Estimating Factor Price Markdowns Using Production Models

Michael Rubens* Yingjie Wu Mingzhi (Jimmy) Xu

June 10, 2025

Abstract

Factor price markdowns are a key object of interest when studying monopsony power. In this article, we test the performance of "production approaches" to estimate factor price markdowns, which have been used increasingly often in the literature. We evaluate the performance of these estimators under various data-generating processes using Monte Carlo simulations. We discuss the commonly made assumptions in this class of estimators, and we address the methodological challenges involved with relaxing these assumptions, such as departing from Hicks neutrality, allowing for nonsubstitutable inputs, and allowing for various types of labor market conduct.

Keywords: Monopsony, Markdowns, Factor-Biased Technological Change, Production Function Estimation
JEL Codes: L11, J42, O33

*Michael Rubens, Department of Economics, UCLA, email: rubens@econ.ucla.edu; Yingjie Wu, Department of Economics, UCLA; Mingzhi Xu, INSE, Peking University. This paper is based on a presentation given at the 2024 EARIE Annual Conference in Amsterdam. It combines new results with a discussion of existing published and unpublished work, as cited throughout the paper. We thank the editor, Myrto Kalouptsidi, and one anonymous referee for their helpful feedback.

1 Introduction

Interest is growing in the monopsony power of firms on their labor and other factor markets. Production functions are increasingly used to estimate "factor price markdowns," a key object of interest when studying monopsony power.[1] However, there is scant evidence on the performance of these "production approach" estimators. To fill this gap, we conduct Monte Carlo simulations in which we generate oligopsonistic labor market equilibria. We use these simulated data to estimate factor price markdowns using existing production approaches. Comparing these estimates to the true markdown distribution enables us to assess how well these production-approach estimators recover true markdowns under a variety of data-generating processes (DGPs).

We start by discussing existing production-approach markdown estimators that extend the markup-estimation approach of Hall (1988) and De Loecker and Warzynski (2012) to allow for endogenous factor prices (Dobbelaere & Mairesse, 2013; Morlacco, 2017; Mertens, 2019; Brooks, Kaboski, Li, & Qian, 2021; Yeh, Hershbein, & Macaluso, 2022; Rubens, 2023; Mertens & Schoefer, 2024; Delabastita & Rubens, 2025). These approaches share an assumption that the price of at least one variable input is exogenous, and they normalize the markup expressions from the cost-minimization first-order conditions of all other variable inputs compared to the variable input with the exogenous input price. This allows for recovering the markdowns for all other inputs.

We test this class of production approaches to markdown estimation by simulating data that are generated under a discrete-choice labor supply model in the spirit of Berry (1994) and Card, Cardoso, Heinig, and Kline (2018), in which we let firms compete oligopsonistically à la Nash-Bertrand. In contrast to Card et al. (2018), who impose monopsonistic competition, we allow for granular employers, similarly to Berger, Herkenhoff, and Mongey (2022) and Azar, Berry, and Marinescu (2022). On the labor demand side, we assume cost-minimizing firms that produce following a Cobb-Douglas production function with two variable inputs: labor and materials. Given the oligopsonistic labor supply side, labor wages are endogenous to individual firms, meaning that the residual labor supply curves are upward-sloping. In contrast, we assume that material prices are taken as given by the individual firms. We run a Monte Carlo simulation in which we sample 250 firms that operate in 50 different markets over 10 years, then we solve for labor market equilibrium in each of these markets. We resample this simulation exercise for 200 random draws. We find that under the Hicks-

[1] See Syverson (2024) for an excellent survey of the literature on markups and markdowns.

neutral DGP, the production-based markdown estimator delivers precise, consistent estimates of wage markdowns and of the production-function coefficients.

Next, we argue that Hicks neutrality is the key assumption needed for markdown identification, as discussed in more detail in Rubens, Wu, and Xu (2024).[2] We relax this assumption by allowing for random coefficients in the production function. Such random coefficients arise if there is unobserved technological heterogeneity between firms, as is likely to occur in many applications. We rerun our Monte Carlo simulation under an identical labor supply side, but with the non-Hicks-neutral production function on the labor demand side. We find that in contrast to the Hicks-neutral DGP, both the production-function coefficients and the wage markdowns are estimated with considerable bias. The main reason for this bias is that latent technological heterogeneity and wage markdowns are not separately identified using the cost-minimization first-order conditions.

We test an alternative production-function estimator proposed by Rubens et al. (2024), which adapts the production-function estimator of Doraszelski and Jaumandreu (2018) to allow for imperfect factor market competition. We find that this estimator delivers consistent estimates of the production function and of markdowns when there is unobserved heterogeneity in the production-function coefficients. We also simulate the performance of this estimator when the true production function is Hicks-neutral: we find that the estimator of Rubens et al. (2024) is less efficient than the Hicks-neutral markdown estimator in this case, but that it still recovers the true production-function coefficients with minimal bias.

We conclude by discussing how other commonly made assumptions have been relaxed in the literature. We discuss allowing for nonsubstitutable inputs in production, labor-market-conduct assumptions other than Nash-Bertrand, differentiated goods and inputs, adjustment frictions, and cases in which none of the factor prices can be reasonably assumed to be competitive.

The rest of this paper is structured as follows. In Section 2, we set up the primitives and behavioral assumptions of our model. In Section 3.2, we estimate markdowns using various simulated DGPs to test and compare various production-approach markdown estimators. Section 4 discusses further extensions. Section 5 concludes.

[2] Another paper that studies markdowns while departing from Hicks neutrality is Azzam, Jaumandreu, and Lopez (2025), which studies U.S. meatpacking firms. That paper, in contrast to ours, does not rely on estimating a factor supply curve to separately identify factor price markdowns from factor-augmenting productivity differences.

