Identification of Production Function Parameters with Noisy Proxy Variables: A Semiparametric Control-Function Analysis

Jorge de la Cal Medina*

December 21, 2025

Abstract

This paper studies identification and estimation of production function parameters when the proxy used in control-function methods is noisy and productivity is only partially predictable from observables. We allow intermediate-input demand to contain an unobserved error term, so that the canonical scalar-unobservable condition fails and productivity cannot be perfectly recovered from the proxy. Identification of the structural coefficients nonetheless goes through under two weaker requirements: conditional mean independence of innovation terms and a forecast-sufficiency condition under which the conditional expectation of future productivity depends on observables only through a one-dimensional index. Casting the problem as a semiparametric conditional moment model, we characterize local identification via the full rank of an orthogonalized Jacobian matrix and discuss corresponding GMM estimators.

1 Introduction

Proxy-variable methods such as Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg et al. (2015) address the simultaneity problem in production function estimation by assuming that investment or intermediate input choices are strictly increasing in firm productivity. This "scalar unobservable" condition implies that productivity can be perfectly recovered from observables, enabling control-function identification strategies.

*University of Amsterdam. Email: j.delacalmedina@uva.nl

Empirically, however, proxy variables are noisy. Input choices reflect not only productivity but also frictions, adjustment costs, shocks, and measurement error. This paper shows that full recoverability of productivity is not necessary for identification of the production-function coefficients. We allow intermediate-input demand to include an additive noise term, so that productivity is only partially predictable given the proxy. Our framework permits a realistic form of misspecification in the proxy structure while still enabling identification and estimation of structural parameters.

The structural coefficients β in the production function are identified via a semiparametric conditional moment restriction. The key idea is to exploit the fact that the proxy remains informative about productivity, even if it does not perfectly reveal it. Our approach uses the framework of Chen et al. (2014), treating the evolution of productivity as a nuisance function and applying orthogonalization to partial out its influence. Local identification is achieved whenever an orthogonalized Jacobian matrix has full column rank, isolating variation in inputs not captured by the proxy index.

Recent work by Doraszelski and Li (2025b) also relaxes the scalar–unobservable condition. Within the OP/LP/ACF proxy–variable framework, they allow for a nonparametric law of motion for productivity and develop tests for invertibility, bias-eliminating moments in special cases, and bias-reducing Neyman-orthogonal moments in the general non-invertible case. Identification is taken up more explicitly in their subsequent paper Doraszelski and Li (2025a), which analyzes a generalized control-function model with both the production function and the law of motion treated as fully nonparametric, under a completeness assumption.

Our paper is similar in that it allows productivity to be only partially predictable from the proxy and uses orthogonal moments with flexible first–step estimation, but differs by developing a semiparametric identification result that treats the productivity forecast function as a nonparametric nuisance.

We develop a semiparametric identification and estimation strategy that does not rely on invertibility, completeness, or full support. Instead, it requires mean independence and a forecast-sufficiency condition on the proxy, making the framework suited for empirical settings with measurement error or frictions in input adjustment.

2 Model

We briefly introduce the control function approach following OP/LP/ACF. Consider the following production function in logs for firm i at time t:[1]

$$y_{it} = x_{it}^{\top}\beta_0 + \omega_{it} + \varepsilon_{it} \tag{1}$$

y is the log of output, x is the vector of the log inputs, which includes a constant to account for an intercept and transformations of labor and capital, ε denotes unexpected productivity shocks, and ω captures firm-specific productivity. Note that implicitly we assume a "value-added specification", intermediate inputs m do not enter (1) directly but serve as a proxy for productivity, as we will see below. Latin letters denote observed variables; Greek letters denote variables unobserved to the researcher. We are interested in identifying and estimating the production function parameter β_0 based on a panel of firms $i = 1, \ldots, N$ over time periods $t = 1, \ldots, T$; with $N \to \infty$ and fixed T.

Behind the identification strategy is a discrete time model of optimizing firms. At time t, firm i chooses inputs conditional on observing a productivity draw ω_{it}. At the end of the period output y_{it} is produced depending on the chosen inputs and to a noise term ε_{it}, unexpected productivity. We allow for part of the input vector x_{it} to be predetermined at $t-1$, denoted at x_{it}^p and part to be flexibly determined at t, denoted as x_{it}^f, so $x_{it} = (x_{it}^p, x_{it}^f)$. The firm also flexibly chooses upon intermediate inputs m_{it}. Hence the choice at t is upon next period's predetermined inputs, x_{it+1}^p, and this period's flexible inputs and intermediate inputs, x_{it}^f, m_{it}. OP/LP/ACF typically treat capital as predetermined and labor as flexible.

