other studies find weak or insignificant effects. Acemoglu et al. (2014) find little evidence of faster productivity growth in IT-intensive U.S. manufacturing industries outside the computer-producing sector, and studies of broadband adoption report similarly mixed results (Bertschek et al., 2015; Haller and Lyons, 2015; DeStefano et al., 2018).

More recent work uses broader measures of digital adoption. Gal et al. (2019) relate industry adoption of digital technologies to firm-level multifactor productivity in European countries, Nucci et al. (2023) use matching and difference-in-differences on Italian firms, and Cusolito (2020) study firms in developing countries. These studies measure productivity as a single residual. They therefore cannot tell whether digitalization makes labor more efficient or makes all inputs more efficient. Much of this literature also measures digitalization through a single indicator such as ICT spending or IT capital (Brynjolfsson and Hitt, 1996b; Jorgenson, 2001; Bertschek and Kaiser, 2004), an approach whose limitations earlier work documents (Black and Lynch, 2001). Following Añón Higón and Bonvin (2024), we combine 13 variables from the ESEE into an overall index and four component indices, allowing us to examine the channels of different technologies.

The second literature shows that technological change is often not Hicks-neutral. Doraszelski and Jaumandreu (2018) recover labor-augmenting productivity from the ratio of input first-order conditions and find that technological

change in Spanish manufacturing is biased toward labor. Raval (2019) and Zhang (2019) document substantial firm-level differences in labor-augmenting productivity and show that these differences matter for factor demand and the labor share. This literature generally assumes perfect competition in the output market or treats the markup as known. We extend it by allowing markups to vary across firms and over time, and by asking how digitalization shifts labor-augmenting productivity relative to Hicks-neutral productivity once market power is accounted for.

The third literature estimates markups from production data. De Loecker and Warzynski (2012) recover the markup as the ratio of a flexible input's output elasticity to its revenue expenditure share, without specifying the form of product-market competition. These estimates are only as reliable as the estimated output elasticities. Raval (2023) shows that markups estimated from labor and from materials are negatively correlated across firms and that labor-augmenting productivity differences can explain this pattern. Demirer (2026) shows that ignoring labor-augmenting productivity biases output elasticities and therefore markups, and proposes a control-function estimator.

Our estimator extends the two-productivity framework of Doraszelski and Jaumandreu (2019) and Malikov et al. (2024) by identifying the technology scale from dynamic quantity-side moments before recovering markups, instead of imposing the returns-to-scale restriction in Flynn et al. (2019) or using the price-based and control-function approaches discussed by Malikov et al. (2024) and Demirer (2026).

The fourth literature links digital technologies to market power. Aggregate markups have risen in many economies (De Loecker et al., 2020), and part of this rise has been attributed to the reallocation of activity toward large, highly productive firms (Autor et al., 2020). Calligaris et al. (2018) find that markups are higher in digitally intensive sectors across 26 countries over 2001–2014 and that the gap relative to other sectors widened over the period. These studies compare markup levels across firms, sectors, or time. We instead estimate how markups respond to additional digital adoption within firms, while allowing productivity to be factor-biased. A smaller markup response in digitally intensive industries can therefore coexist with higher markup levels in those industries.

3 General Model under Imperfect Competition

We develop a production model in which firms have market power in the output market and take input prices as given. The technology accommodates Hicks-neutral and labor-augmenting productivity without imposing constant returns to scale or a particular model of competition. An output shock is realized after input choices. We first derive technology and input conditions under IPC, then obtain PC by restricting the markup to unity.

3.1 Technology

Firm i produces gross output at time t using predetermined capital K_{it}, freely variable labor L_{it}, and freely variable materials M_{it}. The production technology is given by

$$Y_{it} = Y_{it}^{*} e^{\eta_{it}} = F(K_{it}, e^{\omega_{L,it}} L_{it}, M_{it})\, e^{\omega_{N,it}+\eta_{it}}, \tag{1}$$

where η_{it} is a productivity shock unobserved by all agents, $\omega_{L,it}$ is labor-augmenting productivity, and $\omega_{N,it}$ is Hicks-neutral productivity. Both $\omega_{N,it}$ and $\omega_{L,it}$ are known to the firm when inputs are chosen, whereas η_{it} is realized after inputs are chosen. We assume $\mathbb{E}[\eta_{it} \mid \mathcal{I}_{it}] = 0$ and define

$$\theta \equiv \mathbb{E}[e^{\eta_{it}} \mid \mathcal{I}_{it}]\,, \tag{2}$$

which is assumed to be constant across firms and over time.

We use the lower-case letters as natural logarithms of their upper-case counterparts, and write $l_{it}^* \equiv l_{it} + \omega_{L,it}$ for effective labor in logs. The firm's information set $\mathcal{I}_{it}$ contains K_{it}, $\omega_{L,it}$, $\omega_{N,it}$, all input prices, and the entire history of these variables.

We assume $\ln F(\cdot)$ to be translog:

$$\begin{aligned}\ln F(K_{it}, L_{it}^*, M_{it}) = \alpha_0 + \beta_K k_{it} + \beta_L l_{it}^* + \beta_M m_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \tfrac{1}{2}\beta_{LL}(l_{it}^*)^2 + \tfrac{1}{2}\beta_{MM} m_{it}^2 \\ + \beta_{KL} k_{it} l_{it}^* + \beta_{KM} k_{it} m_{it} + \beta_{ML} m_{it} l_{it}^*. \end{aligned} \tag{3}$$

Following Doraszelski and Jaumandreu (2019) and Malikov et al. (2024) (MZZ), we impose separability of capital from the two variable inputs, which requires $\beta_{KL} = \beta_{KM} = 0$. We further impose homogeneity of degree ν in effective labor and materials, which requires $\beta_{LL} + \beta_{ML} = 0$ and $\beta_{MM} + \beta_{ML} = 0$. Denoting $\delta \equiv \beta_{ML}$, these restrictions imply $\beta_{LL} = \beta_{MM} = -\delta$, and we define the degree of homogeneity as

$$\nu \equiv \beta_L + \beta_M. \tag{4}$$

Under these restrictions, the technology can be written as

$$y_{it}^* = \alpha_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \beta_M m_{it} + \beta_L (l_{it} + \omega_{L,it}) - \tfrac{1}{2}\delta (m_{it} - l_{it} - \omega_{L,it})^2 + \omega_{N,it}. \tag{5}$$

3.2 Output elasticities

Differentiating (5) at fixed $\omega_{L,it}$ gives the observation-specific output elasticities

$$\varepsilon_{it}^L = \beta_L + \delta(m_{it} - l_{it} - \omega_{L,it}), \tag{6}$$

$$\varepsilon_{it}^M = \beta_M - \delta(m_{it} - l_{it} - \omega_{L,it}), \tag{7}$$

$$\varepsilon_{it}^K = \beta_K + \beta_{KK} k_{it}. \tag{8}$$

Homogeneity implies that the elasticity with respect to a proportional change in labor and materials is a constant,

$$\varepsilon_{it}^L + \varepsilon_{it}^M = \nu. \tag{9}$$

3.3 Input demand, revenue shares, and markups

The general model retains the multiplicative output-shock structure introduced above. The firm chooses labor and materials before η_{it} is realized, using the expected profit maximization behavior under the conditional expectation $\theta = \mathbb{E}[e^{\eta_{it}} \mid \mathcal{I}_{it}]$.

Under imperfect competition, the firm faces inverse demand $P_{it}^Y = P_{it}^Y(Y_{it}^*, z_{it})$, where z_{it} denotes demand shifters, and profit is

$$\Pi = P_{it}^Y(Y_{it}^*, z_{it}) F(K_{it}, e^{\omega_{L,it}} L_{it}, M_{it})\, e^{\omega_{N,it} + \eta_{it}} - P_{it}^L L_{it} - P_{it}^M M_{it} \tag{10}$$

Because the multiplicative shock is realized after variable inputs are chosen, expected-profit maximization introduces the same factor θ under the maintained timing assumptions. Writing $P_{it} \equiv P_{it}^Y$ for the output price, MC_{it} for

marginal cost, and $\mu_{it} = P_{it}/MC_{it}$ for the markup, the variable-input first-order conditions are

$$\theta MC_{it}\frac{\partial Y_{it}^*}{\partial j_{it}} = P_{it}^j, \qquad j \in \{L, M\}. \tag{11}$$

Multiplying equation (11) by $j_{it}/(P_{it}Y_{it}^*)$ gives

$$\theta\frac{MC_{it}}{P_{it}}\left(\frac{\partial Y_{it}^*}{\partial j_{it}}\frac{j_{it}}{Y_{it}^*}\right) = \frac{P_{it}^j j_{it}}{P_{it}Y_{it}^*}. \tag{12}$$

Thus, the observed revenue shares are

$$s_{it}^j \equiv \frac{P_{it}^j j_{it}}{P_{it}Y_{it}} = \theta\frac{\varepsilon_{it}^j}{\mu_{it}}e^{-\eta_{it}}, \qquad j \in \{L, M\}. \tag{13}$$

Consequently, either variable input provides a noisy signal of the same markup ($\mu_{it}e^{\eta_{it}}$):

$$\frac{\varepsilon_{it}^L}{s_{it}^L} = \frac{\varepsilon_{it}^M}{s_{it}^M} = \frac{\mu_{it}}{\theta}e^{\eta_{it}}. \tag{14}$$

Adding the two observed shares gives

$$R_{it} \equiv s_{it}^L + s_{it}^M = \theta\frac{\nu}{\mu_{it}}e^{-\eta_{it}}. \tag{15}$$

Therefore,

$$RR_{it} \equiv \ln(\theta\nu) - \ln R_{it} = \ln\mu_{it} + \eta_{it}. \tag{16}$$

3.4 Variable-input cost shares and productivity inversion

The shares of labor and materials in total variable expenditure are

$$SC_{it}^L = \frac{s_{it}^L}{s_{it}^L + s_{it}^M}, \qquad SC_{it}^M = 1 - SC_{it}^L. \tag{17}$$

The ratio of the labor and materials conditions removes μ_{it} and η_{it}:

$$\frac{s_{it}^L}{s_{it}^M} = \frac{\varepsilon_{it}^L}{\varepsilon_{it}^M} = \frac{\beta_L + \delta(m_{it} - l_{it} - \omega_{L,it})}{\beta_M - \delta(m_{it} - l_{it} - \omega_{L,it})}. \tag{18}$$

Equation (18) gives

$$\omega_{L,it} = m_{it} - l_{it} + \frac{\beta_L}{\delta} - \frac{\nu}{\delta}SC_{it}^L. \tag{19}$$

Using the degree of homogeneity defined in equation (4), the output elasticities can be written as

$$\varepsilon_{it}^L = \nu SC_{it}^L, \qquad \varepsilon_{it}^M = \nu SC_{it}^M. \tag{20}$$

3.5 Concentrated production equation

Substituting the labor-augmenting productivity inversion in equation (19) into the restricted technology gives the concentrated production equation used to identify technology:

$$y_{it}^{*} = \kappa_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \nu m_{it} - \frac{\nu^2}{2\delta}(SC_{it}^{L})^2 + \omega_{N,it}. \tag{21}$$

Here κ_0 absorbs constants from the inversion; Appendix A gives the full algebra and the restricted PC estimator.

3.6 Perfect competition as a restricted case

Perfect competition imposes $\mu_{it} = 1$ on the general model. The ratio of the two variable-input first-order conditions and the labor-augmenting productivity inversion remain unchanged. The sum of revenue shares no longer contains a persistent markup component. This restricted specification provides a benchmark for measuring the consequences of ignoring market power. The full PC derivations and estimation procedure are in Appendix A.

4 Identification and Estimation under Imperfect Competition

The estimator proceeds in four stages: Stage 1 estimates technology and Hicks-neutral dynamics from the dynamic production equation; Stage 2 estimates labor-augmenting dynamics conditional on that technology; Stage 3 estimates markup dynamics; and Stage 4 predicts latent states and transitory shocks. The output-contribution comparisons use technology and productivity-transition estimates from the first two stages, while the fourth-stage decomposition requires the additional covariance restrictions in Appendix D. The restricted PC estimator is described in Appendix A.[1]

4.1 Stage 1: Hicks-neutral technology and dynamics

Combining equation (21) with $y_{it} = y_{it}^{*} + \eta_{it}$ gives

$$y_{it} = \kappa_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \nu m_{it} - \left(\frac{\nu^2}{2\delta}\right)(SC_{it}^{L})^2 + \omega_{N,it} + \eta_{it}. \tag{22}$$

Hicks-neutral productivity is assumed to follow

$$\omega_{N,it} = \pi_N + \rho_N \omega_{N,i,t-1} + \boldsymbol{\gamma}_N' \mathbf{D}_{i,t-1} + \xi_{N,it}, \qquad \mathbb{E}[\xi_{N,it} \mid \mathcal{I}_{i,t-1}] = 0. \tag{23}$$

Quasi differencing equation (22) gives

$$\begin{aligned} u_{N,it} = {} & y_{it} - \pi_0 - \rho_N y_{i,t-1} - \beta_K(k_{it} - \rho_N k_{i,t-1}) \\ & - \tfrac{1}{2}\beta_{KK}(k_{it}^2 - \rho_N k_{i,t-1}^2) - \nu(m_{it} - \rho_N m_{i,t-1}) \\ & + \left(\frac{\nu^2}{2\delta}\right)\left[(SC_{it}^{L})^2 - \rho_N (SC_{i,t-1}^{L})^2\right] - \boldsymbol{\gamma}_N' \mathbf{D}_{i,t-1}, \end{aligned} \tag{24}$$

[1] Local identification assumes valid instruments, support, and a full-rank population moment Jacobian; numerical moment counts, rank checks, and root selection do not establish global uniqueness or strong identification.

where $\pi_0 = \pi_N + (1-\rho_N)\kappa_0$ and the composite disturbance is

$$u_{N,it} = \xi_{N,it} + \eta_{it} - \rho_N \eta_{i,t-1}. \tag{25}$$

The Stage 1 parameter vector is

$$\boldsymbol{\theta}_N = (\pi_0, \beta_K, \beta_{KK}, \nu, \delta, \rho_N, \gamma_N)'. \tag{26}$$

We estimate this block by nonlinear GMM using

$$\mathbb{E}[\mathbf{Z}_{N,it} u_{N,it}] = \mathbf{0}, \tag{27}$$

where $\mathbf{Z}_{N,it}$ contains predetermined and sufficiently dated variables that are orthogonal to $\xi_{N,it}$, η_{it}, and $\eta_{i,t-1}$. In particular, the presence of $\eta_{i,t-1}$ in equation (25) requires the timing of the instruments to be chosen accordingly.

The common instrument block uses capital dated $t-1$ and its square, together with the $t-2$ labor cost share, digitalization variables, and materials. Under the stated timing assumptions, lagged capital and its square are orthogonal to both η_{it} and $\eta_{i,t-1}$ and to the current Hicks-neutral innovation $\xi_{N,it}$.

$$\mathbf{Z}^0_{N,it} = \left(1,\ \left(SC^L_{i,t-2}\right)^2,\ \mathbf{D}'_{i,t-2},\ m_{i,t-2},\ k_{i,t-1},\ k^2_{i,t-1}\right)', \tag{28}$$

and $\mathbf{Z}_{N,it}$ stacks $\mathbf{Z}^0_{N,it}$ with the sample-specific instruments in Table 1.

