Department of Economics

Athens University of Economics and Business

WORKING PAPER no. 11-2026

Production Function Estimation under Input Distortions: Identification Using Heteroscedasticity-Based Instruments

Dimitris K. Christopoulos and Elias Tzavalis

September 2026

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

Production Function Estimation under Input Distortions: Identification Using Heteroscedasticity-Based Instruments

Dimitris K. Christopoulos and Elias Tzavalis

Athens University of Economics and Business, Athenss 104 34, Greece*

Abstract

Estimating production functions is challenging when firms' input choices are jointly shaped by productivity and distortion shocks arising from credit constraints, taxes, subsidies, or regulatory wedges. In such environments, observed input variation reflects both technological productivity and distortion-driven adjustments, weakening the monotonicity and orthogonality conditions underlying standard proxy-variable and lag-based estimators. This paper develops a two-step estimator for production functions under distortions that identifies production elasticities using heteroskedasticity-based internal instruments. The approach does not require proxy inversion, external instruments, or restrictive one-to-one mappings between inputs and productivity. We establish identification, consistency, and asymptotic normality of the estimator. In the second step, the residual system is embedded in a linear state-space framework and estimated using Kalman filter–based Gaussian quasi-maximum likelihood to recover latent productivity and distortion processes. Monte Carlo evidence shows that the estimator delivers low-bias and low-RMSE elasticity estimates under distortion shocks, persistent measurement error, and feedback from distortions to productivity, while commonly used proxy-variable and lag-based estimators can become substantially biased. An application to manufacturing data yields plausible production elasticities and stable returns to scale. The estimated state-space system indicates that productivity is highly persistent, whereas distortions are more transitory and mean reverting, with statistically significant feedback effects on productivity.

JEL Classification Codes: C26, D24, C33, O47.

Keywords: Production function estimation; Input distortions; Heteroskedasticity-based identification; State-space productivity and distortions; Misallocation

Acknowledgments. We are grateful to Ulrich Doraszelski, Arthur Lewbel, Miguel León-Ledesma, Shigang Li, Jiawei Mo and Yuya Sasaki for helpful comments and suggestions on earlier versions of this paper.

*D.K. Christopoulos, Department of International and European Economic Studies; dchristop@aueb.gr

E. Tzavalis, Department of Economics, etzavalis@aueb.gr

1. Introduction

Estimating production functions in the presence of unobserved productivity is a central problem in applied microeconometrics. Because firms observe at least part of their productivity when choosing inputs, observed input choices are generally correlated with unobserved productivity shocks. This simultaneity problem renders ordinary least squares estimates of production elasticities inconsistent and has motivated a large literature on control-function and instrumental-variable approaches.

A major strand of this literature relies on proxy-variable methods. Following Olley and Pakes (1996), and Ackerberg et al., (2015), these estimators use observed firm decisions, such as investment or intermediate inputs, to control latent productivity. Their central insight is that, under appropriate timing and regularity conditions, firms' input policies can be inverted to recover productivity as a function of observables. This inversion-based approach has become widely used because it addresses simultaneity without requiring external instruments.

The validity of the inversion, however, depends on restrictive conditions. In particular, the proxy variable must be generated by a scalar unobservable and must be strictly monotone in that unobservable conditional on observed state variables. If input choices respond simultaneously to additional firm-level factors, such as credit constraints, taxes, subsidies, regulation, input-price wedges, optimization frictions, or demand shocks, the same observed proxy may correspond to multiple combinations of productivity and distortions. In that case, the proxy no longer delivers a one-to-one mapping to productivity. Ackerberg et al. (2015) further show that standard implementations of inversion-based estimators may suffer from first-stage functional-dependence problems, while Hu et al. (2020) emphasize that proxy variables contaminated by additional unobservables may fail to fully control for latent productivity.

This issue is not merely theoretical. A large literature documents pervasive distortions affecting firms' input allocation, particularly in developing and transition economies; see, among others, Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Midrigan and Xu (2014). When such distortions are present, variation in observed inputs partly reflects distortion-driven wedges rather than technological productivity. Standard proxy-variable estimators may therefore attribute distortion-induced input variation to productivity, potentially biasing estimates of production elasticities. More broadly, the consequences of distortions for estimated production elasticities are closely related to the literature emphasizing factor substitution, endogenous productivity dynamics, and biased technical change in production-function estimation; see, for example, Doraszelski and Jaumandreu (2013), León-Ledesma et al. (2010), and León-Ledesma and Satchi (2019).

An alternative interpretation of these distortions is that firms optimize with respect to latent effective input prices rather than observed market prices. Under this interpretation, distortions operate as wedges in firms' marginal input costs, so that the relevant flexible-input price faced by the firm differs from the price observed by the econometrician. Consequently, firms' input-demand policies depend jointly on technological productivity, effective input costs, and other latent state variables affecting optimization. This perspective connects the present framework to recent work emphasizing cost minimization, latent demand conditions, and imperfectly observed input-price wedges in production-function estimation; see, for example, Grieco et al. (2016), Hu et al. (2020), Doraszelski and Li (2025), and Doraszelski and Li (2026).

These concerns have motivated alternative approaches that relax or avoid exact proxy inversion. Hu et al. (2020) develop a framework based on two conditionally independent proxy variables treated as contaminated measures of latent productivity. Li and Mo (2026) avoid separating structural productivity from unexpected productivity shocks in the first stage and instead derive moment conditions using lagged instruments and composite productivity terms. Doraszelski and Li (2025) characterize the bias arising from imperfect inversion in proxy-variable estimators and show how enriched conditioning sets and Neyman-orthogonal moment conditions can mitigate the resulting errors. Doraszelski and Li (2026) further generalize the control-function approach by allowing productivity to evolve jointly with additional latent state variables, including demand shocks and other unobserved factors affecting firms' optimization problems.

These contributions substantially extend the production-function literature, but they also shift the identifying burden to alternative assumptions. Proxy-based methods require either exact invertibility or sufficiently informative first-stage conditioning sets to avoid bias in estimated elasticities. Approaches based on lagged instruments or composite productivity terms relax exact inversion but instead rely on orthogonality restrictions between lagged observables and productivity or measurement-error innovations, together with assumptions governing the dynamics of latent states. When distortions are persistent, correlated with productivity, or jointly evolve with productivity over time, these assumptions may themselves become difficult to justify empirically.

This paper proposes a complementary source of identification. We develop a two-step estimator for production functions under input distortions that uses heteroskedasticity-based internal instruments, in the spirit of Lewbel (1997, 2012), to identify production elasticities. Rather than treating distortions solely as a source of misspecification, the approach exploits the conditional heteroskedasticity generated by distortion shocks in flexible-input demand. Under timing-based orthogonality conditions, these variance shifts generate internal instruments for endogenous flexible inputs. The estimator therefore does not require proxy inversion, external instruments, or a restrictive one-to-one mapping between observed inputs and productivity.

The first contribution of the paper is to introduce Lewbel-type heteroskedasticity-based instruments into production-function estimation under distortions. To our knowledge, this source of identification has not previously been used in this context. The approach is particularly useful precisely in environments where flexible-input choices are jointly shaped by productivity and distortions. In such settings, the same forces that undermine proxy monotonicity generate the heteroskedastic variation needed for identification. Unlike lag-instrument approaches, identification does not rely primarily on lagged inputs or lagged composite productivity terms as instruments for the production equation.

The second contribution is to combine the first-step elasticity estimates with a state-space framework for productivity and distortions. Once the production elasticities are estimated, the residual system from the production and input-demand equations can be used to recover latent productivity and distortion components. The resulting state-space representation allows estimation of persistence, volatility, covariance, and feedback effects between productivity and distortions. This decomposition is useful not only for measuring productivity and distortions over time, but also for studying how violations of the identifying assumptions underlying alternative estimators translate into bias.

The paper uses this framework to analyze the consequences of failed inversion, invalid lag instruments, persistent measurement error, and dynamic feedback from distortions to productivity. Monte Carlo evidence shows that the proposed estimator remains centered near the true production elasticities across environments with distortion shocks, persistent measurement error, and feedback effects. By contrast, proxy-variable and lag-based estimators can become substantially biased when their identifying assumptions are weakened. In the empirical application to manufacturing data, the estimator yields plausible production elasticities and stable returns to scale. The state-space decomposition indicates that productivity is highly persistent, whereas distortions are more transitory and mean reverting, with limited but statistically significant feedback effects.

The remainder of the paper is organized as follows. Section 2 presents the conceptual framework and shows how input distortions undermine proxy inversion. Section 3 develops the heteroskedasticity-based identification strategy and the state-space estimator. Section 4 presents Monte Carlo evidence, including comparisons with proxy-variable and lag-instrument estimators. Section 5 applies the method to manufacturing data. Section 6 concludes.

2. Economic framework

We consider firms operating in an environment with potential resource misallocation that affects input choices. Firms produce output using predetermined capital and flexible inputs such as labor or intermediate materials.

Let firm j at time t produce output according to the technology

$$Y_{jt} = A_{jt} F(K_{jt}, M_{jt}; \beta), \tag{1}$$

where $A_{jt} \in \mathbb{R}_+$ denotes firm-specific productivity, $K_{jt} \in \mathbb{R}_+$is predetermined capital, $M_{jt} \in \mathbb{R}^{d_m}$ is a vector of flexible inputs chosen within the period, and β is the vector of production elasticities. We assume that firms observe A_{jt} before choosing M_{jt}, so that flexible inputs can respond immediately to productivity conditions specific to the firm, whereas capital reflects decisions taken in previous periods. This timing structure is standard in the empirical production function literature (see Olley and Pakes (1996), and Ackerberg et al. (2015), inter alia).

We allow productivity to contain both an aggregate and a firm-specific component. Aggregate shocks may reflect business-cycle conditions, industry-wide technological change, input-price movements, or broader macroeconomic disturbances. Denoting the aggregate component by D_t, total log productivity is written as

$$\log A_{jt} = D_t + \omega_{jt}^{y},$$

where $\omega_{jt}^{y} \in \mathbb{R}$ represents the firm-specific productivity component.

Beyond productivity, firms may face distortions that drive wedges between marginal products and effective input costs. These distortions may arise from credit constraints, tax policies, subsidies, political connections, regulatory frictions, or preferential access to resources. A large literature shows that such wedges generate dispersion in marginal products and reduce aggregate productivity.

Let $\omega_{jt}^{m} \in \mathbb{R}^{d_m}$ denote firm-specific input wedges affecting the effective marginal cost of flexible inputs. If ω_{jt}^{m} is interpreted as a log wedge, the firm's flexible-input choice can be represented as

$$\max_{M_{jt}} \left\{ P_{jt} A_{jt} F(K_{jt}, M_{jt}; \beta) - \left[\exp(\omega_{jt}^{m}) \odot P_{jt}^{M}\right]' M_{jt} \right\}, \tag{2}$$

where P_{jt}^{M} is the vector of observed flexible-input prices, and $\odot$ denotes the Hadamard (elementwise) product, that is, element-by-element multiplication. Thus, $\exp\left(\omega_{jt}^{m}\right) \odot P_{jt}^{M}$ is the vector of distortion-adjusted input prices.

The econometrician observes the market input-price vector P_{jt}^{M}, whereas firms optimize with respect to the latent effective input-price vector

$$P_{jt}^{M^*} = \exp(\omega_{jt}^{m}) \odot P_{jt}^{M}.$$

Thus, the distortion component ω_{jt}^{m} operates as an unobserved wedge between observed and effective marginal input costs. This interpretation connects the present framework to the literature that models distortions as latent shadow-cost wedges affecting firms' input choices.

This structure has two econometric implications. First, flexible inputs are endogenous because firms observe productivity when making input decisions. Second, input demand reflects not only productivity and observed state variables, but also unobserved wedges. The next subsection shows how this feature undermines the exact proxy-inversion logic used by standard proxy-variable estimators and motivates the alternative source of identification developed in Section 3.

2.1 Proxy identification and the role of distortions

Standard proxy-variable estimators rely on the ability to invert a proxy input variable to recover productivity. In the absence of distortions, the structural input policy can be written as

$$M_{jt} = \phi\left(\omega_{jt}^{y}, K_{jt}\right), \quad \frac{\partial \phi}{\partial \omega_{jt}^{y}} > 0,$$

so that productivity can be recovered through the inverse policy function

$$\omega_{jt}^{y} = \phi^{-1}(M_{jt}, K_{jt}).$$

This inversion step forms the basis of the proxy-variable approaches developed by Olley and Pakes (1996), and Ackerberg et al. (2015).

When distortions affect firms' effective input costs, however, the structural flexible-input policy becomes

$$M_{jt} = \tilde{\phi}\left(\omega_{jt}^{y}, \omega_{jt}^{m}, s_{jt}\right),$$

Let s_{jt}denote the vector of observable variables available to the econometrician that influence flexible-input demand or the distribution of input-demand disturbances. This vector always includes predetermined capital, K_{jt}, and may additionally contain lagged capital, observed output and input prices, aggregate conditions, industry-time controls, financial variables, and other observable firm characteristics. Formally,

$$s_{jt} = \left(K_{jt}, K_{j,t-1}, P_{jt}^{Y}, P_{jt}^{M}, D_{t}, \text{Ƶ}_{jt}\right)',$$

where P_{jt}^{Y} and P_{jt}^{M} denote output and flexible-input prices, $D_t$collects aggregate conditions, and Ƶ_{jt} denotes a vector of additional observed firm-level characteristics.

The conditioning vector $s_{jt}$is broader than the set of productive inputs entering the production technology. While only K_{jt} and M_{jt} directly enter the production function, the remaining elements of s_{jt} affect firms' input-demand decisions or the conditional variance of input-demand disturbances. In the empirical implementation, s_{jt} serves both as the conditioning set in the reduced-form flexible-input equation and as the source of observable shifters used to construct the heteroskedasticity-based instruments.

Firms optimize with respect to the latent effective input-price vector

$$P_{jt}^{M^*} = \exp(\omega_{jt}^{m}) \odot P_{jt}^{M},$$

while the econometrician observes P_{jt}^{M}. The proxy input therefore depends on multiple latent states. Even if the input policy is strictly increasing in productivity for a fixed realization of ω_{jt}^{m}, variation in ω_{jt}^{m} shifts the input-demand schedule. Holding productivity and observable state variables fixed, firms facing different realizations of the distortion process choose different flexible-input levels because distortions alter effective marginal input costs. Consequently, identical observed flexible-input choices may arise from different combinations of productivity and distortion states, so that the proxy input no longer uniquely indexes productivity.

The vector s_{jt} contains predetermined variables such as K_{jt}, but these variables should not be interpreted as strictly exogenous. Capital used in period t may have been chosen in earlier periods based on lagged productivity, lagged distortions, and other investment-relevant information. Thus, s_{jt} may be correlated with lagged latent states. The identification strategy developed below therefore does not rely on exact scalar invertibility of the proxy policy. Instead, it exploits heteroskedasticity-generated internal instruments constructed from observable state variables under the timing and orthogonality restrictions stated in Section 3.

The following proposition formalizes this failure of exact proxy inversion.

Proposition 1 (*Nonexistence of exact proxy inversion under distortions*). *Suppose the flexible-input policy satisfies*

$$M_{jt} = \tilde{\phi}(\omega_{jt}^{y}, \omega_{jt}^{m}, s_{jt}),$$

where $\tilde{\phi}(\cdot)$ *is strictly increasing in* ω_{jt}^{y} *for each fixed realization of* $(\omega_{jt}^{m}, s_{jt})$. *Assume further that* ω_{jt}^{m} *has non-degenerate support conditional on* s_{jt}, *that the policy depends nontrivially on* ω_{jt}^{m}, *and that the induced input-policy ranges overlap for at least two distortion realizations on a set with positive probability. Then, under the induced distribution of* (M_{jt}, s_{jt}), *there does not generally exist a single-valued function* $d(\cdot)$ *such that*

$$\omega_{jt}^{y} = d(M_{jt}, s_{jt})$$

almost surely.

Proof: See Appendix A1.

Proposition 1 shows that, when unobserved wedges shift input demand, the proxy input is not generally a sufficient statistic for productivity. This does not imply that production elasticities are necessarily unidentified, nor that proxy-based estimators must always be inconsistent. Rather, it shows that nonparametric productivity recovery through scalar proxy inversion requires additional restrictions. The econometric framework developed below therefore distinguishes between the failure of exact inversion and the identification of production elasticities, and proposes an alternative method based on heteroskedasticity-generated internal instruments.

2.2 Input choice, timing, and predetermined state variables

The noninevitability result in Proposition 1 arises because firms' flexible-input choices depend jointly on productivity, distortions, and other state variables affecting optimization. To clarify the economic environment underlying the identification strategy, this subsection describes the timing of firms' input decisions in the presence of latent distortions.

At the beginning of period t, firm j observes current productivity and the current flexible input wedge, together with the relevant state variables affecting input demand. The econometrician observes the subset s_{jt}, defined above, which contains observable variables entering the flexible-input policy or shifting the distribution of input-demand disturbances. Conditional on this information, the firm chooses flexible inputs such as materials and labor. Production then takes place according to the structural production function described above. After current-period production and profits are realized, the firm chooses investment, which determines next period's capital stock according to

$$K_{j,t+1} = (1 - \delta_K)K_{jt} + I_{jt},$$

where δ_K denotes the depreciation rate and I_{jt} denotes investment.

Let $\mathcal{F}_{j,t-1}$ denote the firm's information set at the end of period $t - 1$, after the investment decision determining K_{jt} has been made. We write this information set schematically as

$$\mathcal{F}_{j,t-1} = \{K_{jt}, K_{j,t-1}, M_{j,t-1}, Y_{j,t-1}, P^M_{j,t-1}, P^Y_{j,t-1}, \omega^y_{j,t-1}, \omega^m_{j,t-1}, \dots\}.$$

This notation is meant to capture the information available to the firm, not necessarily the information observed by the econometrician. In particular, $\mathcal{F}_{j,t-1}$ may contain lagged productivity and distortion states that are latent from the econometrician's perspective. Capital predeterminedness means that current capital is measurable with respect to this lagged information set:

$$K_{jt} \in \mathcal{F}_{j,t-1}.$$

Thus, current-period capital is predetermined but not necessarily exogenous. It may be chosen optimally in previous periods on the basis of lagged productivity, past input choices, and other investment-relevant information (see, e.g, Olley and Pakes (1996, Ackerberg et al (2015) and, more recently, Doraszelski and Li (2025, 2026)). Consequently, K_{jt} may be correlated with lagged latent states.