2 Model

2.1 Primitives

We start by providing a model of factor demand and supply under oligopsonistic competition, which serves as a DGP to test the various approaches to markdown estimation. In what follows, we discuss the model primitives: the production function and the factor supply model.

Production Function

Let firms be indexed by f and time periods by t. We assume that firms use two factors of production: labor L_{ft} and materials M_{ft}, which are transformed into a scalar output level Q_{ft} following a production function $H(.)$. Hicks-neutral productivity is indicated by Ω_{ft}, measurement error in output is indicated by ε_{ft}, and there is possibly unobserved heterogeneity in the production-function coefficients β_{ft}:

$$Q_{ft} = H(L_{ft}, M_{ft}, \beta_{ft})\Omega_{ft} \exp(\varepsilon_{ft}) \tag{1a}$$

Next, we highlight three assumptions.[3] Assumption 1 rules out perfect complementarities between inputs, such as in a Leontief production function. We relax this assumption in Section 4.1.

***Assumption* 1** *The production function $H(.)$ is twice differentiable.*

Assumption 2 assumes that both the produced good and the inputs are undifferentiated. We discuss how to relax this assumption in Section 4.4.

***Assumption* 2** *Both the good Q_{ft} and the inputs L_{ft} and M_{ft} are homogeneous.*

For the simulations, we impose a simple Cobb-Douglas production function, which in logs gives Equation (1b), but any production function that satisfies Assumption 1 can be used. The output elasticities of labor and materials are denoted as β^l_{ft} and β^m_{ft}:

$$q_{ft} = \beta^l_{ft} l_{ft} + \beta^m_{ft} m_{ft} + \omega_{ft} + \varepsilon_{ft} \tag{1b}$$

[3] As a convention throughout the paper, we list the assumptions that are used to estimate the production function and markdowns, whereas other assumptions that are used to simulate the data but are not key to estimating the production function are simply mentioned, but are not listed separately.

We impose an AR(1) transition process for the sum of the log of Hicks-neutral productivity and measurement error, with serial correlation ρ and i.i.d. productivity shocks e_{ft}. This assumption is useful—because it allows for estimating the production function using a dynamic panel approach—but is not strictly necessary.

$$\omega_{ft} + \varepsilon_{ft} = \rho(\omega_{ft-1} + \varepsilon_{ft-1}) + e_{ft} \tag{2}$$

Finally, we assume that both labor and materials are variable, static inputs, meaning that they are not subject to adjustment frictions and fully depreciate during every period.[4]

***Assumption* 3** *Labor and materials are variable, static inputs.*

Labor Supply

We impose a discrete choice model of labor supply with oligopsonistic competition in the spirit of Berry (1994) and Card et al. (2018), to simulate an environment in which markdowns vary between firms, and in which firms set wages strategically. Firms pay per-unit wages W_{ft}^{l} to workers i, who choose their employment between a set of firms, $\mathcal{F}_t$, with $f = 0$ indicating the outside option of being unemployed. We assume that firms are not able to wage-discriminate between workers. We assume that the utility of a worker i who works at firm f depends on the log of the wage W_{ft}^{l}, an unobserved amenity ξ_{ft}, and an i.i.d. type-I distributed firm-worker error term υ_{ift}:

$$U_{ift} = \underbrace{\gamma \ln(W_{ft}^{l}) + \xi_{ft}}_{\equiv \delta_{ft}} + \upsilon_{ift} \tag{3}$$

We denote mean utility as δ_{ft}, which we normalize to zero for the outside option, as usual: $\delta_{0t} = 0$. Using the logit formula, the labor market share $s_{ft} = \frac{L_{ft}}{\sum_{g \in \mathcal{F}_t} L_{gt}}$ is given by

$$s_{ft} = \frac{\exp(\delta_{ft})}{\sum_{g \in \mathcal{F}_t} \exp(\delta_{gt})}$$

[4]Fixed inputs, such as capital, can be added to the model, but they need to be solved using a dynamic investment model, rather than the static cost-minimization problem for the variable inputs. However, fixed inputs cannot be used to identify markdowns in this approach, as it requires normalizing first-order conditions from the static cost-minimization approach.

Denoting the labor force as $\overline{L}$, the labor supply function is given by

$$L_{ft} = \frac{\exp(\gamma \ln(W^l_{ft}) + \xi_{ft})}{\sum_{g \in \mathcal{F}_t} \exp(\gamma \ln(W^l_{ft}) + \xi_{ft})} \overline{L} \tag{4}$$

We denote the ψ^l_{ft} and ψ^m_{ft} as the inverse of the supply elasticities of labor and materials, or in short, "inverse supply elasticities"[5]:

$$\psi^l_{ft} \equiv \frac{1}{\frac{\partial L_{ft}}{\partial W^l_{ft}} \frac{W^l_{ft}}{L_{ft}}} \qquad \psi^m_{ft} \equiv \frac{1}{\frac{\partial M_{ft}}{\partial W^m_{ft}} \frac{W^m_{ft}}{M_{ft}}} \tag{5}$$

The imposed labor supply model implies the following inverse labor supply elasticity[6] ψ^l_{ft}:

$$\psi^l_{ft} = \frac{1}{\gamma(1 - s_{ft})} \tag{6}$$

We assume that materials are supplied perfectly elastically, meaning that $\psi^m_{ft} = 0$. Therefore, intermediate-input prices are exogenous to individual firms.