Table 1: Hicks-neutral productivity with sample-specific instruments

	Sample-specific instruments added to $\mathbf{Z}^0_{N,it}$	Moments / free parameters	
Sample		Composite	Four components
Overall	$R_{i,t-2}$, $ICT_{i,t-2}$	8 / 7	11 / 10
Small firms	$SC^L_{i,t-2}$, $R_{i,t-2}$	8 / 7	11 / 10
Large firms	$s^L_{i,t-2}$, ln(effective hours$_{i,t-2}$)	8 / 7	11 / 10
Low-digital industries	$R_{i,t-2}$, $ICT_{i,t-2}$, $\mathbf{1}\{\text{metal products}\}_i$	9 / 7	12 / 10
High-digital industries	$s^L_{i,t-2}$, $k_{i,t-1}m_{i,t-2}$	8 / 7	11 / 10

Notes: Every sample uses the common instruments $\mathbf{Z}^0_{N,it}$ in (28). $R_{i,t-2} = s^L_{i,t-2} + s^M_{i,t-2}$ is the lagged sum of the labor and materials expenditure shares of revenue; $s^L_{i,t-2}$ is the lagged labor expenditure share of revenue; $ICT_{i,t-2}$ is the lagged ICT index built from the IT-technology, digital-human-capital and stakeholder-facing items of the digitalization index; ln(effective hours$_{i,t-2}$) is lagged log total effective hours; $\mathbf{1}\{\text{metal products}\}_i$ indicates the metal and metal products industry within the low-digital sample. In the component specification $\mathbf{D}_{i,t-2}$ contains the four lagged component indices and the sample-specific instruments are unchanged.

After estimating the technology parameters, we construct the firm-year Hicks-neutral productivity signal

$$H_{it} \equiv y_{it} - \left[\widehat{\kappa}_0 + \widehat{\beta}_K k_{it} + \tfrac{1}{2}\widehat{\beta}_{KK} k^2_{it} + \widehat{\nu} m_{it} - \left(\frac{\widehat{\nu}^2}{2\widehat{\delta}}\right)(SC^L_{it})^2\right]. \tag{29}$$

By construction,

$$H_{it} = \omega_{N,it} + \eta_{it}. \tag{30}$$

Stage 1 identifies technology and constructs a Hicks-neutral productivity signal containing the transitory disturbance. Signal extraction in Stage 4 supplies the reported decomposition into persistent productivity and transitory noise.

4.2 Stage 2: Labor-augmenting dynamics

Stage 2 uses the Stage 1 Hicks-neutral estimates $(\widehat{\nu}, \widehat{\delta})$ as fixed inputs when estimating labor-augmenting productivity dynamics. Labor-augmenting productivity follows the controlled first-order Markov process

$$\omega_{L,it} = \rho_L \omega_{L,i,t-1} + \boldsymbol{\gamma}_L' \mathbf{D}_{i,t-1} + \xi_{L,it}, \qquad \mathbb{E}[\xi_{L,it} \mid \mathcal{I}_{i,t-1}] = 0. \tag{31}$$

Evaluating the input-condition inversion at the Stage 1 estimates of ν and δ gives

$$\omega_{L,it} = (m_{it} - l_{it}) + \frac{\beta_L}{\widehat{\delta}} - \frac{\widehat{\nu}}{\widehat{\delta}} SC_{it}^L, \tag{32}$$

Substituting this inversion at t and $t-1$ into the labor-augmenting transition, and collecting the constant $\beta_L/\widehat{\delta}$, yields the quasi-differenced estimating equation

$$\begin{aligned}(m_{it} - l_{it}) - \frac{\widehat{\nu}}{\widehat{\delta}} SC_{it}^L = & -(1-\rho_L)\frac{\beta_L}{\widehat{\delta}} + \rho_L \left[(m_{i,t-1} - l_{i,t-1}) - \frac{\widehat{\nu}}{\widehat{\delta}} SC_{i,t-1}^L\right] \\ & + \boldsymbol{\gamma}_L' \mathbf{D}_{i,t-1} + \xi_{L,it}.\end{aligned} \tag{33}$$

Solving (33) for the innovation gives the residual

$$\begin{aligned}\xi_{L,it} = & (m_{it} - l_{it}) - \rho_L (m_{i,t-1} - l_{i,t-1}) + (1-\rho_L)\frac{\beta_L}{\widehat{\delta}} \\ & - \frac{\widehat{\nu}}{\widehat{\delta}}(SC_{it}^L - \rho_L SC_{i,t-1}^L) - \boldsymbol{\gamma}_L' \mathbf{D}_{i,t-1},\end{aligned} \tag{34}$$

which depends on the parameter vector

$$\boldsymbol{\theta}_L = (\beta_L, \rho_L, \boldsymbol{\gamma}_L')'. \tag{35}$$

We use the instrument vector

$$\mathbf{Z}_{L,it} = \left(1,\ (m_{i,t-1} - l_{i,t-1}) - \frac{\widehat{\nu}}{\widehat{\delta}} SC_{i,t-1}^L,\ \mathbf{D}_{i,t-1}'\right)', \tag{36}$$

with moment conditions

$$\mathbb{E}[\mathbf{Z}_{L,it} \xi_{L,it}] = \mathbf{0}. \tag{37}$$

4.3 Stage 3: Markup dynamics and the second signal

We allow the log markup to follow

$$\ln \mu_{it} = \pi_\mu + \rho_\mu \ln \mu_{i,t-1} + \boldsymbol{\gamma}_\mu' \mathbf{D}_{i,t-1}, \qquad \mathbb{E}[\xi_{\mu,it} \mid \mathcal{I}_{i,t-1}] = 0. \tag{38}$$

Using the Stage 1 estimate of ν, equation (16) implies the markup signal $RR_{it} = \ln(\theta\nu) - \ln R_{it} = \ln \mu_{it} + \eta_{it}$. Because θ is constant, the empirically unadjusted signal is

$$\widetilde{RR}_{it} \equiv \ln \nu - \ln R_{it} = \ln \mu_{it} + \eta_{it} - \ln \theta \tag{39}$$

Stage 3 estimates the markup transition by exactly identified linear instrumental variables, conditional on the

Stage 1 technology estimates. Write $r_{it}^{\mu} = \widetilde{RR}_{it}$ for the observed markup signal and $a = \widetilde{\pi}_{\mu} = \pi_{\mu} - (1-\rho_{\mu})\ln\theta$. The observed-signal residual is

$$u_{\mu,it} = r_{it}^{\mu} - a - \rho_{\mu} r_{i,t-1}^{\mu} - \boldsymbol{\gamma}_{\mu}' \mathbf{D}_{i,t-1} = \xi_{\mu,it} + \eta_{it} - \rho_{\mu}\eta_{i,t-1}. \tag{40}$$

The lagged observed signal contains $\eta_{i,t-1}$ and is endogenous in this transition. The instrument vector is

$$\mathbf{Z}_{\mu,it} = \left(1,\ r_{i,t-2}^{\mu},\ \mathbf{D}_{i,t-2}'\right)'. \tag{41}$$

The instruments are orthogonal to the current markup innovation and to the current and lagged transitory shocks.

The Hicks-neutral and markup blocks therefore provide two observed signals with the same transitory output shock:

$$H_{it} = \omega_{N,it} + \eta_{it}, \qquad \widetilde{RR}_{it} = \left(\ln\mu_{it} - \ln\theta\right) + \eta_{it}. \tag{42}$$

Their difference removes the common error,

$$H_{it} - \widetilde{RR}_{it} = \omega_{N,it} - \ln\mu_{it} + \ln\theta. \tag{43}$$

4.4 Stage 4: Signal Extraction

The preceding stages estimate technology and the productivity and markup transitions. Stage 4 applies a univariate Kalman filter and Rauch–Tung–Striebel disturbance smoother to the markup signal alone, using the estimated markup transition and disturbance variances. This post-estimation step supplies firm-year markup predictions and a transitory-shock prediction. Hicks-neutral productivity is then recovered.

Write $\widetilde{\lambda}_{it}^{\mu} = \ln\mu_{it} - \ln\theta$ for the normalized log markup, distinct from log materials m_{it}. The scalar state and measurement equations are

$$\widetilde{\lambda}_{it}^{\mu} = \widetilde{\pi}_{\mu} + \rho_{\mu}\widetilde{\lambda}_{i,t-1}^{\mu} + \gamma_{\mu} D_{i,t-1} + \xi_{\mu,it}, \tag{44}$$

$$\widetilde{RR}_{it} = \widetilde{\lambda}_{it}^{\mu} + \eta_{it}. \tag{45}$$

Under Assumption D.1, the markup residual's variance and lag-one covariance identify the two variance components as in Proposition D.1. The filter uses these scalar variances, the markup transition, and the markup signal.

Let $\mathcal{R}_i$ contain the markup signals in a consecutive-year spell for firm i, conditional on the observed transition regressors. The scalar smoother supplies

$$\widehat{\widetilde{\lambda}}_{it}^{\mu} = \mathrm{Proj}_{\mathrm{lin}}(\widetilde{\lambda}_{it}^{\mu} \mid \mathcal{R}_i). \tag{46}$$

The subsequent recovery is

$$\widehat{\eta}_{it} = \widetilde{RR}_{it} - \widehat{\widetilde{\lambda}}_{it}^{\mu}, \qquad \widehat{\omega}_{N,it}^{SE} = H_{it} - \widehat{\eta}_{it}, \qquad \widehat{y}_{it}^{*} = y_{it} - \widehat{\eta}_{it}. \tag{47}$$

Exactly one $\widehat{\eta}_{it}$ is computed for each firm-year and reused in both equations:

$$H_{it} - \widehat{\omega}_{N,it}^{SE} = \widetilde{RR}_{it} - \widehat{\widetilde{\lambda}}_{it}^{\mu} = y_{it} - \widehat{y}_{it}^{*} = \widehat{\eta}_{it}. \tag{48}$$

Labor-augmenting productivity is recovered directly from the input-condition ratio using the preceding tech-

nology estimates:

$$\widehat{\omega}_{L,it} = m_{it} - l_{it} + \frac{\widehat{\beta}_L}{\widehat{\delta}} - \frac{\widehat{\nu}}{\widehat{\delta}} SC^L_{it}. \tag{49}$$

Appendix D specifies variance estimation, initialization, filtering, smoothing, and bootstrap implementation.

5 Data and Digitalization Measures

We use the Encuesta Sobre Estrategias Empresariales (ESEE), an annual panel of Spanish manufacturing firms covering 2000–2019. The survey reports output, inputs, digital technology variables, changes in input and output prices, and other firm characteristics. After cleaning and applying the lag requirements, the overall PC estimation sample with the overall index contains 2,213 firms and 17,087 firm-year observations. We retain firms with at least three usable observations in the relevant estimation sample. Both models use the same underlying ESEE data. We classify firms with fewer than 200 workers as small and firms with at least 200 workers as large. Table 3 reports the corresponding numbers of observations and firms. Subgroup samples are cleaned and lagged separately before imposing usable-history requirements. Firm-size changes can interrupt subgroup histories, so subgroup counts need not sum to pooled counts.

The ESEE reports changes in output and materials prices rather than price levels. Following Jaumandreu and Lin (2018), we construct Paasche price indices for each firm. Separate indices are constructed for output and intermediate inputs. We use these indices to convert nominal production and intermediate input expenditures into real quantities.

Real output is nominal production of goods and services divided by the firm-specific output price index. We construct real capital using the perpetual inventory method with an annual depreciation rate d_K of 0.15. With PI_t equal to the investment-price index divided by 100, we have

$$K_{it} = (1 - d_K)K_{i,t-1} + \frac{I_{i,t-1}}{PI_t}.$$

Investment made in period $t-1$ enters capital used in period t. We deflate lagged investment by the current-year investment-price index.

Materials are measured by total intermediate consumption. This includes raw materials, energy, and purchased services. We deflate nominal expenditures by the firm-specific intermediate input price index. The resulting measure captures changes in real materials rather than changes in expenditure caused only by prices.

Labor is measured by total hours worked. The hourly wage equals total labor compensation divided by total hours. This separates the quantity of labor from the price of labor services. We do not use the wage as a measure of labor efficiency. Labor efficiency is recovered from the model under the maintained identification assumptions.

5.1 Digitalization Indices

We construct a multidimensional digitalization index following Añón Higón and Bonvin (2024). The index contains 13 variables grouped into four components. The IT technology component includes investment in computer equipment and the use of external software programming services. The digital human capital component measures external expenditure on employee training in computing and information technologies. The automation component covers the use of industrial robots, computer-aided design, flexible manufacturing systems, local area networks, and CNC machine tools. The stakeholder interaction component includes ownership of an internet domain, a company webpage,

online sales to businesses and consumers, and purchases from suppliers through the internet. We include CNC machine tools to reflect the ESEE manufacturing setting.

We construct each digitalization measure according to the type of survey variable. Monetary variables are converted to real values, while categorical variables are converted into indicator variables. The two monetary variables are investment in computer equipment and external expenditure on IT training. The ESEE reports computer equipment investment as a percentage of total capital goods investment. We convert this percentage into an absolute flow by multiplying it by total capital goods investment. We then convert both monetary flows into stocks using the perpetual inventory method with a 15% depreciation rate. Each stock is divided by its industry and year mean. We assign the resulting relative stock to a decile and rescale it as $(d-1)/9$, where $d \in \{1, \ldots, 10\}$. The resulting scores lie between zero and one.

The categorical variables are survey adoption codes rather than ordered measures of intensity. Six variables are observed every four years: robots, CAD, flexible manufacturing systems, LAN, CNC machine tools, and external software programming services. After converting these variables to binary indicators, we carry an observed response forward for at most three years within a firm. This fills the years between survey waves. It does not replace an explicit missing response or nonresponse code. The five stakeholder variables are observed annually and remain binary after recoding.

The overall index is the unweighted sum of the 13 transformed variables. At the final aggregation stage, we set a missing transformed component to zero. We then normalize the raw sum to a scale from zero to one using its minimum and maximum values. A higher value indicates a broader and more intensive digital presence. We construct the four component indices in the same way.

We classify industries as low or high digital intensity using Calvino et al. (2018).[2] This classification reflects a firm's industry, not its own digitalization score. Firm-size groups use the fewer-than-200-worker threshold defined above. We estimate the model separately for these groups and for the pooled sample.

5.2 Summary Statistics

Table 2 presents summary statistics for output, inputs, and prices across 17,087 firm-year observations. Mean real gross output is 51.3 million euros, compared with a median of 10.1 million euros and a standard deviation of 102.9 million euros. Real capital and materials average 12.4 million and 36.2 million euros, respectively, with corresponding medians of 2.0 million and 6.5 million euros. Annual labor input averages approximately 314 thousand hours per firm, with a median of 101 thousand hours and a standard deviation of 483.5 thousand hours. Across these variables, means substantially exceed medians, and standard deviations exceed means, indicating considerable heterogeneity in production scale and suggesting right-skewed distributions.

Price indices show less relative dispersion than output and input quantities. The output price index averages 1.008, with a standard deviation of 0.094, whereas the materials price index averages 1.038, with a standard deviation of 0.153. Nominal hourly wages average 19.29 euros, with a median of 18.35 euros. The middle half of observations lies between 13.68 and 23.72 euros per hour. The capital investment-price index averages 0.831, with a standard deviation of 0.089. These statistics describe variation across the pooled firm-year sample and therefore reflect both differences across firms and changes over time.

[2]The low-digital group comprises metals and metal products, non-metallic minerals, chemical products, food, beverages and tobacco, and textiles, leather and footwear. The high-digital group comprises agricultural and industrial machinery, electrical goods, transport equipment, timber and furniture, and paper and printing. Other manufacturing is excluded from the industry comparisons.