The observable vector s_{jt}, defined in Section 2.1, may contain both predetermined variables and contemporaneously observed controls. Let $x_{jt} \subseteq s_{jt}$ denote the predetermined observable subvector of s_{jt}. This subvector contains variables that are measurable with respect to $\mathcal{F}_{j,t-1}$.[1]

Let ξ_{jt} denote the current innovation vector in the latent productivity-distortion processes, formally specified in Section 3. The primitive timing restriction is

$$E[\xi_{jt} \mid \mathcal{F}_{j,t-1}] = 0.$$

[1]The vector x_{jt} is a predetermined subset of the broader conditioning vector s_{jt}. For example, if $s_{jt} = (K_{jt}, K_{j,t-1}, P_{jt}^{Y}, P_{jt}^{M}, D_t, \dots)'$, then a possible choice is $x_{jt} = (K_{jt}, K_{j,t-1}, M_{j,t-1}, Y_{j,t-1}, P_{j,t-1}^{M}, P_{j,t-1}^{Y}, D_t, \dots)'$. More generally, x_{jt}contains only variables that are predetermined with respect to the period-t innovations. Variables that may respond contemporaneously to those innovations can still belong to s_{jt}, but should enter x_{jt}only through lagged values. Thus, s_{jt} represents the full observable information set used to model flexible-input demand, whereas x_{jt} denotes the predetermined component used in constructing the baseline heteroskedasticity-based instruments. When all elements of s_{jt} are predetermined, the two vectors coincide, so that $x_{jt} = s_{jt}$.

This restriction says that current innovations to productivity and distortions are not predictable from the information available when current capital was predetermined. Since $x_{jt}$is measurable with respect to $\mathcal{F}_{j,t-1}$, the law of iterated expectations implies

$$E[x_{jt}\xi_{jt}] = 0.$$

More generally, any measurable function of predetermined observables is orthogonal to current innovations, provided the relevant moments exist.

Importantly, this implication applies to current innovations, not to persistent latent state levels. The timing restriction does not imply $E[\omega^{y}_{j,t-1} \mid x_{jt}] = 0$, nor does it imply $E[\omega^{m}_{j,t-1} \mid x_{jt}] = 0$. Indeed, because capital may be chosen based on lagged productivity and other investment-relevant information, $K_{jt} \in x_{jt}$ may be correlated with $\omega^{y}_{j,t-1}$. This distinction is important in environments with persistent distortions. Predeterminedness alone is not sufficient for the input-demand residual to satisfy $E\left[u^{m}_{jt} \mid x_{jt}\right] = 0$ when that residual contains a persistent distortion state. Section 3 therefore states the additional restrictions on the distortion dynamics and on the predetermined observable vector x_{jt} under which the input-demand residual is conditionally mean independent of the variables used to construct the heteroskedasticity-based instruments.

3. Econometric framework

3.1 Estimation and identification

This section develops a two-step estimator. The first step identifies production elasticities in the presence of endogenous inputs and distorted input choices using heteroskedasticity-based internal instruments, in the spirit of Lewbel (1997, 2012). The second step uses the residuals from the production and flexible-input equations to estimate a linear state-space system for latent productivity and distortions by quasi-maximum likelihood (QML) and Kalman smoothing.

We begin with the log production equation

$$y_{jt} = f(k_{jt}, m_{jt}; \beta) + u^{y}_{jt}, \qquad u^{y}_{jt} = \omega^{y}_{jt} + \varepsilon^{y}_{jt}, \tag{3}$$

where lowercase letters denote logarithms of the corresponding level variables. Thus, $y_{jt} \in R$ denotes log output, $k_{jt} = logK_{jt}$ denotes predetermined log capital, $m_{jt} \in R^{d_m}$ is the vector of log flexible inputs, and $\beta \in R^p$ is the vector of production elasticities. The composite production disturbance $u_{jt}^y \in R$ consists of a persistent productivity component, $\omega_{jt}^y \in R$, and a transitory disturbance, $\varepsilon_{jt}^y \in R$. When capital or flexible-input quantities appear in information sets or conditioning vectors, the level variables K_{jt}, M_{jt} , and their logarithmic counterparts k_{jt}, m_{jt}, , including lagged values, carry the same information sets, defined in Section 2.

A central difficulty is that predetermined variables are generally not valid conventional instruments for equation (3). Let $x_{jt} \in R^{d_x}$ denote the predetermined observable subvector of the broader observable conditioning vector $s_{jt} \in R^{d_s}$, as defined in Section 2.2. Predeterminedness implies orthogonality to current innovations, but not to persistent productivity levels. To see this, suppose productivity evolves according to

$$\omega_{jt}^y = \rho_{yy}\omega_{j,t-1}^y + \rho_{ym}'\omega_{j,t-1}^m + \xi_{jt}^y, \tag{4}$$

where $\rho_{yy} \in R$, $\rho_{ym} \in R^{d_m}$, and $\xi_{jt}^y \in R$. Even if $x_{jt}$is predetermined so that $E(\xi_{jt}^y \mid x_{jt}) = 0$, this does not imply $E(\omega_{jt}^y \mid x_{jt}) = 0$. The latter would be a much stronger condition. Current capital K_{jt}, equivalently current log capital k_{jt}, may have been chosen on the basis of lagged productivity information, and lagged flexible inputs $M_{jt},$ equivalently $m_{jt},$ may be correlated with lagged flexible-input distortions. Hence, variables such as k_{jt}, $k_{j,t-1}$, or $m_{jt},$ are hard to defend as ordinary instruments for the production equation when productivity is persistent and may depend on lagged distortions.

In equation (4), the coefficient ρ_{yy} captures standard productivity persistence, while the cross-effect ρ_{ym} allows distortions to affect productivity dynamically. Lagged distortions may influence future productivity through input misallocation, capacity utilization, investment, or learning. This feedback structure is consistent with evidence and models in the misallocation and financial-frictions literatures, where wedges are often persistent and systematically related to firm characteristics (see, e.g., Midrigan and Xu (2014), and David and Venkateswaran (2019)).

The strategy below therefore does not use x_{jt} directly as a level instrument for equation (3). Instead, x_{jt} is used to generate heteroskedasticity-based instruments from a reduced-form flexible-input equation. For flexible inputs, we write

$$m_{jt} = \Pi' s_{jt} + u_{jt}^m, \qquad u_{jt}^m = \omega_{jt}^m + \varepsilon_{jt}^m, \tag{5}$$

where $m_{jt} \in R^{d_m}$, $s_{jt} \in R^{d_s}$, $\Pi \in R^{d_s \times d_m}$, $u_{jt}^m \in R^{d_m}$, $\omega_{jt}^m \in R^{d_m}$, and $\varepsilon_{jt}^m \in R^{d_m}$. Equation (5) is interpreted as a reduced form or local projection of the structural flexible-input policy discussed in Section 2.[2] It is not the structural input policy itself. The vector s_{jt} may include capital variables, expressed in logs in the econometric specification, together with observed output and input prices, aggregate controls, industry-time controls, and other observed variables that affect flexible-input demand or the volatility of input-demand disturbances. The distortion component ω_{jt}^m represents wedges in the flexible-input demand condition, including regulatory distortions, tax wedges, subsidies, or other forms of resource misallocation that cause flexible-input usage to deviate from the frictionless optimum. A large literature documents the importance of such distortions for firms' input allocation and measured productivity. See, among others, Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Bartelsman et al. (2013), David and Venkateswaran (2019), and Li and Mo (2026).

The reduced-form input equation is useful because its residual contains variation in flexible-input distortions. The Lewbel instruments are constructed by interacting the residual u_{jt}^m with predetermined observables x_{jt}. For this construction to be valid, the input-demand residual must satisfy a mean-independence condition with respect to x_{jt}:

$$E(u_{jt}^m \mid x_{jt}) = 0.$$

Since $u_{jt}^m = \omega_{jt}^m + \varepsilon_{jt}^m$, this condition requires restrictions on the latent distortion state. In the baseline, the productivity-distortion process is specified as upper triangular. In particular, if

[2] In the structural model, optimal input demand depends on both productivity and distortions through the firm's first-order condition. Rather than estimating this nonlinear policy function directly, we approximate it locally by a linear function of observable state variables x_{jt}.

flexible-input distortions are persistent, the input-demand residual is not automatically mean independent of x_{jt}. Assume that ω_{jt}^{m} follows the AR(1) process

$$\omega_{jt}^{m} = \rho_{my}\omega_{j,t-1}^{y} + \rho_{mm}\omega_{j,t-1}^{m} + \xi_{jt}^{m}. \tag{6}$$

Then, predeterminedness gives $E(\xi_{jt}^{m} \mid x_{jt}) = 0$, but it does not generally imply $E(\omega_{j,t-1}^{m} \mid x_{jt}) = 0$, meaning conditional mean-independence between the lagged flexible-input distortion state and x_{jt}. Therefore, if flexible-input distortions are persistent, the input-demand residual is not automatically mean independent of x_{jt}. This is why, in the baseline, we impose the upper-triangular restriction $\rho_{my} = 0$ and require either $E(\omega_{j,t-1}^{m} \mid x_{jt})$ =0, or, in a special case, transitory distortions. Under upper-triangular dynamics, if $\rho_{mm} = 0$, then distortions are transitory, $\omega_{jt}^{m} = \xi_{jt}^{m}$, and the timing condition $E(\xi_{jt}^{m} \mid x_{jt}) = 0$, together with $E(\varepsilon_{jt}^{m} \mid x_{jt}) = 0$, is sufficient for $E(u_{jt}^{m} \mid x_{jt}) = 0$.

The restriction $\rho_{my} = 0$ means that the flexible-input wedge follows its own process and may be persistent but is not directly driven by lagged productivity. We allow lagged flexible-input distortions to affect future productivity through ρ_{ym}, but rule out the reverse dynamic channel from lagged productivity to future flexible-input distortions. This asymmetry is economically meaningful. A flexible-input wedge can reflect short-run effective input-cost disturbances, procurement frictions, working-capital constraints, tax or regulatory wedges, supplier disruptions, or temporary financing conditions. These forces may affect the cost or availability of flexible inputs without being mechanically generated by the firm's lagged productivity state. The remaining conditional mean-independence condition, $E(\omega_{j,t-1}^{m} \mid x_{jt}) = 0$, requires that the predetermined variables used to construct the Lewbel instruments are not systematically related to firms' lagged flexible-input wedges. Economically, this means that these predetermined variables should not be systematically driven by past flexible-input wedges. This condition is plausible when $\omega_{j,t}^{m}$ is interpreted as a flexible-input wedge or effective input-cost disturbance, rather than as a persistent investment wedge or capital-policy state.

The second ingredient for identification is conditional heteroskedasticity in the input-demand residual. The Lewbel construction requires that the variance of u_{jt}^{m}, but not its covariance with the production residual, varies systematically with the predetermined variables x_{jt}. The residual-level heteroskedasticity and relevance conditions are stated in Assumption 1 below, and a primitive state-dependent innovation process that generates them is discussed immediately after the assumption.

The identifying assumptions are summarized as follows.

Assumption 1 *(Timing, input-residual orthogonality, and heteroskedasticity)***.** For all firms $j$and periods t, assume:

(i) Predetermined observables and current innovations: The vector $x_{jt} \in R^{d_x}$ is predetermined with respect to period-t innovations. That is, x_{jt} is measurable with respect to the lagged firm information set $\mathcal{F}_{j,t-1}$, and $E(\xi_{jt} \mid \mathcal{F}_{j,t-1}) = 0$. Hence,

$$E(\xi_{jt}^{y} \mid x_{jt}) = 0 \qquad \text{and} \qquad E(\xi_{jt}^{m} \mid x_{jt}) = 0.$$

(ii) Upper triangular state dynamics: The latent productivity-distortion process satisfies the upper triangular restriction introduced above. Thus, lagged flexible-input distortions may affect future productivity, but lagged productivity does not directly enter the law of motion for current flexible-input distortions.

(iii) Input-residual mean independence: The reduced-form input-demand residual satisfies

$$\mathrm{E}(u_{jt}^{m} \mid x_{jt}) = 0.$$

Given $u_{jt}^{m} = \omega_{jt}^{m} + \varepsilon_{jt}^{m}$, a sufficient condition is $\mathrm{E}\big(\varepsilon_{jt}^{m} \mid x_{jt}\big) = 0$, together with either $\mathrm{E}\big(\omega_{j,t-1}^{m} \mid x_{jt}\big) = 0$ for persistent distortions, or $\rho_{mm} = 0$ for transitory distortions, in which case $\omega_{jt}^{m} = \xi_{jt}^{m}$.

(iv) Production transitory disturbance: The transitory production disturbance satisfies $E(\varepsilon_{jt}^{y} \mid x_{jt}) = 0$, and the transitory disturbances are serially uncorrelated and orthogonal to the relevant state innovations.

(v) Conditional covariance invariance: The covariance between the input-demand residual and the production residual is invariant with respect to the predetermined variables used to generate the Lewbel instruments:

$$E(u_{jt}^{m} u_{jt}^{y} \mid x_{jt}) = C_{my},$$

where $C_{my} \in R^{d_m}$ is a constant vector, not necessarily equal to zero. At the innovation level, we can allow the more general condition

$$E\left(\xi_{jt}^{m} \xi_{jt}^{y} \mid x_{jt} \right) = \sigma_{\xi^{y} \xi^{m}} \neq 0.$$

Thus, contemporaneous productivity and distortion innovations may be correlated. The identifying restriction is that this covariance is not systematically shifted by x_{jt}.

(vi) Conditional heteroskedasticity and relevance: The input-demand residual is conditionally heteroskedastic with respect to x_{jt}, and this heteroskedasticity is relevant:

$$E[\tilde{x}_{jt}(u_{jt}^{m} \odot u_{jt}^{m})'] \neq 0,$$

where $\tilde{x}_{jt} = x_{jt} - E(x_{jt}) \in R^{d_x}$. Thus, $\tilde{x}_{jt}(u_{jt}^{m} \odot u_{jt}^{m})' \in R^{d_x \times d_m}$. Equivalently, for at least one component $x_{jt,k}$, $k = 1, \dots, d_x$, and one component $u_{jt,d}^{m}$, $d = 1, \dots, d_m$, $Cov[x_{jt,k}, (u_{jt,d}^{m})^2] \neq 0$.

Assumption 1(ii) avoids requiring predetermined capital to be orthogonal to lagged productivity, since capital may be chosen using lagged productivity information. Assumption 1(iii) instead imposes the conditional mean-independence condition discussed above, $E(\omega_{j,t-1}^{m} \mid x_{jt}) = 0$, the variables used to generate the Lewbel instruments should not be systematically related to firms' lagged flexible-input wedges.

A sufficient simple example clarifies conditions (v) and (vi). Suppose that the heteroskedastic component of the input-demand residual is orthogonal to the production disturbance, while any covariance between the persistent productivity and distortion components is invariant with respect to x_{jt}. For instance, let

$$\xi_{jt}^{m} = h(x_{jt}; \gamma)^{1/2} \Sigma_e^{1/2} e_{jt},$$

where $0 < c_h \le h(x_{jt};\gamma) \le C_h < \infty, \Sigma_e \in R^{d_m \times d_m}$ is positive definite, $E(e_{jt} \mid x_{jt}) = 0$, and $e_{jt}$is orthogonal to ξ_{jt}^y and ε_{jt}^y conditional on x_{jt}. Then x_{jt} shifts the conditional variance of the input-demand residual but not its conditional covariance with the production residual. Thus, $Var\left(u_{jt}^m \mid x_{jt} \right)$ may vary with x_{jt}, while $E\left(u_{jt}^m u_{jt}^y \mid x_{jt} \right)$ remains constant.

This heteroskedasticity specification is economically natural. Financial frictions, regulatory burdens, tax exposure, policy interventions, procurement frictions, and collateral constraints need not affect all firms with equal volatility. Firms with different size, capital intensity, leverage, ownership structure, input intensity, supplier exposure, or regulatory exposure may therefore experience systematically different volatility in effective flexible-input wedges. Accordingly, $h(x_{jt};\gamma)$ can be interpreted as a firm-specific distortion-volatility index that scales the intensity of wedge innovations according to predetermined firm characteristics. This interpretation is consistent with empirical evidence showing substantial heterogeneity in firm-level distortions and financing conditions across firms and industries; see, for example, Moll (2014) and David and Venkateswaran (2019). In addition, models of financial amplification imply that shocks propagate asymmetrically depending on firms' balance-sheet strength and collateral constraints; see Kiyotaki and Moore (1997) and Ottonello and Winberry (2020). These arguments support a conditional heteroskedasticity specification in which predetermined firm characteristics affect volatility, but not necessarily the conditional mean, of flexible-input wedge innovations.

Under the maintained assumptions made for the innovations vector $\xi_{jt} = \begin{bmatrix} \xi_{jt}^y \\ \xi_{jt}^m \end{bmatrix} \in \mathbb{R}^{1+d_m}$, its variance-covariance matrix conditional on x_{jt} is given as

$$\Sigma_{\xi,jt} \equiv Var(\xi_{jt} \mid x_{jt}) = \begin{bmatrix} \sigma_{\xi^y}^2 & \sigma_{\xi^y\xi^m}{}' \\ \sigma_{\xi^m\xi^y} & \Sigma_{\xi^m,jt} \end{bmatrix},$$

where $\Sigma_{\xi^m,jt} = h(x_{jt};\gamma)\Sigma_e \in \mathbb{R}^{d_m \mathrm{x} d_m}$. We assume that the parameter space is restricted so that $\Sigma_{\xi,jt}$ is positive definite for all admissible values of x_{jt}. These restrictions ensure that the state innovation covariance matrix is well defined and uniformly nonsingular. The block-diagonal case, obtained by imposing $\sigma_{\xi^y\xi^m} = 0$, is a simplifying special case.

Under Assumption 1, and in particular the input-residual mean-independence, covariance-invariance, and heteroskedasticity-relevance conditions, the internal Lewbel-type instruments are constructed from the residuals of the reduced-form input equation. Define

$$Z_{jt}^{L} = \tilde{x}_{jt} \otimes u_{jt}^{m}, \tag{7}$$

where $Z_{jt}^{L} \in R^{d_x d_m}$. The predetermined variables x_{jt} are used to generate Z_{jt}^{L}. They are not used as conventional level instruments for the production equation. This is essential because x_{jt}, and especially capital, may be correlated with persistent productivity levels. The identifying moment is therefore based on the generated instruments Z_{jt}^{L}, not on x_{jt} itself.