***Assumption* 4** *Residual intermediate-input supply is perfectly price elastic:* $\psi^m_{ft} = 0$.

2.2 Behavior and Equilibrium

Cost Minimization

Producers choose inputs in every period to minimize variable costs, forming expectations about output Q. Given that the shock ε_{ft} is assumed to be classical measurement error, firms' expectations of output are $E[Q_{ft} \exp(\varepsilon_{ft})] = Q_{ft}$. Therefore, firms choose inputs such that $Q_{ft} = H(.)\Omega_{ft}$. Denoting marginal costs as λ_{ft}, the cost-minimization problem is given by Equation (7):

$$\min_{W^l_{ft}, M_{ft}} \Big[W^m_{ft} M_{ft} + W^l_{ft} L_{ft} - \lambda_{ft}\big(Q_{ft} - H(.)\Omega_{ft}\big) \Big] \tag{7}$$

[5]This inverse elasticity is not to be confounded with the elasticity of inverse supply $\frac{\partial W^l_{ft}}{\partial L_{ft}} \frac{L_{ft}}{W_{ft}}$, which does not have the same value under oligopsonistic competition.

[6]We derive this in Appendix A.2.

Output is sold at a price P_{ft} and firms set a price markup $\mu^p_{ft} \equiv (P_{ft} - \lambda_{ft})/\lambda_{ft}$. Solving the cost-minimization problem delivers the following labor demand and material demand[7]:

$$\underbrace{\frac{P_{ft}}{(\mu^p_{ft}+1)}\beta^l L_{ft}^{\beta^l-1} M_{ft}^{\beta^m}\Omega_{ft}}_{MRPL_{ft}} = \underbrace{W^l_{ft}(1+\psi^l_{ft})}_{MCL_{ft}} \tag{8}$$

$$\underbrace{\frac{P_{ft}}{(\mu^p_{ft}+1)}\beta^m L_{ft}^{\beta^l} M_{ft}^{\beta^m-1}\Omega_{ft}}_{MRPM_{ft}} = \underbrace{W^m_{ft}}_{MCM_{ft}} \tag{9}$$

Firms set wages and material quantities such that the marginal revenue product of each input ($MRPL$ and $MRPM$) equates to its marginal cost (MCL and MCM). For materials, the marginal cost is simply the materials price, as this price is exogenous, but for labor the marginal cost includes the inverse labor supply elasticity.

We define equilibrium as the input quantity and prices vector $(L_{ft}, M_{ft}, W^l_{ft})$ that is the solution of the system of equations consisting of (4), (8), and (9), which are labor supply, labor demand, and materials demand. Materials supply, being perfectly elastic, does not enter this system of equations, as the materials price W^m is exogenous.

Markups and Markdowns

Rearranging the first-order conditions, the markup μ^p_{ft} can be expressed as a function of output elasticities, input revenue shares, and input supply elasticities, as also done in De Loecker, Goldberg, Khandelwal, and Pavcnik (2016):

$$\mu^p_{ft} = \frac{\beta^j_{ft}}{\alpha^j_{ft}(1+\psi^j_{ft})\exp(\varepsilon_{ft})} - 1 \quad \forall j = l, m \tag{10}$$

in which α^j_{ft} denotes the expenditure on input j as a share of sales, $\alpha^l_{ft} \equiv W^l_{ft}L_{ft}/P_{ft}Q_{ft}$ and $\alpha^m_{ft} \equiv W^m_{ft}M_{ft}/P_{ft}Q_{ft}$.

The markdown of the wage below the marginal revenue product of labor is denoted as $\mu^w_{ft} \equiv (MRPL_{ft} - W^l_{ft})/MRPL_{ft}$, and can be expressed as a function of the inverse elasticity of labor supply:

$$\mu^w_{ft} = \frac{\psi^l_{ft}}{1+\psi^l_{ft}} \tag{11}$$

[7]We derive these in Appendix A.1.

The more inelastic the labor supply curve (larger ψ), the greater a firm's ability to exercise monopsony power and suppress wages, which would result in a larger wage markdown.

3 Identification and Estimation

3.1 Hicks-Neutral Approach

Markdown Identification

We start by discussing estimation of the inverse labor supply elasticity ψ^l_{ft} through the production approach. Following Dobbelaere and Mairesse (2013), Morlacco (2017), Brooks et al. (2021), and Yeh et al. (2022), the inverse labor supply elasticity can be expressed by weighting the ratio of input expenditures by the respective output elasticities of both inputs:

$$\psi^l_{ft} = \frac{\beta^l_{ft}}{\beta^m_{ft}} \frac{\alpha^m_{ft}}{\alpha^l_{ft}} - 1 \tag{12}$$

Equation (12) makes clear that as soon as there is no unobserved heterogeneity in the production-function coefficients β^l_{ft} and β^m_{ft}, the relative markdown of labor wages compared to the material price markdown is identified if the production-function coefficients are identified. Therefore, this approach requires imposing the additional assumption of Hicks neutrality, which imposes that there is no unobserved heterogeneity in the coefficients β^l_{ft} and β^m_{ft}.

Assumption 5 (Hicks neutrality): *There is a scalar unobservable in production, Ω_{ft}. This implies that the coefficients β^l_{ft} and β^m_{ft} are fully observed.*

Under a Cobb-Douglas functional form, Assumption 5 implies homogeneous output elasticities, $\beta^l_{ft} = \beta^l$ and $\beta^m_{ft} = \beta^m$. Other functional forms for the production function, such as a translog or a CES production function, can allow for heterogeneity in output elasticities across firms and time. However, the key assumption implied by Hicks neutrality is that there is no unobserved heterogeneity in these output elasticities, that the only source of unobserved heterogeneity in production is the scalar productivity residual Ω_{ft}.

Under Assumption 5, the production-approach markdown expression (12) is point-identified as soon as the output elasticities β^l and β^m are identified.