2.1 Model Assumptions

The model is assumed to satisfy the following assumptions. The identifying moment conditions will be based on the previously explained timing assumptions. It is thus helpful to introduce the information $\mathcal{I}_{it}$, which includes all variables known to firm i at t.

Assumption 1. *$\mathcal{I}_{it}$ includes all current and past realizations of ω_{it}, $\{\omega_{i\tau}\}_{\tau \le t} \subset \mathcal{I}_{it}$, and moreover satisfies $\mathbb{E}[\varepsilon_{it}|\mathcal{I}_{it}] = 0$.*

Consequently, every input decision the firm chooses as a function of ω_{it}, is also contained in $\mathcal{I}_{it}$. Thus $(x_{it+1}^p, x_{it+1}^f, m_{it}) \in \mathcal{I}_{it}$.

[1] This specification accommodates the commonly used Cobb-Douglas and translog production functions

Assumption 2. *The productivity process ω_{it} follows a first-order Markov process:*

$$g(\omega_{it+1} \mid \mathcal{I}_{it}) = g(\omega_{it+1} \mid \omega_{it}), \tag{2}$$

where $g(\cdot \mid \omega_{it})$ denotes the conditional density, which is known to firms and is stochastically increasing in ω_{it}.

Assumptions 1–2 imply that ω_{it} can be written as its conditional expectation at time $t-1$ plus an innovation term, with a slight abuse of notation using $g(\cdot)$ for the conditional mean:

$$\omega_{it} = g(\omega_{it-1}) + \xi_{it}, \tag{3}$$

where innovation ξ_{it} satisfies $\mathbb{E}[\xi_{it}|\mathcal{I}_{it-1}] = 0$.

The fundamental assumption of the control function approach is that firm productivity ω_{it} can be proxied by observables. Specifically, it is assumed that there exists a demand function for a flexibly chosen intermediate input, m_{it}, which depends on productivity and predetermined inputs:

$$m_{it} = h(\omega_{it}, x_{it}^p).$$

Inverting $h(\omega, x^p)$ with respect to ω yields a control function in terms of observables,

$$\omega_{it} = h^{-1}(m_{it}, x_{it}^p).$$

The canonical OP/LP/ACF framework imposes the so-called *scalar unobservable assumption*, under which the firm's choice of m_{it} is fully determined by ω_{it}, conditional on the predetermined input x_{it}^p. Equivalently, productivity is perfectly recoverable from the proxy variables:

$$\omega_{it} = \mathbb{E}[\omega_{it} \mid m_{it}, x_{it}^p].$$

This assumption implies that the proxy variables (m_{it}, x_{it}^p) are perfectly correlated with the productivity term and contain no independent variation. However, arguably such a requirement is highly demanding and rarely plausible in empirical applications. It rules out any form of noise in the proxy arising from, for example, measurement or optimization errors, unexpected supply- or demand-side shocks, or heterogeneity in input prices and technologies. For a detailed discussion of such proxy errors, see Hu et al. (2020).

To relax this overly restrictive assumption, we allow for an unexplained component in the intermediate-input demand function. This admits measurement error in m and unobserved heterogeneity in firms' intermediate input decisions. We will show that we don't require exact recovery of productivity from observables for identification; instead, identification relies on *mean independence* of the resulting forecast errors with respect to the firm's information set, as formalized in Assumption 3 below. This point is also made in Doraszelski and Li (2025b).

Assumption 3. *Intermediate input demand function with error*

$$m_{it} = h(\omega_{it}, x^p_{it}, \nu_{it}), \tag{4}$$

$h(\omega, x^p, \nu)$ *is strictly monotone in* ω.

The error ν_{it} may capture, for example, measurement error, adjustment frictions, or taste shocks in intermediate-input choice. Since the mapping $\omega \mapsto h(\omega, x^p_{it}, \nu_{it})$ is invertible, we can write

$$\omega_{it} = h^{-1}(m_{it}, x^p_{it}, \nu_{it}). \tag{5}$$

However, since ν_{it} is unobserved to the econometrician, productivity cannot be recovered from (m_{it}, x^p_{it}) alone, and the scalar-unobservable condition of OP/LP/ACF fails. Setting $\nu_{it} = 0$ collapses Assumption 3 to the usual scalar-unobservable case in which ω_{it} is perfectly recoverable from (m_{it}, x^p_{it}). Note that setting $\nu_{it} = 0$ collapses Assumption 3 to the usual scalar-unobservable case of OP/LP/ACF.