Under Assumption 1, these instruments are orthogonal to the production disturbance:

$$E(Z_{jt}^{L} u_{jt}^{y}) = 0. \tag{8}$$

To see this, use the residual-level covariance-invariance condition:

$$E(Z_{jt}^{L} u_{jt}^{y}) = E[\tilde{x}_{jt} \otimes E(u_{jt}^{m} u_{jt}^{y} \mid x_{jt})] = E[\tilde{x}_{jt} \otimes C_{my}] = 0,$$

since $E(\tilde{x}_{jt}) = 0$, by construction. It does not require $C_{my} = 0$. Hence, nonzero contemporaneous covariance between productivity and distortion shocks is compatible with the exclusion condition for Z_{jt}^{L}, provided that this covariance is not systematically shifted by the predetermined variables used to generate the instruments.

Thus, the Lewbel-type instruments remain valid even if productivity and input-demand distortions are related through the latent-state system. The exclusion restriction does not require u_{jt}^{m} and u_{jt}^{y} to be uncorrelated.[3] It requires only that their covariance does not vary systematically with the predetermined variables used to generate the instruments. By contrast, Assumption 1(vi) requires the variance of the input-demand residual to vary with those variables. This asymmetry between conditionally invariant covariance and conditionally varying variance is the source of heteroskedasticity-based identification.

[3] For instance, a sudden flexible-input wedge shock, such as a temporary supply disruption or working-capital shock, may reduce capacity utilization or input quality within the period, generating contemporaneous correlation between distortion and productivity innovations. Conversely, an unexpected productivity innovation may temporarily affect effective input wedges by changing scale, capacity utilization, input demand, or working-capital needs.

The instruments are relevant because conditional heteroskedasticity generates correlation between Z_{jt}^L and the endogenous production inputs. The flexible-input equation implies that this correlation depends on moments involving

$$E[\tilde{x}_{jt}(u_{jt}^m \odot u_{jt}^m)'] \in R^{d_x \times d_m}.$$

However, relevance for the full vector of production elasticities requires a rank condition. Formally, if $\beta_0 \in R^p$ denotes the true parameter vector, identification requires

$$rank\ \left(E\left[Z_{jt}^L \frac{\partial f(k_{jt}, m_{jt}; \beta_0)}{\partial \beta'}\right]\right) = p,$$

where $\frac{\partial f(K_{jt}, m_{jt}; \beta_0)}{\partial \beta'} \in R^{1 \times p}$, and therefore $Z_{jt}^L \frac{\partial f(k_{jt}, m_{jt}; \beta_0)}{\partial \beta'} \in R^{d_x d_m \times p}$. See Appendix A2.

In the Cobb-Douglas case, where the production regressors are $(k_{jt}, m_{jt}')' \in R^{1+d_m}$, this requires the matrix

$$E[Z_{jt}^L(k_{jt}, m_{jt}')]$$

to have full column rank $1 + d_m$. Hence, identification of the capital elasticity does not rely on treating capital as exogenous. It relies on the validity and relevance of the generated heteroskedasticity-based instruments.

The production moment condition is therefore

$$E[Z_{jt}^L\{y_{jt} - f(k_{jt}, m_{jt}; \beta)\}] = 0. \tag{9}$$

Estimation of β can be carried out by GMM or IV using the sample analogue of this moment condition. In practice, the Lewbel-type instruments are constructed by replacing u_{jt}^m with the least-squares residuals from the reduced-form input equation and replacing $E(x_{jt})$ with the sample mean $\bar{x}$. Thus,

$$\hat{Z}_{jt}^L = (x_{jt} - \bar{x}) \otimes \hat{u}_{jt}^m,$$

where $\hat{Z}_{jt}^L \in R^{d_x d_m}$. Under standard first-stage consistency conditions, we have $\hat{Z}_{jt}^L = Z_{jt}^L + o_p(1)$. Inference should account for the generated-instrument error through heteroskedasticity-robust or bootstrap standard errors. Inference should account for the generated-instrument error through an appropriately constructed heteroskedasticity-robust or bootstrap procedure.

This completes the first step of the procedure. Production elasticities are identified without requiring proxy-input monotonicity, external instruments, or exact recovery of productivity. The next section uses the generated residuals from the production and flexible-input equations to estimate the latent productivity-distortion system.

3.2 State-space estimation of the productivity and distortion components

Having obtained consistent estimates of the production elasticities, the second step uses the generated residuals from the production and flexible-input equations to estimate the latent productivity-distortion system. The purpose of this step is not to re-identify the production elasticities, but to decompose residual variation into persistent productivity, persistent input distortions, and transitory disturbances.

Conditional on the first-step estimates, the state-space parameters are estimated by maximizing the Gaussian quasi-likelihood based on the Kalman filter. The nonlinearity of $h(x_{jt};\gamma)$ affects only the observed, time-varying innovation covariance and does not make the state-space system nonlinear (see also Appendix A3). Conditional on x_{jt} and the parameter vector, the innovation covariance matrix is known at each step of the Kalman recursion. The parameters estimated in this step include the persistence and feedback parameters in ρ, the innovation variances and covariances, the transitory variances in $\Sigma_{\xi,jt}$, the heteroskedasticity parameters in $h(x_{jt};\gamma)$ and the transitory variances $\sigma^2_{\varepsilon^y}, \Sigma_{\varepsilon^m}$. The boundedness and positive-definiteness restrictions stated before ensure that the covariance matrices entering the Kalman likelihood are well defined and nonsingular. Denote the resulting estimator by

$$\hat{\theta}_{IVSS} = \left(\hat{\rho}, \hat{\alpha}', \hat{\sigma}^2_{\xi^y}, \sigma_{\xi^y\xi^m}, \hat{\gamma}', \hat{\Sigma}_e, \hat{\sigma}^2_{\varepsilon^y}, \hat{\Sigma}_{\varepsilon^m}\right)'.$$

The Kalman filter and fixed-interval smoother are applied to the generated residual vector

$$\hat{u}_{jt} = \begin{bmatrix} \hat{u}^y_{jt} \\ \hat{u}^m_{jt} \end{bmatrix} \in \mathbb{R}^{1+d_m}.$$

They provide firm-time filtered and smoothed estimates of the latent productivity and distortion components,

$$\widehat{\omega}^{\eta}_{jt|t} \quad \text{and} \quad \widehat{\omega}^{\eta}_{jt|T}, \quad \eta \in \{y, m\},$$

along with their associated conditional covariance matrices.

This two-step structure has two advantages. First, production elasticities are identified in the first step using Lewbel-type internal instruments, without requiring proxy inversion, external instruments, or exact recovery of productivity. Second, conditional on these elasticity estimates, the state-space step exploits the dynamic structure of the generated residuals to recover latent productivity and distortion paths, estimate their persistence, covariance, and feedback effects, and estimate the heteroskedasticity function $h(x_{jt}; \gamma)$. Thus, the first step provides robust elasticity estimates under relatively weak moment restrictions, while the second step uses additional dynamic structure to study the evolution and volatility of productivity and distortions.

In particular, the Kalman-based likelihood decomposes $\hat{u}^{y}_{jt}$ and $\hat{u}^{m}_{jt}$ into persistent and transitory components and yields estimates of the conditional innovation variance

$$\hat{\Sigma}_{\xi,jt} = \Sigma_{\xi,jt}\big(x_{jt}; \hat{\gamma}, \hat{\alpha}, \hat{\sigma}^2_{\xi^y}, \hat{\sigma}_{\xi^y \xi^m}, \hat{\Sigma}_e\big),$$

thereby linking distortion volatility directly to observable firm characteristics.

The following proposition establishes the asymptotic normality of the estimator $\hat{\theta}_{IVSS}$.

Proposition 2 (Asymptotic normality of the IVSS estimator, $\hat{\theta}_{IVSS}$**).** *Let $\theta \in \Theta \subset \mathbb{R}^{d_\theta}$ denote the vector of second-step state-space parameters. Suppose Assumption 1 holds. In addition, assume that: (i) T is fixed and $N \to \infty$, with independence across firms, (ii) the relevant fourth moments exist, (iii) the first-step estimators used to construct the generated residuals are $\sqrt{N}$-consistent and asymptotically linear and (iv) the second-step Gaussian quasi-likelihood satisfies standard smoothness, identification, and nonsingularity conditions.*

Then the two-step IVSS estimator $\hat{\theta}_{IVSS}$, obtained from the second-step score equations based on the generated residuals, satisfies

$$\sqrt{N}(\hat{\theta}_{IVSS} - \theta_0) \xrightarrow{d} N(0, V_\theta), \tag{10}$$

where $V_\theta = H(\theta_0)^{-1}\Omega H(\theta_0)^{-1\prime}, H(\theta_0) = -E\left[\frac{\partial S_j(\theta_0;u_j)}{\partial\theta'}\right]$, *and*

$$u_{jt} \equiv (u_{jt}^{y}, u_{jt}^{m\prime})' \in \mathbb{R}^{1+d_m}, and\ \ u_j \equiv (u_{j1}', \dots, u_{jT}')' \in \mathbb{R}^{T(1+d_m)},$$

$S_j(\theta; u_j)$ *denotes firm j's contribution to the Gaussian quasi-score associated with the state-space likelihood based on the stacked disturbance vector* u_j*, and Ω is the asymptotic variance matrix of the second-step score, augmented to account for the use of generated residuals in the first step. A closed-form expression for Ω is given in Appendix A4.*

Proof: See Appendix A4.[4]

For inference, consistent estimation of V_θ should account for heteroskedasticity, within-firm dependence, and the use of generated residuals in the second step. One approach is to construct a plug-in sandwich estimator using firm-level score contributions, augmented by the influence-function contribution from the first-step estimates. This estimator can be made robust to arbitrary heteroskedasticity and within-firm dependence by clustering at the firm level. Alternatively, inference may be based on a firm-level wild bootstrap that preserves the within-firm time-series structure. In each bootstrap replication, the firm-level contributions are reweighted by independent mean-zero random variables, such as Rademacher weights, with the first- and second-step estimation uncertainty carried through the bootstrap procedure. Under fixed *T* and large *N*, both the appropriately constructed firm-cluster-robust sandwich estimator and the cluster wild bootstrap provide consistent inference for the IVSS estimator. In the empirical application in Section 5, we report firm-cluster-robust sandwich standard errors.

3.3 Bias of proxy and lag-based estimators under the state process

This subsection examines the behavior of proxy-based estimators under the latent-state framework developed above. Proposition 1 established that when flexible-input demand depends jointly on productivity and distortions, the proxy variable generally cannot be inverted to recover firm-level productivity exactly. Importantly, however, failed invertibility does not automatically imply

[4] Although the full system could in principle be estimated by full-information maximum likelihood, we adopt the two-step procedure because it separates identification of the production elasticities from latent-state estimation, is computationally much more tractable, and relies on weaker parametric assumptions.

inconsistent estimation of production elasticities. The key issue is whether the second-stage moment conditions remain valid when the productivity control is only an imperfect predictor of the underlying productivity state.

To address this question, we characterize the population bias of the estimators of Olley and Pakes (1996) and Ackerberg, Caves, and Frazer (2015) (ACF) under our state-space process and derive conditions under which these estimators remain consistent. We also consider the adjustment proposed by Doraszelski and Li (2025) (DL), which augments the first-stage information set to mitigate the prediction-error bias generated by failed invertibility. DL shows that failed invertibility creates a first-stage prediction error that contaminates the second-stage moment conditions, and that enriching the first-stage information set with variables that later serve as second-stage instruments can eliminate or substantially reduce this source of bias.

Let X_{jt}^P denote the information set used by a proxy estimator to construct the productivity control function. A generic proxy specification takes the form

$$X_{jt}^P = \{k_{jt}, m_{jt}, t, \dots\},$$

although the exact composition of the information set depends on the estimator and application. Some specifications may additionally include labor or other flexible inputs, i.e., $X_{jt}^P = \{k_{jt}, m_{jt}, l_{jt}, t, \dots\}$. More generally, the productivity control function is approximated flexibly using the variables contained in X_{jt}^P. We also consider the DL modification, which augments the proxy-information set with the lead of capital. Denoting the enriched information set by X_{jt}^D,

$$X_{jt}^D = X_{jt}^P \cup \{k_{j,t+1}\}.$$

The inclusion of $k_{j,t+1}$ expands the information available in the first-stage projection and is intended to reduce the prediction-error component arising from failed invertibility.

When exact inversion holds, the proxy control recovers the productivity state. When inversion fails, the first-stage proxy recovers only the best predictor of productivity based on the available information,

$$\widetilde{\omega}_{jt}^y = E(\omega_{jt}^y \mid X_{jt}^P),$$

with associated prediction error

$$\zeta_{jt} = \omega_{jt}^{y} - \widetilde{\omega}_{jt}^{y}.$$

To illustrate the mechanism, consider a linear-Gaussian approximation in which, after controlling for the observed determinants of flexible-input demand, the relevant proxy information is summarized by the flexible-input residual u_{jt}^{m}. This special case is used solely for intuition; Proposition 3 itself does not rely on linearity or Gaussianity. In the scalar case, $u_{jt}^{m} \in \mathbb{R}$, the linear projection is

$$\widetilde{\omega}_{jt}^{y} = E(\omega_{jt}^{y} \mid u_{jt}^{m}) = \pi u_{jt}^{m},$$

where

$$\pi = \frac{\mathrm{Cov}(\omega_{jt}^{y}, u_{jt}^{m})}{\mathrm{Var}(u_{jt}^{m})} = \frac{\mathrm{Cov}(\omega_{jt}^{y}, \omega_{jt}^{m})}{\mathrm{Var}(\omega_{jt}^{m}) + \mathrm{Var}(\varepsilon_{jt}^{m})},$$

using $u_{jt}^{m} = \omega_{jt}^{m} + \varepsilon_{jt}^{m}$ and assuming that ε_{jt}^{m} is orthogonal to the latent states. Thus, the proxy variable is informative about productivity only through the covariance structure linking the productivity and distortion states. This covariance is governed by the transition. Proposition 3 generalizes this intuition to the multivariate setting and shows how these latent-state dependencies generate population bias in proxy-based estimators.

Proposition 3 (Bias of proxy-based estimators). *Under the maintained assumptions of the model and the latent VAR(1) state process, consider a proxy-based estimator that uses the proxy control* $\widetilde{\omega}_{jt}^{y}$ *and the second-stage instrument vector* Z_{jt}^{P}*, containing predetermined and lagged variables. Let* $r_{jt}(\beta_0) = \frac{\partial f(k_{jt}, m_{jt}; \beta_0)}{\partial \beta} \in \mathbb{R}^{p}$ *and* $Z_{jt}^{P} \in \mathbb{R}^{q}$*. Suppose that* $E(Z_{jt}^{P} \xi_{jt}^{y}) = 0$*. Then, in a just-identified linear IV representation, with* $q = p$*, the proxy estimator satisfies*

$$\beta_{IV}^{P} - \beta_0 = \left(E\left(Z_{jt}^{P} r_{jt}(\beta_0)'\right)\right)^{-1} \left(\rho_{yy} E(Z_{jt}^{P} \zeta_{j,t-1}) + E\left(Z_{jt}^{P} \left(\omega_{j,t-1}^{m}\right)'\right) \rho_{ym}\right). \qquad \textit{(11a)}$$

For an overidentified GMM estimator, with $q \geq p$*, the analogous local approximation is*

$$\beta_{GMM}^{P} - \beta_0 \simeq (D'WD)^{-1} D'W \left(\rho_{yy} E\left(Z_{jt}^{P} \zeta_{j,t-1}\right) + E\left(Z_{jt}^{P} \left(\omega_{j,t-1}^{m}\right)'\right) \rho_{ym}\right), \qquad \textit{(11b)}$$

where $D = E\left(Z_{jt}^{P} r_{jt}(\beta_0)'\right) \in \mathbb{R}^{q\times p}$ and $W \in \mathbb{R}^{q\times q}$ is the GMM weighting matrix.

Proof. See Appendix A5.

The results of Proposition 3 clarify the conditions under which proxy-based estimators and the DL adjustment provide consistent estimates of β. The proposition shows that the asymptotic bias of a proxy estimator contains two distinct components. The first component,

$$\rho_{yy} E(Z_{jt}^{P} \zeta_{j,t-1}),$$

is the proxy prediction-error channel. It arises because failed inversion implies that the first-stage proxy recovers $\widetilde{\omega}_{jt}^{y}$ rather than ω_{jt}^{y}. This is the channel emphasized by Doraszelski and Li (2025). If the resulting prediction error is correlated with the instruments used in the second stage, the moment conditions become invalid.

The second component,

$$E\left[Z_{jt}^{P}\left(\omega_{j,t-1}^{m}\right)'\right]\rho_{ym},$$

is the latent-distortion-state channel. This term arises because the state-space process allows current productivity to depend on lagged distortions. The channel is absent in standard scalar productivity processes but emerges naturally when productivity and distortions evolve jointly. Consequently, even if the prediction-error channel is eliminated, a proxy estimator may remain inconsistent whenever the instruments are exposed to the lagged distortion state.

Equations (11a) and (11b) imply that proxy-based estimators are consistent when both bias channels vanish. A strong sufficient condition is

$$\rho_{yy} = \rho_{ym} = 0,$$

together with the maintained timing restrictions $E(Z_{jt}^{P} \xi_{jt}^{y}) = E(Z_{jt}^{P} \varepsilon_{jt}^{y}) = 0$, and the usual rank condition. If $\rho_{yy} = 0$, lagged proxy prediction errors do not propagate into current productivity. If $\rho_{ym} = 0$, lagged distortions do not affect current productivity.

A less restrictive and more relevant case allows productivity to be persistent, so that $\rho_{yy} \neq 0$. In this case, consistency requires the proxy prediction error to be orthogonal to the second-stage instruments, $E(Z_{jt}^{P} \zeta_{j,t-1}) = 0$, and also requires the distortion-state channel to vanish, i.e.,

$E\left[Z_{jt}^{P}\left(\omega_{j,t-1}^{m}\right)'\right]\rho_{ym}$. Thus, failed proxy inversion does not by itself imply inconsistency. The proxy estimator can remain consistent if the unrecovered component of productivity is orthogonal to the instruments and either lagged distortions do not enter current productivity, $\rho_{ym} = 0$, or the instruments are orthogonal to the lagged distortion state. The first condition is the channel emphasized by Doraszelski and Li (2025); enriching the proxy-information set can reduce or eliminate the prediction-error component by making $\zeta_{j,t-1}$ orthogonal to the second-stage instruments. This may occur, for example, when the second-stage instruments are contained in the lagged proxy-information set.