Identifying the Production Function

To identify the production function, we combine imposing timing assumptions on input choices, as in Ackerberg, Caves, and Frazer (2015), with relying on the law of motion for productivity from Equation (2), as in Blundell and Bond (2000). Taking ρ-differences, the productivity shock can be written as

$$e_{ft} = q_{ft} - \rho q_{ft-1} - \beta^l(l_{ft} - \rho l_{ft-1}) - \beta^m(m_{ft} - \rho m_{ft-1})$$

Assuming that labor and materials are both variable inputs, the following moment conditions are formed for lags $r = 1$ up to $r = T - 1$, with the panel length being denoted as T. As in Ackerberg et al. (2015), the identifying assumption is that the variable inputs (in our case, materials and labor) are chosen after the firm observes the observable component of the productivity shock e_{ft}[8]:

$$\mathbb{E}\Big[e_{ft}(\rho, \beta^l, \beta^m)| \begin{pmatrix} L_{ft-r} \\ M_{ft-r} \end{pmatrix} \Big]_{r=1}^{T-1} = 0 \tag{13}$$

We estimate the production-function coefficients (β^l, β^m) using these moment conditions, including the first and second lag of labor and materials as instruments.[9] We use the resulting estimates $(\hat{\beta}^l, \hat{\beta}^m)$ to estimate the inverse supply elasticity ψ^l_{ft} using Equation (12). Hence, the inverse labor supply elasticity is estimated from the production function alone, without needing to estimate the labor supply parameters γ and ξ_{ft}.

Monte Carlo Simulation

We simulate a dataset of 50 independent labor markets that each contain five firms, which are observed over 10 years. Hence, the simulated dataset contains 250 firms and 2,500 observations. We parametrize the true output elasticities of labor and materials[10] at $\beta^l = 0.5$ and $\beta^m = 0.3$. We let intermediate input prices W^m_{ft} in the first year be distributed as a normal distribution $W^m_{f1} \sim \mathcal{N}(5, 0.05)$ and let it evolve by firm-level shocks that are $\mathcal{N}(0, 0.01)$ distributed. Similarly, we let the initial log productivity distribution be normally distributed

[8] The part of e_{ft} that is caused by measurement error ε_{ft} is unobservable to firms and iid, and is thus orthogonal to current and past input choices by assumption. However, firms observe the part of e_{ft} that is due to shocks to total factor productivity ω_{ft}.

[9] Given that we need to identify three parameters but have four moment conditions, our estimator is overidentified.

[10] We let the labor coefficient be larger than the material coefficient because we allow for heterogeneity in the labor coefficient, and because we want to avoid labor coefficients being close to or below zero.

$\omega_{f1} \sim \mathcal{N}(1, 0.1)$ and let the productivity shocks be $\mathcal{N}(0, 0.001)$ distributed. The serial correlation in productivity is set at $\rho = 0.6$. The resulting distribution of log productivity has a mean of 1/4 and a standard deviation of 1/3. We draw measurement error ε_{ft} from a uniform distribution on the $[0, \frac{1}{1000}]$ interval. We normalize the total labor market size to one, and we assume exogenous product prices, which we also normalize to one, $P = 1$. Given the exogenous prices assumption, we set the price markup to $\mu = 0$.

We conduct a Monte Carlo simulation with 200 independent draws. For each iteration, we numerically solve the model by finding the equilibrium wages and market shares of all firms $(L_{ft}, M_{ft}, W^l_{ft})$ such that the system of equations consisting of (4), (8), and (9) are solved for every firm in every year, and such that labor supply is equal to labor demand at the market level.

Estimation

We estimate the production-function parameters ρ, β^l, and β^m on the resulting dataset using the moment conditions from Equation (13), and then we plug these estimates into Equation (12) to estimate the inverse residual labor supply elasticities ψ^l_{ft} at all firms in every year.