Let w_{it} be the vector of inputs chosen at by firm i at t, $w_{it} = ({x^p_{it+1}}^\top, {x^f_{it}}^\top, {m_{it}}^\top)^\top$. Moreover define the vector of input decisions up to time t

$$z_{it} = (w_{it}^\top, w_{it-1}^\top, \ldots, w_{i1}^\top)^\top$$

The moment conditions implied by Assumptions 1–3 are defined with respect to the firm's information set $\mathcal{I}_{it}$. To link them to the researcher's observables, we impose the following assumption. Note that it holds immediately if z_{it} is $\mathcal{I}_{it}$-measurable, that is, if $z_{it} \subset \mathcal{I}_{it}$.

Assumption 4. *For all* i, t *and for* $s \leq t$*,*

$$\mathbb{E}[\varepsilon_{it} \mid z_{is}] = 0,$$
$$\mathbb{E}[\xi_{it} \mid z_{is-1}] = 0.$$

Since, as laid out above, we cannot recover ω entirely, we use an *index* based on the researcher's observables. To ensure identification using in the OP/LP/ACF moment conditions is still possible, we need to to make some further assumptions on that index. As control function we use the regression function $\mathbb{E}[\omega_{it} \mid z_{it} = z]$, with which we can define the index $\widetilde{\omega}_{it} := \mathbb{E}[\omega_{it} \mid z_{it}]$. We then have

$$\mathbb{E}[\omega_{it} \mid z_{it}] \;=\; \widetilde{\omega}_{it}.$$

The proxy error $\eta_{it} := \omega_{it} - \widetilde{\omega}_{it}$ by construction satisfies $\mathbb{E}[\eta_{it} \mid z_{it}] = 0$.

For identification, we require a forecast sufficiency property, which holds that the forecast of next period's productivity, $\mathbb{E}[\omega_{it+1} \mid z_{it}]$, depends only on $\widetilde{\omega}_{it}$.

Assumption 5. *There exists a function* Γ_0 *such that*

$$\mathbb{E}[\omega_{it+1} \mid z_{it}] \;=\; \Gamma_0\left(\widetilde{\omega}_{it}\right).$$

This allows us to define the forecast error as

$$\zeta_{it+1} := \omega_{it+1} - \Gamma_0\big(\widetilde{\omega}_{it}\big),$$

which by construction satisfies $\mathbb{E}[\zeta_{it+1} \mid z_{it}] = 0$. This moment condition will be crucial for identification.[2]

Note that Assumption 5 does not hold in general. For example, if $g(\omega) = \omega^2$, then

$$\mathbb{E}[\omega_{it+1} \mid z_{it}] = \mathbb{E}[\omega_{it}^2 \mid z_{it}] = \tilde{\omega}_{it}^2 + \mathrm{Var}(\eta_{it} \mid z_{it}),$$

so the forecast depends on the conditional variance and cannot be written as a function of $\tilde{\omega}_{it}$ alone. A linear $g(\omega)$, as in the previous section satisfies Assumption 5 immediately.

For nonlinear g, Assumption 5 does not hold automatically: the conditional forecast $\mathbb{E}[\omega_{it+1} \mid z_{it}]$ may depend not only on $\tilde{\omega}_{it}$ but also on higher-order moments of the proxy error. A sufficient condition is a location–error structure: suppose $\omega_{it} = \tilde{\omega}_{it} + \eta_{it}$, where η_{it}

[2] Doraszelski and Li (2025b) state an analogous restriction to Assumption 5 in their Theorem 1 using two instrument/proxy vectors sets: x_{it-1}, the covariates used to construct the first-stage control/index, and z_{it}, the instrument set used in the second-stage moment restriction. If one *aligns the first-stage conditioning set with the instrument set* by setting $x_{it-1} = z_{it}$, then the Doraszelski–Li index becomes $\mathbb{E}[\omega_{it-1} \mid x_{it-1}] = \mathbb{E}[\omega_{it-1} \mid z_{it}]$, and their condition (9) reduces to the same forecast-sufficiency statement as Assumption 5.

is independent of z_{it} and its distribution is stable across i and t. Then

$$\mathbb{E}[\omega_{it+1} \mid z_{it}] = \Gamma_0(\tilde{\omega}_{it}),$$

for $\Gamma_0(u) := \mathbb{E}[g(u + \eta)]$. As a simple illustration, if $g(\omega) = \omega^2$ and $\mathrm{Var}(\eta_{it}) = \sigma_\eta^2$, then $\Gamma_0(u) = u^2 + \sigma_\eta^2$.