The DL adjustment is therefore effective against the prediction-error channel, but Proposition 3 shows that it does not, in general, eliminate the distortion-feedback channel. Once the prediction-error channel is removed, the remaining bias is

$$\beta_{IV}^{DL} - \beta_0 = \left[E(Z_{jt}^{P}X_{jt}')\right]^{-1} E\left[Z_{jt}^{P}\left(\omega_{j,t-1}^{m}\right)'\right]\rho_{ym}.$$

Hence, the DL estimator is consistent if either $\rho_{ym} = 0$, so that lagged distortions do not enter current productivity, or $E\left[Z_{jt}^{P}\left(\omega_{j,t-1}^{m}\right)'\right] = 0$. The latter restriction may fail when the instrument set includes lagged flexible inputs or other variables that inherit persistence from flexible-input distortions. Thus, enriching the proxy-information set can remove the prediction-error channel, but it does not by itself remove the distortion-feedback channel.

4. Monte Carlo evidence

4.1 Finite-sample performance of the production-elasticity estimators

In this section, Monte Carlo experiments are used to evaluate the finite-sample performance of the proposed heteroskedasticity-based instrumental variables–state-space estimator, denoted IVSS. The simulations are designed to study environments in which flexible-input choices are affected by latent input distortions and, as a result, the one-to-one mapping between flexible inputs and productivity that underlies standard proxy-control estimators may fail.

The main Monte Carlo design is based on a two-shifter data-generating process. The first shifter is predetermined capital k_{jt} and the second is an additional predetermined shifter, q_{jt}, which affects the conditional variance of the flexible-input distortion innovation and is used to construct the Lewbel-type instruments. Thus, the observed shifter vector is $x_{jt} = (k_{jt}, q_{jt})'$. The shifter q_{jt} affects flexible-input demand, but it satisfies the conditional mean-independence condition with respect to the lagged flexible-input distortion state.

Output is generated from the Cobb–Douglas production function

$$y_{jt} = \beta_k k_{jt} + \beta_m m_{jt} + \omega_{jt}^{y} + \varepsilon_{jt}^{y}.$$

The true production elasticities are set to $\beta_k = 0.300$, $\beta_m = 0.500$. The transitory production disturbance is generated as $\varepsilon_{jt}^{y} \sim IIDN\left(0, \sigma_{\varepsilon^y}^2\right)$, with $\sigma_{\varepsilon^y} = 0.18$. The capital-related shifter evolves according to a productivity-dependent autoregressive process,

$$k_{jt} = \rho_k k_{j,t-1} + \kappa_y \omega_{j,t-1}^{y} + u_{jt}^{k},$$

where $\rho_k = 0.400$, $\kappa_y = 0.030$, and $u_{jt}^{k} \sim IIDN\left(0, \sigma_{u^k}^2\right)$, with $\sigma_{u^k} = 0.447$. This specification allows capital to be predetermined but not strictly exogenous, since current capital may depend on lagged productivity.

The second predetermined shifter follows the following process:

$$q_{jt} = \rho_q q_{j,t-1} + u_{jt}^{q},$$

where $q_{jt}$is generated independently of the productivity and distortion innovations. The value of ρ_q is set to $\rho_q = 0.4$. The innovation u_{jt}^{q} is generated independently of the productivity and distortion innovations.

Flexible input demand is generated as

$$m_{jt} = \theta_1 k_{jt} + \theta_2 k_{j,t-1} + \theta_3 q_{jt} + \omega_{jt}^{m} + \varepsilon_{jt}^{m}.$$

The slope coefficients in the flexible-input equation are set to $\theta_1 = 0.6$, $\theta_2 = 0.3$ and $\theta_3 = 0.8$. The transitory flexible-input disturbance satisfies $\varepsilon_{jt}^{m} \sim IIDN(0, \sigma_{\varepsilon^m}^2)$, with $\sigma_{\varepsilon^m} = 0.45$.

The latent productivity and distortion components follow the upper-triangular state process

$$\begin{bmatrix}\omega_{jt}^{y}\\ \omega_{jt}^{m}\end{bmatrix}=\begin{bmatrix}\rho_{yy} & \rho_{ym}\\ 0 & \rho_{mm}\end{bmatrix}\begin{bmatrix}\omega_{j,t-1}^{y}\\ \omega_{j,t-1}^{m}\end{bmatrix}+\begin{bmatrix}\xi_{jt}^{y}\\ \xi_{jt}^{m}\end{bmatrix},$$

with the persistence parameters are calibrated as $\rho_{yy} = 0.7$ and $\rho_{mm} = 0.4$. The coefficient ρ_{ym} measures the dynamic feedback from lagged flexible-input distortions to current productivity. The baseline value is $\rho_{ym} = 0.25$. This parameter is central to the Monte Carlo design. When $\rho_{ym} \neq 0$, lagged distortions become part of the productivity process. Consequently, estimators that rely on lagged inputs, lagged proxy variables, or lagged composite productivity as instruments may become biased if those variables are correlated with the lagged distortion state.

Productivity innovation is generated as $\xi_{jt}^{y} \sim IIDN(0, \sigma_{\xi^y}^2)$, with $\sigma_{\xi^y} = 0.5$, while the distortion innovation is conditionally heteroskedastic. Specifically,

$$\xi_{jt}^{m} = h(k_{jt}, q_{jt}; \gamma)^{1/2} e_{jt},$$

where $e_{jt} \sim IIDN(0,1)$, $e_{jt} \perp \xi_{jt}^{y}$, and

$$h(k_{jt}, q_{jt}; \gamma) = \exp\left(\gamma_0 + \gamma_k k_{jt} + \gamma_q q_{jt}\right).$$

These assumptions imply a block-diagonal innovation covariance matrix $\Sigma_{\xi,jt}$ for the simulations, i.e., $\sigma_{\xi^m\xi^y} = 0$.

Throughout the simulations, we assume $\gamma_0 = 0$ and the benchmark heteroskedasticity specification uses $\gamma_k = 0.25$ and $\gamma_q = 0.5$. For $k_{jt} = 0$ and $q_{jt} = 0$, this normalization implies $\mathrm{Var}\left(\xi_{jt}^{m} \mid k_{jt} = 0, q_{jt} = 0\right) = 1$.

Let $\hat{u}_{jt}^{m}$ denote the residual from this reduced-form input equation. The generated Lewbel instruments are

$$z_{jt}^{k} = \left(k_{jt} - \bar{k}\right)\hat{u}_{jt}^{m} \quad \text{and} \quad z_{jt}^{q} = \left(q_{jt} - \bar{q}\right)\hat{u}_{jt}^{m},$$

where $\bar{k}$ and $\bar{q}$ are the sample means of k_{jt} and q_{jt}, respectively, The Monte Carlo experiment consists of 500 replications of panels with $N = 500$ firms and $T = 8$ periods. In each replication, we estimate the production elasticities using IVSS and compare its performance with three alternative benchmark estimators discussed in Section 3: the proxy-control estimator of Ackerberg et al. (2015), denoted ACF; the composite-productivity estimator of Li and Mo (2026), denoted LM; and the enriched proxy-control estimator of Doraszelski and Li (2025), denoted DL.

In the Monte Carlo design, the proxy-information set used by the proxy-based estimators is $X_{jt}^{P} = \{k_{jt}, m_{jt}, q_{jt}\}$. Time effects are included in the implementation but suppressed from the notation. Since flexible input demand depends jointly on productivity, ω_{jt}^{y}, and the latent distortion state, ω_{jt}^{m}, the same observed flexible input may correspond to different combinations of productivity and distortions. This weakens the proxy-inversion logic and generates the proxy prediction-error channel discussed in Proposition 3.

The LM estimator avoids proxy inversion by using composite productivity and moment conditions based on lagged variables. However, as discussed in Section 3, these moments can become invalid when lagged distortions enter current productivity through

$$\omega_{jt}^{y} = \rho_{yy}\omega_{j,t-1}^{y} + \rho_{ym}\omega_{j,t-1}^{m} + \xi_{jt}^{y},$$

because lagged inputs and lagged composite productivity may be correlated with $\omega_{j,t-1}^{m}$when $\rho_{ym} \neq 0$.

The DL estimator addresses failed invertibility by enriching the first-stage proxy-information set. In the Monte Carlo implementation, DL augments X_{jt}^{P} with the lead of capital, $k_{j,t+1}$, so that

$$X_{jt}^{D} = X_{jt}^{P} \cup \{k_{j,t+1}\} = \{k_{jt}, m_{jt}, q_{jt}, k_{j,t+1}\}.$$

This enrichment is designed to mitigate the proxy prediction-error channel. The difference from IVSS is that DL remains an enriched proxy-control/GMM estimator, whereas IVSS identifies production elasticities using heteroskedasticity-generated internal instruments based on the input-equation residual. Thus, DL reduces the consequences of failed proxy inversion, while IVSS avoids proxy inversion and does not rely on lagged flexible-input or lagged composite-productivity instruments. This distinction matters when $\rho_{ym} \neq 0$, because persistent distortions may still contaminate lagged instruments even if the proxy prediction-error channel is reduced.

The Monte Carlo analysis considers four sets of scenarios. First, the baseline design sets $\rho_{ym} = 0.25$, so that lagged distortions affect current productivity. Second, we consider the no-feedback case, $\rho_{ym} = 0$, in which distortions do not enter the productivity process. This scenario is useful because it shuts down the dynamic distortion channel and therefore provides a benchmark in which the identifying restrictions of the competing estimators are less likely to fail. Third, we vary the feedback coefficient by considering $\rho_{ym} = -0.25$ and $\rho_{ym} = 0.5$. These cases allow distortions to have either negative or stronger positive effects on future productivity. Finally, we vary the strength of heteroskedasticity by setting $\gamma_q = 0.15$ for weak heteroskedasticity and $\gamma_q = 0.8$ for strong heteroskedasticity.

Table 1 reports the baseline general design (see Panel A), where the distortion-to-productivity feedback parameter is set to $\rho_{ym} = 0.25$. In this case, lagged flexible-input distortions enter current productivity, so the latent distortion state directly affects the productivity process. The IVSS estimator remains close to the true production elasticities. The biases are 0.042 for the capital elasticity, β_k, and 0.029 for the flexible-input elasticity, β_m. By contrast, the ACF estimator displays a small bias for β_k, equal to -0.011, but overestimates β_m, with a bias of 0.102. The LM estimator is substantially biased, overestimating β_k by 0.191 and underestimating β_m by -0.423. The DL estimator reduces the magnitude of the capital-elasticity bias relative to LM, but still displays sizeable bias, especially for β_m, where the bias is 0.265. These results support the theoretical mechanism discussed in Section 3: when persistent distortions affect productivity dynamics, proxy-based and lag-based moment conditions may become contaminated by the latent distortion state.

Panel B of Table 1 reports the second-step IVSS estimates of the latent state-space parameters. These estimates are not used as competing production-elasticity estimates; rather, they assess how well the second step recovers the latent productivity-distortion dynamics after the production elasticities have been estimated. The dynamic parameters are recovered accurately. The productivity persistence parameter, ρ_{yy}, has bias $-0.003$and RMSE 0.029, while the distortion persistence parameter, ρ_{mm}, has bias $-0.017$and RMSE 0.039. The feedback coefficient from distortions to future productivity, ρ_{ym}, is also estimated precisely, with bias $0.002$and RMSE 0.019. The heteroskedasticity parameters are somewhat less precisely estimated, with RMSE values of 0.111 for γ_k and 0.102 for γ_q, reflecting the greater difficulty of recovering conditional variance parameters in a short panel. Overall, the results show that the second-step state-space estimator recovers both the latent dynamics and the main variance structure well.

Table 1. Monte Carlo results– Baseline scenario

Panel A: Estimates of production elasticities by alternative methods

	IVSS		**ACF**		**LM**		**DL**	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.042	0.029	-0.011	0.102	0.191	-0.423	-0.088	0.265
RMSE	0.061	0.050	0.060	0.104	0.196	0.424	0.108	0.267

Panel B: IVSS estimates of the state-space parameters

	ρ_{yy}	ρ_{mm}	ρ_{ym}	σ_{ξ^y}	σ_{ε^y}	γ_k	γ_q	σ_{ξ^m}
True	0.70	0.40	0.25	0.25	0.447	0.20	0.80	0.50
Bias	-0.003	-0.017	0.002	-0.001	0.005	-0.036	-0.063	-0.025
RMSE	0.029	0.039	0.019	0.029	0.045	0.111	0.102	0.066

Notes: The table reports the true parameter values together with the bias and root mean squared error (RMSE) of the estimates obtained from the Monte Carlo simulations. Bias and RMSE are computed across replications for each estimator. The heteroscedasticity-based IV state-space (IVSS) estimator combines Lewbel-type internal instruments, predetermined observable variables and exploiting heteroskedasticity in distortion shocks. The ACF proxy estimator uses the flexible input directly as a proxy and includes capital linearly, allowing both input elasticities to be identified. The LM estimator relies on orthogonality between unexpected productivity shocks and lagged flexible inputs. The DL estimator is an enriched proxy-control/GMM estimator, which augments the first-stage proxy-information set with the lead of capital, $k_{j,t+1}$. Panel A reports estimates of the production elasticities β_k and β_m under the alternative estimators. Panel B reports estimates of the state-space system parameters obtained from the IVSS estimator. Results are based on 500 replications with N = 500 firms and T = 8 periods. The Monte Carlo DGP imposes $\sigma_{\xi^m\xi^y} = 0$; hence the reported state-space parameters correspond to the block-diagonal case of $\Sigma_{\xi,jt}$.

Table 2 considers the no-feedback scenario, $\rho_{ym} = 0$. This case shuts down the dynamic channel through which lagged distortions enter current productivity. As expected, all estimators improve substantially. IVSS remains accurate, with biases of 0.037 for β_k and 0.002 for β_m. ACF also performs well, with biases of 0.038 and 0.001. LM becomes essentially unbiased, with biases of

0.000 and 0.001, and has the lowest RMSE values in this scenario. DL also performs well, with biases of 0.048 and 0.004. These results are important because they confirm that the poor performance of the competing estimators in the baseline design is not mechanical. It arises mainly when distortions feed dynamically into productivity and thereby contaminate the lagged variables or proxy controls used for identification.

Table 2. Monte Carlo results – No distortion scenario ($\rho_{ym} = 0$)

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.037	0.002	0.038	0.001	0.000	0.001	0.048	0.004
RMSE	0.052	0.036	0.056	0.015	0.023	0.018	0.066	0.032

Notes: The table reports Monte Carlo bias and root mean squared error (RMSE) for the production elasticities β_k and β_m under the no-distortion-feedback scenario, $\rho_{ym} = 0$. In this case, lagged flexible-input distortions do not enter the productivity process. Results are reported for the IVSS, ACF, LM, and DL estimators. Bias and RMSE are computed across 500 Monte Carlo replications with $N = 500$ firms and $T = 8$ periods.

Table 3 examines robustness to the strength of heteroskedasticity. Panel A considers a less strong, but still present, degree of heteroskedasticity, with $\gamma_1 = 0.15$. Even in this case, IVSS performs well. Its biases are 0.034 for β_k and 0.039 for β_m, showing that the estimator does not require very strong heteroskedasticity to remain informative. Panel B considers stronger heteroskedasticity, with $\gamma_1 = 0.8$. The IVSS biases remain very similar, equal to 0.036 for β_k and 0.04 for β_m, while the RMSE values improve relative to the less strong heteroskedasticity case. This pattern is consistent with the role of heteroskedasticity in strengthening the Lewbel-type instruments.

The competing estimators show little change between the less strong and stronger heteroskedasticity cases, as expected. Their identification does not rely directly on the conditional variance shifts exploited by IVSS. ACF continues to overestimate the flexible-input elasticity, with a bias around 0.10. LM remains strongly biased, with positive bias for β_k and negative bias for β_m. DL also remains biased, especially for β_m. Hence, the main source of bias for ACF, LM, and DL is not the strength of heteroskedasticity itself, but the interaction between persistent distortions, proxy inversion, and lag-based moment conditions.

Table 3. Monte Carlo results– Alternative degrees of heteroskedasticity

Panel A: γ_1 = 0.15

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.034	0.039	-0.010	0.100	0.187	-0.417	-0.092	0.269
RMSE	0.096	0.124	0.050	0.101	0.191	0.418	0.110	0.270

Panel B: $\boldsymbol{\gamma_1}$ = 0.80

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.036	0.040	-0.012	0.105	0.198	-0.433	-0.087	0.265
RMSE	0.052	0.049	0.050	0.107	0.202	0.434	0.108	0.266

Notes: The table reports Monte Carlo bias and RMSE of the estimated production elasticities for alternative degrees of heteroskedasticity in the distortion variance function, depending on parameter γ_1. Results are based on 500 replications with N = 500 firms and T = 8 periods. The benchmark case γ_1 = 0.50 is reported in Table 1.

Table 4 reports results for alternative values of the feedback parameter ρ_{ym}. When feedback is negative, $\rho_{ym} = -0.25$, IVSS remains comparatively stable, with biases of 0.086 for $\beta_k$and 0.035 for β_m. ACF becomes more biased, overestimating β_k by 0.131 and underestimating β_m by -0.098. LM displays large biases with the opposite sign, underestimating β_k by -0.216 and overestimating β_m by 0.433. DL performs poorly in this case, especially for the capital elasticity, with a bias of -0.297.

When feedback is positive and stronger, $\rho_{ym} = 0.5$, IVSS again remains relatively stable, with biases of 0.076 for β_k and 0.077 for β_m. ACF now displays a large upward bias in the flexible-input elasticity, equal to 0.203. LM becomes severely biased, overestimating β_k by 0.345 and underestimating β_m by -0.793. DL reduces the bias relative to LM in this positive-feedback case, but its RMSE remains sizeable, especially for β_m. These results show that the magnitude and sign of ρ_{ym}matter greatly for estimators that rely on proxy recovery or lagged moments. Once lagged distortions enter productivity, the identifying assumptions behind ACF, LM, and DL become sensitive to the dynamic distortion channel.

Table 4. Monte Carlo results – Different degree of distortion feedback effects

Panel A: $\rho_{ym} = -0.25$

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.086	0.035	0.131	-0.098	-0.216	0.433	-0.297	0.243
RMSE	0.096	0.058	0.140	0.100	0.221	0.435	0.300	0.246

Panel B: ρ_{ym} =0.50

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.076	0.077	-0.018	0.203	0.345	-0.793	-0.120	0.136
RMSE	0.097	0.098	0.069	0.200	0.352	0.795	0.149	0.612

Notes: The table reports Monte Carlo bias and RMSE of the estimated production elasticities β_k and β_m under alternative values of the distortion-to-productivity feedback parameter, ρ_{ym}. Panel A sets $\rho_{ym} = -0.25$, so that lagged distortions reduce current productivity, while Panel B sets $\rho_{ym} = 0.50$, implying stronger positive feedback from lagged distortions to current productivity. Bias and RMSE are computed across 500 Monte Carlo replications with $N = 500$ firms and $T = 8$ periods.