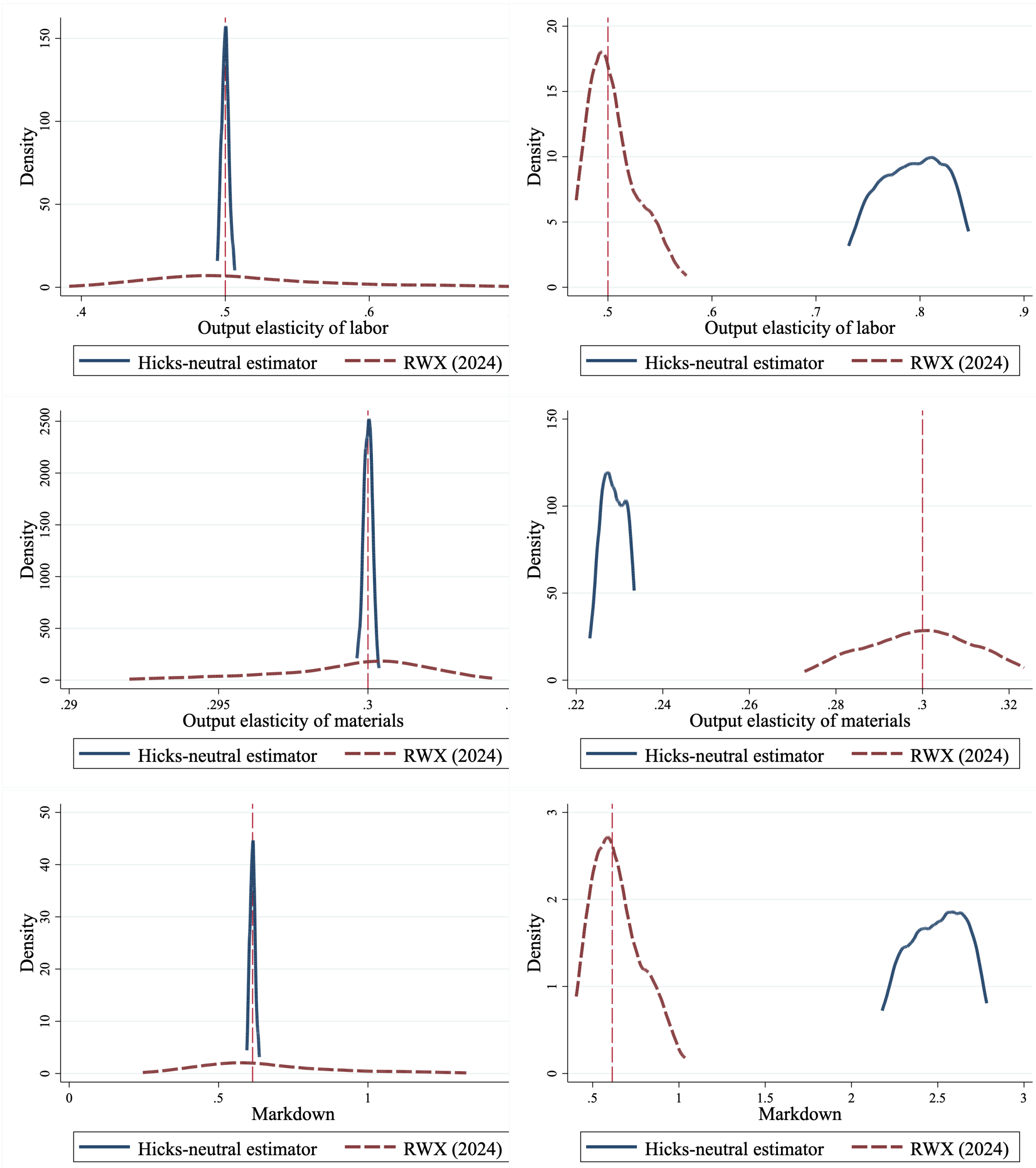
Results

The distribution of the resulting estimates are visualized in the solid blue lines in Panel A of Figure 1. We find that the Hicks-neutral production-function estimator yields precise, consistent estimates of the output elasticities of labor and materials. As summarized in Panel A of Table 1, the output elasticities of labor and materials are estimated at their true values of 0.5 and 0.3, with the standard deviation of these estimates across bootstrap iterations being very small, at 0.003 for labor and below 0.001 for materials. Hence, the production function is identified even if the labor market is imperfectly competitive. Moreover, we find that the production-function estimator delivers a consistent estimate of the inverse labor supply elasticity ψ^l_{ft}, which is on average estimated at its true average of 0.615, with a standard deviation across draws of merely 0.009.

The assumptions imposed on the labor supply model and on conduct were only necessary to simulate the dataset; they were not used to estimate production function: we estimated markdowns correctly using the production function while remaining agnostic about the model of competition on the labor market, the functional form for labor utility, and the distribution of wage markdowns.

Figure 1: Monte-Carlo Simulations

A: Hicks-Neutral DGP | **B: Factor-Biased DGP**



3.2 Introducing Unobserved Technological Heterogeneity

We now revisit the identification approach outlined in Section 2 by relaxing Assumption 5, Hicks neutrality. This is a departure from prior production-approach markdown estimators, including Morlacco (2017), Mertens (2019), Brooks et al. (2021), Yeh et al. (2022), Rubens (2023), Mertens and Schoefer (2024), and Delabastita and Rubens (2025).

Instead of the Cobb-Douglas production function with constant output elasticities from Equation (1b), we allow for unobserved random coefficients $(\beta^l_{ft}, \beta^m_{ft})$. We retain the Cobb-Douglas functional form for simplicity, but we refer to Rubens et al. (2024) for estimation of a Constant Elasticity of Substitution production function under imperfect labor market competition, and for an empirical application in the context of the Chinese nonferrous metals industry.

Identification

Equation (12) makes clear that to identify the wage markdown, it is crucial to fully estimate the random coefficients β^m_{ft} and β^l_{ft}. Although there is a literature on estimating production functions with nonscalar unobservables, such as Doraszelski and Jaumandreu (2018) and Demirer (2019), these estimators rely on the assumption of perfect factor market competition, which imposes $\psi^l_{ft} = \psi^m_{ft} = 0$. In contrast, Rubens et al. (2024) develop an estimator that allows for both nonscalar unobservables in production and imperfect factor market competition. In this section, we lay out this identification strategy in the context of our production model.

Rubens et al. (2024) rely on jointly estimating the labor supply curve and the production function. Using the discrete-choice labor supply model imposed above, the labor supply equation to be estimated is given by

$$s_{ft} - s^0_{it} = \gamma \ln(W_{ft}) + \xi_{ft} \tag{14}$$

As was derived in Equation (6), an estimate of the inverse labor supply elasticity, $\hat{\psi}^l_{ft}$, can be recovered as a function of the estimated wage coefficient in labor supply $\hat{\gamma}$ and the observed labor market share s_{ft}. Then, from Equation (12), one can express the output elasticity of labor as a function of the estimated wage markdown $\hat{\psi}^l_{ft}$, the observed revenue shares α^l_{ft} and α^m_{ft}, and the yet-to-be-estimated materials coefficient β^m. We opt to impose a homogeneous coefficient β^m, but to allow for heterogeneous returns to scale[11]:

$$\hat{\beta}^l_{ft} = \frac{\hat{(}\psi^l_{ft} + 1)\alpha^l_{ft}\beta^m}{\alpha^m_{ft}} \tag{15}$$