More generally, the distribution of η_{it} may vary across firms or time, due to heteroskedasticity or latent heterogeneity. In this case, the forecast function must also vary across observations, $\Gamma_{0,it}$. To account for this, one may allow Γ_0 to depend flexibly on firm and time effects, for example by modeling $\Gamma_{0,it}(u) = \Gamma_0(u; \alpha_i, \tau_t)$ or by partialing out such effects nonparametrically. Bonhomme et al. (2025) provide methods for consistent estimation under standard regularity conditions. This framework accommodates a wider class of data-generating processes while preserving identification of β_0 under Assumption 5.

3 Identification

3.1 Two-stage Identification

The canonical control–function strategy proceeds in two stages.

First stage

Substituting (1) with the definition of the control function gives

$$y_{it} = x_{it}^\top \beta_0 + \mathbb{E}[\omega_{it} \mid z_{it} = z] + \eta_{it} + \varepsilon_{it}. \tag{6}$$

Assumption 4 implies

$$\mathbb{E}[\eta_{it} + \varepsilon_{it} \mid z_{it}] = 0. \tag{7}$$

Because x_{it} may also enter $\mathbb{E}[\omega_{it} \mid z_{it} = z]$ linearly, β_0 is not identified from (6) directly. We therefore collect terms

$$\Phi_0(z_{it}) := x_{it}^\top \beta_0 + \mathbb{E}[\omega_{it} \mid z_{it} = z] = \mathbb{E}[y_{it} \mid z_{it}],$$

which is nonparametrically identified from the joint distribution of (y_{it}, z_{it}).[3]

Second Stage

By Assumption 5, we can write (1) as

$$y_{it} = x_{it}^\top \beta_0 + \Gamma_0(\widetilde{\omega}_{it-1}) + \zeta_{it} + \varepsilon_{it}. \tag{8}$$

Assumptions 4–5 imply

$$\mathbb{E}[\zeta_{it} + \varepsilon_{it} \mid z_{it-1}] = 0. \tag{9}$$

This implies that the following second-stage regression function is satisfied at the truth (β_0, Γ_0):

$$\begin{aligned} \mathbb{E}[y_{it} \mid z_{it-1}] &= \mathbb{E}[x_{it}^\top \beta_0 \mid z_{it-1}] \; + \; \Gamma_0\big(\widetilde{\omega}_{it-1}\big) \\ &= \mathbb{E}[x_{it}^\top \beta_0 \mid z_{it-1}] \; + \; \Gamma_0\big(\Phi_0(z_{it-1}) - x_{it-1}^\top \beta_0\big), \end{aligned} \tag{10}$$

where in the last line we used the first-stage relation $\widetilde{\omega}_{it-1} = \Phi_0(z_{it-1}) - x_{it-1}^\top \beta_0$.

The second stage model (8)-(9) is a conditional moment model with a finite–dimensional parameter β and an infinite–dimensional nuisance function Γ. Following the semiparametric identification framework of Chen et al. (2014), identification of β_0 proceeds by partialling out the nuisance space induced by functions of the index. In fact our model is similar to their "single IV index model". In order to show identification we will thus use their results.

Define the second–stage residual

$$\rho_{it}(\beta, \Gamma, \Phi) := y_{it} - x_{it}^\top \beta - \Gamma\big(\Phi(z_{it-1}) - x_{it-1}^\top \beta\big).$$

Assuming Φ_0 is identified from the first stage, the conditional moment restriction can be written as

$$m(\beta, \Gamma) \; := \; \mathbb{E}\big[\rho_{it}(\beta, \Gamma, \Phi_0) \mid z_{it-1}\big].$$

with $m(\beta_0, \Gamma_0) = 0$.