Overall, the general-design Monte Carlo evidence delivers three main conclusions. First, IVSS provides stable production-elasticity estimates across the baseline, no-feedback, heteroskedasticity, and feedback scenarios. Second, when $\rho_{ym} = 0$, the competing estimators perform much better, confirming that their biases are driven primarily by the distortion-feedback channel. Third, when $\rho_{ym} \neq 0$, ACF is affected by failed proxy inversion, LM is affected by the invalidity of lagged composite-productivity and lagged-input moments, and DL mitigates but does not fully eliminate the consequences of persistent distortions. The advantage of IVSS is that it identifies the production elasticities using heteroskedasticity-generated internal instruments based on the input-equation residual, rather than relying on proxy inversion or lagged flexible-input instruments.

4.2 Robustness design using current and lagged capital as shifters

We next consider an alternative Monte Carlo design motivated by empirical applications in which an additional excluded predetermined shifter is not available. Instead of using an autonomous q_{jt}, we set the observed shifter vector equal to current and lagged capital,

$$\mathrm{x}_{jt} = (k_{jt}, k_{j,t-1})'.$$

That is, in this design the second shifter is $q_{jt} = k_{j,t-1}$. This specification is valid under the maintained conditional mean-independence restriction $E(\omega_{j,t-1}^{m} \mid k_{jt}, k_{j,t-1}) = 0$. It is considered to assess the performance of the IVSS estimator when an independent predetermined shifter is unavailable. All other elements of the data-generating process are calibrated as in the previous Monte Carlo exercise. In particular, the production function, the law of motion for capital, the upper-triangular productivity-distortion process, and the variances of the transitory disturbances and state innovations are kept unchanged. The only change concerns the reduced-form flexible-input equation. Flexible input is now generated as

$$m_{jt} = \theta_1 k_{jt} + \theta_2 k_{j,t-1} + \omega_{jt}^{m} + \varepsilon_{jt}^{m},$$

where $\theta_1 = 0.6$ and $\theta_2 = 0.3$. Thus, flexible input demand depends on current capital, lagged capital, the latent flexible-input distortion component ω_{jt}^{m}, and the transitory input disturbance ε_{jt}^{m}.

The distortion innovation remains conditionally heteroskedastic. Since the observed shifter vector is now $\left(k_{jt}, k_{j,t-1}\right)'$, the heteroskedasticity function is specified as

$$h(k_{jt}, k_{j,t-1}; \gamma) = \exp\left(\gamma_0 + \gamma_k k_{jt} + \gamma_{k,-1} k_{j,t-1}\right),$$

where the baseline heteroskedasticity parameters are set as in the previous exercise: $\gamma_0 = 0$, $\gamma_k = 0.25$ and $\gamma_{k,-1} = 0.50$. At the centered values of the shifters, the conditional variance is normalized to one. The heteroskedasticity robustness exercises also follow the previous design. The less strong heteroskedasticity case sets $\gamma_k = 0.150$, while the stronger heteroskedasticity case sets $\gamma_k = 0.80$.

The Lewbel-type instruments are constructed from the generated residual of the reduced-form flexible-input equation as follows:

$$z_{jt}^{k} = \left(k_{jt} - \bar{k}\right)\hat{u}_{jt}^{m} \quad \text{and} \quad z_{jt}^{k,-1} = \left(k_{j,t-1} - \bar{k}_{-1}\right)\hat{u}_{jt}^{m},$$

where $\hat{u}_{jt}^{m}$ are the residuals of the input − equation, and $\bar{k}_{-1}$ is the sample mean of $k_{j,t-1}$.

The results of this Monte Carlo exercise are presented in Tables 5–8. These tables consider the same scenarios examined in Tables 1–4 for the main two-shifter design. Specifically, Table 5 reports the baseline case with $\rho_{ym} = 0.25$, Table 6 considers the no-feedback specification with $\rho_{ym} = 0$, Table 7 examines weaker and stronger forms of heteroskedasticity, and Table 8 studies alternative feedback parameters with $\rho_{ym} = -0.25$ and $\rho_{ym} = 0.5$. The only modification relative to the main design is that the additional predetermined shifter $q_{jt}$is replaced by lagged capital, $k_{j,t-1}$. As a result, the heteroskedasticity-based instruments are generated exclusively from current and lagged capital. This specification is particularly relevant for empirical applications where an independent predetermined shifter is unavailable. It therefore provides a useful assessment of whether the IVSS estimator can continue to achieve reliable identification when the available heteroskedastic variation originates solely from predetermined capital variables.

A number of conclusions emerge from Tables 5–8. First, replacing the additional predetermined shifter q_{jt} with lagged capital $k_{j,t-1}$ does not materially affect the performance of the IVSS estimator. Across all scenarios, the estimated biases and RMSE values remain similar to those obtained in the main two-shifter design, indicating that current and lagged capital provide sufficient variation for heteroskedasticity-based identification. The effect on the comparison estimators is less uniform. While the performance of ACF remains broadly similar to that observed in Tables 1–4, both LM and DL become somewhat more sensitive to the distortion process when lagged capital replaces the independent shifter. Nevertheless, the overall ranking of the estimators remains unchanged, with IVSS continuing to exhibit the smallest bias and RMSE across most scenarios.

Second, IVSS continues to deliver accurate estimates of the production elasticities. In the baseline specification, the bias is only 0.017 for β_k and 0.061 for β_m, with corresponding RMSE values of 0.043 and 0.078. These values are comparable to those obtained under the main design and remain substantially smaller than those of the competing estimators. IVSS also performs well in the no-feedback case ($\rho_{ym} = 0$) and remains robust under both weaker and stronger heteroskedasticity. These results confirm that Lewbel-type instruments constructed solely from current and lagged capital continue to provide informative identifying variation even in the absence of an additional predetermined shifter.

Table 5. Monte Carlo results for $x_{jt} = (k_{jt}, k_{j,t-1})'$ – Baseline scenario

Panel A: Estimates of production elasticities by alternative methods

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.30	0.50	0.30	0.50	0.30	0.50	0.30	0.50
Bias	0.017	0.061	-0.011	0.099	0.319	-0.560	-0.431	0.631
RMSE	0.043	0.078	0.050	0.101	0.313	0.563	0.435	0.634

Panel B: IVSS estimates of the state-space parameters

	ρ_{yy}	ρ_{mm}	ρ_{ym}	σ_{ξ^y}	σ_{ε^y}	γ_k	$\gamma_{k,-1}$	σ_{ε^m}
True	0.70	0.40	0.25	0.25	0.447	0.20	0.50	0.50
Bias	-0.001	-0.023	0.004	-0.003	0.009	-0.037	-0.019	-0.039
RMSE	0.028	0.042	0.019	0.029	0.045	0.121	0.128	0.068

Notes: *The table reports Monte Carlo bias and root mean squared error (RMSE) for the estimated production elasticities β_k and β_m. The design uses $x_{jt} = (k_{jt}, k_{j,t-1})'$, so that $q_{jt} = k_{j,t-1}$. All other simulation settings and estimator definitions are as in Table 1.*

Third, the comparison with ACF, LM, and DL broadly reinforces the conclusions obtained in the main Monte Carlo design. When the distortion-feedback channel is absent ($\rho_{ym} = 0$), all estimators perform reasonably well, although IVSS remains among the most accurate. Once lagged distortions affect current productivity ($\rho_{ym} \neq 0$), substantial differences emerge. ACF continues to perform relatively well for the capital elasticity but exhibits increasing bias in the flexible-input elasticity. LM becomes highly sensitive to the distortion-feedback mechanism, generating sizeable biases for both elasticities. DL is similarly affected and, in several cases, performs even worse than in the main design. This suggests that replacing the independent predetermined shifter with lagged capital strengthens the interaction between distortions, input choices, and productivity dynamics, making identification more challenging for estimators that do not explicitly model the distortion process. By contrast, IVSS remains comparatively stable across all feedback specifications.

Table 6. Monte Carlo results for $\mathrm{x}_{jt} = (k_{jt}, k_{j,t-1})'$ **– No distortion scenario** ($\rho_{ym} = 0$)

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.300	0.500	0.300	0.500	0.300	0.500	0.300	0.500
Bias	0.034	0.009	0.038	0.000	-0.000	0.000	0.046	0.005
RMSE	0.047	0.042	0.057	0.016	0.025	0.023	0.071	0.048

Notes: *The table reports Monte Carlo bias and root mean squared error (RMSE) for the estimated production elasticities* β_k*and* β_m*under the no-distortion-feedback scenario,* $\rho_{ym} = 0$*. The design uses* $x_{jt} = (k_{jt}, k_{j,t-1})'$*. All other simulation settings and estimator definitions are as in Table 2.*

Table 7. Monte Carlo results for $\mathrm{x}_{jt} = (k_{jt}, k_{j,t-1})'$**– Alternative degrees of heteroskedasticity**

Panel A: γ_k = 0.15

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.300	0.500	0.300	0.500	0.300	0.500	0.300	0.500
Bias	0.013	0.067	-0.011	0.098	0.316	-0.515	-0.431	0.629
RMSE	0.048	0.100	0.060	0.100	0.320	0.518	0.430	0.638

Panel B: γ_k = 0.80

	IVSS		ACF		LM		DL	
	β_x	β_m	β_x	β_m	β_x	β_m	β_x	β_m
True	0.300	0.500	0.300	0.500	0.300	0.500	0.300	0.500
Bias	0.025	0.053	-0.012	0.103	0.324	-0.569	-0.430	0.632
RMSE	0.043	0.064	0.050	0.105	0.328	0.571	0.434	0.636

Notes: *The table reports Monte Carlo bias and root mean squared error (RMSE) for the estimated production elasticities* β_k *and* β_m*under alternative degrees of heteroskedasticity in the distortion innovation. The design uses* $x_{jt} = (k_{jt}, k_{j,t-1})'$*. Panel A considers the less strong heteroskedasticity case,* $\gamma_k = 0.15$*, while Panel B considers the stronger heteroskedasticity case,* $\gamma_k = 0.80$*. All other simulation settings and estimator definitions are as in Table 3.*

Fourth, the heteroskedasticity experiments indicate that IVSS performs well even when the degree of heteroskedasticity is relatively weak. Increasing heteroskedasticity generally improves the precision of the IVSS estimates, as stronger variance shifts enhance the information content of the generated instruments. In the baseline feedback specification, for example, the RMSE of $\beta_k$declines from 0.048 when $\gamma_k = 0.15$ to 0.043 when $\gamma_k = 0.80$, while the RMSE of β_m falls from 0.10 to 0.064. In contrast, the performance of ACF, LM, and DL changes little across the heteroskedasticity scenarios, consistent with the fact that their identification strategies do not exploit conditional variance shifts. These results further support the central role of heteroskedasticity-based identification in the IVSS estimator.

Finally, the alternative feedback experiments confirm that the dynamic interaction between distortions and productivity remains the principal challenge for the comparison estimators. As the magnitude of the feedback parameter increases, the biases of ACF, LM, and DL generally become larger, particularly for the flexible-input elasticity. The deterioration is especially pronounced for LM and DL, whose RMSE values increase substantially when ρ_{ym}moves away from zero. In contrast, although the performance of IVSS is also affected by stronger feedback, the resulting biases and RMSE values remain considerably smaller than those of the competing estimators. These findings indicate that explicitly modelling the distortion process provides substantial gains when distortions have dynamic effects on future productivity.

Table 8. Monte Carlo results for $\mathrm{x}_{jt} = (k_{jt}, k_{j,t-1})'$ **– Different degree of distortion effects**

Panel A: $\rho_{ym} = -0.25$

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.300	0.500	0.300	0.500	0.300	0.500	0.300	0.500
Bias	0.085	-0.030	0.131	-0.099	-0.337	0.579	-0.508	0.745
RMSE	0.094	0.059	0.141	0.101	0.342	0.582	0.513	0.754

Panel B: ρ_{ym} =0.50

	IVSS		ACF		LM		DL	
	β_k	β_m	β_k	β_m	β_k	β_m	β_k	β_m
True	0.300	0.500	0.300	0.500	0.300	0.500	0.300	0.500
Bias	0.043	0.119	-0.018	0.200	0.579	-1.062	-0.702	1.062
RMSE	0.068	0.137	0.078	0.201	0.603	1.066	0.706	1.067

Notes: *The table reports Monte Carlo bias and root mean squared error (RMSE) for the estimated production elasticities* β_k*and* β_m*under alternative values of the distortion-to-productivity feedback parameter,* ρ_{ym}*. The design uses* $\boldsymbol{x}_{jt} = (k_{jt}, k_{j,t-1})'$. *Panel A sets* $\rho_{ym} = -0.25$*, so that lagged distortions reduce current productivity, while Panel B sets* $\rho_{ym} = 0.50$*, implying stronger positive feedback from lagged distortions to current productivity. All other simulation settings and estimator definitions are as in Table 4.*

5. Empirical application

We apply the proposed IVSS estimator to the firm-level Chinese manufacturing data used by Li and Mo (2026), based on the Annual Survey of Industrial Enterprises. The sample covers 1998–2007 and forms a large, unbalanced panel of manufacturing firms. To maintain comparability with Li and Mo, we use their cleaned sample and variable definitions and focus on a value-added specification in logs, where $k_{it} = \log K_{it}$ is treated as the predetermined input and $l_{it} = \log L_{it}$ as the flexible input. The estimated coefficients therefore have the interpretation of output elasticities with respect to capital and labor.

In the empirical implementation, the reduced-form labor equation includes current and lagged log capital together with year fixed effects. For the construction of the heteroskedasticity-based instruments, we use $\mathrm{x}_{jt} = (k_{jt}, k_{j,t-1})'$. Thus, the Lewbel-type instruments are $z_{jt}^{k} = \left(k_{jt} - \bar{k}\right)\hat{u}_{jt}^{m}$ and $z_{jt}^{k,-1} = \left(k_{j,t-1} - \bar{k}_{-1}\right)\hat{u}_{jt}^{m}$. All specifications include year fixed effects. Unlike Li and Mo, who report estimates separately for low- and high-distortion industries, we estimate the model on the full sample. This is because the second-step state-space system recovers the latent productivity and distortion components, providing evidence on their persistence, interaction, and dynamic feedback effects, which are central to understanding how distortions affect firm-level input allocation.

Tables 9(a)–9(b) summarize the empirical results. Table 9(a) reports the production elasticities obtained from the alternative estimators, while Table 9(b) presents the state-space parameters obtained in the second step of the IVSS estimator. For the state-space model, we report the specification that allows feedback from distortions to productivity, $\rho_{ym} \neq 0$. For the conditional distortion-innovation variance, we specify $Var(\xi_{jt}^{m} \mid x_{jt}) = h\left(k_{jt}; \gamma\right) = exp(\gamma_0 + \gamma_1 \tilde{k}_{jt} + \gamma_2 \tilde{k}_{j,t-1})$, where $\tilde{k}_{it}$ and $\tilde{k}_{i,t-1}$ denote centered current and lagged log capital, k_{jt} respectively. We normalize the scale parameter $\sigma_e^2 = 1$ and estimate γ_0 freely. Because the heteroskedasticity shifters are centered, $\exp(\gamma_0)$ represents the baseline distortion-innovation variance evaluated at the sample means of current and lagged capital. Standard errors reported in both tables are clustered at the firm level.[5]

[5] Firm-level clustering allows for arbitrary heteroskedasticity and within-firm dependence over time, while treating firms as independent clusters for inference; see Cameron and Miller (2015).

To implement the ACF-type proxy estimator, we use the flexible input as the proxy variable and include predetermined capital in the production function. The second-stage instruments are based on lagged inputs, following the standard proxy-control implementation. For the DL estimator of Doraszelski and Li (2025), we enrich the first-stage proxy-information set by adding the lead of capital, $k_{j,t+1}$, as discussed in Section 3. For the LM estimator of Li and Mo (2026), we use lagged composite productivity and lagged input variables as instruments. Finally, for the IVSS estimator, the first step uses the heteroskedasticity-based instruments constructed from the centered current and lagged capital shifters described above.

Table 9(a): Production function elasticities

	β_k	β_l	RTS	Nobs
DL	0.063 (0.009)	0.937 (0.031)	1.001 (0.024)	45066
ACF	0.078 (0.008)	1.009 (0.028)	1.087 (0.022)	65442
LM	0.036 (0.013)	1.080 (0.039)	1.117 (0.030)	44839
IVSS	0.322 (0.004)	0.490 (0.007)	0.812 (0.006)	65442
	$F_{w-iv} = 35.83$ $[p = 0.001]$ and $Het_F = 42.934$ $[p = 0.001]$			

Notes: *The table reports production-function elasticities estimated using the IVSS, ACF, DL, and LM estimators. All specifications include year fixed effects. Standard errors (in parentheses) are clustered at the firm level. RTS denotes returns to scale (β_k+β_l).For IVSS, F_{w-IV} indicates strong instrument relevance, while Het_F strongly rejects homoskedasticity and supports the heteroskedasticity-based identification strategy.*

As diagnostics for the IVSS production-function estimates, Table 9(a) reports the Kleibergen–Paap weak-identification statistic, denoted by $F_{w\text{-}IV}$. This statistic provides a heteroskedasticity-robust diagnostic for instrument relevance. The table also reports an F-type heteroskedasticity diagnostic, denoted by Het_F, based on the first-stage residuals $\hat{u}_{jt}^m$ from the reduced-form flexible-input equation. The statistic is obtained from an auxiliary regression of $\left(\hat{u}_{jt}^m\right)^2$ on the variables assumed to enter the heteroskedasticity function, $i.e.,$ k_{jt} and k_{jt-1}, including year fixed effects. The null hypothesis is that the coefficients on the heteroskedasticity shifters are jointly equal to zero, implying conditional homoskedasticity of the flexible-input residual with respect to k_{jt} and $k_{j,t-1}$. We test this joint restriction using a firm-cluster-robust F-statistic, with inference based on standard errors clustered at the firm level. Thus, Het_F is a panel-data analogue of the Breusch–Pagan heteroskedasticity diagnostic adapted to allow for within-firm dependence over time.[6] Rejection of the null provides evidence of systematic conditional variance variation with k_{jt} and $k_{j,t-1}$, supporting the source of heteroskedastic variation exploited by the IVSS instruments.

The results in Table 9(a) indicate that the IVSS estimates are economically plausible and internally consistent. The estimated capital elasticity is $\hat{\beta}_k = 0.322$, while the estimated labor elasticity is $\hat{\beta}_l = 0.49$. The implied returns to scale are therefore $RTS = \hat{\beta}_k + \hat{\beta}_l = 0.812$, indicating decreasing returns to scale. The first-stage relevance statistic is large, indicating that the excluded instruments are strongly correlated with the endogenous labor input. The test statistic Het_F strongly rejects the null of conditional homoskedasticity, providing evidence of the heteroskedasticity exploited by the Lewbel identification strategy.