Substituting this output elasticity of labor into the production function results in Equation (16), in which the term $a_{ft} \equiv \frac{\hat{\psi}^l_{ft}\alpha^l_{ft}l_{ft}}{\alpha^m_{ft}} + m_{ft}$ is composed solely of observed and estimated terms. Hence, the error term in the production function is again reduced to a scalar unobservable ω_{ft} and measurement error ε_{ft}:

$$q_{ft} = \beta^m \underbrace{[\frac{\hat{(}\psi^l_{ft} + 1)\alpha^l_{ft}l_{ft}}{\alpha^m_{ft}} + m_{ft}]}_{a_{ft}} + \omega_{ft} + \varepsilon_{ft} \tag{16}$$

Again using the equation of motion for productivity, we isolate the productivity shock e_{ft} as

$$e_{ft} = q_{ft} - \rho q_{ft-1} - \beta^m(a_{ft} - \rho a_{ft-1})$$

The moment conditions to estimate the parameters (β^m, ρ) are given by

$$\mathbb{E}\Big[e_{ft}(\rho, \beta^m)| \begin{pmatrix} L_{ft-r} \\ M_{ft-r} \end{pmatrix} \Big]_{r=1}^{T-1} = 0 \tag{17}$$

We again estimate the production-function parameters taking up to two lags. Using the estimated materials coefficient $\hat{\beta}^m$, the full distribution of the output elasticities of labor β^l_{ft} can be recovered using Equation (15), which is now a function of data and estimated parameters from labor supply (γ) and the production function (β^m).

Monte Carlo Simulations

We keep the same parametrization from the Monte Carlo simulation in Section 2, with the only difference that we now allow for unobserved heterogeneity in the output elasticity of labor. We parametrize this unobserved heterogeneity as $\beta^l_{ft} \sim \mathcal{U}[\frac{1}{3}, \frac{2}{3}]$. Although we assume that the error terms v^l_{ft} and v^m_{ft} are idiosyncratic in the Monte Carlo simulation, one could allow for persistent differences and a time trend in this technological heterogeneity, in or-

[11] Alternatively, one could easily allow for heterogeneity in β^m_{ft} while imposing homogeneous returns to scale.

der to incorporate factor-biased technological change over time. We solve for labor market equilibrium using the same procedure that was outlined in Section 3.1.

Estimation

We estimate the production function and markdowns twice. First, as a means of comparison, we follow the estimation procedure outlined in Section 3.1, which assumes Hicks neutrality. Second, we estimate the production function using the estimation procedure from Rubens et al. (2024) that was outlined above. We start by estimating Equation (14). Given the latent firm amenities ξ_{ft}, we need to find an instrument for wages that is excluded from the error term ξ_{ft}. We assume that a labor demand shifter z_{ft} is available, which we construct as a variable that is correlated with productivity but uncorrelated to the amenity firm ξ_{ft}. We parametrize this labor demand shifter as the sum of TFP and an error term u_{ft}, which is normally distributed with a zero mean and standard deviation of 0.01:

$$z_{ft} = \frac{\Omega_{ft}}{2} + u_{ft}$$

With this labor demand shifter at hand, we estimate the labor supply curve (14) using 2SLS. Using the estimated parameter $\hat{\gamma}$ and the observed labor market share s_{ft}, we compute wage markdowns following Equation (6). Finally, we substitute the markdown estimate ψ^l_{ft} into Equation (16) and form the moment conditions (17) to estimate the production-function parameters β^m and ρ. The full distribution of the output elasticity β^l_{ft} can then be recovered using Equation (16).

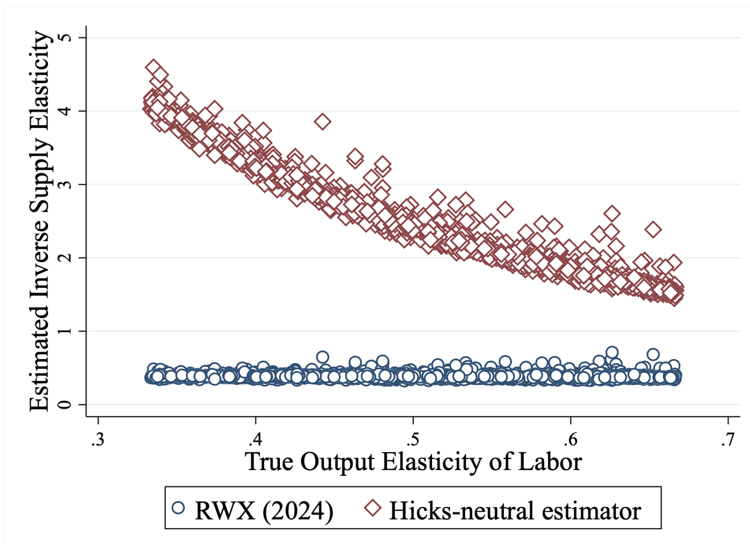
Results Under the Factor-Biased Data-Generating Process

We visualize the production-function estimates for the DGP with random coefficients in the production function in Panel B of Figure 1. The solid blue lines in Figure 1 report the estimates using the Hicks-neutral production-function estimator. It is clear that the markdown estimator that relies on Hicks neutrality does a poor job of estimating the production-function coefficients: the labor coefficient is estimated at 0.79, which is 58% above its true value, whereas the materials coefficient is estimated at 0.23, which is 24% below the true value. As a result, the Hicks-neutral model estimates ψ^l at 2.494 on average, which is four times larger than the true average value of 0.613. This leads the econometrician to believe that wages are marked down 71% below the marginal revenue product of labor,[12] whereas wages are in reality marked down 38%.

[12]Using Equation (11), $\mu^w = (2.494/(1 + 2.494))$.