Following Chen et al. (2014), for identification we will use the orthogonalized Jacobian

[3] See Chernozhukov et al. (2022) for a general treatment of nonparametric identification in conditional moment models.

matrix. For this we need the derivative of $m(\beta, \Gamma_0)$ with respect to β at β_0

$$m'_\beta = -\mathbb{E}[x_{it}^\top \mid z_{it-1}] + \mathbb{E}\big[\Gamma'_0(\widetilde{\omega}_{it-1})\, x_{it-1}^\top \mid z_{it-1}\big]$$

and the Gateaux derivative of $m(\beta_0, \Gamma)$ with respect to Γ at Γ_0 in direction l

$$m'_\Gamma l = -\,\mathbb{E}[l(\widetilde{\omega}_{it-1}) \mid z_{it-1}].$$

Let $\mathcal{M}$ denote the closure of the linear span of $m'_\Gamma(\Gamma_0 - \Gamma)$

$$\mathcal{M} \;:=\; \overline{\left\{\; \mathbb{E}[l(\widetilde{\omega}_{it-1}) \mid z_{it-1}] \;:\; \mathbb{E}[l(\widetilde{\omega}_{it-1})^2] < \infty \;\right\}}.$$

$\mathcal{M}$ is also called the nuisance tangent space.

For k-th unit vector in $\mathbb{R}^{d_x}$, e_k, define

$$\tau_k^* = \arg\min_{\tau \in \mathcal{M}} \; \mathbb{E}\left[(m'_\beta e_k - \tau)^2\right].$$

Define the $d_x \times d_x$ matrix Π with entries

$$\Pi_{jk} := \mathbb{E}\left[\left(m'_\beta e_j - \tau_j^*\right)\left(m'_\beta e_k - \tau_k^*\right)\right], \qquad j, k = 1, \ldots, d_x.$$

For identification it is necessary that Π has full rank. Economically, the full-rank condition on Π requires that variation in (x_{it}, x_{it-1}) contains a component not spanned by functions of the index $\tilde{\omega}_{it-1}$.

Proposition 1 (Local Identification)**.** *Suppose Assumptions 1–5 hold. Assume further that Γ_0 is continuously differentiable with Lipschitz derivative, and that the regressors satisfy $\mathbb{E}\|x_{it}\|^4 < \infty$ for all t. If the orthogonalized Jacobian matrix Π is nonsingular, then the structural parameter β_0 is locally identified.*

Sketch of proof. We apply Theorem 7 of Chen et al. (2014), which establishes local identification under three main conditions: (i) linear and bounded Fréchet differentiability of the moment function, (ii) full rank of the orthogonalized Jacobian Π, and (iii) uniform smoothness of the score operator with respect to the structural parameter. The rank condition is directly assumed to hold, so we verify the other two conditions (see Appendix for full proof).

Fréchet differentiability. The moment function $m(\beta, \Gamma)$ is Fréchet differentiable at (β_0, Γ_0) with derivative

$$m'(\theta - \theta_0) = m'_\beta(\beta - \beta_0) + m'_\Gamma(\Gamma - \Gamma_0),$$

which is linear by construction. Boundedness follows from the Lipschitz continuity of Γ_0' and the moment condition $\mathbb{E}\|x_{it}\|^4 < \infty$.

Uniform smoothness. Uniform smoothness of $m'_\beta(\beta, \Gamma)$ follows from the Lipschitz continuity of Γ_0' and the fourth-moment condition on x_{it}, together with standard envelope and domination arguments over small neighborhoods of (β_0, Γ_0).

These conditions imply the local identification of β_0 by Theorem 7 of Chen et al. (2014).

$\square$

3.2 Single-stage Identification via a Residual Index

We now present a sufficient condition under which the identification moment condition can be expressed in terms of the observable residual

$$v_{it-1}(\beta) := y_{it-1} - x_{it-1}^\top \beta,$$

rather than the latent productivity index $\widetilde{\omega}_{it-1}$. This leads to a single-index structure that underlies a one-step estimator. A key advantage of this approach is that it avoids the need to separately estimate the functions Φ_0 and Γ_0, thereby reducing the nonparametric dimensionality of the model and improving robustness in finite samples.

Assumption 6. *There exists a measurable function Λ_0 such that, at the true parameter β_0,*

$$\mathbb{E}[\omega_{it+1} \mid z_{it}] = \mathbb{E}\left[\Lambda_0\big(v_{it}(\beta_0)\big) \mid z_{it}\right]. \tag{11}$$

Assumption 6 is stronger than Assumption 5. Note that we do not require uniqueness of Λ_0. It implies the following regression:

$$\mathbb{E}[y_{it} \mid z_{it-1}] = \mathbb{E}[x_{it}^\top \beta_0 \mid z_{it-1}] \;+\; \mathbb{E}\big[\Lambda_0\big(v_{it-1}(\beta_0)\big) \,\big|\, z_{it-1}\big]. \tag{12}$$

Identification of β_0 in this single-stage formulation requires an appropriate rank condition, analogous to Proposition 1 for the two-stage identification above.