Table 2: Summary statistics for output, inputs, and prices

Variable	Observations	Mean	SD	P25	Median	P75
Real gross output (million)	17,087	51.326	102.912	2.396	10.083	47.380
Real capital (million)	17,087	12.412	25.570	0.433	2.042	11.004
Labor (thousand hours)	17,087	313.972	483.489	41.000	101.000	378.833
Real materials (million)	17,087	36.206	74.528	1.400	6.495	32.338
Output price index	17,087	1.008	0.094	0.953	0.996	1.051
Materials price index	17,087	1.038	0.153	0.948	1.014	1.112
Nominal hourly wage (hour)	17,087	19.287	7.458	13.678	18.347	23.724
Capital investment-price index	17,087	0.831	0.089	0.749	0.862	0.912

Notes: Statistics use the overall PC estimation sample. Real output and materials are nominal values divided by their firm-specific price indices. The hourly wage is total personnel cost divided by total hours. The final row reports the investment-price index for the same observations. Monetary values are displayed in the units indicated in each row.

Table 3 reports summary statistics for the digitalization measures together with sample coverage. The table reports the corresponding counts for each group. Differences across specifications reflect sample cleaning, data availability, and lag requirements.

The mean overall digitalization index is 0.356 in the overall sample. The mean is 0.291 for small firms and 0.526 for large firms. Large firms also have higher means for all four component indices. The largest difference is in digital human capital, with means of 0.151 for small firms and 0.611 for large firms. Differences across industry groups are smaller. The mean overall index is 0.334 in low-digital industries and 0.394 in high-digital industries. These comparisons are descriptive and are based on the relevant estimation samples.

The component scores combine binary adoption measures with transformed expenditure stocks. A zero score for digital human capital does not necessarily imply that a firm provides no training. It may reflect the lowest relative stock category or the treatment of a missing component. The scores therefore measure relative digitalization rather than physical technology stocks or adoption probabilities.

Table 3: Digitalization summary statistics and estimation-sample coverage

Measure	Observations	Firms	Mean	SD	P25	Median	P75
Overall							
Overall index	17,087	2,213	0.356	0.208	0.187	0.346	0.508
IT technologies	15,874	2,140	0.431	0.325	0.143	0.375	0.714
Digital human capital	15,874	2,140	0.276	0.369	0.000	0.000	0.500
Automation	15,874	2,140	0.388	0.324	0.200	0.400	0.600
Stakeholder-facing	15,874	2,140	0.328	0.232	0.200	0.400	0.400
Small firms							
Overall index	12,246	1,689	0.291	0.181	0.154	0.275	0.421
IT technologies	11,389	1,625	0.334	0.284	0.100	0.250	0.556
Digital human capital	11,389	1,625	0.151	0.273	0.000	0.000	0.250
Automation	11,389	1,625	0.314	0.293	0.000	0.200	0.600
Stakeholder-facing	11,389	1,625	0.297	0.224	0.200	0.200	0.400
Large firms							
Overall index	4,343	546	0.526	0.183	0.385	0.529	0.663
IT technologies	3,988	525	0.691	0.282	0.444	0.786	0.944
Digital human capital	3,988	525	0.611	0.385	0.286	0.714	1.000
Automation	3,988	525	0.580	0.323	0.400	0.600	0.800
Stakeholder-facing	3,988	525	0.414	0.247	0.200	0.400	0.600
Low-digital industries							
Overall index	10,610	1,374	0.334	0.204	0.171	0.321	0.485
IT technologies	9,871	1,328	0.425	0.325	0.143	0.375	0.714
Digital human capital	9,871	1,328	0.270	0.368	0.000	0.000	0.500
Automation	9,871	1,328	0.348	0.315	0.000	0.200	0.600
Stakeholder-facing	9,871	1,328	0.315	0.229	0.200	0.400	0.400
High-digital industries							
Overall index	5,883	789	0.394	0.208	0.240	0.387	0.538
IT technologies	5,459	762	0.442	0.325	0.167	0.389	0.750
Digital human capital	5,459	762	0.288	0.367	0.000	0.000	0.600
Automation	5,459	762	0.459	0.327	0.200	0.400	0.800
Stakeholder-facing	5,459	762	0.351	0.236	0.200	0.400	0.400

Notes: Digital indices are winsorized within each sample and year at the 2.5th and 97.5th percentiles. Nominal expenditures, price indices, and real quantities follow the definitions in Section 5. Firms may appear in more than one subgroup over time.

6 Empirical Results

Investment in digital technologies improves both productivity dimensions, with a stronger labor-augmenting response among small firms. The component results distinguish pooled responses from those among small firms, and the IPC estimates also identify an overall markup response. We first examine productivity and markup dynamics, then distinguish the contributions of individual digitalization dimensions. Section 6.4 compares the productivity responses between PC and IPC on a common estimation sample and documents the significant reduction in the Hicks-neutral response when market power is allowed. The final subsection separates persistent Hicks-neutral productivity from transitory shocks.

Throughout this section, responses refer to a one-point increase in the relevant digitalization index, conditional on lagged productivity or markups. The indices range from zero to one, so this increase spans their full normalized range. Percentage interpretations use the first-order approximation of 100 times the transition coefficient. Differences between these approximate percentage responses are expressed in percentage points.

6.1 Digitalization and Productivity

In both perfectly and imperfectly competitive environments, digitalization improves both labor-augmenting and Hicks-neutral productivity under both specifications (Table 4). Under PC, a one-point increase in the index implies

approximately 3.7% higher next-period labor-augmenting productivity and 6.2% higher Hicks-neutral productivity. The corresponding IPC responses are approximately 5.5% and 4.3%. Both models therefore relate adoption with productivity gains. On the common estimation sample used for the direct comparison in Section 6.4, the PC responses are approximately 6.5% for Hicks-neutral productivity and 4.8% for labor-augmenting productivity. The paired tests use these common-sample estimates.

The clearest firm-size difference concerns labor efficiency. Small firms have labor-augmenting responses of approximately 4.6% under PC and 5.9% under IPC. The direct subgroup tests show that these responses exceed those of large firms by 5.6 and 6.1 percentage points, respectively (Table 5). The positive small-firm response and the size difference are statistically supported under both specifications. Hicks-neutral gains extend to both firm-size groups. Their small-minus-large difference is 3.4 percentage points under PC, but the IPC estimates do not establish a size difference in this channel. The results thus identify a consistent small-firm advantage in the labor-augmenting response, alongside broader Hicks-neutral improvements.

Industry patterns depend more on the specification. Under PC, the labor-augmenting response is approximately 6.4% in high-digital-intensity industries and 1.6% in low-digital-intensity industries, a difference of 4.8 percentage points. Under IPC, the corresponding responses are 6.2% and 5.0%, and their difference is imprecise.

Table 4: Digitalization and productivity under PC and IPC

Outcome	Overall	Small	Large	Low	High
A. Perfect competition (PC)					
Labor-augmenting productivity	0.037***	0.046***	−0.010	0.016	0.064***
	(0.007)	(0.009)	(0.015)	(0.011)	(0.012)
Hicks-neutral productivity	0.062***	0.067***	0.033**	0.061***	0.062***
	(0.008)	(0.011)	(0.014)	(0.011)	(0.014)
Observations	17,087	12,246	4,343	10,610	5,883
Firms	2,213	1,689	546	1,374	789
B. Imperfect competition (IPC)					
Labor-augmenting productivity	0.055***	0.059***	−0.002	0.050***	0.062***
	(0.007)	(0.008)	(0.011)	(0.010)	(0.017)
Hicks-neutral productivity	0.043***	0.042***	0.054**	0.049***	0.019
	(0.008)	(0.010)	(0.026)	(0.010)	(0.017)
Observations	13,291	9,266	3,415	8,278	4,523
Firms	1,816	1,324	457	1,120	656

Notes: Firm-block-bootstrap standard errors are in parentheses. The bootstrap repeats all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. All columns use 499 replications. The digitalization index ranges from zero to one. Coefficients are log-productivity responses to a one-point index increase; multiply by 100 to obtain approximate percentage responses to that increase. Small and Large refer to firm size; Low and High refer to industry digital intensity. PC uses one lag and the minimal exactly identified MZZ labor instrument set. IPC uses two lags and estimates Hicks-neutral technology before labor dynamics conditional on that technology.

Table 5: Differences in productivity responses by firm size and industry digital intensity

Outcome	Small	Large	Small–large	High	Low	High–low
A. Perfect competition (PC)						
Labor-augmenting productivity	0.046***	−0.010	0.056***	0.064***	0.016	0.048***
	(0.009)	(0.015)	(0.017)	(0.012)	(0.011)	(0.016)
Hicks-neutral productivity	0.067***	0.033**	0.034**	0.062***	0.061***	0.001
	(0.011)	(0.014)	(0.017)	(0.014)	(0.011)	(0.017)
B. Imperfect competition (IPC)						
Labor-augmenting productivity	0.059***	−0.002	0.061***	0.062***	0.050***	0.012
	(0.008)	(0.011)	(0.015)	(0.017)	(0.010)	(0.020)
Hicks-neutral productivity	0.042***	0.054**	−0.012	0.019	0.049***	−0.030
	(0.010)	(0.026)	(0.028)	(0.017)	(0.010)	(0.020)

Notes: Firm-block-bootstrap standard errors are in parentheses. The bootstrap repeats all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. All differences use 499 replications from a joint union-firm bootstrap, preserving covariance when firms appear in both groups. Small and Large refer to firm size; High and Low refer to industry digital intensity, which does not classify individual firms by their own adoption scores. Reported contrasts compare the marginal effects of digitalization. Coefficients and contrasts are in log units; multiplying by 100 gives approximate percentage responses and differences in those responses in percentage points, respectively, for a one-point index increase.

6.2 Digitalization and Markup Dynamics

Digitalization has a positive estimated markup response under IPC. A one-point increase in the index implies approximately 0.7% higher next-period markups, with responses of 1.3% among small firms and 1.2% in low-digital-intensity industries (Table 6). These positive responses are statistically significant.

The principal markup difference is across industries: the response is 1.8 percentage points smaller in high-digital-intensity industries than in low-digital-intensity industries. This contrast provides evidence that the pricing response to further adoption varies with industry digital intensity, even though the high-digital response itself is imprecise.

Table 6: Digitalization and markup responses under IPC

Outcome	Overall	Small	Large	Low	High
A. Log-markup transition coefficients					
Log-markup transition	0.007**	0.013**	0.003	0.012***	−0.007
	(0.003)	(0.005)	(0.007)	(0.004)	(0.004)
Observations	13,291	9,266	3,415	8,278	4,523
Firms	1,816	1,324	457	1,120	656

Outcome	Small	Large	Small–large	Low	High	High–low
B. Differences across subgroups						
Log markup	0.013**	0.003	0.010	0.012***	−0.007	−0.018***
	(0.005)	(0.007)	(0.008)	(0.004)	(0.004)	(0.006)

Notes: Firm-block-bootstrap standard errors are in parentheses. The bootstrap repeats all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. The index ranges from zero to one. Panel A reports log-markup responses to a one-point increase in lagged digitalization, using 499 replications in every column. Multiply coefficients by 100 to obtain approximate percentage responses; multiply coefficient contrasts by 100 to obtain differences in those responses in percentage points. IPC estimates markup dynamics after Hicks-neutral technology and labor dynamics. Panel B uses a joint union-firm bootstrap to compare subgroup coefficients, preserving covariance when firms appear in both groups; the size contrast uses 498 usable draws and the industry contrast uses 499. Small and Large refer to firm size; High and Low refer to industry digital intensity.

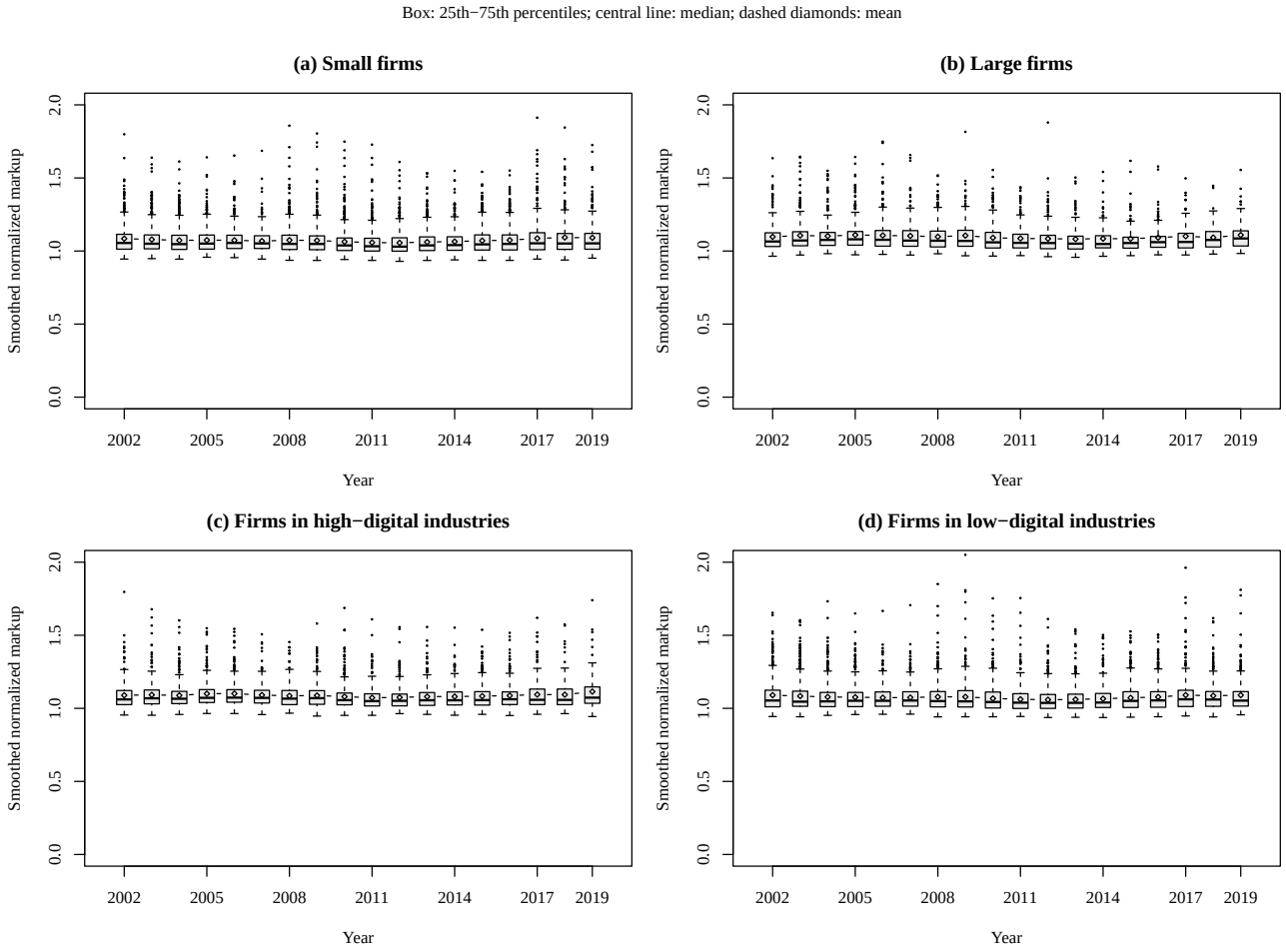


Figure 1: Annual markup distributions by firm size and industry digital intensity

Figure 1 shows a widening high-minus-low digital-industry markup gap in the normalized annual distributions. From 2002 to 2019, mean markups move from 1.085 to 1.092 among small firms and from 1.096 to 1.110 among large firms; the high-digital mean moves from 1.090 to 1.117, while the low-digital mean moves from 1.090 to 1.090. Over the overlapping 2002–2014 window, the high-digital mean falls from 1.090 to 1.082 and the low-digital mean from 1.090 to 1.066, so the gap widens even though both group means decline. This pattern is consistent with the widening industry-level gap documented by Calligaris et al. (2018), although the samples, normalization, and observation windows differ. These distributional comparisons concern markup levels; the transition estimates concern the conditional response to further adoption, which can be smaller in high-digital industries despite their higher markup levels.