The comparison estimators display substantial dispersion. DL yields a low capital elasticity, 0.063, and a high labor elasticity, 0.937, implying approximately constant returns to scale. ACF produces an even higher labor elasticity, 1.009, and returns to scale above one. LM estimates a very small capital elasticity, 0.036, and a labor elasticity of 1.080, also implying increasing returns to scale. By contrast, IVSS assigns a larger role to capital and a more moderate role to labor, producing decreasing returns to scale. This pattern is consistent with the Monte Carlo evidence, where proxy-based and lag-based estimators become sensitive when flexible-input choices are affected by latent distortions and the proxy-inversion or lagged-moment conditions are weakened.

Our results display a pattern similar to that documented by Li and Mo (2026) in industries characterized by relatively high resource misallocation. In their high-distortion food and beverage industries, Li and Mo find that their distortion-robust estimator yields a higher capital elasticity and a substantially lower labor elasticity than ACF, with the estimates changing from to for capital and from to for labor. By contrast, the two estimators produce very similar coefficients in their

[6] See Breusch and Pagan (1979), Koenker (1981), and Wooldridge (2010).

low-distortion industries. In our full-sample application, we obtain the same directional contrast. Relative to ACF, IVSS assigns a substantially larger role to capital and a smaller role to labor. This similarity should be interpreted as a qualitative rather than a direct numerical comparison, since Li and Mo estimate separate production functions for selected low- and high-distortion industries, whereas we estimate a common production function on the full sample and address latent distortions through heteroskedasticity-based internal instruments.

Table 9(b) reports the state-space estimates obtained in the second step of the IVSS procedure. The estimated productivity persistence parameter is $\hat{\rho}_{yy} = 0.867$, indicating that productivity shocks are highly persistent. In contrast, the distortion component exhibits rapid reversal dynamics, with $\hat{\rho}_{mm} = -0.487$. The negative autoregressive coefficient implies that distortion shocks tend to overshoot and subsequently reverse rather than accumulate over time, consistent with firms gradually adjusting inputs toward desired levels.

Table 9(b): State-space estimates

ρ_{yy}	ρ_{mm}	ρ_{ym}	$\sigma^2_{\xi^y}$	γ_0	γ_1	γ_2	σ_{ε^y}	σ_{ε^m}
0.867	-0.487	0.186	0.110	-4.140	-2.098	1.872	0.324	0.286
(0.005)	(0.031)	(0.054)	(0.004)	(0.131)	(0.188)	(0.220)	(0.006)	(0.004)

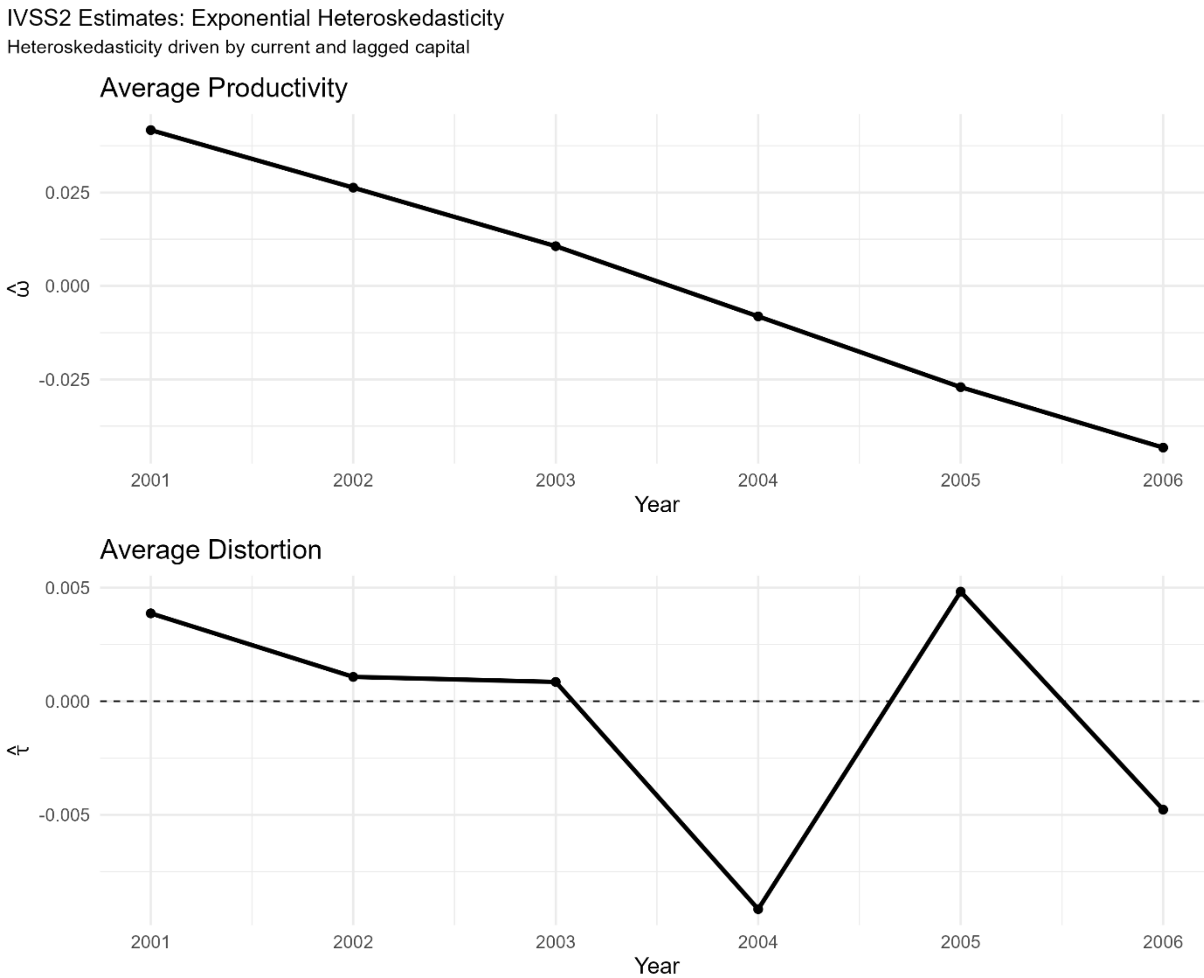
Notes: *The table reports the estimated parameters of the latent state-space system obtained in the second step of the IVSS procedure. The parameters ρ_{yy} and ρ_{mm} characterize the dynamic evolution of latent productivity and distortion, respectively, while ρym captures the dynamic transmission from distortions to future productivity. The state-innovation variances $\sigma^2_{\xi^y}$ and σ^2_e correspond to the productivity and distortion innovations, respectively. The covariance between the two state innovations, $\sigma_{\xi^y\xi^m}$, is restricted to zero because its unrestricted estimate is very close to zero and statistically insignificant. The contemporaneous covariance between the measurement disturbances is also restricted to zero. The conditional variance of the distortion innovation is specified as $h(k_{jt};\gamma) = exp(\gamma_0 + \gamma_1 \tilde{k}_{jt} + \gamma_2 \tilde{k}_{j,t-1})$ where $\tilde{k}_{it}$ and $\tilde{k}_{i,t-1}$ denote centered current and lagged log capital, k_{jt} respectively. The scale parameter σ^2_e is normalized to one. The parameters σ_{ε^y} and σ_{ε^m} denote the standard deviations of the transitory measurement disturbances in the productivity and distortion measurement equations, respectively. Standard errors (in parentheses) are clustered at the firm level; the delta method is used for nonlinear parameter transformations.*

The estimated transmission parameter from distortions to future productivity is positive and statistically significant, $\hat{\rho}_{ym} = 0.186$. This finding should not be interpreted as evidence that distortions enhance productivity. Rather, it is consistent with dynamic adjustment models in which firms face adjustment frictions and cannot instantaneously attain their optimal factor allocations. In such environments, temporary deviations from frictionless input choices can affect measured productivity during the adjustment process. Asker et al. (2014) show that dynamic adjustment frictions can generate substantial dispersion in marginal products even when firms behave optimally, while Baqaee and Farhi (2020) emphasize that measured productivity reflects both technological and allocative conditions. Consequently, temporary distortions may influence future measured productivity even when the distortions themselves are short-lived.

The innovation variance estimates indicate that productivity shocks are substantially more volatile than distortion shocks at the baseline values of the heteroskedasticity shifters. The estimated productivity innovation variance is sizeable, $\hat{\sigma}^2_{\xi y} = 0.112$, while the baseline distortion-innovation variance, evaluated at the sample means of current and lagged capital, is $\exp(\hat{\gamma}_0) = \exp(-4.14) \approx 0.016$. However, the heteroskedasticity parameters provide strong evidence of state-dependent distortion volatility, consistent with the test statistic rejecting conditional homoskedasticity in the first-stage flexible-input residuals. Current and lagged capital enter the conditional variance function with opposite signs, indicating offsetting contemporaneous and lagged effects of firms' capital positions on distortion volatility. In the log-linear variance specification, the effect of a sustained change in capital is governed by the sum of the two coefficients, $\gamma_1 + \gamma_2$. The estimated sum is -0.226. A cluster-robust Wald test strongly rejects the null hypothesis $H_0\colon \gamma_1 + \gamma_2 = 0$ $(\mathrm{F} = 7.342, \mathrm{p} = 0.007)$ indicating a statistically significant negative long-run effect. Thus, once both current and lagged capital have adjusted, persistently higher capital is associated with lower distortion-innovation variance. Together with the rejection of conditional homoskedasticity, these results support the heteroskedastic specification underlying the IVSS estimator and indicate systematic dynamic variation in distortion volatility with firms' capital positions.

Figure 1 reports the smoothed estimates of the latent productivity and distortion components recovered in the second step of the IVSS procedure, together with the implied output losses associated with distortions. The productivity component exhibits a gradual downward trend throughout the sample period, declining from positive values in 2001 to negative values by 2006. This pattern is consistent with the high persistence estimate $\hat{\rho}_{yy} = 0.867$ reported in Table 9(b), indicating that productivity shocks dissipate only slowly over time. Although distortions affect future productivity through the estimated feedback channel $\hat{\rho}_{ym} = 0.186$, the relatively modest magnitude of this coefficient suggests that the evolution of productivity is driven primarily by its own persistent dynamics and innovation process.

Figure 1. Estimated productivity and distortion components



Notes: The figure reports the cross-sectional averages of the smoothed latent productivity component, ω_{jt}^{y}, and the distortion component, ω_{jt}^{m}, obtained from the second step of the IVSS estimator.

By contrast, the distortion component displays no clear trend and instead fluctuates around a relatively stable negative mean. These short-run fluctuations are consistent with the strong reversal dynamics implied by $\rho_{mm} = -0.487$, indicating that distortion shocks are largely transitory and tend to be corrected over time rather than accumulate. The observed variability of the distortion component is also consistent with the state-dependent heteroskedasticity documented in Table 9(b). In particular, the estimates $\gamma_1 = -2.098$ and $\gamma_2 = 1.872$ imply that the conditional variance of distortion innovations varies systematically with firms' current and lagged capital positions. Conditional on lagged capital, higher current capital is associated with lower distortion-innovation variance, whereas, conditional on current capital, higher lagged capital is associated with greater distortion-innovation variance. The opposite signs of the two coefficients therefore point to a dynamic relationship between firms' capital positions and distortion volatility, generating substantial heterogeneity in distortion dynamics across firms.

Overall, Figure 1 highlights a sharp distinction between the two latent states. Productivity evolves smoothly and persistently over time, whereas the distortion component exhibits substantially greater short-run variation and frequent reversals. Together with the estimated positive dynamic transmission coefficient, these results indicate that temporary distortions can affect both contemporaneous input allocation and subsequent productivity dynamics, underscoring the importance of distinguishing persistent productivity differences from transitory allocation wedges.

6. Conclusion

This paper develops an estimator for production functions that remains valid when input choices are jointly shaped by productivity and distortion shocks. In such environments, observed input variation reflects both technological efficiency and distortion-driven adjustments, complicating identification of production elasticities. The proposed approach addresses this problem by exploiting heteroskedasticity in flexible-input demand to construct internal (Lewbel-type) instruments, thereby identifying the component of input variation that is orthogonal to productivity innovations. We show that this approach delivers consistent estimates without relying on proxy inversion, external instruments, or restrictive assumptions on the mapping between inputs and latent productivity.

The second step embeds the model in a state-space framework, allowing the joint evolution of productivity and distortions to be recovered from the residual structure of the production and input equations. This provides a unified decomposition that separates persistent productivity from transitory distortions and measurement noise using Kalman filter–based likelihood methods. As a result, the approach not only identifies production elasticities but also provides evidence on the persistence, interaction, and dynamic feedback effects of the latent productivity and distortion components.

Monte Carlo evidence shows that the estimator remains accurate across a range of challenging settings, including distortion-driven input endogeneity, persistent measurement error in flexible inputs, and feedback effects from distortions to productivity. By contrast, proxy-variable estimators and lag-based instrumental-variable estimators can become substantially biased when monotonicity or orthogonality conditions are weakened, especially when lagged inputs are contaminated by persistent measurement error or distortion-driven feedback.

The empirical application further illustrates that the method delivers economically plausible elasticities and a stable characterization of returns to scale. The state-space estimates reveal highly persistent productivity alongside more transitory, mean-reverting distortions. Importantly, although the estimated distortion-to-productivity feedback is relatively modest, it is positive and statistically significant, indicating that temporary distortions can affect not only contemporaneous input allocation but also subsequent productivity dynamics. This finding underscores the importance of distinguishing persistent productivity differences from transitory allocation wedges, while allowing for dynamic interactions between them.

Appendix A.

In this Appendix, we provide the proofs of the main results.

Appendix A1. Proof of Proposition 1

For simplicity, fix the observable state vector at $s_{jt} = s$, where s collects the observable state variables entering the flexible-input policy. We present the argument for a scalar flexible input used as a proxy. The same logic applies componentwise when M_{jt} is a vector and one of its components is used as the proxy input.

The flexible-input policy is $M_{jt} = \tilde{\phi}(\omega_{jt}^{y}, \omega_{jt}^{m}, s)$. By assumption, the policy depends nontrivially on the distortion component ω_{jt}^{m}. Therefore, there exist a productivity value $\bar{\omega}^{y}$ and two distinct distortion realizations $\omega_1^m \neq \omega_2^m$ in the support of ω_{jt}^{m} such that

$$\tilde{\phi}(\bar{\omega}^{y}, \omega_1^m, s) \neq \tilde{\phi}(\bar{\omega}^{y}, \omega_2^m, s).$$

Here, $\bar{\omega}^{y}$ denotes the fixed productivity value at which the input policy differs across the two distortion realizations. Without loss of generality, suppose that

$$\tilde{\phi}(\bar{\omega}^{y}, \omega_1^m, s) < \tilde{\phi}(\bar{\omega}^{y}, \omega_2^m, s).$$

Define the function

$$c(\omega^{y}) = \tilde{\phi}(\omega^{y}, \omega_2^m, s),$$

where ω^{y} belongs to the support of productivity. Since $\tilde{\phi}(\omega^{y}, \omega_2^m, s)$ is strictly increasing in ω^{y}, the function $c(\cdot)$ is strictly increasing and therefore invertible on its image. In particular, suppose that the value of $\tilde{\phi}(\bar{\omega}^{y}, \omega_1^m, s)$ lies in the image of $c(.)$. Then, there exists a unique productivity value $\widetilde{\omega}^{y}$ such that

$$\tilde{\phi}(\bar{\omega}^{y}, \omega_1^m, s) = \tilde{\phi}(\widetilde{\omega}^{y}, \omega_2^m, s).$$

Because

$$\tilde{\phi}(\bar{\omega}, \omega_2^m, s) > \tilde{\phi}(\bar{\omega}, \omega_1^m, s)$$

and $c(\cdot)$ is strictly increasing, this value satisfies $\widetilde{\omega}^{y} < \bar{\omega}$.

Define the common observed flexible-input realization

$$M^* = \tilde{\phi}(\bar{\omega}^y, \omega_1^m, s) = \tilde{\phi}(\tilde{\omega}^y, \omega_2^m, s).$$

Then the same observed flexible-input value M^*, conditional on the same observable state vector s, is consistent with two distinct productivity-distortion pairs: $(\bar{\omega}^y, \omega_1^m)$ and $(\tilde{\omega}^y, \omega_2^m)$, with $\bar{\omega}^y \neq \tilde{\omega}^y$. Therefore, conditional on s, the observed flexible input M^* does not uniquely determine ω_{jt}^y. Equivalently, the set

$$S(M^*; s) = \{\omega^y : M_{jt} = \tilde{\phi}\big(\omega^y, \omega^m, s_{jt}\big) \quad for\ some\ \omega^m\}$$

contains more than one element on a positive-probability set. Hence, there does not generally exist a single-valued function $d(\cdot)$ such that

$$\omega_{jt}^y = d(M_{jt}, s_{jt})$$

almost surely. Thus, productivity cannot generally be recovered through scalar proxy inversion when the flexible-input policy depends nontrivially on latent distortions. This proves Proposition 1.

Appendix A2: Identification of production elasticities without proxy monotonicity

Proposition A1. Suppose Assumption 1 holds. Even if flexible-input demand is not monotone in productivity because of latent distortions, the production elasticity vector $\beta \in R^p$ is locally point identified by the moment condition

$$E[Z_{jt}^L\{y_{jt} - f(K_{jt}, m_{jt}; \beta)\}] = 0,$$

provided that

$$rank\left(E\left[Z_{jt}^L \frac{\partial f(K_{jt}, m_{jt}; \beta_0)}{\partial \beta'}\right]\right) = p,$$

where $Z_{jt}^L = \tilde{x}_{jt} \otimes u_{jt}^m \in R^{d_x d_m}$, $\tilde{x}_{jt} = x_{jt} - E(x_{jt})$, and β_0 denotes the true parameter vector. The rank condition ensures that the heteroskedasticity-generated instruments induce linearly independent variation in the production function with respect to each parameter in β.