To illustrate Assumption 6, we revisit the previous examples of g and show how Λ can be constructed explicitly. Note that

$$v_{i,t-1}(\beta_0) = y_{i,t-1} - x_{i,t-1}^\top \beta_0 = \omega_{i,t-1} + \varepsilon_{i,t-1}.$$

This clarifies the role of Λ_0: conditional on $z_{i,t-1}$, the contribution of the measurement error

$\varepsilon_{i,t-1}$ is “purged” in expectation, so that

$$\mathbb{E}[\Lambda_0(\omega_{i,t-1} + \varepsilon_{i,t-1}) \mid z_{i,t-1}] = \mathbb{E}[g(\omega_{i,t-1}) \mid z_{i,t-1}].$$

For a linear law of motion, $g(\omega) = \rho\,\omega$, Assumption 6 holds with $\Lambda_0(v) = \rho v$ as long as $\mathbb{E}[\varepsilon_{i,t-1} \mid z_{i,t-1}] = 0$, which holds by Assumption 4. If $g(\omega) = \omega^2$, then under restrictions on the second moments $\mathbb{E}[\omega_{i,t-1}\varepsilon_{i,t-1} \mid z_{i,t-1}] = 0$ and $\mathbb{E}[\varepsilon_{i,t-1}^2 \mid z_{i,t-1}] = \sigma_\varepsilon^2$, we have that

$$\Lambda_0(v) = v^2 - \sigma_\varepsilon^2$$

satisfies Assumption 6.

For general g, if we are willing to assume $\varepsilon_{it} \perp (\omega_{it}, z_{it})$ with density f_ε, then if there exists a measurable function Λ_0 such that

$$g(\omega_{it-1}) = \mathbb{E}[\Lambda_0(\omega_{it-1} + \varepsilon_{it-1}) \mid \omega_{it-1}] = \int \Lambda_0(\omega_{it-1} + u)\, f_\varepsilon(u)\, du, \tag{13}$$

also satisfies Assumption 6. So Λ_0, can be viewed as a (generally non-unique) deconvolution of g through the error distribution. Existence of a measurable Λ solving the Fredholm convolution equation (13) is not automatic. This is a standard issue in linear inverse problems; see Carrasco et al. (2007) for discussion of range conditions and regularization in deconvolution settings, and Newey and Powell (2003) for implementation via regularization.

If the distribution of ε_{it} varies with i and t, then a common Λ_0 may fail to exist. A natural relaxation is to allow $\Lambda_{0,it}$:

$$\mathbb{E}[\Lambda_{0,it}(v_{i,t-1}(\beta_0)) \mid z_{i,t-1}] = \mathbb{E}[\omega_{it} \mid z_{i,t-1}].$$

For example, if $g(\omega) = \omega^2$ and we have heteroskedasticity $\mathbb{E}[\varepsilon_{i,t-1}^2 \mid z_{i,t-1}] = \sigma_{\varepsilon,it}^2$, one can take $\Lambda_{0,it}(v) = v^2 - \sigma_{\varepsilon,it}^2$. Implementing such heterogeneity introduces additional nuisance structure, which can be handled using orthogonalization methods for nonlinear panels (Bonhomme et al., 2025).

4 Estimation

This section presents two estimators for the structural parameter β. Both approaches are based on the conditional moment restriction developed in Section 3, but they differ in how they treat the infinite-dimensional nuisance functions.

In both cases we use GMM (Hansen, 1982). For a given parameter vector θ and corresponding sample moment function

$$f_{nT}(\theta) := \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=2}^{T} g_{it}(\theta),$$

where $r_{it}(\theta)$ is the residuals and vector z_{it} collects the instruments, the GMM estimator is defined as

$$\hat{\theta} := \arg\min_{\theta \in \Theta} \left\| f_{nT}(\theta) \right\|_W^2,$$

with $\|v\|_W^2 := v^\top W v$ denoting the quadratic form induced by a symmetric positive semi-definite weighting matrix W.

4.1 Two-Step Estimator

The first approach estimates the conditional expectation Φ_0 and the productivity evolution function Γ in separate steps, followed by GMM estimation of the structural parameters.