6.3 Digitalization Dimensions, Productivity, and Markups

The pooled labor-augmenting responses are positive and significant for IT technologies, digital human capital, and automation under both specifications, and for stakeholder-facing technologies under IPC. Among small firms, the IPC estimates support the first three components, so the relevant list depends on the sample and specification. We enter the four dimensions described in Section 5 jointly in each transition; each response corresponds to a one-point increase in one component index, holding the other three constant. Tables 7 and 8 report the PC and IPC estimates.

Digital human capital has a particularly consistent labor-augmenting response across specifications and indus-

try groups. A one-point increase implies approximately 1.9% higher next-period labor-augmenting productivity under PC and 2.2% under IPC. IT technologies and automation also have positive overall labor-augmenting responses under both specifications. Among small firms under IPC, the responses are approximately 1.6% for IT technologies, 2.5% for digital human capital, and 1.9% for automation. Together, these estimates connect the small-firm labor-efficiency gains to several dimensions of adoption.

Automation is associated with gains in both productivity dimensions. Its overall response in labor-augmenting productivity is approximately 1.0% under PC and 1.1% under IPC, while its Hicks-neutral response is 1.8% and 1.7%, respectively. The Hicks-neutral response is also evident among small firms and in low-digital-intensity industries. Other Hicks-neutral patterns are more sensitive to the specification: stakeholder-facing digitalization has a positive overall response under PC but a less precise response under IPC, while digital human capital has a positive overall response under IPC.

Automation also has the clearest reported component-level markup response under IPC. A one-point increase in its index implies approximately 0.9% higher next-period markups overall, 1.2% among small firms, and 1.3% in low-digital industries. The component results therefore associate automation with responses in productivity and pricing, while digital human capital is consistently linked to labor-efficiency improvements.

Table 7: Digitalization components and productivity effects under PC

Component	Overall	Small	Large	Low	High
Labor-augmenting productivity					
IT technologies	0.010**	0.008	0.002	0.006	0.019***
	(0.005)	(0.006)	(0.010)	(0.006)	(0.007)
Digital human capital	0.019***	0.021***	0.008	0.021***	0.020***
	(0.004)	(0.006)	(0.007)	(0.005)	(0.007)
Automation	0.010**	0.019***	−0.007	0.005	0.018**
	(0.004)	(0.005)	(0.007)	(0.006)	(0.007)
Stakeholder-facing digitalization	0.007	0.009	0.000	−0.007	0.016
	(0.007)	(0.008)	(0.011)	(0.009)	(0.010)
Hicks-neutral productivity					
IT technologies	0.011*	0.007	0.017*	0.005	0.023***
	(0.005)	(0.007)	(0.009)	(0.007)	(0.009)
Digital human capital	0.009	0.008	0.010	0.003	0.019*
	(0.006)	(0.008)	(0.009)	(0.007)	(0.010)
Automation	0.018***	0.022***	0.002	0.027***	0.000
	(0.005)	(0.006)	(0.009)	(0.007)	(0.008)
Stakeholder-facing digitalization	0.021***	0.026***	0.005	0.017*	0.028**
	(0.007)	(0.009)	(0.012)	(0.009)	(0.011)
Observations	15,874	11,389	3,988	9,871	5,459
Firms	2,140	1,625	525	1,328	762

Notes: Firm-block-bootstrap standard errors are in parentheses, from 499 replications repeating all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.
Component coefficients refer to a one-point increase in the relevant index, holding the other components constant; multiply by 100 for an approximate percentage response.

Table 8: Digitalization components and their effects on productivity and markups under IPC

Component	Overall	Small	Large	Low	High
Labor-augmenting productivity					
IT technologies	0.018***	0.016***	−0.004	0.013**	0.021**
	(0.004)	(0.006)	(0.019)	(0.006)	(0.009)
Digital human capital	0.022***	0.025***	0.012	0.026***	0.024***
	(0.004)	(0.006)	(0.008)	(0.005)	(0.009)
Automation	0.011***	0.019***	−0.004	0.008	0.015*
	(0.004)	(0.005)	(0.027)	(0.006)	(0.009)
Stakeholder-facing digitalization	0.011**	0.009	0.004	0.011	0.004
	(0.006)	(0.007)	(0.034)	(0.007)	(0.010)
Hicks-neutral productivity					
IT technologies	0.003	0.002	0.008	0.004	0.005
	(0.004)	(0.005)	(0.017)	(0.006)	(0.008)
Digital human capital	0.009**	0.011*	0.015	0.010**	0.000
	(0.004)	(0.006)	(0.014)	(0.005)	(0.009)
Automation	0.017***	0.017***	0.016	0.022***	0.007
	(0.004)	(0.006)	(0.020)	(0.005)	(0.007)
Stakeholder-facing digitalization	0.010	0.013*	0.002	0.008	0.011
	(0.007)	(0.008)	(0.022)	(0.009)	(0.011)
Log-markup transition					
IT technologies	−0.003	−0.005	0.009	−0.003	0.000
	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)
Digital human capital	0.002	0.005	0.002	0.003	−0.004
	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)
Automation	0.009***	0.012***	0.003	0.013***	0.004
	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)
Stakeholder-facing digitalization	−0.003	0.000	−0.008	−0.002	−0.005
	(0.004)	(0.005)	(0.006)	(0.005)	(0.006)
Observations	12,161	8,487	3,085	7,607	4,126
Firms	1,717	1,250	424	1,068	617

Notes: Firm-block-bootstrap standard errors are in parentheses, from 499 replications repeating all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Four component indices enter each transition jointly. Each coefficient refers to a one-point increase in the relevant component index, holding the others constant; multiply by 100 for an approximate percentage response. Bootstrap replications used for productivity, in column order: 499, 499, 499, 499, 499. Bootstrap replications used for markup: 499, 499, 497, 499, 499. The two excluded large-firm component draws have nonpositive technology scale estimates.

6.4 Market power and the composition of productivity gains

Table 9, Panel B, reports the IPC-minus-PC differences using a common sample. PC is re-estimated on the IPC sample, and each of the 499 paired firm-block bootstrap draws uses the same resampled firms in both models. Panel A reports within-model gaps on each model's main estimation sample. The samples differ because of lag requirements. Table 10 reports the corresponding output contributions on the common sample.

On the common sample, the PC responses are approximately 6.5% for Hicks-neutral productivity and 4.8% for labor-augmenting productivity. Using unrounded common-sample estimates, the labor-augmenting-minus-Hicks-neutral response gap rises by 3.0 percentage points between PC and IPC, with a paired-test p-value of 0.015 (Panel B). This paired change differs from the comparison implied by the main PC estimation-sample figures because the underlying firm-year observations differ. The Hicks-neutral coefficient falls significantly, while the change in the labor-augmenting coefficient is imprecisely estimated. The evidence therefore supports a smaller Hicks-neutral response under IPC and a significant shift in the relative composition of the responses between models.

Panel A provides descriptive context. In the main PC estimation sample, the Hicks-neutral response exceeds the labor-augmenting response by approximately 2.5 percentage points. In the main IPC estimation sample, the labor-augmenting response is larger in the point estimates, but the within-model gap is imprecisely estimated. The central finding concerns the significant between-model change on the common sample.

The output comparison weights labor-efficiency improvements by their output elasticity. For a one-point increase in the index, the Hicks-neutral contribution to log output is γ_N, and the labor-efficiency contribution is

$\gamma_L \overline{\varepsilon^L}$, where $\overline{\varepsilon^L} = \nu \overline{SC^L}$. Each bootstrap draw uses its own technology and transition estimates and resampled mean labor cost share. Table 10 reports these first-order contributions in percent and their differences in percentage points.

On the common sample, the overall Hicks-neutral output contribution is 6.51% under PC and 4.31% under IPC, while the labor-efficiency contribution is 1.11% and 1.38%, respectively. The Hicks-neutral-minus-labor-efficiency gap narrows from 5.40 to 2.93 percentage points. Its change of −2.47 percentage points is significant. The paired contribution changes are a 2.20-percentage-point reduction for Hicks-neutral productivity and a 0.28-percentage-point increase for labor efficiency. The labor-efficiency contribution change in Table 10 is statistically significant, although the labor-augmenting transition-coefficient change in Table 9 is imprecisely estimated. These are different quantities. The output contribution weights each model's labor-augmenting coefficient by its own estimated labor elasticity, with uncertainty in both quantities propagated through the paired bootstrap. The IPC specification therefore assigns a larger relative contribution to labor efficiency, while retaining a larger overall output contribution from Hicks-neutral productivity.

This narrowing is also evident among small firms and in high-digital-intensity industries, where the contribution gaps decline by 3.28 and 4.86 percentage points, respectively. The low-digital decline of 2.19 percentage points and the large-firm increase of 2.83 percentage points are imprecise. Within IPC, the high-digital contribution gap is only 0.25 percentage points and is not distinguishable from zero. The sensitivity of the estimated composition of gains is thus particularly apparent among small firms and in high-digital industries.

Table 9: Market power and productivity transition coefficients

		Firm size		Industry digital intensity	
Statistic	Overall	Small	Large	Low	High
A. Labor-augmenting minus Hicks-neutral effect					
PC, full PC sample	-0.025^{**}	−0.021	-0.044^{**}	-0.045^{***}	0.002
	(0.011)	(0.015)	(0.021)	(0.016)	(0.018)
IPC	0.012	0.016	-0.057^{*}	0.001	0.043^{*}
	(0.011)	(0.013)	(0.030)	(0.015)	(0.024)
B. IPC minus PC on the common sample, paired draws					
Labor-augmenting effect	0.008	−0.002	−0.009	0.017^{*}	0.007
	(0.005)	(0.006)	(0.010)	(0.009)	(0.016)
Hicks-neutral effect	-0.022^{**}	-0.032^{**}	0.027	−0.017	-0.046^{**}
	(0.010)	(0.013)	(0.024)	(0.013)	(0.021)
Labor-augmenting minus Hicks-neutral	0.030^{**}	0.030^{**}	−0.035	0.034^{**}	0.052^{**}
	(0.012)	(0.015)	(0.027)	(0.017)	(0.026)
Scale parameter ν	0.068^{***}	0.067^{***}	0.088	0.068^{***}	0.080
	(0.009)	(0.016)	(0.093)	(0.014)	(0.078)

Notes: Firm-block-bootstrap standard errors are in parentheses, using 499 replications. Stars indicate two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.
Panel A reports labor-augmenting-minus-Hicks-neutral transition coefficients within each model's main estimation sample. These samples derive from the same underlying ESEE data but differ because of lag requirements; the main PC overall sample contains 17,087 observations. Panel B reports IPC-minus-PC differences on the common IPC estimation sample. PC is re-estimated on that sample, and both models use identical firm-year observations and the same resampled firms in each paired bootstrap draw.
All differences are computed draw by draw before calculating standard errors and tests. Small and Large refer to firm size; Low and High refer to industry digital intensity.

Table 10: Market power and the output contributions of productivity gains

Statistic	Overall	Firm size		Industry digital intensity	
		Small	Large	Low	High
A. Output contributions on the common sample					
PC: labor-efficiency contribution	1.11^{***}	1.54^{***}	0.12	0.73^{***}	1.37^{***}
	(0.19)	(0.25)	(0.26)	(0.25)	(0.28)
PC: Hicks-neutral contribution	6.51^{***}	7.46^{***}	2.78^{*}	6.64^{***}	6.48^{***}
	(0.95)	(1.29)	(1.66)	(1.26)	(1.60)
PC: Hicks-neutral minus labor efficiency	5.40^{***}	5.93^{***}	2.66	5.91^{***}	5.11^{***}
	(0.97)	(1.32)	(1.70)	(1.31)	(1.64)
IPC: labor-efficiency contribution	1.38^{***}	1.59^{***}	-0.05	1.19^{***}	1.67^{***}
	(0.17)	(0.23)	(0.25)	(0.23)	(0.51)
IPC: Hicks-neutral contribution	4.31^{***}	4.24^{***}	5.45^{**}	4.92^{***}	1.92
	(0.78)	(0.97)	(2.60)	(0.96)	(1.70)
IPC: Hicks-neutral minus labor efficiency	2.93^{***}	2.64^{***}	5.49^{**}	3.73^{***}	0.25
	(0.82)	(1.01)	(2.64)	(1.02)	(1.84)
B. IPC minus PC output contributions, paired draws					
Labor-efficiency contribution change	0.28^{**}	0.06	-0.16	0.46^{**}	0.31
	(0.12)	(0.15)	(0.21)	(0.21)	(0.47)
Hicks-neutral contribution change	-2.20^{**}	-3.23^{**}	2.67	-1.72	-4.56^{**}
	(1.03)	(1.34)	(2.43)	(1.29)	(2.14)
Hicks-neutral dominance gap change	-2.47^{**}	-3.28^{**}	2.83	-2.19	-4.86^{**}
	(1.07)	(1.38)	(2.45)	(1.35)	(2.21)
Dominance-gap change p-value	0.020	0.017	0.247	0.104	0.028
Observations (IPC sample)	13,291	9,266	3,415	8,278	4,523

Notes: Firm-block-bootstrap standard errors are in parentheses, using 499 replications. Stars indicate two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, and $^{*}p < 0.10$.
Panel A reports local, first-order output responses to a one-point increase in the digitalization index at fixed physical inputs. Contributions are expressed in percent; gaps and the changes in Panel B are expressed in percentage points. The Hicks-neutral contribution is $100\gamma_N$, and the labor-efficiency contribution is $100\gamma_L\overline{\varepsilon^L}$, where $\overline{\varepsilon^L} = \nu\overline{SC^L}$.
Each draw uses its own coefficients and resampled firm-year mean labor cost share, preserving uncertainty in the elasticity. The dominance gap is the Hicks-neutral contribution minus the labor-efficiency contribution; a negative change in Panel B indicates that this gap narrows under IPC. All differences are computed draw by draw before calculating standard errors and tests.
Small and Large refer to firm size; Low and High refer to industry digital intensity.

6.5 Persistent productivity and transitory shocks

Unadjusted production residuals and revenue shares confound persistent firm fundamentals with transitory output shocks, η_{it}, such as measurement error, temporary machine downtime, or unanticipated demand spikes. Table 11 compares the unadjusted signals with the smoothed series after filtering out these transitory disturbances (see Appendix D for the state-space smoothing procedure).