PROOF: Define the population moment function

$$g(\beta) = E[Z_{jt}^{L}\{y_{jt} - f(K_{jt}, m_{jt}; \beta)\}].$$

At the true parameter vector β_0, $y_{jt} = f(K_{jt}, m_{jt}; \beta_0) + u_{jt}^{y}$. Therefore,

$$g(\beta_0) = E(Z_{jt}^{L} u_{jt}^{y}).$$

By Assumption 1, the Lewbel-type instruments are orthogonal to the production disturbance, so

$$E(Z_{jt}^{L} u_{jt}^{y}) = 0.$$

Hence, $g(\beta_0) = 0$. The Jacobian of the moment function is $\frac{\partial g(\beta)}{\partial \beta'} = -E\left[Z_{jt}^{L} \frac{\partial f(K_{jt}, m_{jt}; \beta)}{\partial \beta'}\right]$. If

$$E\left[Z_{jt}^{L} \frac{\partial f(K_{jt}, m_{jt}; \beta_0)}{\partial \beta'}\right]$$

has full column $rank\ p = \dim(\beta)$, then the Jacobian has full column rank at β_0. By standard local GMM identification arguments, the equation $g(\beta) = 0$has a unique local solution at β_0. Hence, β is locally point identified.

Remark (Linear Cobb-Douglas case). If the production function is linear in logs,

$$f(K_{jt}, m_{jt}; \beta) = \beta_K k_{jt} + m_{jt}' \beta_m,$$

with $\beta = (\beta_K, \beta_m')' \in R^p, p = 1 + d_m$, then the production-regressor vector is

$$w_{jt} = (k_{jt}, m_{jt}')' \in R^{1+d_m}.$$

The moment condition becomes

$$E[Z_{jt}^{L}\{y_{jt} - w_{jt}' \beta\}] = 0.$$

Local identification requires the standard IV/GMM rank condition

$$rank\left(E[Z_{jt}^{L} w_{jt}']\right) = 1 + d_m.$$

Since $Z_{jt}^{L} \in R^{d_x d_m}$, the order condition is

$$d_x d_m \geq 1 + d_m.$$

This condition says that the number of Lewbel instruments generated must be at least as large as the number of production elasticities to be estimated. The rank condition further requires these instruments to be sufficiently correlated with the production inputs.

Appendix A3: Kalman filter with heteroscedastic distortion-state noise

This appendix presents the state-space representation of the model and the Kalman filter recursions used to evaluate the likelihood under the two-step estimation procedure proposed in the paper. Special attention is given to the role of conditional heteroskedasticity in distortion innovation. In addition to Assumption 1, the state-space representation requires finite second moments, serial independence of the state innovations and measurement disturbances, and the positive-definiteness conditions stated below. The Kalman filter recursions depend on the conditional first and second moments of the state-space system and therefore remain well defined under the time-varying innovation covariance structure considered here. For estimation, we use the Gaussian likelihood as a quasi-likelihood. Although the conditional innovation variance is a nonlinear function of the observed shifters , the state transition and measurement equations remain linear in the latent states. Conditional on and the parameter vector, the innovation covariance matrix is known at each step of the recursion. The standard linear Kalman filter therefore applies exactly with a time-varying covariance matrix, and no nonlinear filtering approximation is required. Conditional Gaussianity makes the resulting criterion the exact conditional Gaussian likelihood; more generally, we interpret it as a Gaussian quasi-likelihood based on the correctly specified conditional first and second moments.

Consider the measurement equations:

$$u_{jt}^{y} = \omega_{jt}^{y} + \varepsilon_{jt}^{y}, \qquad u_{jt}^{m} = \omega_{jt}^{m} + \varepsilon_{jt}^{m},$$

with state transition equation

$$\omega_{jt} = \rho\omega_{j,t-1} + \xi_{jt},$$

where $\omega_{jt} = \begin{bmatrix} \omega_{jt}^{y} \\ \omega_{jt}^{m} \end{bmatrix}$, $\rho = \begin{bmatrix} \rho_{yy} & \rho_{ym} \\ 0 & \rho_{mm} \end{bmatrix}$ and $\xi_{jt} = \begin{bmatrix} \xi_{jt}^{y} \\ \xi_{jt}^{m} \end{bmatrix}$, with innovation variance-covariance (conditional on x_{jt}) given as

$$\Sigma_{\xi,jt} \equiv Var(\xi_{jt} \mid x_{jt}) = \begin{bmatrix} \sigma_{\xi^y}^2 & \sigma'_{\xi^y \xi^m} \\ \sigma_{\xi^m \xi^y} & \Sigma_{\xi^m,jt} \end{bmatrix},$$

where $\Sigma_{\xi^m,jt} = h\left(x_{jt};\gamma\right)\Sigma_e$. If a parametric specification is adopted for $h\left(x_{jt};\gamma\right)$, the nuisance parameter vector γ is included in the overall parameter vector θ.

Stacking the measurement equations yields

$$u_{jt} = \begin{bmatrix} u_{jt}^{y} \\ u_{jt}^{m} \end{bmatrix} = \Gamma\omega_{jt} + \varepsilon_{jt}, \qquad \Gamma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where Γ is the measurement matrix and $\varepsilon_{jt} = \begin{bmatrix} \varepsilon_{jt}^{y} \\ \varepsilon_{jt}^{m} \end{bmatrix}$ is the measurement noise which has variance-covariance

$$\Sigma_\varepsilon \equiv \mathrm{Var}(\varepsilon_{jt}) = \begin{bmatrix} \sigma_{\varepsilon^y}^2 & 0 \\ 0 & \Sigma_{\varepsilon^m} \end{bmatrix}.$$

Throughout the Kalman recursions, we impose the boundedness condition $0 < c_h \leq h\left(x_{jt};\gamma\right) \leq C_h < \infty$, and we require both $\Sigma_{\xi,jt}$ and Σ_ε to be positive definite for all admissible values of x_{jt}. These conditions ensure that the predicted state covariance and the forecast-error covariance matrix are well defined and positive definite.

For each j, the Kalman filter evaluates the likelihood recursively over time. The procedure consists of initialization, prediction, innovation, and updating steps. The algorithm begins with an initialization of the state variables mean and its variance-covariance:

$$\mathrm{a}_{j,0|0} = E\left(\omega_{j0}\right) \qquad P_{j,0|0} = \mathrm{Var}\left(\omega_{j0}\right),$$

which holds under the stationarity conditions for ω_{jt}^{y} and ω_{jt}^{m}. We set $\mathrm{a}_{j,0|0} = 0$. Because $\Sigma_{\xi,jt}$ is time-varying through its dependence on x_{jt}, the state process does not admit a strictly time-invariant unconditional variance. To initialize the Kalman filter, we therefore construct a time-invariant prior for the state variance-covariance based on the unconditional mean of the innovation covariance matrix, defined as $\Sigma_\xi = E(\Sigma_{\xi,jt})$, where the expectation is taken with respect to the stationary distribution of the predetermined process x_{jt}. The initial state covariance $P_{j,0|0}$ is then obtained as the stationary solution to the Lyapunov equation

$$P_{j,0|0} = \rho P_{j,0|0}\rho' + \Sigma_\xi.$$

This provides a coherent unconditional prior for the variance–covariance matrix of the state vector ω_{j0}. After initialization, the Kalman recursion proceeds using the time-varying innovation covariance matrix $\Sigma_{\xi,jt}$ for $t \geq 1$, so that the predicted state covariance evolves consistently with the heteroskedastic structure implied by the model.

Given $a_{j,t-1|t-1}$ and $P_{j,t-1|t-1}$, the prediction step computes the prior (one-step-ahead) estimates:

$$\mathrm{a}_{j,t|t-1} = \rho \mathrm{a}_{j,t-1|t-1} \quad P_{j,t|t-1} = \rho P_{j,t-1|t-1}\rho' + \Sigma_{\xi,jt}.$$

Because the variance-covariance matrix $\Sigma_{\xi,jt}$ depends on x_{jt} through the heteroskedasticity function $h(x_{jt};\gamma)$, the prediction covariance matrix is time-varying even though matrix Γ is constant. Conditional heteroskedasticity therefore directly affects the evolution of uncertainty about the latent states.

Using the predicted state, the innovation (forecast error) is computed as

$$v_{jt} = u_{jt} - \mathrm{a}_{j,t|t-1},$$

which can be written as $v_{jt} = \left(\omega_{jt} - \mathrm{a}_{j,t|t-1}\right) + \varepsilon_{jt}$, using the measurement equation $u_{jt} = \omega_{jt} + \varepsilon_{jt}$.

Because the measurement matrix is I, the variance-covariance of v_{jt} reduces to

$$F_{jt} = P_{j,t|t-1} + \Sigma_\varepsilon = P_{j,t|t-1} + \Sigma_\varepsilon.$$

Since $P_{j,t|t-1}$ depends on $\Sigma_{\xi,jt}$, F_{jt} (reflecting estimation uncertainty) inherits the time variation generated by the heteroskedastic state innovation.

The updating step combines prior information with the current observation. The Kalman gain is

$$K_{jt} = P_{j,t|t-1}F_{jt}^{-1},$$

and the filtered state mean and covariance are

$$\mathrm{a}_{j,t|t} = \mathrm{a}_{j,t|t-1} + K_{jt}v_{jt}, \qquad P_{j,t|t} = P_{j,t|t-1} - K_{jt}F_{jt}K'_{jt}.$$

Given a parameter vector θ, the Kalman filter delivers the filtered state moments $\{\mathrm{a}_{j,t|t}(\theta), P_{j,t|t}(\theta)\}_{t=1}^{T}$ and the one-step-ahead moments $\{\mathrm{a}_{j,t|t-1}(\theta), P_{j,t|t-1}(\theta)\}_{t=1}^{T}$. Evaluating these objects at the second-step QML estimate $\hat{\theta}_{IVSS}$ (see Appendix A4) yields the fitted (filtered) latent state process $\widehat{\omega}_{jt|t} = \mathrm{a}_{j,t|t}(\hat{\theta}_{IVSS})$and associated conditional covariance $P_{j,t|t}(\hat{\theta}_{IVSS})$. Applying the fixed-interval Kalman smoother further produces smoothed state moments $\{\mathrm{a}_{j,t|T}(\hat{\theta}_{IVSS}), P_{j,t|T}(\hat{\theta}_{IVSS})\}_{t=1}^{T}$, which we use as estimates of ω_{jt} based on the generated residual series $\hat{u}_{j1}, \dots, \hat{u}_{jT}$, obtained in the first step.

The conditional heteroskedastic distortion innovation variance is obtained by plugging the estimate $\hat{\gamma}_{IVSS}$ (the γ-subvector of $\hat{\theta}_{IVSS}$) into $h(x_{jt};\gamma)$, so that the time-varying innovation covariance used in prediction is $\hat{\Sigma}_{\xi,jt} \equiv \Sigma_{\xi,jt}(x_{jt};\hat{\gamma}_{IVSS})$. Through its effect on the prediction covariance $P_{j,t|t-1}$ and hence $F_{jt} = P_{j,t|t-1} + \Sigma_{\varepsilon}$, conditional heteroskedasticity affects the Kalman gain $K_{jt} = P_{j,t|t-1}F_{jt}^{-1}$. When $h(x_{jt};\hat{\gamma}_{IVSS})$is large, the predicted variance of ξ_{jt}^{m} increases, raising the corresponding element of $P_{j,t|t-1}$ and therefore the Kalman gain. The filter consequently adjusts the distortion state more strongly in response to the observed u_{jt}^{m}.

Appendix A4: Proof of Proposition 2 (Asymptotic normality of $\hat{\boldsymbol{\theta}}_{IVSS}$)

The asymptotic distribution of the two-stage $\hat{\theta}_{IVSS}$ estimator follows from (i) standard IV/GMM asymptotics for the first-stage estimator of $b \equiv (\beta', \pi')' \in \mathbb{R}^{d_b}$ under Assumption 1 and (ii) standard quasi-maximum-likelihood regularity conditions for the second-stage state-space QML estimator, including smoothness, identification, finite moments, and nonsingularity of the expected Hessian.

In addition to Assumption 1, we impose standard moment and identification conditions. We assume the instrument vector Z_{jt} is a measurable function of the predetermined variables x_{jt} and the following finite fourth moments hold uniformly in t:

$$E(\| x_{jt} \|^4) < \infty, E(\| \varepsilon_{jt} \|^4) < \infty, E(\| \xi_{jt} \|^4) < \infty.$$

Under stationarity, these conditions imply finite fourth moments for ω_{jt} and u_{jt}. With fixed $T$and independence across j, they ensure that the cross-sectional law of large numbers and central limit theorem apply in the cross-sectional dimension (over j) to the first-step GMM/LS moment functions and to the second-step likelihood score contributions.

Asymptotic results for the first-step estimators of $b \equiv (\beta', \pi')' \in \mathbb{R}^{d_b}$: Assume the influence function representation for j (fixed T, $N \to \infty$):

$$\sqrt{N}(\hat{b} - b_0) = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \psi_{b,j} + o_p(1), \quad \psi_{b,j} = \begin{pmatrix} \psi_{\beta,j} \\ \psi_{\pi,j} \end{pmatrix},$$

so that

$$\sqrt{N}\left(\hat{b} - b_0\right) \xrightarrow{d} \mathcal{N}(0, \Sigma_b), \quad with \quad \Sigma_b = \mathrm{E}\left(\psi_{b,j} \psi'_{b,j}\right). \tag{A1}$$

For the GMM estimator of β, the influence function $\psi_{\beta,j}$ becomes

$$\psi_{\beta,j} = -(G'WG)^{-1} G'W \, g_j(\beta_0) = -(G'WG)^{-1} G'W \left(\frac{1}{T} \sum_{t=1}^{T} Z_{jt} \, u_{jt}^{y}\right),$$

where $\quad g_j(\beta) \equiv \frac{1}{T} \sum_{t=1}^{T} Z_{jt} \, u_{jt}^{y}, \quad G \equiv \mathrm{E}\left(\frac{\partial g_j(\beta_0)}{\partial \beta'}\right)$ and $W \equiv \left(E\left(g_j(\beta_0) g_j(\beta_0)'\right)\right)^{-1}$.

For the GMM estimator of β, the asymptotic result (A1) holds under conditions (i)–(v) of Assumption 1, the existence of finite fourth moments for the disturbances and instruments as previously stated, and the full column rank condition of $E(z_{jt} x'_{jt})$, which ensures identification.

For the LS estimator of π, the influence function $\psi_{\pi,j}$ becomes

$$\mathrm{vec}\left(\psi_{\pi,j}\right) = \left(I_M \otimes Q_x{}^{-1}\right) \mathrm{vec}\left(\frac{1}{T} \sum_{t=1}^{T} x_{jt} \, u_{jt}^{m\prime}\right), \ \ where \ \ Q_x \equiv \mathrm{E}\left(x_{jt} x'_{jt}\right).$$

For this estimator, the asymptotic result (A1) likewise holds under analogous conditions to the GMM case, now requiring full rank of Q_x. Note that in the scalar case $M = 1$, $\mathrm{vec}\left(\psi_{\pi,j}\right)$ reduces to

$$\psi_{\pi,j} = Q_x{}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} x_{jt}\, u_{jt}^{m}\right).$$

Asymptotic results for the IVSS estimator: Based on the Kalman filter recursions (given in Appendix A3), the Gaussian QML likelihood *j*-contribution is written as

$$\ell_j(\theta; u_j) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log|F_{jt}| + v_{jt}' F_{jt}^{-1} v_{jt}\right) + \text{const},$$

where $u_j = \left(u'_{j1}, u'_{j2}, \dots, u'_{jT}\right)'$, $v_{jt} = u_{jt} - \mathrm{a}_{j,t|t-1}$ *and* $F_{jt} = P_{j,t|t-1} + \Sigma_\varepsilon$. For j, the score is

$$S_j(\theta; u_j) = \frac{\partial \ell_j(\theta; u_j)}{\partial \theta}.$$

The IVSS estimator of the vector of parameters $\theta = \left(\rho, \sigma_{\xi y}^2, \gamma', \sigma_{\xi^m \xi^y}, \sigma_{\varepsilon y}^2, \sigma_{\varepsilon^m}^2\right)' \in \mathbb{R}^{d_\theta}$ satisfies

$$\frac{1}{N}\sum_{j=1}^{N} S_j\left(\hat{\theta}_{IVSS}; \widehat{u_j}\right) = 0,$$

where $\widehat{u_j}$ denotes the vector series of the residuals based on the first-step estimates of $\hat{b} = (\hat{\beta}', \hat{\pi}')'$.

Mean-value expansion of $\frac{1}{N}\sum_{j=1}^{N} S_j\left(\hat{\theta}_{IVSS}; \widehat{u_j}\right)$ around the true vector of parameters θ_0 yields

$$0 = \frac{1}{N}\sum_{j=1}^{N} S_j\left(\theta_0; \widehat{u_j}\right) + \left[\frac{1}{N}\sum_{j=1}^{N}\frac{\partial S_j\left(\widetilde{\theta}\,; \widehat{u_j}\right)}{\partial \theta'}\right]\left(\hat{\theta} - \theta_0\right),$$

for some $\widetilde{\theta}$ between $\hat{\theta}_{2SML}$ and θ_0.

Under the regularity conditions (smoothness, fixed T, independence across j, nonsingular expected Hessian and generated-residual equicontinuity),[7] as $N \to \infty$,

$$\frac{1}{N}\sum_{j=1}^{N}\frac{\partial S_j(\theta_0;\widehat{u_j})}{\partial \theta'} \xrightarrow{p} -H(\theta_0), \text{ where } H(\theta_0) = -\mathrm{E}\left(\frac{\partial S_j(\theta_0;u_j)}{\partial \theta'}\right).$$

Hence,

$$\sqrt{N}(\hat{\theta}_{2SML} - \theta_0) = H(\theta_0)^{-1}\tfrac{1}{\sqrt{N}}\sum\nolimits_{j=1}^{N} S_j\,(\theta_0;\widehat{u_j}) + o_p(1). \tag{A2}$$

The generated-residual equicontinuity condition requires that the score gradient is Lipschitz in u_j in a neighborhood of the true residual vector and that $N^{-1}\sum\nolimits_{j=1}^{N}\left\|\widehat{u_j} - u_j\right\|^2 = O_p(N^{-1})$.

This ensures that replacing by in the Hessian contributes only .

Because $\widehat{u_j}$ depends on first-step estimates $\hat{b} = (\hat{\beta}', \hat{\pi}')'$, we expand $S_j(\theta_0;\widehat{u_j})$ around u_j:

$$S_j(\theta_0;\widehat{u_j}) = S_j(\theta_0;u_j) + \frac{\partial S_j(\theta_0;u_j)}{\partial {u_j}'}\left(\widehat{u_j} - u_j\right) + o_p(N^{-1/2}).$$

From the first-order expansion of $\hat{u}_j$, we have

$$\left(\widehat{u_j} - u_j\right) = B_j(\hat{b} - b_0) + o_p(N^{-1/2}),$$

where $B_j = \frac{\partial u_j(b_0)}{\partial b'} \in \mathbb{R}^{T(1+d_m)\times d_b}$.