Step 1. We estimate the conditional expectation function

$$\Phi_0(z_{it}) := \mathbb{E}[y_{it} \mid z_{it}]$$

nonparametrically, for example by series regression. Let $\hat{\Phi}_0$ denote the resulting estimator.

Step 2. Using (8), define the second-step residual

$$r_{2,it}(\beta, \gamma) := y_{it} - x_{it}^\top \beta - \Gamma_J\left(\hat{\Phi}_0(z_{it-1}) - x_{it-1}^\top \beta; \gamma\right),$$

where we approximate Γ by a sieve $\Gamma_J(v;\gamma) := \sum_{j=1}^{J} \gamma_j \phi_j(v)$, where $\{\phi_j\}_{j=1}^{J}$ is a given set of basis functions and $\gamma \in \mathbb{R}^J$.

Define the second-step sample moment

$$f_{2,nT}(\beta, \gamma) := \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=2}^{T} r_{2,it}(\beta, \gamma)\, z_{it-1}.$$

Under standard regularity conditions on the sieve approximation, the resulting estimator is consistent and asymptotically normal.[4]

4.2 One-Step Estimator

The second approach avoids separate estimation of Φ_0 and Γ_0 by working directly with the production residual

$$v_{it}(\beta) := y_{it} - x_{it}^{\top}\beta,$$

and exploiting the single-index conditional moment restriction in Assumption 6. We treat Λ as an unknown nuisance function of a scalar argument, and approximate it by a sieve.

Define the one-step residual

$$r_{1,it}(\beta,\lambda) := v_{it}(\beta) - \Lambda_J\big(v_{i,t-1}(\beta);\lambda\big) = y_{it} - x_{it}^{\top}\beta - \Lambda_J\big(y_{i,t-1} - x_{i,t-1}^{\top}\beta;\lambda\big),$$

where $\lambda \in \mathbb{R}^J$. The sample moment is

$$f_{1,nT}(\beta,\lambda) := \frac{1}{n(T-1)}\sum_{i=1}^{n}\sum_{t=2}^{T} r_{1,it}(\beta,\lambda)\, z_{i,t-1}.$$

Again, under regularity conditions the resulting estimator is consistent and asymptotically normal (Newey and McFadden, 1994).

5 Conclusion

This paper has provided a semiparametric identification analysis of production function models with noisy proxy variables, relaxing the scalar–unobservable condition underlying the OP/LP/ACF framework. When intermediate-input demand contains an unobserved error and productivity is only partially predictable from observables, exact recovery of productivity is no longer feasible, yet identification of the structural coefficients remains possible under conditional mean independence and a forecast-sufficiency condition. Casting the problem as a semiparametric conditional moment model in the sense of Chen et al. (2014), we show that local identification is governed by the full rank of an orthogonalized Jacobian matrix that strips out the nuisance tangent space generated by functions of the index. We also outline two GMM estimators—a flexible two-step series estimator and a more structured one-step estimator based on an additional single-index restriction—illustrating how control-

[4]See, e.g., Newey and McFadden (1994) for series GMM asymptotics.

function methods can be extended to empirically realistic environments with measurement error, frictions, and imperfect proxies without relying on strong invertibility or completeness assumptions.

References

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

Bonhomme, S., Jochmans, K., and Weidner, M. (2025). A neyman-orthogonalization approach to the incidental parameter problem.

Carrasco, M., Florens, J.-P., and Renault, E. (2007). Chapter 77 linear inverse problems in structural econometrics estimation based on spectral decomposition and regularization. In *Handbook of Econometrics*, volume 6, pages 5633–5751. Elsevier B.V.

Chen, X., Chernozhukov, V., Lee, S., and Newey, W. K. (2014). Local identification of nonparametric and semiparametric models. *Econometrica*, 82(2):785–809.

Chernozhukov, V., Escanciano, J. C., Ichimura, H., Newey, W. K., and Robins, J. (2022). Locally robust semiparametric estimation. *Econometrica*, 90(4):1501–1535.

Doraszelski, U. and Li, L. (2025a). A generalized control function approach to production function estimation. AEA Papers and Proceedings, forthcoming.

Doraszelski, U. and Li, L. (2025b). Production function estimation without invertibility: Imperfectly competitive environments and demand shocks. Working paper, 2025.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054.

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

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using intermediate inputs to control for unobservables. *Review of Economic Studies*, 70(2):317–341.

Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In Engle, R. F. and McFadden, D., editors, *Handbook of Econometrics*, volume 4, chapter 36, pages 2111–2245. Elsevier.

Newey, W. K. and Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578.

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

Appendix

A.1 Proof of Proposition 1

We verify the sufficient conditions for local identification in Theorem 7 of Chen et al. (2014) by establishing their Assumptions 4–5. Let $\|\cdot\|$ denote the Euclidean norm.

Fréchet differentiability (Assumption 4). The Gateaux derivative of the moment condition $m(\beta, \Gamma)$ is given by

$$m'(\theta - \theta_0) = m'_\beta(\beta - \beta_0) + m'_\Gamma(\Gamma - \Gamma_0),$$

where

$$m'_\beta v = -\mathbb{E}[x_{it} - \Gamma'_0(\widetilde{\omega}_{it-1})x_{it-1} \mid z_{it-1}], \quad m'_\Gamma l = -\mathbb{E}[l(\widetilde{\omega}_{it-1}) \mid z_{it-1}].$$

Linearity is immediate. Boundedness follows from the L^2 contraction of conditional expectation and the assumptions:

$$\|m'_\Gamma l\| \leq \|l(\widetilde{\omega}_{it-1})\|_{L^2}, \quad \|m'_\beta v\| \leq \left(\sqrt{\mathbb{E}\|x_{it}\|^2} + \|\Gamma'_0\|_\infty \sqrt{\mathbb{E}\|x_{it-1}\|^2}\right) \|v\|.$$

Rank condition. By assumption, the orthogonalized Jacobian matrix Π is nonsingular.

Uniform smoothness (Assumption 5). Fix $\varepsilon > 0$ and define neighborhoods

$$\mathcal{B}_r = \{\beta : \|\beta - \beta_0\| \leq r\}, \quad \mathcal{N}_\Gamma^\delta = \{\Gamma : \sup_v |\Gamma'(v) - \Gamma'_0(v)| \leq \delta\}.$$

Define

$$V(\beta, \Gamma) := \Gamma'(\widetilde{\omega}_{it-1} + x_{it-1}^\top(\beta_0 - \beta))x_{it-1}.$$

Then

$$\begin{aligned}\|m'_\beta(\beta,\Gamma) - m'_\beta(\beta_0,\Gamma_0)\|^2 &= \|\mathbb{E}[V(\beta,\Gamma) - V(\beta_0,\Gamma_0) \mid z_{it-1}]\|^2 \\ &\leq 2\|\mathbb{E}[V(\beta,\Gamma) - V(\beta,\Gamma_0) \mid z_{it-1}]\|^2 \\ &\quad + 2\|\mathbb{E}[V(\beta,\Gamma_0) - V(\beta_0,\Gamma_0) \mid z_{it-1}]\|^2.\end{aligned}$$

For the first term, Jensen's inequality and the definition of $\mathcal{N}_\Gamma^\delta$ yield:

$$\sup_{\Gamma\in\mathcal{N}_\Gamma^\delta} \mathbb{E}\left[\sup_{\beta\in\mathcal{B}_r} \|\mathbb{E}[V(\beta,\Gamma) - V(\beta,\Gamma_0) \mid z_{it-1}]\|^2\right] \leq \delta^2 \mathbb{E}\|x_{it-1}\|^2.$$

For the second term, using Lipschitz continuity of Γ'_0 (constant C) and Cauchy–Schwarz:

$$\mathbb{E}\left[\sup_{\beta\in\mathcal{B}_r} \|\mathbb{E}[V(\beta,\Gamma_0) - V(\beta_0,\Gamma_0) \mid z_{it-1}]\|^2\right] \leq C^2 r^2 \mathbb{E}\|x_{it-1}\|^4.$$

Combining both terms,

$$\sup_{\Gamma\in\mathcal{N}_\Gamma^\delta} \left(\mathbb{E}\left[\sup_{\beta\in\mathcal{B}_r} \|m'_\beta(\beta,\Gamma) - m'_\beta(\beta_0,\Gamma_0)\|^2\right]\right)^{1/2} \leq \left(2\delta^2 \mathbb{E}\|x_{it-1}\|^2 + 2C^2 r^2 \mathbb{E}\|x_{it-1}\|^4\right)^{1/2}.$$

Choosing δ, r small enough ensures this is below ε.

Conclusion. With Assumptions 4–5 verified and Π nonsingular, Theorem 7 of Chen et al. (2014) implies that β_0 is locally identified.