Filtering out transitory noise shows that most of the observed variation in productivity and markups reflects persistent structural differences across firms rather than temporary fluctuations. In the overall sample, removing η_{it} reduces the standard deviation of Hicks-neutral productivity from 0.1618 in the unadjusted signal to 0.1454 in the recovered series. This corresponds to a 19.2% reduction in pooled variance, while the mean remains virtually unchanged (0.0645 versus 0.0647). Because the smoothing procedure filters a noisy signal, this variance reduction summarizes the dispersion of the recovered series rather than an exact orthogonal decomposition. The estimated transitory shock variance is 0.0048, while the predicted shocks have a standard deviation of 0.0516 and a model-based prediction RMSPE of 0.0467, indicating nontrivial filtering uncertainty at the individual firm-year level.

Comparing subsamples reveals two additional economic patterns. First, transitory output volatility is concentrated among smaller and more digitally intensive firms. The transitory variance in small firms of 0.0058 is more than double that of large firms 0.0026, indicating that small manufacturers face far greater year-to-year idiosyncratic output volatility. By contrast, large firms exhibit higher persistent markup innovation variance ($\widehat{\sigma}^2_{\xi_\mu} = 0.0035$ versus 0.0025), consistent with more durable shifts in pricing power among larger producers. Second, even after purging transitory shocks, the cross-sectional dispersion of labor-augmenting productivity of 0.4686 is more than three times larger than that of Hicks-neutral productivity of 0.1454 and over five times larger than that of log markups, 0.0863. Thus, persistent heterogeneity across Spanish manufacturers is primarily driven by differences in labor efficiency

rather than neutral technology or market power.

Table 11: Transitory-shock extraction under imperfect competition

Outcome	Overall	Small	Large	Low	High
$\widehat{\sigma}^2_{\eta}$	0.0048	0.0058	0.0026	0.0044	0.0060
	(0.0005)	(0.0006)	(0.0004)	(0.0005)	(0.0007)
$\widehat{\sigma}^2_{\xi_\mu}$	0.0028	0.0025	0.0035	0.0027	0.0010
	(0.0008)	(0.0011)	(0.0008)	(0.0005)	(0.0006)
Shock prediction RMSPE	0.0467	0.0490	0.0389	0.0450	0.0436
	(0.0018)	(0.0031)	(0.0019)	(0.0015)	(0.0024)
Firm-year means					
Labor-augmenting productivity	0.3173	0.2443	−0.0869	0.3468	0.2305
	(0.0351)	(0.0326)	(0.1101)	(0.0659)	(0.1615)
Hicks-neutral signal	0.0647	0.0621	0.0683	0.0708	0.0605
	(0.0059)	(0.0068)	(0.0205)	(0.0120)	(0.0168)
Recovered Hicks-neutral productivity	0.0645	0.0621	0.0681	0.0705	0.0604
	(0.0059)	(0.0068)	(0.0205)	(0.0120)	(0.0168)
Predicted transitory shock	0.0001	0.0001	0.0001	0.0002	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Smoothed normalized log markup	0.0669	0.0660	0.0883	0.0682	0.0811
	(0.0098)	(0.0167)	(0.0704)	(0.0145)	(0.0731)
Firm-year standard deviations					
Labor-augmenting productivity	0.4686	0.4410	0.3697	0.5069	0.3739
	(0.0136)	(0.0146)	(0.0797)	(0.0346)	(0.1123)
Hicks-neutral signal	0.1618	0.1635	0.1551	0.1648	0.1535
	(0.0033)	(0.0051)	(0.0791)	(0.0348)	(0.0631)
Recovered Hicks-neutral productivity	0.1454	0.1441	0.1453	0.1500	0.1347
	(0.0031)	(0.0050)	(0.0797)	(0.0354)	(0.0652)
Predicted transitory shock	0.0516	0.0598	0.0333	0.0489	0.0662
	(0.0044)	(0.0056)	(0.0046)	(0.0043)	(0.0070)
Smoothed normalized log markup	0.0863	0.0845	0.0966	0.0864	0.0808
	(0.0038)	(0.0051)	(0.0080)	(0.0034)	(0.0041)

Notes: Parentheses report firm-block-bootstrap standard errors from 499 attempts repeating all estimation stages. Firm-year summaries and shock prediction RMSPE use 499, 499, 499, 499, 483 draws in column order; coefficient and variance estimates use all 499 draws. The 16 high-digital draws with a negative implied markup-innovation variance omit the scalar decomposition. Shock prediction RMSPE is the square root of the average model-based disturbance prediction-error variance, conditional on fitted parameters; it is distinct from the SD of predicted shocks and its bootstrap SE. The same predicted shock is used in the share and output recoveries.

7 Diagnostics and Robustness

The following checks assess the technology, instruments, and sensitivity of the transition estimates used in the output comparison. They qualify the maintained specifications rather than independently validating their assumptions.

7.1 Technology and persistence estimates

7.1.1 Perfect competition

Under PC, median returns to scale range from 0.949 to 0.987 across samples. Mean elasticity of the inputs is positive in every sample. The PC persistence estimates range from 0.796 to 0.903 across the two productivity processes and samples. The estimates indicate persistence in both productivity processes.

Table 12: Technology, persistence and estimation diagnostics under PC

Parameter	Overall	Small	Large	Low	High
ρ_L	0.891***	0.884***	0.903***	0.901***	0.862***
	(0.005)	(0.005)	(0.011)	(0.005)	(0.010)
ρ_N	0.840***	0.853***	0.796***	0.837***	0.838***
	(0.011)	(0.013)	(0.027)	(0.014)	(0.021)
ν	0.872***	0.871***	0.876***	0.872***	0.873***
	(0.002)	(0.002)	(0.004)	(0.002)	(0.003)
δ	−0.129***	−0.139***	−0.111***	−0.123***	−0.137***
	(0.003)	(0.004)	(0.009)	(0.004)	(0.006)
β_K	−0.129***	−0.159***	−0.058	−0.141***	−0.072
	(0.041)	(0.062)	(0.194)	(0.053)	(0.066)
β_{KK}	0.0156***	0.0174***	0.0099	0.0168***	0.0102**
	(0.0029)	(0.0047)	(0.0110)	(0.0038)	(0.0045)
Median returns to scale	0.969	0.949	0.987	0.977	0.950
Mean capital elasticity	0.098	0.078	0.113	0.106	0.078
Share with negative capital elasticity	0.000	0.000	0.000	0.000	0.000
Observations	17,087	12,246	4,343	10,610	5,883
Firms	2,213	1,689	546	1,374	789

Notes: Firm-block-bootstrap standard errors are in parentheses, from 499 replications repeating all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Inference is conditional on the sample construction and instrument choices. Step 2 is exactly identified; Step 3 is nonlinear least squares. The point selects the feasible labor root with positive variable-input coefficients and the smallest Step-3 residual sum of squares, falling back to all feasible roots if none has positive coefficients. Bootstrap draws match the feasible labor-persistence root nearest the point estimate within a window of 0.15, clipped to the maintained bounds.

7.1.2 IPC

The IPC Hansen reference p-values range from 0.123 to 1.000. They do not reject the overidentifying restrictions at conventional levels; these reference tests alone do not establish instrument validity. Mean IPC returns to scale range from 0.971 to 1.003. The large-firm estimate is close to one and slightly above it, while the other point estimates are below one. Mean output elasticity of the inputs are positive in every IPC sample. IPC labor-augmenting persistence ranges from 0.906 to 0.942, Hicks-neutral persistence from 0.510 to 0.819, and markup persistence from 0.836 to 0.892. The large-firm Hicks-neutral persistence estimate of 0.510 is lower than those in the other samples.

Table 13: Technology, persistence and estimation diagnostics under IPC

Parameter	Overall	Small	Large	Low	High
ρ_L	0.934***	0.919***	0.940***	0.942***	0.906***
	(0.005)	(0.006)	(0.014)	(0.007)	(0.016)
ρ_N	0.777***	0.817***	0.510***	0.788***	0.819***
	(0.029)	(0.034)	(0.172)	(0.033)	(0.097)
ρ_μ	0.846***	0.859***	0.846***	0.836***	0.892***
	(0.019)	(0.024)	(0.038)	(0.022)	(0.032)
ν	0.939***	0.938***	0.961***	0.941***	0.951***
	(0.009)	(0.016)	(0.092)	(0.014)	(0.078)
δ	−0.206***	−0.209***	−0.176***	−0.203**	−0.198
	(0.006)	(0.008)	(0.043)	(0.089)	(0.217)
β_K	−0.016	−0.022	0.554	−0.021	0.017
	(0.030)	(0.078)	(1.640)	(0.242)	(0.231)
β_{KK}	0.0038*	0.0041	−0.0297	0.0039	0.0010
	(0.0021)	(0.0060)	(0.0923)	(0.0165)	(0.0170)
Mean returns to scale	0.978	0.971	1.003	0.978	0.982
Mean capital elasticity	0.039	0.034	0.042	0.037	0.031
Share with negative capital elasticity	0.000	0.000	0.102	0.000	0.000
Hansen reference p	0.686	0.539	1.000	0.522	0.123
Mean $\widehat{\nu}/R$	1.078	1.077	1.100	1.078	1.092
	(0.011)	(0.019)	(0.106)	(0.016)	(0.089)
Observations	13,291	9,266	3,415	8,278	4,523
Firms	1,816	1,324	457	1,120	656

Notes: Firm-block-bootstrap standard errors are in parentheses, from 499 replications repeating all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

7.2 Alternative materials-demand proxy

The sensitivity specification replaces the main Hicks-neutral block with a lagged materials-demand proxy adjusted using a forecast of lagged markup. Appendix C gives its construction and additional exclusion restrictions. Table 14 reports estimates for the same IPC samples and capital measure.

Table 14: Sensitivity using a markup-adjusted materials proxy

Outcome	Overall	Small	Large	Low	High
Labor-augmenting productivity	0.061***	0.053***	0.006	0.033***	0.055***
	(0.010)	(0.013)	(0.014)	(0.012)	(0.013)
Hicks-neutral productivity	0.039***	0.025**	0.005	0.016	0.025
	(0.014)	(0.013)	(0.011)	(0.010)	(0.018)
Log-markup transition	0.007**	0.013***	0.003	0.012***	−0.007
	(0.003)	(0.005)	(0.006)	(0.004)	(0.004)
ν	0.118	0.367	1.004	0.909	0.833
ρ_N	0.987	0.988	0.921	0.947	0.963
Draws at $\nu = 0.05$	107	39	41	6	8
Draws at $\rho_N = 0.995$	2	39	5	0	6
Observations	13,291	9,266	3,415	8,278	4,523
Firms	1,816	1,324	457	1,120	656

Notes: Firm-block-bootstrap standard errors are in parentheses, from 499 replications repeating all estimation stages. Stars use two-sided, unadjusted normal-reference tests: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$. Inference is conditional on the sample construction and instrument choices. This comparison uses the IPC composite-index sample and the alternative estimation order in Appendix C. It imposes $0.05 \leq \nu \leq 3$, $-0.95 \leq \rho_N \leq 0.995$, and capital elasticity at least 10^{-6} on the retained support. Capital curvature is estimated in all five groups. Normal-reference stars are not boundary-calibrated.

The materials-proxy specification preserves positive, significant labor-augmenting responses in the overall, small-firm, low-digital, and high-digital samples, while the large-firm response remains imprecise. The overall labor-augmenting coefficient is 0.061, compared with 0.055 in the baseline IPC specification in Table 4, Panel B. Hicks-neutral responses are significant at the five percent level only in the overall and small-firm samples. The small-firm Hicks-neutral coefficient falls from 0.042 to 0.025, while its variable-input scale parameter falls from 0.938 to 0.367; the corresponding overall scale parameter falls from 0.939 to 0.118. Markup slopes are unchanged because the scale parameter enters the markup signal through a constant absorbed by the transition intercept.

These technology changes matter for the identification of ν. In the pooled alternative specification, 107 of 499 bootstrap draws (21.4%) reach the lower bound $\nu = 0.05$, and 20 pooled draws admit multiple feasible zero-moment roots. An optimizer audit reproduces the pooled point estimate across 75 combinations of starting values and lower bounds; a competing zero-moment root near $\nu = 0.955$ violates capital positivity. The low estimate is therefore a reproducible solution of the alternative moments, while the boundary mass and multiple admissible bootstrap roots reveal sensitivity to the restrictions and root selection. These features are consistent with weakly informative alternative moments or tension between the moments and the admissibility constraints, but do not distinguish those explanations or constitute a formal weak-identification test. Ordinary normal-reference bootstrap inference is not calibrated for this boundary mass.

The baseline does not display the same low-ν pattern in the saved runs. Its pooled estimate is approximately 0.939, and its 499 bootstrap estimates range from approximately 0.917 to 0.964; none of the stored bootstrap estimates across the five baseline composite-index samples is at or below 0.05. This comparison concerns boundary behavior, not proof of strong identification. The baseline also conditions on a numerical root-selection rule, and some subgroup replications require its fallback; neither numerical full rank nor an interior pooled bootstrap distribution establishes global uniqueness or strong identification.

We retain the quantity-side baseline because it identifies technology from the dynamic production equation without a generated markup forecast or the additional materials-demand exclusion restrictions in Appendix C. The

alternative additionally requires its forecast error to be orthogonal to the instruments and lagged digitalization to have no direct materials-demand effect conditional on the included states. The sensitivity exercise therefore evaluates a different set of maintained restrictions and cannot independently validate the baseline technology scale. We interpret the paired baseline comparison conditional on its own timing, support, exclusion, and local-rank assumptions.

8 Policy Implications

The results have three implications for evaluating digital adoption. First, evaluations should measure the channels separately. The common-sample comparison yields a significantly smaller Hicks-neutral response under IPC than under PC. Evaluations based on the competitive benchmark may therefore place too much weight on Hicks-neutral gains relative to the IPC specification. This sensitivity matters when interpreting the channels through which adoption affects firms, workers, and consumers. Labor-augmenting gains raise the marginal product of labor and therefore bear on wages and labor demand, whereas Hicks-neutral gains lower unit costs across all inputs. The estimated markup coefficient is approximately one-eighth of the labor-augmenting coefficient. This ratio compares responses of different outcomes; it does not measure the share of digitalization's benefits attributable to market power or provide a welfare or surplus decomposition.

Second, small firms have lower average digitalization and significantly larger labor-augmenting responses than large firms under both specifications. Large firms, which are the most digitalized, show no detectable labor-efficiency response. This pattern suggests that smaller firms deserve attention in studies of the returns to digital adoption, but does not establish the benefits of a subsidy. Under IPC, the small-firm responses are significant for IT technologies, digital human capital, and automation, so programmes that combine equipment with implementation assistance and IT training are consistent with the evidence. Among large firms the measurable response operates through Hicks-neutral productivity, which suggests that the case for support there rests on cost reduction rather than on labor efficiency.

Third, competition concerns about digitalization point to automation specifically. Automation is the only component with a significant markup response, while IT technologies, digital human capital, and stakeholder-facing technologies show no statistically detectable markup response. The markup response is also significantly smaller in digitally intensive industries than in other industries. Higher markup levels in those industries therefore need not imply a larger markup response to further adoption. Competition analysis may benefit from tracking automation adoption together with marginal costs and consumer prices, since a higher markup accompanied by lower marginal cost need not raise prices.

These implications rest on conditional relationships within the production framework. They identify which channels, firms, and technologies merit attention, but returns to a specific programme require evidence on its costs, on the additional adoption it induces, and on wages, employment, and prices.

9 Conclusion

This paper separates the effects of digitalization on labor-augmenting productivity, Hicks-neutral productivity, and markups without imposing a particular model of competition. The common-sample comparison shows that accounting for market power significantly reduces the estimated Hicks-neutral response. It also shifts the relative composition of productivity gains toward labor efficiency. The change in the labor-augmenting coefficient alone is imprecisely estimated. When labor efficiency is weighted by its output elasticity, Hicks-neutral productivity still makes the larger

contribution to output, but the gap narrows significantly under IPC.