Figure 2 shows the source of the identification problem by plotting the estimated inverse labor supply elasticity against the true output elasticity of labor, β^l_{ft} across observations in a single bootstrap iteration (the first of 200 iterations), for both estimators. In the Hicks-neutral model, the latent variation in the output elasticity of labor is interpreted as wage-markdown variation: firms with high output elasticities of labor are estimated to set a low wage markdown, because their cost share of labor is higher than average. In contrast, our estimator delivers elasticities of inverse labor supply that are independent of the output elasticity of labor, as is true in the underlying DGP.

Figure 2: Estimated Inverse Labor Supply Elasticity vs. Output Elasticity of Labor



The red dashed lines in Panel B of Figure 1 plot the estimates using the method of Rubens et al. (2024) for the random-coefficients DGP. The inverse labor supply elasticity is estimated with a small bias, at 0.669 compared to the true value of 0.613, which is due to the small-sample properties of the instrumental variables estimator of labor supply. Turning to the production-function coefficients, we find that our estimator delivers consistent output elasticity estimates. Hence, the production function can be estimated even with random coefficients and imperfect labor market competition, but it needs to be estimated jointly with the labor supply curve.

Results Under the Hicks-Neutral Data-Generating Process

How does the estimator of Rubens et al. (2024) perform if there is in reality no unobserved heterogeneity in the output elasticities? Panel A of Table 1 shows that the output elasticities of labor is still estimated reasonably close to the truth, at 0.516, which implies an upward bias of 3.2%, whereas the materials elasticity is estimated consistently. The standard errors on these estimates—0.071 and 0.003 for labor and materials, respectively—are much higher than when using the Hicks-neutral estimator, but still relatively precise. The full distribution of the output elasticity and markdown estimates are visualized as the red lines in Panel A of Figure 1.

Assuming Exogenous Input Prices

Finally, we reestimate the production function under both DGPs using the method of Rubens et al. (2024), but assuming exogenous input prices. This corresponds to using the estimator of Doraszelski and Jaumandreu (2018) for a random-coefficients Cobb-Douglas model. We find that imposing exogenous input prices when the true DGP is oligopsonistic and Hicks-neutral results in a serious bias in the materials coefficient, which is estimated at 0.492, whereas the true β^m is 0.3, as can be seen in the middle columns of Panel A in Table 1. The estimates are very similar when the DGP is factor-biased, as shown in Panel B of Table 1.

4 Further Extensions

4.1 Nonsubstitutable Inputs

So far, we have assumed a gross-output production function, through Assumption 1. However, a lot of industries feature production processes in which it is hard to substitute between materials and the other factors of production. Rubens (2023) relaxes Assumption 1 by allowing materials to be a perfect complement to a composite term of labor and capital, in order to study monopsony power of cigarette manufacturing firms in China. The production function is given by Equation (18), which reflects that when producing cigarettes, tobacco leaves M cannot be substituted with either capital or labor:

$$Q_{ft} = \min\{L_{ft}^{\beta^l} K_{ft}^{\beta^k}; \beta^m m_{ft}\}\Omega_{ft} \exp(\varepsilon_{ft}) \tag{18}$$

Thus, the markup now takes on a different form, which reflects that marginal costs are additive in labor and materials and which incorporates inverse factor supply elasticities[13]:

$$\mu_{ft} = (\frac{\alpha_{ft}^l}{\beta^l}\psi_{ft}^l + \alpha_{ft}^m \psi_{ft}^m)^{-1} \tag{19}$$

In contrast to the models in Sections 2 and 3.2, the first-order conditions for labor and ma-

[13]This formula diverges from De Loecker and Scott (2022) by allowing for endogenous input prices.

Table 1: Monte Carlo Simulations: Summary

(a) DGP 1: Hicks-neutral		Hicks-neutral estimator		RWX(2024) with exo. wage		RWX(2024) with endo. wage	
		Est.	S.E.	Est.	S.E.	Est.	S.E.
mean(β^l)	true = 0.5	0.500	0.003	0.508	0.000	0.516	0.071
sd(β^l)	true = 0	0.000	.	0.006	<0.001	0.002	0.002
β^m	true = 0.3	0.300	0.000	0.492	<0.001	0.299	0.003
ψ^l	true = .614	0.614	0.009	0.000	.	0.670	0.248

(b) DGP 2: Random coefficients		Hicks-neutral estimator		RWX(2024) with exo. wage		RWX(2024) with endo. wage	
		Est.	S.E.	Est.	S.E.	Est.	S.E.
mean(β^l)	true = 0.5	0.792	0.048	0.503	0.001	0.509	0.060
sd(β^l)	true = 0.096	0.000	.	0.050	0.001	0.120	0.303
β^m	true = 0.3	0.229	0.004	0.497	0.001	0.298	0.028
ψ^l	true = .613	2.494	0.256	0.000	.	0.669	0.249

Notes: This table reports the results of the Monte Carlo simulations, which were carried out with 200 iterations. Panel A reports the estimates when the true DGP is Hicks-neutral. The first two columns report the Hicks-neutral estimator. The final four columns report the estimator of Rubens et al. (2024), both when assuming exogenous wages (columns 3–4) and when allowing for endogenous wages (columns 5–6). Panel B does the same but covers the case in which the true DGP is not Hicks-neutral, but instead features unobserved random coefficients in production.

terials are no longer linearly independent; rather, there is a single first-order condition that takes into account both input prices and input supply elasticities. The reason for this reduction in the number of first-order conditions is that firms do not choose labor and materials separately, as one input choice determines the other input quantity as well. This is a problem for identification of either input-price markdown, as one can no longer divide the two first-order conditions by each other to express the markdown in function of output elasticities and revenue shares.

Of course, it is always possible that there is a third variable input, such as energy. If this third input is substitutable with the input over which monopsony power is exerted (so far, labor), then the markdown on that substitutable input can still be identified by solving for the energy first-order condition and the markup expression (19).