Define

$$A_j \equiv \frac{\partial S_j(\theta_0;u_j)}{\partial {u_j}'}B_j \in \mathbb{R}^{d_\theta \times d_b}.$$

[7] The generated-residual equicontinuity condition requires that the score gradient is Lipschitz in u_j in a neighborhood of the true residual vector and that $N^{-1}\sum\nolimits_{j=1}^{N}\left\|\widehat{u_j} - u_j\right\|^2 = O_p(N^{-1})$. This ensures that replacing u_j by $\widehat{u_j}$ in the Hessian contributes only $o_p(1)$.

Then

$$S_j(\theta_0;\hat{u}_j) = S_j(\theta_0;u_j) + A_j(\hat{b}-b_0) + o_p(N^{-1/2}).$$

Substituting the last result and $\sqrt{N}(\hat{b}-b_0) = \frac{1}{\sqrt{N}}\sum_{j=1}^{N}\psi_{b,j} + o_p(1)$ into (A1) yields

$$\frac{1}{\sqrt{N}}\sum_{j=1}^{N} S_j\left(\theta_0;\hat{u}_j\right) = \frac{1}{\sqrt{N}}\sum_{j=1}^{N} S_j\left(\theta_0;u_j\right) + A\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\psi_{b,j} + o_p(1),$$

where $\frac{1}{N}\sum_{j=1}^{N} A_j \overset{p}{\rightarrow} A \equiv E(A_j)$. Hence,

$$\sqrt{N}(\hat{\theta}_{2SML}-\theta_0) = H(\theta_0)^{-1}\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\left(S_j\left(\theta_0;u_j\right) + A\,\psi_{b,j}\right) + o_p(1),$$

which, by the CLT, implies

$$\sqrt{N}(\hat{\theta}_{IVSS}-\theta_0) \overset{d}{\rightarrow} \mathcal{N}(0,V_\theta),$$

where

$$V_\theta = H(\theta_0)^{-1}\Omega H(\theta_0)^{-1'},$$

and

$$\Omega = \mathrm{E}\left[\left(S_j{+}A\psi_{b,j}\right)\left(S_j{+}A\psi_{b,j}\right)'\right],$$

where $S_j \equiv S_j(\theta_0;u_j)$ for notation convenience.

Decomposing Ω,

$$\Omega = E\left(S_jS_j'\right) \quad + \quad A\,\Sigma_b\,A' \quad + \quad A\,\mathrm{E}\left(\psi_{b,j}S_j'\right) + \mathrm{E}\left(S_j\psi_{b,j}'\right)A',$$

shows that (i) the term $\mathrm{E}\left(S_jS_j'\right)$ is the variance that would arise if the true disturbances u_{jt} were observed, (ii) the term $A\,\Sigma_b\,A'$ captures additional variability due to using generated residuals $\hat{u}_{jt}$ and (iii) the term $A\,\mathrm{E}\left(\psi_{b,j}S_j'\right) + \mathrm{E}\left(S_j\psi_{b,j}'\right)A'$ captures covariance between first- and second-step j- contributions.

Appendix A5: Proof of Proposition 3 (Bias of proxy-based estimators)

From the production residual decomposition,

$$u_{jt}^{y} = \omega_{jt}^{y} + \varepsilon_{jt}^{y},$$

where

$$\omega_{jt}^{y} = \rho_{yy}\omega_{j,t-1}^{y} + (\omega_{j,t-1}^{m})'\rho_{ym} + \xi_{jt}^{y},$$

by the first row of the latent VAR(1) process.

The proxy-based second stage uses $\widetilde{\omega}_{j,t-1}^{y}$ instead of the true $\omega_{j,t-1}^{y}$. The proxy-based residual evaluated at β_0 is

$$\tilde{u}_{jt}^{y}(\beta_0) = u_{jt}^{y} - \rho_{yy}\widetilde{\omega}_{j,t-1}^{y}.$$

Substituting the state equation gives

$$\tilde{u}_{jt}^{y}(\beta_0) = \rho_{yy}\omega_{j,t-1}^{y} + (\omega_{j,t-1}^{m})'\rho_{ym} + \xi_{jt}^{y} + \varepsilon_{jt}^{y} - \rho_{yy}\widetilde{\omega}_{j,t-1}^{y}.$$

Using

$$\omega_{j,t-1}^{y} = \widetilde{\omega}_{j,t-1}^{y} + \zeta_{j,t-1},$$

we obtain

$$\tilde{u}_{jt}^{y}(\beta_0) = \xi_{jt}^{y} + \varepsilon_{jt}^{y} + \rho_{yy}\zeta_{j,t-1} + (\omega_{j,t-1}^{m})'\rho_{ym}.$$

Multiplying by Z_{jt}^{P} and taking expectations,

$$E[Z_{jt}^{P}\tilde{u}_{jt}^{y}(\beta_0)] = \rho_{yy}E[Z_{jt}^{P}\zeta_{j,t-1}] + E[Z_{jt}^{P}(\omega_{j,t-1}^{m})']\rho_{ym},$$

since $E(Z_{jt}^{P}\xi_{jt}^{y}) = 0, E(Z_{jt}^{P}\varepsilon_{jt}^{y}) = 0$.

The just-identified IV moment condition is

$$E\left[Z_{jt}^{P}\left(y_{jt} - f(k_{jt}, m_{jt}; \beta) - \rho_{yy}\widetilde{\omega}_{j,t-1}^{y}\right)\right] = 0.$$

For the linear case (or after linearizing around β_0), this condition gives

$$0 = E[Z_{jt}^P \tilde{u}_{jt}^y(\beta_0)] - E[Z_{jt}^P X_{jt}'](\beta_{IV}^P - \beta_0).$$

Substituting $E[Z_{jt}^P \tilde{u}_{jt}^y(\beta_0)]$ into the last equation yields (11a).

For the overidentified GMM case, define

$$g(\beta) = E[Z_{jt}^P \tilde{u}_{jt}^y(\beta)].$$

A first-order expansion around $\beta_0$gives

$$g(\beta) \simeq g(\beta_0) - D(\beta - \beta_0).$$

The GMM estimator minimizes $g(\beta)'Wg(\beta)$. The first-order condition implies

$$D'W\{g(\beta_0) - D(\beta_{GMM}^P - \beta_0)\} = 0.$$

Thus,

$$\beta_{GMM}^P - \beta_0 \simeq (D'WD)^{-1} D'Wg(\beta_0),$$

and substituting the expression for $g(\beta_0)$ gives (11b). This completes the proof.

References

Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6), 2411–2451.

Asker, J., Collard-Wexler, A., & De Loecker, J. (2014). Dynamic inputs and resource (mis)allocation. *Quarterly Journal of Economics*, 129(4), 1711–1753.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1), 135–171.

Baqaee D.R. and E. Farhi 2020. Productivity and Misallocation in General Equilibrium. *Quarterly Journal of Economics*, 135(1), 105–163.

Bartelsman, E., J.Haltiwanger, and S.Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review* 103 (1): 305–334.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3), 623–685.

Blundell, R., & Bond, S. (2000). GMM estimation with persistent panel data: An application to production functions. *Econometric Reviews*, 19(3), 321–340.

Buera, F. J., J.P. Kaboski, and Y. Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review* 101 (5): 1964–2002.

Cameron, A. C., and D.L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2): 317–372

Cooper, R. W., & Haltiwanger, J. C. (2006). On the nature of capital adjustment costs. *The Review of Economic Studies*, 73(3), 611–633.

David, J. M., and V.Venkateswaran. 2019. "The Sources of Capital Misallocation." *American Economic Review* 109 (7): 2531–2567.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review* 102 (6): 2437–2471.

Doraszelski, U., & Jaumandreu, J. (2013). R&D and productivity: Estimating endogenous productivity. *Review of Economic Studies*, 80(4), 1338–1383.

Doraszelski, U. and L. Li. (2026). A Generalized Control Function Approach to Production Function Estimation. *AEA Papers and Proceedings* 116: 483–489.

Doraszelski, U., & Lixiong, L. (2025). Production Function Estimation without Invertibility: Imperfectly Competitive Environments and Demand Shocks. *NBER Working Paper No. 33939*. National Bureau of Economic Research.

Durbin, J., & Koopman, S. J. (2012). *Time Series Analysis by State Space Methods* (2nd ed.). Oxford University Press.

Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1), 394–425.

Gandhi, A., Navarro, S., & Rivers, D. (2020). On the identification of gross output production functions. *Journal of Political Economy*, 128(8), 2973–3016.

Grieco, P. L. E., S. Li, and H. Zhang. (2016). "Production Function Estimation with Unobserved Input Price Dispersion." *International Economic Review* 57 (2): 665–690.

Griliches, Z, & J.A. Hausman (1986). Errors in variables in panel data. *Journal of Econometrics,* 31, 93-118.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4), 1403–1448.

Hu, Y., Huang, G., & Sasaki, Y. (2020). Estimating production functions with robustness against errors in the proxy variables. *Journal of Econometrics, 215(2)*, 375–398.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45.

Kim, K. I., Petrin, A., & Song, S. (2016)**.** Estimating production functions with control functions when capital is measured with error. *Journal of Econometrics, 190*(2), 267–279.

Kiyotaki, N., and J.Moore. 1997. "Credit Cycles." *Journal of Political Economy* 105 (2): 211–248.

Kleibergen, F., & Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1), 97–126.

León-Ledesma, M. A., & Satchi, M. (2019). Appropriate technology and balanced growth. The *Review of Economic Studies*, 86(2), 807–835.

León-Ledesma, M. A., McAdam, P., & Willman, A. (2010). Identifying the elasticity of substitution with biased technical change. *American Economic Review*, 100(4), 1330–1357.

Lewbel, A. (1997). Constructing instruments for regressions with measurement error when no additional data are available. *Econometrica*, 65(5), 1201–1213.

Lewbel, A. (2012). Using heteroskedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1), 67–80.

Li, S., & Mo, J. (2026). Production Function Estimation With Resource Misallocation. *Journal of Applied Econometrics*, forthcoming.

Midrigan, V., & Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2), 422–458.

Moll, B. 2014. “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?” *American Economic Review* 104 (10): 3186–3221.

Olley, G. S., & Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6), 1263–1297.

Ottonello, P. and Winberry, T. (2020). “Financial Heterogeneity and the Investment Channel of Monetary Policy.” *Econometrica* 88 (6): 2473–2502.

Restuccia, D., & Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4), 707–720.

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics*, 85(4), 777–792.

Song, Z., Storesletten, K., & Zilibotti, F. (2011). Growing like China. *American Economic Review*, 101(1), 196–233.

Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147–162.

Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3), 112–114.

Department of Economics
Athens University of Economics and Business

List of Recent Working Papers

2024

01-24 Market Timing & Predictive Complexity, Stelios Arvanitis, Foteini Kyriazi, Dimitrios Thomakos
02-24 Multi-Objective Frequentistic Model Averaging with an Application to Economic Growth, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, Nikolas Topaloglou
03-24 State dependent fiscal multipliers in a Small Open Economy, Xiaoshan Chen, Jilei Huang, Petros Varthalitis
04-24 Public debt consolidation: Aggregate and distributional implications in a small open economy of the Euro Area, Eleftherios-Theodoros Roumpanis
05-24 Intangible investment during the Global Financial Crisis in the EU, Vassiliki Dimakopoulou, Stelios Sakkas and Petros Varthalitis
06-24 Time will tell! Towards the construction of instantaneous indicators of different agent-types, Iordanis Kalaitzoglou, Stelios Arvanitis
07-24 Norm Constrained Empirical Portfolio Optimization with Stochastic Dominance: Robust Optimization Non-Asymptotics, Stelios Arvanitis
08-24 Asymptotics of a QLR-type test for optimal predictive ability, Stelios Arvanitis
09-24 Strongly Equitable General Equilibrium Allocations, Angelos Angelopoulos
10-24 The Greek macroeconomy: A note on the current situation and future outlook, Apostolis Philippopoulos
11-24 Evolution of Greek Tax System, A Survey of Legislated Tax Changes from 1974 to 2018, Panagiotis Asimakopoulos
12-24 Macroeconomic Impact of Tax Changes, The case of Greece from 1974 to 2018, Panagiotis Asimakopoulos
13-24 `Pareto, Edgeworth, Walras, Shapley' Equivalence in a Small Economy, Angelos Angelopoulos
14-24 Stimulating long-term growth and welfare in the U.S, James Malley and Apostolis Philippopoulos
15-24 Distributionally Conservative Stochastic Dominance via Subsampling, Stelios Arvanitis
16-24 Before and After the Political Transition of 1974. Institutions, Politics, and the Economy of Post-War Greece, George Alogoskoufis
17-24 Gaussian Stochastic Volatility, Misspecified Volatility Filters and Indirect Inference Estimation, Stelios Arvanitis, Antonis Demos
18-24 Endogenous Realtor Intermediation and Housing Market Liquidity, Miroslav Gabrovski, Ioannis Kospentaris, Victor Ortego-Marti
19-24 Universal Choice Spaces and Expected Utility: A Banach-type Functorial Fixed Point, Stelios Arvanitis, Pantelis Argyropoulos, Spyros Vassilakis

Πατησίων 76, 104 34 Αθήνα. Τηλ.: 210 8203303-5. E-mail: econ@aueb.gr / www.aueb.gr
76, Patission Street, Athens 104 34 Greece. Tel.: (+30) 210 8203303-5

2025

01-25 Democracy, redistribution, and economic growth: Some evidence from post-1974 Greece, George C. Bitros

02-25 Assessing Downside Public Debt Risks in an Environment of Negative Interest Rates Growth Differentials, Yiannis Dendramis, Georgios Dimitrakopoulos and Elias Tzavalis

03-25 Wicksellian General Equilibrium, Angelos Angelopoulos

04-25 Institutions and Ethics, Angelos Angelopoulos

05-25 General Equilibrium in a Capitalist Economy, Angelos Angelopoulos

06-25 Public investment multipliers revisited: The role of production complementarities, Vasiliki Dimakopoulou, George Economides, and Apostolis Philippopoulos

07-25 The heterogeneous causal effects of the EU's Cohesion Fund, Angelos Alexopoulos, Ilias Kostarakos, Christos Mylonakis, and Petros Varthalitis

08-25 Gaussian Invariant Markov Chain Monte Carlo, Michalis K. Titsias, Angelos Alexopoulos , Siran Liu, and Petros Dellaportas

09-25 Financial Intermediation, Investment, and Monetary Policy, Ioannis Kospentaris and Florian Madison

10-25 Temperature as a premium factor. Insurance incorporating climate change. A comparative case study between Europe and North America, Nomikou-Lazarou Irene

11-25 How High will our Buildings Become? Angelos Angelopoulos

12-25 Emergency Room General Equilibrium, Angelos Angelopoulos

13-25 Updating Density Estimates using Conditional Information Projection: Stock Index Returns and Stochastic Dominance Relations, Stelios Arvanitis, Richard McGee, and Thierry Post

14-25 Measuring Textual Cohesion via Graph Curvature, Weighted Dominance, and Embedding-Based Coherence, Stelios Arvanitis

15-25 Universal Preference Spaces for Expected Utility via a Banach-type Functorial Fixed Point, Stelios Arvanitis, Pantelis Argyropoulos, and Spyros Vassilakis

16-25 Planet Earth's General Equilibrium, Angelos Angelopoulos

17-25 Electable and Stable Insiders' Coalition Governments, Tryphon Kollintzas and Lambros Pechlivanos

18-25 Long Memory from Cheeger Bottlenecks and Long Cycles in Network Dynamics, Stelios Arvanitis

19-25 Robust Multiobjective Model Averaging in Predictive Regressions, Stelios Arvanitis, Mehmet Pinar, Thanasis Stengos, and Nikolas Topaloglou

20-25 On robust Bayesian causal inference, Angelos Alexopoulos and Nikolaos Demiris

21-25 Leverage Effect, Volatility Feedback and the Influence of Jumps, Orestis Agapitos, Ioannis Papantonis, Leonidas Rompolis, and Elias Tzavalis

22-25 Leadership: The other deficit of post-1974 Greece, George C. Bitros

2026

01-26 The Tell-Tale Clock! Speed and Agent Composition in High Frequency Trading, Stelios Arvanitis

02-26 Sentiment Utopia Index: Rhetorical Structure and Corpus-Relative Financial Sentiment, Stelios Arvanitis

03-26 Draghi's report a year later: Beneficial but no remedy for EU challenges, George C. Bitros, Anastasios G. Malliaris, and Mark S. Rzepczynski

04-26 Tax Structure Reform and Macroeconomic Strategy in Greece: Econometric Evidence and Policy Implications, Panagiotis Asimakopoulos

05-26 A Perspective on International Monetary Leadership. From Silver and Gold to the Dollar Standard, George Alogoskoufis

06-26 Behavioral Balanced Scorecard Framework for Tax Administrations: Integrating Tax Morale, Institutional Trust, and Compliance Behavior, Panagiotis Asimakopoulos

07-26 Education and governance: The Russell–Hayek Conundrum through the lenses of ancient Sparta and the "Sparta of the East", George C. Bitros

08-26 Overhauling Democracy: Layered State Authority, Corporate-Style Accountability, and the Free Market Economy, George C. Bitros and Steve Bakalis

09-26 ECONOMICS 200. The Economics of transition to a Steady State Economy, Theodore P. Lianos

10-26 Empirical Likelihood Tests for Pairwise Stochastic Dominance, Stelios Arvanitis, Richard McGee, and Thierry Post

11-26 Production Function Estimation under Input Distortions: Identification Using Heteroscedasticity-Based Instruments, Dimitris K. Christopoulos and Elias Tzavalis

Department of Economics
Athens University of Economics and Business

The Department is the oldest Department of Economics in Greece with a pioneering role in organising postgraduate studies in Economics since 1978. Its priority has always been to bring together highly qualified academics and top quality students. Faculty members specialize in a wide range of topics in economics, with teaching and research experience in world-class universities and publications in top academic journals.

The Department constantly strives to maintain its high level of research and teaching standards. It covers a wide range of economic studies in micro-and macroeconomic analysis, banking and finance, public and monetary economics, international and rural economics, labour economics, industrial organization and strategy, economics of the environment and natural resources, economic history and relevant quantitative tools of mathematics, statistics and econometrics.

Its undergraduate program attracts high quality students who, after successful completion of their studies, have excellent prospects for employment in the private and public sector, including areas such as business, banking, finance and advisory services. Also, graduates of the program have solid foundations in economics and related tools and are regularly admitted to top graduate programs internationally. Three specializations are offered:1. Economic Theory and Policy, 2. Business Economics and Finance and 3. International and European Economics. The postgraduate programs of the Department (M.Sc and Ph.D) are highly regarded and attract a large number of quality candidates every year.

For more information:

https://www.dept.aueb.gr/en/econ/