The remaining results show where these gains arise. Labor-efficiency gains are larger among small firms, while large firms, which are the most digitalized, show no detectable labor-augmenting response. In the pooled IPC estimates, all four digitalization components have positive, statistically significant labor-augmenting responses, and automation is the only component with a statistically significant markup response. The effects of digitalization therefore cannot be captured by a single productivity measure, and attributing its gains requires separating factor-biased technology from market power.

These findings bear on the labor share because the two productivity channels affect it differently. Hicks-neutral gains leave relative factor demand unchanged, whereas labor-augmenting gains alter it in a direction that depends on the substitutability of labor and materials. The distinction connects our results to the discussion of labor-complementing technologies in Acemoglu et al. (2026). Our estimates concern productivity and markups at fixed inputs, so they do not yet measure how digitalization changes labor demand, wages, or prices. Combining the estimated technology with input-demand responses to quantify these effects is a natural next step.

References

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., and Price, B. (2014). Return of the Solow paradox? IT, productivity, and employment in US manufacturing. *American Economic Review: Papers & Proceedings*, 104(5):394–399.

Acemoglu, D., Johnson, S., and Autor, D. (2026). Pro-worker AI: Concept and evidence. Working Paper.

Añón Higón, D. and Bonvin, D. (2024). Digitalization and trade participation of smes. *Small Business Economics*, 62:857–877.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2):645–709.

Bertschek, I., Briglauer, W., Hüschelrath, K., Kauf, B., and Niebel, T. (2015). Broadband infrastructure and economic growth. *Journal of Economic Surveys*, 29(5):1042–1065.

Bertschek, I. and Kaiser, U. (2004). Productivity effects of organizational change: Microeconometric evidence. *Management Science*, 50(3):394–404.

Black, S. E. and Lynch, L. M. (2001). How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics*, 83(3):434–445.

Brynjolfsson, E. and Hitt, L. M. (1996a). Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Science*, 42(4):541–558.

Brynjolfsson, E. and Hitt, L. M. (1996b). Productivity, business profitability, and consumer surplus: Three different measures of information technology value. *MIS Quarterly*, 20(2):121–142.

Brynjolfsson, E. and Hitt, L. M. (2000). Beyond computation: Information technology, organizational transformation and business performance. *Journal of Economic Perspectives*, 14(4):23–48.

Calligaris, S., Criscuolo, C., and Marcolin, L. (2018). Mark-ups in the digital era. OECD Science, Technology and Industry Working Papers 2018/10, OECD Publishing, Paris.

Calvino, F., Criscuolo, C., Marcolin, L., and Squicciarini, M. (2018). A taxonomy of digital intensive sectors. *OECD Science, Technology and Industry Working Papers*, (2018/14).

Cusolito, A. P. (2020). Firms, technology, and digital transformation. World Bank Working Paper.

De Loecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *The Quarterly Journal of Economics*, 135(2):561–644.

De Loecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102(6):2437–2471.

Demirer, M. (2026). Production function estimation with factor-augmenting technology: An application to markups. forthcoming, Econometrica.

DeStefano, T., Kneller, R., and Timmis, J. (2018). Broadband and firm performance. Working Paper.

Doraszelski, U. and Jaumandreu, J. (2018). Measuring the bias of technological change. *Journal of Political Economy*, 126(3):1027–1084.

Doraszelski, U. and Jaumandreu, J. (2019). Using cost minimization to estimate markups. Working paper, Boston University and University of Pennsylvania.

Draca, M., Sadun, R., and Van Reenen, J. (2008). ICT, innovation and productivity. Working Paper.

Flynn, Z., Gandhi, A., and Traina, J. (2019). Measuring markups with production data. Working paper, June 13.

Gal, P., Nicoletti, G., Renault, T., Sorbe, S., and Timiliotis, C. (2019). Digitalisation and productivity: In search of the holy grail—firm-level empirical evidence from EU countries. Technical Report 1533, OECD Publishing.

Gartner (2026). Survey shows 45% of CFOs say their AI investments lean toward productivity, while 20% say these investments lean toward decision quality. Technical report, Gartner Research, Stamford, CT.

Haller, S. A. and Lyons, S. (2015). Broadband adoption and firm productivity: Evidence from Irish manufacturing firms. *Telecommunications Policy*, 39(1):1–13.

Jaumandreu, J. and Lin, S. (2018). Prices under innovation: Evidence from manufacturing firms. Discussion Paper 13146, CEPR.

Jorgenson, D. W. (2001). Information technology and the US economy. *American Economic Review*, 91(1):1–32.

Malikov, E., Zhao, S., and Zhang, J. (2024). A system approach to structural identification of production functions with multi-dimensional productivity. In Parmeter, C. F., Tsionas, M. G., and Wang, H.-J., editors, *Essays in Honor of Subal Kumbhakar*, volume 46 of *Advances in Econometrics*, pages 211–263. Emerald Group Publishing Limited.

Nucci, F., Puccioni, C., and Ricchi, O. (2023). Digital technologies and productivity: A firm-level investigation. *Economic Modelling*, 128:106524.

PwC (2026). 29th annual global CEO survey: Leading through uncertainty in the age of AI. Technical report, PricewaterhouseCoopers, London, UK.

Raval, D. (2023). Testing the production approach to markup estimation. *The Review of Economic Studies*, 90(5):2592–2611.

Raval, D. R. (2019). The micro elasticity of substitution and non-neutral technology. *The RAND Journal of Economics*, 50(1):147–167.

Stiroh, K. J. (2005). Reassessing the impact of IT in the production function: A meta-analysis and sensitivity tests. *Annales d'Économie et de Statistique*, 79/80:529–561.

Zhang, H. (2019). Non-neutral technology, firm heterogeneity, and labor demand. *Journal of Development Economics*, 140:145–168.

A Perfect-Competition Benchmark: Derivation and Estimation

The benchmark is a perfectly competitive (PC) model in which the firm takes the output price P_{it}^Y as given. Our estimation of this benchmark follows Malikov et al. (2024) (MZZ).

Under perfect competition, the firm's profit is:

$$\pi = P_{it}^Y F(K_{it}, e^{\omega_{L,it}} L_{it}, M_{it})\, e^{\omega_{N,it}+\eta_{it}} - P_{it}^L L_{it} - P_{it}^M M_{it} \tag{50}$$

Since η affects realized output, the first-order conditions (FOCs) of expected profit maximization over L_{it} and M_{it} imply

$$P_{it}^Y\, \theta\, \varepsilon_{it}^L \frac{e^{y_{it}^*}}{L_{it}} = P_{it}^L, \tag{51}$$

$$P_{it}^Y\, \theta\, \varepsilon_{it}^M \frac{e^{y_{it}^*}}{M_{it}} = P_{it}^M. \tag{52}$$

We now define the revenue shares and their sum,

$$s_{it}^L = \frac{P_{it}^L L_{it}}{P_{it}^Y Y_{it}}, \qquad s_{it}^M = \frac{P_{it}^M M_{it}}{P_{it}^Y Y_{it}}, \qquad R_{it} = s_{it}^L + s_{it}^M. \tag{53}$$

Dividing (51) and (52) by revenue $P_{it}^Y Y_{it}$, adding, and using (9) gives the share equation

$$\ln R_{it} = \ln(\theta\nu) - \eta_{it}. \tag{54}$$

Taking the ratio of (51) and (52) gives

$$\frac{P_{it}^L L_{it}}{P_{it}^M M_{it}} = \frac{\varepsilon_{it}^L}{\varepsilon_{it}^M} = \frac{\beta_L + \delta(m_{it} - l_{it} - \omega_{L,it})}{\beta_M - \delta(m_{it} - l_{it} - \omega_{L,it})}. \tag{55}$$

The output price, Hicks-neutral productivity, θ, and η_{it} all cancel.

Solving (55) gives

$$\delta(m_{it} - l_{it} - \omega_{L,it}) = \beta_M SC_{it}^L - \beta_L SC_{it}^M = \nu SC_{it}^L - \beta_L, \tag{56}$$

and hence the labor-augmenting-productivity term can be expressed as

$$\omega_{L,it} = m_{it} - l_{it} + \frac{\beta_L}{\delta} - \frac{\nu}{\delta} SC_{it}^L. \tag{57}$$

where labor share of total variable cost is

$$SC_{it}^{L} = \frac{P_{it}^{L} L_{it}}{P_{it}^{L} L_{it} + P_{it}^{M} M_{it}}, \Rightarrow SC_{it}^{M} = 1 - SC_{it}^{L}. \tag{58}$$

Substituting (56) into (6)–(7) yields the identities

$$\varepsilon_{it}^{L} = \nu SC_{it}^{L}, \qquad \varepsilon_{it}^{M} = \nu SC_{it}^{M}. \tag{59}$$

Equations (57) and (59) are implications of the restricted translog with two variable inputs. They hold in both market environments studied below, because a markup common to the two conditions cancels from the ratio. This common inversion is the point at which our extension connects to the competitive benchmark.

Substituting (56) into (5) and collecting terms concentrates effective labor out of the production function:

$$y_{it}^{*} = \kappa_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \nu\, m_{it} - \frac{\nu^2}{2\delta}\left(SC_{it}^{L}\right)^2 + \omega_{N,it}, \qquad \kappa_0 = \alpha_0 + \frac{\beta_L^2}{2\delta}. \tag{60}$$

Equation (60) expresses output in terms of ν, δ, and the composite intercept κ_0.

To estimate the competitive benchmark model outlined above, we adopt the following three-step procedure.[3]

A.1 Step 1: Estimation of θ, ν and prediction of η_{it}

Let N_{obs} denote the number of firm-years of observations. Because $\mathbb{E}[\eta_{it}] = 0$, the sample analogues of (54) are

$$\widehat{\eta}_{it} = \overline{\ln R} - \ln R_{it}, \tag{61}$$

$$\widehat{\theta} = \frac{1}{N_{\text{obs}}} \sum_{it} \exp(\widehat{\eta}_{it})\,, \tag{62}$$

$$\widehat{\nu} = \frac{\exp\left(\overline{\ln R}\right)}{\widehat{\theta}}. \tag{63}$$

The first step therefore identifies $\nu = \beta_L + \beta_M$ and θ and gives a prediction of η_{it}. Under PC, the markup restriction is $\mu_{it} = 1$, so the normalized log-markup state is the constant $-\ln\theta$. The observed share equation then gives

$$\widetilde{RR}_{it}^{PC} = -\ln\widehat{\theta} + \widehat{\eta}_{it}, \qquad H_{it}^{PC} = \widehat{\omega}_{N,it} + \widehat{\eta}_{it}. \tag{64}$$

The same firm-year shock recovered from the competitive share equation is subtracted from the production signal when recovering Hicks-neutral productivity. Because the markup state is fixed under PC, this direct inversion replaces the scalar markup filtering step used under IPC.

A.2 Step 2: Prediction of labor-augmenting productivity

Let $\mathbf{D}_{it}$ collect the firm's digitalization measures. We assume that labor-augmenting productivity follows the following law of motion

$$\omega_{L,it} = \rho_L \omega_{L,i,t-1} + \boldsymbol{\gamma}_L' \mathbf{D}_{i,t-1} + \xi_{L,it}, \qquad \mathbb{E}[\xi_{L,it} \mid \mathcal{I}_{i,t-1}] = 0. \tag{65}$$

The equation contains no separate intercept because a constant is not separately identified from β_L/δ. Substituting the proxy (57) at t and $t-1$ produces the term $(1-\rho_L)\beta_L/\delta$, which plays the role of the intercept.

[3] For details on the estimation of the restricted translog model with factor-biased productivity, see Malikov et al. (2024).

Step 2 estimates the law of motion in (65). Since $\omega_{L,it}$ is unobserved, we replace it using the labor productivity proxy in (57). Substituting the proxy at t and $t-1$ gives

$$\begin{aligned}\xi_{L,it} = \; & (m_{it} - l_{it}) + \frac{\beta_L}{\delta} - \frac{\nu}{\delta} SC^L_{it} \\ & - \rho_L \left[(m_{i,t-1} - l_{i,t-1}) + \frac{\beta_L}{\delta} - \frac{\nu}{\delta} SC^L_{i,t-1} \right] - \boldsymbol{\gamma}_L' \mathbf{D}_{i,t-1}. \end{aligned} \tag{66}$$

This equation serves as a proxy for the innovation term. We estimate the parameter vector $\left(\frac{\beta_L}{\delta}, \frac{\nu}{\delta}, \rho_L, \boldsymbol{\gamma}_L'\right)'$ by nonlinear GMM. The instrument vector is

$$\mathbf{Q}_{L,it} = \left(1,\, m_{i,t-1} - l_{i,t-1},\, SC^L_{i,t-1},\, \mathbf{D}'_{i,t-1}\right)', \tag{67}$$

which contains a constant, the lagged input ratio, the lagged labor cost share, and the lagged digitalization variables. The corresponding moment conditions are

$$\mathbb{E}\left[\mathbf{Q}_{L,it}\, \xi_{L,it}\left(\frac{\beta_L}{\delta}, \frac{\nu}{\delta}, \rho_L, \boldsymbol{\gamma}_L \right) \right] = \mathbf{0}. \tag{68}$$

The number of instruments equals the number of parameters in this block. Step 2 is therefore exactly identified under the maintained exclusion and rank conditions.

Since $\xi_{L,it}$ does not contain Y_{it}, the variables dated $t-1$ in (67) are valid instruments under the maintained information assumptions. Thus Step 2 provides estimates of β_L/δ and ν/δ, together with ρ_L and $\boldsymbol{\gamma}_L$. Step 1 has already identified $\nu = \beta_L + \beta_M$. We can therefore recover the production function effects as

$$\widehat{\delta} = \frac{\widehat{\nu}}{\widehat{(\nu/\delta)}}, \qquad \widehat{\beta}_L = \widehat{(\beta_L/\delta)}\widehat{\delta}, \qquad \widehat{\beta}_M = \widehat{\nu} - \widehat{\beta}_L. \tag{69}$$

Thus, the first two steps identify δ, β_L, and β_M, as well as the persistence and digitalization parameters in the law of motion for labor-augmenting productivity.

A.3 Step 3: Hicks-neutral productivity

Hicks-neutral productivity follows

$$\omega_{N,it} = \rho_{0N} + \rho_N \omega_{N,i,t-1} + \boldsymbol{\gamma}_N' \mathbf{D}_{i,t-1} + \xi_{N,it}, \qquad \mathbb{E}[\xi_{N,it} \mid \mathcal{I}_{i,t-1}] = 0. \tag{70}$$

Two representations of $\omega_{N,it}$ are available. From the production function,

$$y^{\dagger}_{it} \equiv y_{it} - \beta_M m_{it} - \beta_L(l_{it} + \omega_{L,it}) + \tfrac{1}{2}\delta(m_{it} - l_{it} - \omega_{L,it})^2 = \alpha_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \omega_{N,it} + \eta_{it}. \tag{71}$$

Using (71) we express

$$\omega_{N,it} = y^{\dagger}_{it} - [\alpha_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \eta_{it}] \tag{72}$$

Alternatively, we can use materials demand to proxy for unobserved Hicks-neutral productivity. Solving (52) for

$\omega_{N,it}$ gives the materials-demand inversion

$$m_{it}^{\dagger} \equiv \left(p_{it}^{M} - p_{it}^{Y}\right) - \ln\theta - \ln[\beta_M - \delta(m_{it} - l_{it} - \omega_{L,it})] + (1-\beta_M)\, m_{it} \\ - \beta_L(l_{it} + \omega_{L,it}) + \tfrac{1}{2}\delta(m_{it} - l_{it} - \omega_{L,it})^2 = \alpha_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \omega_{N,it}, \tag{73}$$

where p_{it}^{M} and p_{it}^{Y} are log input and output prices. We can use (73) to express $\omega_{N,it}$ which is

$$\omega_{N,it} = m_{it}^{\dagger} - [\alpha_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2]. \tag{74}$$

The inversion (74) does not contain Y_{it} and is therefore free of η_{it}, which is precisely why it, rather than (72), is used to replace lagged Hicks-neutral productivity. Substituting $\omega_{N,i,t-1} = m_{i,t-1}^{\dagger} - \alpha_0 - \beta_K k_{i,t-1} - \frac{1}{2}\beta_{KK} k_{i,t-1}^2$ into (70) and then into (71) gives

$$y_{it}^{\dagger} = \pi_0 + \beta_K k_{it} + \tfrac{1}{2}\beta_{KK} k_{it}^2 + \rho_N \left[m_{i,t-1}^{\dagger} - \beta_K k_{i,t-1} - \tfrac{1}{2}\beta_{KK} k_{i,t-1}^2\right] \\ + \boldsymbol{\gamma}_N' \mathbf{D}_{i,t-1} + \xi_{N,it} + \eta_{it}, \tag{75}$$

where $\pi_0 = \rho_{0N} + (1-\rho_N)\alpha_0$. This specification follows in the Doraszelski and Jaumandreu (2019), Malikov et al. (2024). We estimate this equation by profiled nonlinear least squares in $(\beta_K, \beta_{KK}, \pi_0, \rho_N, \boldsymbol{\gamma}_N)$, holding the step-1 and step-2 estimates fixed. Because the composite error is $\xi_{N,it} + \eta_{it}$ and k_{it} is predetermined, no additional instruments are required. In particular, competitive Step 3 has no $\mathbf{Q}_{N,it}$ instrument vector: it is profiled nonlinear least squares, not an IV or GMM. Standard errors for the full sequence are obtained by a firm-level block bootstrap in which all three steps are re-estimated.

The two productivity series are recovered observation by observation as

$$\widehat{\omega}_{L,it} = m_{it} - l_{it} + \frac{\widehat{\beta}_L}{\widehat{\delta}} - \frac{\widehat{\nu}}{\widehat{\delta}} SC_{it}^{L}, \tag{76}$$

$$\widehat{\omega}_{N,it} = \widehat{m}_{it}^{\dagger} - [\widehat{\alpha}_0 + \widehat{\beta}_K k_{it} + \tfrac{1}{2}\widehat{\beta}_{KK} k_{it}^2] \tag{77}$$

B Distributions of Elasticities, Productivities and Markups

Table 15 summarizes the PC estimates. In the overall sample, mean output elasticities are 0.098 for capital, 0.235 for labor, and 0.637 for materials. Labor and materials elasticities sum to the estimated variable-input scale ν by construction. Their equal standard deviations follow from these identities rather than providing independent evidence about technology.

The final two columns of Table 15 report recovered log productivity after subtracting each sample's 2001 median. These are overall firm-year distributions relative to a common within-sample reference point. A negative observation is below that reference median. Labor-augmenting productivity has greater dispersion in its own log units than Hicks-neutral productivity, but these components enter output differently, so their dispersions are not directly comparable output contributions.

Table 15: PC elasticities and normalized productivity distributions

Sample	ε_K	ε_L	ε_M	$\widetilde{\omega}_L$	$\widetilde{\omega}_N$
Overall	0.098 (0.032) [0.073, 0.097, 0.123]	0.235 (0.126) [0.139, 0.214, 0.307]	0.637 (0.126) [0.565, 0.658, 0.732]	0.094 (0.445) [-0.205, 0.101, 0.376]	0.044 (0.194) [-0.061, 0.062, 0.168]
Small firms	0.078 (0.026) [0.059, 0.078, 0.099]	0.254 (0.133) [0.153, 0.235, 0.335]	0.617 (0.133) [0.536, 0.636, 0.718]	0.174 (0.429) [-0.125, 0.177, 0.462]	0.059 (0.186) [-0.049, 0.076, 0.180]
Large firms	0.113 (0.011) [0.105, 0.112, 0.119]	0.187 (0.089) [0.118, 0.181, 0.238]	0.689 (0.089) [0.638, 0.695, 0.757]	0.034 (0.387) [-0.214, 0.016, 0.257]	0.075 (0.170) [-0.025, 0.075, 0.175]
Low-digital industries	0.106 (0.035) [0.078, 0.105, 0.133]	0.225 (0.132) [0.125, 0.200, 0.297]	0.647 (0.132) [0.575, 0.673, 0.747]	0.095 (0.480) [-0.226, 0.088, 0.396]	0.048 (0.207) [-0.061, 0.072, 0.178]
High-digital industries	0.078 (0.020) [0.061, 0.076, 0.094]	0.249 (0.110) [0.169, 0.234, 0.314]	0.624 (0.110) [0.560, 0.640, 0.704]	0.140 (0.383) [-0.114, 0.150, 0.396]	0.044 (0.184) [-0.066, 0.058, 0.167]

Notes: Cells show mean (SD) and [P25, median, P75]. Productivity is centered on the median of the first estimation year in each sample. The materials inversion and output-minus-transitory-shock recovery of Hicks productivity agree to numerical precision after consistent measurement of expenditures, prices, and quantities.

Tables 16 and 17 report IPC distributions. They follow the format of Table 15. The first reports input elasticities and log productivities. The second reports smoothed normalized markups in logs and levels. Standard deviations in these appendix tables describe dispersion across firm-years. They are not bootstrap standard errors.

Table 16: IPC elasticities and normalized productivity distributions

Sample	ε_K	ε_L	ε_M	$\widetilde{\omega}_L$	$\widetilde{\omega}_N$
Overall	0.039 (0.008) [0.033, 0.039, 0.045]	0.249 (0.134) [0.148, 0.227, 0.325]	0.689 (0.134) [0.613, 0.712, 0.791]	0.122 (0.469) [-0.200, 0.111, 0.425]	0.068 (0.145) [-0.023, 0.056, 0.151]
Small firms	0.034 (0.006) [0.029, 0.034, 0.038]	0.271 (0.142) [0.164, 0.249, 0.355]	0.667 (0.142) [0.582, 0.689, 0.774]	0.157 (0.441) [-0.146, 0.142, 0.435]	0.066 (0.144) [-0.022, 0.055, 0.147]
Large firms	0.042 (0.032) [0.023, 0.045, 0.065]	0.203 (0.097) [0.129, 0.197, 0.259]	0.758 (0.097) [0.702, 0.764, 0.832]	0.016 (0.370) [-0.239, 0.013, 0.256]	0.068 (0.145) [-0.030, 0.054, 0.151]
Low-digital industries	0.037 (0.008) [0.031, 0.037, 0.043]	0.238 (0.140) [0.132, 0.212, 0.315]	0.702 (0.140) [0.625, 0.729, 0.809]	0.138 (0.507) [-0.218, 0.116, 0.476]	0.067 (0.150) [-0.026, 0.054, 0.152]
High-digital industries	0.031 (0.002) [0.030, 0.031, 0.033]	0.269 (0.118) [0.183, 0.252, 0.338]	0.682 (0.118) [0.612, 0.699, 0.768]	0.112 (0.374) [-0.137, 0.125, 0.365]	0.071 (0.135) [-0.014, 0.063, 0.151]

Notes: Cells show mean (SD) and [P25, median, P75] across unweighted firm-year observations; parentheses contain distributional standard deviations. Quartiles use the same median-unbiased convention as Table 15. The overall, small, large, low-digital, and high-digital samples contain 13,291, 9,266, 3,415, 8,278, 4,523 observations, respectively. Estimates use the composite digitalization index and lagged-capital instruments.

Table 17: IPC normalized markup distributions

Sample	$\widehat{\widetilde{\lambda}}^{\mu}$	$\exp(\widehat{\widetilde{\lambda}}^{\mu})$
Overall	0.067 (0.086) [0.008, 0.046, 0.101]	1.073 (0.104) [1.008, 1.047, 1.107]
Small firms	0.066 (0.085) [0.008, 0.045, 0.100]	1.072 (0.100) [1.008, 1.046, 1.105]
Large firms	0.088 (0.097) [0.026, 0.062, 0.118]	1.098 (0.124) [1.026, 1.064, 1.126]
Low-digital industries	0.068 (0.086) [0.008, 0.047, 0.105]	1.075 (0.101) [1.008, 1.048, 1.111]
High-digital industries	0.081 (0.081) [0.027, 0.060, 0.111]	1.088 (0.095) [1.028, 1.062, 1.117]

Notes: Cells show mean (SD) and [P25, median, P75] across unweighted firm-year observations; parentheses contain distributional standard deviations. Quartiles use the same median-unbiased convention as Table 15. The overall, small, large, low-digital, and high-digital samples contain 13,291, 9,266, 3,415, 8,278, 4,523 observations, respectively. Estimates use the composite digitalization index and lagged-capital instruments. Here $\widetilde{\lambda}^{\mu}_{it} = \ln \mu_{it} - \ln \theta$. Columns report smoothed normalized log markup and its exponential, the markup relative to the common scale θ.

C Robustness Check Using Lagged Materials Demand as the Hicks-Neutral Proxy

The sensitivity specification uses a materials-demand proxy for lagged Hicks-neutral productivity. Its markup-forecast correction is an additional construction. It begins with the labor-productivity block, estimates the markup forecast next, and then estimates Hicks-neutral technology. This order differs from the main IPC sequence. The production technology and markup definition are maintained.

C.1 Materials-demand inversion under IPC

Following the materials-demand approach discussed by Malikov et al. (2024), this sensitivity specification constructs a proxy for lagged Hicks-neutral productivity and adds the markup-forecast correction derived below.

Let $w_{it}^M \equiv \ln W_{it}^M$, $p_{it} \equiv \ln P_{it}$, and $\mathfrak{m}_{it} \equiv \ln \mu_{it}$. The materials first-order condition can be written in logs as

$$w_{it}^M - p_{it} = \ln \theta - \mathfrak{m}_{it} + \ln \varepsilon_{it}^M + \ln F(K_{it}, L_{it}^*, M_{it}) + \omega_{N,it} - m_{it}, \tag{78}$$

where $\varepsilon_{it}^M = \nu SC_{it}^M$ from the share identities. The constant $\ln \theta$ is retained separately from κ_0; it is carried explicitly here to keep the derivation transparent. η_{it} does not enter this condition because it affects measured output and revenue, not the productivity state observed by the firm. From equation (59), $\varepsilon_{M,it}^* = \nu SC_{it}^M$, where $SC_{it}^M = 1 - SC_{it}^L$. Define the scale-free materials-demand index

$$\mathcal{M}_{it}^*\Big(\nu, \frac{\nu}{\delta}\Big) \equiv (w_{it}^M - p_{it}) - \ln SC_{it}^M + (1-\nu) m_{it} + \left(\frac{\nu^2}{2\delta}\right) (SC_{it}^L)^2. \tag{79}$$

Substitution of the reduced restricted translog technology in equation (60) into equation (78) gives

$$\mathcal{M}_{it}^*\Big(\nu, \frac{\nu}{\delta}\Big) = \kappa_0 + \beta_K k_{it} + \frac{1}{2}\beta_{KK} k_{it}^2 + \omega_{N,it} - \mathfrak{m}_{it} + \ln \nu + \ln \theta. \tag{80}$$

Thus, under IPC, raw materials demand identifies Hicks-neutral productivity net of the log markup. Treating $m_{i,t-1}$ one-for-one as $\omega_{N,i,t-1}$ would impose an incorrect scale and load persistent markup variation into productivity.

The required markup correction follows from the maintained markup block. Because $R_{it} = \theta \nu \mu_{it}^{-1} e^{-\eta_{it}}$, define the scale-free observed log-markup measure

$$b_{it} \equiv -\ln R_{it} = \mathfrak{m}_{it} + \eta_{it} - \ln \nu - \ln \theta. \tag{81}$$

Adding a constant to the dependent variable of equation (38) changes only its intercept. Consequently, b_{it} follows the same transition with intercept $\pi_b = \pi_\mu - (1-\rho_\mu)\ln(\theta\nu)$ and unchanged slopes (ρ_μ, γ_μ). Because $\ln \mu_{i,t-1}$ is latent and the observed $b_{i,t-1}$ contains $\eta_{i,t-1}$, it cannot serve as its own instrument. We estimate the observed-signal transition using a constant, $b_{i,t-2}$, and predetermined digitalization as instruments, and then form the one-step forecast

$$\widehat{b}_{i,t-1|t-2} = \widehat{\pi}_b + \widehat{\rho}_\mu(-\ln R_{i,t-2}) + \widehat{\gamma}_\mu D_{i,t-2}. \tag{82}$$

The lagged materials-demand proxy for the location-normalized Hicks-neutral state is

$$\mathcal{H}_{i,t-1}^M\Big(\boldsymbol{\Theta}, \frac{\nu}{\delta}\Big) \equiv \mathcal{M}_{i,t-1}^*\Big(\nu, \frac{\nu}{\delta}\Big) - \beta_K k_{i,t-1} - \frac{1}{2}\beta_{KK} k_{i,t-1}^2 + \widehat{b}_{i,t-1|t-2}. \tag{83}$$

Under a correctly specified markup transition, $\mathcal{H}_{i,t-1}^M = \kappa_0 + \omega_{N,i,t-1} + e_{i,t-1}^M$, where $e_{i,t-1}^M$ combines the lagged transitory output shock $\eta_{i,t-1}$ and the negative of the markup-forecast error. The constant κ_0 is absorbed by the dynamic intercept below.

C.2 Alternative sequential nonlinear-GMM block

Substituting equation (83) for the lagged state in equation (23) gives the robustness-check residual

$$u^M_{n,it}(\boldsymbol{\Theta}) = y_{it} - \pi_0 - \beta_K k_{it} - \frac{1}{2}\beta_{KK}k^2_{it} - \nu m_{it} + \frac{\nu}{2}\widehat{\left(\frac{\nu}{\delta}\right)}(SC^L_{it})^2 \\ - \rho_N \mathcal{H}^M_{i,t-1}\left(\boldsymbol{\Theta}, \widehat{\left(\frac{\nu}{\delta}\right)}\right) - \gamma_N D_{i,t-1}, \qquad \pi_0 = \pi_N + (1-\rho_N)\kappa_0. \tag{84}$$

As in the main IPC specification, the location of the recovered Hicks-neutral composite is fixed by setting its first-estimation-year median to zero. This normalization does not recover the primitive production-function parameters.

The composite disturbance is $\xi_{N,it} + \eta_{it} - \rho_N e^M_{i,t-1}$. Relative to the benchmark dynamic production equation, lagged observed output no longer appears directly. The generated markup-forecast error enters instead. Conditional on the labor-block estimate of ν/δ, estimate $(\pi_0, \beta_K, \beta_{KK}, \nu, \rho_N, \gamma_N)'$ by two-step nonlinear GMM using

$$\mathbb{E}\big[\mathbf{Z}^M_{n,it} u^M_{n,it}(\boldsymbol{\Theta}_0)\big] = \mathbf{0}, \tag{85}$$

where the baseline instrument vector is

$$\mathbf{Z}^M_{n,it} = \big(1, D_{i,t-1}, m_{i,t-1}, k_{it}, k^2_{it}, p^M_{i,t-1} - p_{i,t-1}\big)'. \tag{86}$$

Capital dated t is predetermined by assumption. The lagged variable inputs and relative input price instrument the current static choices and the materials-demand state.

The alternative estimation begins with a freely estimated labor block, rather than conditioning on the main IPC estimates of ν and δ. Set $A = \beta_L/\delta$, $B = \nu/\delta$, and $q_{it} = m_{it} - l_{it}$. Its actual parameter and instrument vectors are

$$\boldsymbol{\vartheta}^M_L = (A, B, \rho_L, \gamma_L)', \qquad \mathbf{Z}^M_{L,it} = (1, q_{i,t-1}, SC^L_{i,t-1}, D_{i,t-1})'. \tag{87}$$

The four moments use

$$\xi^M_{L,it} = q_{it} - \rho_L q_{i,t-1} + (1-\rho_L)A - B(SC^L_{it} - \rho_L SC^L_{i,t-1}) - \gamma_L D_{i,t-1}, \qquad \mathbb{E}[\mathbf{Z}^M_{L,it}\xi^M_{L,it}] = \mathbf{0}. \tag{88}$$

This is the labor-first ratio estimation in Appendix A; B is jointly estimated with A, persistence, and the digitalization coefficient. Stored branch anchors supply starting values, rather than fixed technology estimates. After the alternative Hicks block estimates ν, recover $\delta = \nu/B$ and $\beta_L = A\delta$. We then estimate the markup transition and construct the forecast in equation (82). Finally, we estimate the Hicks-neutral block in equation (85) and adjust the markup intercept using the estimated technology scale. Each firm-block bootstrap replication repeats this alternative sequence. The estimation samples are the same as those used for the main IPC estimates. Output and materials price indices are winsorized at the 2.5th and 97.5th percentiles.

C.3 Identification conditions

The robustness check requires an interior materials choice, $\nu SC^M_{it} > 0$, correct measurement of relative materials prices, and price taking in the materials market. The fitted markup forecast must remove the conditional markup component of materials demand, and its forecast error must be orthogonal to the instruments in equation (86). Lagged digitalization must not shift materials demand directly after conditioning on Hicks-neutral productivity, the markup process, and the included state variables. These are additional assumptions rather than consequences of

using a lag. If they fail, γ_N combines technological adoption with an omitted input-demand or markup channel.

D Signal Extraction for Hicks-Neutral Productivity and Markups

The reported IPC decomposition uses the markup-only Kalman filter and Rauch–Tung–Striebel smoother in Section 4.4, followed by recovery of Hicks-neutral productivity as $H_{it} - \widehat{\eta}_{it}$. This procedure produces Table 11, Figure 1, and Tables 16–17. The PC distributions in Table 15 use the competitive share inversion in Appendix A. Signal extraction is a post-estimation decomposition and does not re-estimate structural GMM coefficients.

D.1 State and measurement equations

Define

$$\widetilde{RR}_{it} \equiv \ln \widehat{\nu} - \ln R_{it}.$$

Ignoring first-step estimation error in $\widehat{\nu}$ for exposition, equations (38) and (39) imply the state-space representation

$$\text{state:} \qquad \widetilde{\lambda}^{\mu}_{it} = \widetilde{\pi}_{\mu} + \rho_{\mu}\widetilde{\lambda}^{\mu}_{i,t-1} + \gamma_{\mu} D_{i,t-1} + \xi_{\mu,it}, \tag{89}$$

$$\text{measurement:} \qquad \widetilde{RR}_{it} = \widetilde{\lambda}^{\mu}_{it} + \eta_{it}, \tag{90}$$

where $\widetilde{\lambda}^{\mu}_{it} \equiv \ln \mu_{it} - \ln \theta$ is the log markup normalized relative to $\ln \theta$. Because θ is constant, this normalization changes only the transition intercept: $\widetilde{\pi}_{\mu} = \pi_{\mu} - (1 - \rho_{\mu}) \ln \theta$. The persistence parameter, the digitalization effect, and the innovation are unchanged. Hence the Stage 3 composite disturbance is

$$u_{\mu,it} = \xi_{\mu,it} + \eta_{it} - \rho_{\mu}\eta_{i,t-1}. \tag{91}$$

Identification of the two components comes instead from their distinct serial covariance pattern.

D.2 Variance identification

Let

$$\sigma^2_{\xi} \equiv \mathrm{Var}(\xi_{\mu,it}), \qquad \sigma^2_{\eta} \equiv \mathrm{Var}(\eta_{it}),$$

with both variances constant within an estimation sample.

Assumption D.1(Second-moment restrictions). The sequences $\{\xi_{\mu,it}\}$ and $\{\eta_{it}\}$ have mean zero and are serially uncorrelated. They are mutually orthogonal at all leads and lags used below. Their variances are finite and constant within the estimation sample. In addition,

$$0 < |\rho_{\mu}| < 1.$$

The restriction $\rho_{\mu} \neq 0$ is needed to identify σ^2_{η} from the first-order autocovariance. Values of $|\rho_{\mu}|$ close to zero imply weak variance separation in finite samples.

Proposition D.1(Variance separation). Under Assumption D.1 and the Stage 3 transition,

$$\mathrm{Var}(u_{\mu,it}) = \sigma_\xi^2 + (1+\rho_\mu^2)\sigma_\eta^2, \tag{92}$$

$$\mathrm{Cov}(u_{\mu,it}, u_{\mu,i,t-1}) = -\rho_\mu \sigma_\eta^2, \tag{93}$$

$$\mathrm{Cov}(u_{\mu,it}, u_{\mu,i,t-h}) = 0, \qquad h \geq 2. \tag{94}$$

Proof. Using equation (91),

$$u_{\mu,it} = \xi_{\mu,it} + \eta_{it} - \rho_\mu \eta_{i,t-1}.$$

The variance result follows because the three terms are pairwise orthogonal under Assumption D.1. For the lag-one covariance, the only nonzero cross-product is generated by $\eta_{i,t-1}$, which enters $u_{\mu,it}$ with effect $-\rho_\mu$ and $u_{\mu,i,t-1}$ with effect one. This gives

$$\mathrm{Cov}(u_{\mu,it}, u_{\mu,i,t-1}) = -\rho_\mu \sigma_\eta^2.$$

At lags $h \geq 2$, the two composite disturbances contain no common transitory-shock term, and all remaining cross-products are zero by Assumption D.1. This gives equation (94). □

Because $\rho_\mu \neq 0$, equations (92) and (93) identify the two disturbance variances:

$$\sigma_\eta^2 = -\frac{\mathrm{Cov}(u_{\mu,it}, u_{\mu,i,t-1})}{\rho_\mu}, \tag{95}$$

$$\sigma_\xi^2 = \mathrm{Var}(u_{\mu,it}) - (1+\rho_\mu^2)\sigma_\eta^2. \tag{96}$$

D.3 Univariate filtering and smoothing

The baseline filter uses only the markup signal. Filtering is performed separately for each consecutive-year spell, restarting after a firm change or calendar-year gap. A flat initial state prior yields first-observation filtered mean $x_{1|1} = \widetilde{RR}_{i1}$ and variance $P_{1|1} = \widehat{\sigma}_\eta^2$. The recursions below apply from the second observation onward.

Let $x_{t|s}$ denote the scalar linear predictor of $\widetilde{\lambda}_{it}^{\mu}$ using the markup signal through period s, conditional on the observed transition regressors, and let $P_{t|s}$ denote its mean-squared prediction error. The implementation uses the algebraically equivalent Joseph form for the variance update to improve numerical stability.

$$x_{t|t-1} = \widehat{\widehat{\pi}}_\mu + \hat{\rho}_\mu x_{t-1|t-1} + \hat{\gamma}_\mu D_{i,t-1}, \qquad P_{t|t-1} = \hat{\rho}_\mu^2 P_{t-1|t-1} + \widehat{\sigma_\xi^2}, \tag{97}$$

$$K_t = \frac{P_{t|t-1}}{P_{t|t-1} + \widehat{\sigma_\eta^2}}, \qquad x_{t|t} = x_{t|t-1} + K_t\big(\widetilde{RR}_{it} - x_{t|t-1}\big), \tag{98}$$

$$P_{t|t} = (1-K_t)P_{t|t-1}. \tag{99}$$

Because the sample is historical rather than a real-time forecasting exercise, we use the Rauch–Tung–Striebel smoother after the forward pass. For $t = T-1, \ldots, 1$,

$$J_t = \frac{P_{t|t}\hat{\rho}_\mu}{P_{t+1|t}}, \qquad x_{t|T} = x_{t|t} + J_t\big(x_{t+1|T} - x_{t+1|t}\big), \tag{100}$$

$$P_{t|T} = P_{t|t} + J_t^2\big(P_{t+1|T} - P_{t+1|t}\big). \tag{101}$$

The recursion is initialized at the final filtered values $x_{T|T}$ and $P_{T|T}$.

The smoothed predictor of the transitory output shock is

$$\widehat{\eta}_{it} = \widetilde{RR}_{it} - x_{t|T}. \tag{102}$$

With the observed signal treated as fixed, the measurement disturbance and state prediction have the same model-based error variance. The implementation propagates the scalar smoothing covariance and reports

$$\widehat{\mathrm{MSE}}(\widehat{\eta}_{it}) = P_{t|T}, \qquad \widehat{\mathrm{SD}}_{\mathrm{pred},it} = \sqrt{P_{t|T}}, \qquad \mathrm{RMSPE}_{\eta} = \left(n^{-1} \sum_{it} P_{t|T} \right)^{1/2}. \tag{103}$$

These quantities condition on fitted parameters and the maintained scalar model and initialization. They differ from the cross-sectional dispersion of $\widehat{\eta}$ and from bootstrap parameter uncertainty. Under Gaussian disturbances and valid conditional exogeneity they have the conditional prediction-error interpretation. Without Gaussianity, the decomposition is a model-based linear prediction.

D.4 Recovering Hicks-neutral productivity

In the implemented sequential recovery, the production equation supplies the composite

$$\widehat{(\omega_{N,it} + \eta_{it})}$$

from equations (29)–(30). The signal-extraction step yields the firm-year predictor

$$\widehat{\omega}^{SE}_{N,it} = H_{it} - \widehat{\eta}_{it}. \tag{104}$$

D.5 Implementation and interpretation

Each consecutive-year uses the flat-prior first-observation update described above. Scalar filtering and disturbance smoothing are then applied with the fitted nonnegative variances. One predicted shock is stored and reused in the markup, Hicks-neutral and output recoveries in equation (48).

The predictions and error variances treat the observed digitalization path as given. A conditional-mean interpretation additionally requires Gaussian disturbances and the relevant conditional exogeneity; the IV timing restrictions alone do not establish this smoothing condition. More generally, consistent parameter estimation under valid moments and maintained rank conditions permits estimation of the model-based predictor. With $e_{it} = \eta_{it} - \widehat{\eta}_{it}$, both recovered series retain the same extraction error:

$$\widehat{\omega}^{SE}_{N,it} = \omega_{N,it} + e_{it}, \qquad \widehat{\widetilde{\lambda}}^{\mu}_{it} = \widetilde{\lambda}^{\mu}_{it} + e_{it}. \tag{105}$$

The firm-block bootstrap repeats technology estimation, conditional labor estimation, IV markup estimation and covariance-based variance recovery, scalar filtering, disturbance smoothing and common-shock recovery. Bootstrap standard errors summarize estimator variation; the model-based prediction-error variance in equation (103) separately records unresolved firm-year uncertainty.

E Variable Definitions

- **Investment (I_{it}):** The value of current-period investment in equipment goods, excluding buildings, land, and financial assets. Investment made in period t becomes productive in period $t+1$.
- **Capital (K_{it}):** Real capital constructed by the perpetual inventory recursion and industry–size initialization described in Section 5, using reported real net capital when available and an annual depreciation rate d_K of 0.15.
- **Labor (L_{it}):** Total hours worked by all employees of the firm in year t.
- **Materials (M_{it}):** The real value of intermediate consumption, comprising raw materials, energy, and purchased services. Nominal intermediate consumption expenditures are deflated by a firm-specific intermediate consumption price index constructed from price variations reported in the ESEE.
- **Output (Y_{it}):** The real value of goods and services produced, obtained by deflating nominal production values by a firm-specific output price index constructed using the Paasche index formula.
- **Nominal output and the variable-cost ratio:** $P_{it}^{Y}Y_{it}$ is nominal goods-and-services production, the denominator used in the model's expenditure shares. The symbol R_{it} denotes total labor and materials expenditure divided by nominal output, not nominal sales.
- **Wage ($W_{L,it}$):** The hourly wage rate, defined as total labor compensation divided by total effective hours worked.
- **Output Price Index (P_{it}^{Y}):** A firm-specific Paasche price index constructed from price variations of output sales reported in the ESEE, following the methodology of Jaumandreu and Lin (2018).
- **Materials Price Index (P_{it}^{M}):** A firm-specific Paasche price index constructed from price variations of intermediate consumption reported in the ESEE.
- **Labor share ($s_{L,it}$):** The ratio of total labor compensation to total revenue, $s_{L,it} \equiv W_{L,it}L_{it}/(P_{it}Y_{it})$. It is revenue-denominated and therefore differs from the value-added labor share.
- Materials share ($s_{M,it}$): The ratio of total materials expenditure to total revenue, $s_{M,it} \equiv W_{M,it}M_{it}/(P_{it}Y_{it})$.

F Digitalization Index

Table 18: Variables used to construct the digitalization index

Dimension	Variable	Definition
Automation	Use of robots	Whether the company uses robots in the production process (Yes/No)
	Use of CNC machine tools	Whether the company uses computer numerically controlled machine tools in the production process (Yes/No)
	Use of computer-aided design	Whether the company makes use of computer-assisted design (Yes/No)

Dimension	Variable	Definition
	Use of flexible systems	Whether the company combines machine tools controlled by computer, robotics, and/or CAD through a mainframe computer (Yes/No)
	Use of LAN	Whether the company uses a local area network in its manufacturing activities (Yes/No)
Interactions with stakeholders	Ownership of internet domain	Whether the company has its own internet domain (Yes/No)
	Ownership of a webpage	Whether the company's web page is stored on its own servers (Yes/No/NA)
	Sales to companies via internet	Availability of a sales system for other companies through internet (Yes/No/NA)
	Sales to final consumer via internet	Availability of a sales system for final customers through internet (Yes/No/NA)
	Purchases from suppliers online	Whether goods or services are purchased through the internet (Yes/No/NA)
Digital-related human capital	External expenditure on IT training	Value of external expenditure on employee training in computing and information technologies (Euros)
IT technologies	Investment in computer equipment	Share of purchases of information processing equipment in total tangible fixed asset purchases (%)
	Hiring of software programming	Whether the company contracts out software-programming services (Yes/No)

Notes: The global index contains the 13 variables shown. ESEE adoption codes are recoded as 1 for Yes and 0 for No; non-response codes are treated as missing before aggregation. Computer-equipment investment is converted from a percentage of total capital-goods investment into an absolute flow before capitalization. Computer-equipment investment and IT-training expenditure are capitalized with 15% depreciation. Quadrennial indicators are carried forward within firms for at most three years. Section 5.1 describes scaling, normalization, and the treatment of missing components at aggregation.