However, this does not apply if firms exert monopsony power on the nonsubstitutable input. Even if one could add variable inputs that substitute with labor, resulting in additional first-order conditions, this would not allow one to write the inverse material supply elasticity ψ_{ft}^{m} as a function of output elasticities and data, because materials are perfect complements to any of these other variable inputs. In this case, one needs to either estimate or impose a markup, or estimate the factor supply elasticity, as discussed in Rubens (2023). This identification strategy has been implemented in various industries, including Chinese tobacco manufacturing (Rubens, 2023), German car manufacturing (Hahn, 2024), French dairy production (Avignon & Guigue, 2022), and Chinese coal mining (Zheng, 2024).

4.2 Labor Market Conduct

The labor market simulations in Section 3 impose that firms compete à la Nash-Bertrand, which implies oligopsonistic competition. This approach nests models of monopsonistic competition if labor market shares approximate zero. In this subsection, we consider models of labor market competition other than oligopsonistic or monopsonistic competition.

Collusion

Firms might collude on their input markets: coordinating their wage or employment choices rather than making these decisions independently. Delabastita and Rubens (2025) consider markdown estimation when firms potentially collude. Maintaining the assumption of Hicks neutrality, they show that the wage markdown can still be estimated using the production approach, even if firms collude on their labor markets. Next, they combine estimation of a labor supply model with the production estimates to identify conduct on the labor market;

they find that their collusion estimates align with the observed introduction of a cartel in the Belgian coal mining industry.

Bargaining

In many labor market settings, firms and workers bargain over wages, rather than firms merely posting wages (Caldwell, Haegele, & Heining, 2025). This bargaining can either be individual or collective, through a labor union. Rubens (2024) considers production-approach markdown estimation when wages are bargained over. The author's empirical application focuses on Illinois coal operators, which bargain over wages with miner unions. A methodological challenge arises because to identify bargaining parameters, an estimate of the production function is needed, but bargaining parameters need to be known in order to identify the production function. Rubens (2024) addresses this problem using a fixed point estimator, in which production-function estimation is nested in a loop over which bargaining parameters are guessed. In that particular application, the estimation procedure converges quickly toward a stable set of estimates of both bargaining abilities and production-function coefficients.

4.3 Differentiation and Multiproduct Firms

Product and Input Differentiation

Assumption 2 imposed that both goods and inputs are homogeneous, which is clearly a strong assumption in many settings. Although vertical product differentiation can be allowed for using a price control in the production function (De Loecker et al., 2016; Rubens, 2023), most goods and factors are horizontally differentiated as well. Hahn (2024) addresses the challenge of differentiated goods by estimating a hedonic price model for car manufacturers, which incorporates car characteristics in addition to a production function. This model is then used to estimate markdowns and examine bargaining between car manufacturers and parts producers. A distinct challenge arises when the inputs, rather than the products, are differentiated. Lamadon, Mogstad, and Setzler (2022) address this challenge by allowing for heterogeneous worker quality using matched employer-employee data.

Multiproduct Firms

Production-function estimation with multiproduct firms is challenging even if input markets are perfectly competitive, because data on inputs are usually not available at the product level. Various approaches have been developed to address this challenge (De Loecker et al., 2016; Dhyne, Petrin, Smeets, & Warzynski, 2022; Orr, 2022; Valmari, 2023) without allowing for imperfect factor market competition. Avignon and Guigue (2022) estimate factor price markdowns for the French dairy industry while allowing for multiproduct firms. They combine engineering data to assign input costs to the various products with production-function estimation to recover markups and markdowns.

4.4 Adjustment Frictions

Assumption 4 imposes that both materials and labor are variable and static inputs. In many applications, it is reasonable that at least a subset of these inputs will be subject to adjustment frictions, such as hiring or firing costs. Although estimation of the production function is not hampered by such adjustment frictions (only the imposed timing assumptions would change), these frictions do pose a challenge for markdown identification using the production-function approach, because the markup and markdown expressions 10 and 11 are obtained by solving a static cost-minimization problem. Adjustment frictions lead to additional wedges between marginal revenue products and input prices that are unrelated to the exercise of monopsony power. One possibility to separately identify adjustment costs from monopsony distortions is to, again, jointly estimate a labor supply model and a production model and to recover frictions using matched employer-employee data. Chan, Mattana, Salgado, and Xu (2023) implement such a model using Danish data.

4.5 No Competitive Input Market

Finally, Assumption 3 imposes that intermediate input prices are exogenous to firms. This assumption is common in the literature (Morlacco, 2017; Brooks et al., 2021; Yeh et al., 2022; Delabastita & Rubens, 2025); it is needed to point-identify the markdown when using only the production function, as made clear by Equation (11). If all input markets are imperfectly competitive, meaning that no input price is exogenous, there are two potential solutions. First, one could impose a model of imperfect competition and estimate a factor supply curve for one of the inputs, as carried out in Section 3.2, and still identify the markdown of the remaining inputs using the production approach. Alternatively, Treuren (2022) proposes estimating a *revenue* production function, in contrast to the *quantity* production functions used in this paper so far, to identify wage markdowns without having to assume competitive material markets.

Whereas the benefit of allowing for endogenous material prices is clear, using a revenue production function comes at the cost of imposing homogeneous goods demand elasticities between firms, which restricts the set of models of imperfect competition on the product market one can allow for. As with any assumption, the tradeoff between imposing additional restrictions on product market competition while relaxing the assumptions in terms of input market competition is specific to the empirical application, and depends on the type of industry at hand.

5 Conclusions

In this paper, we review production approaches to estimate factor price markdowns. We discuss the commonly made assumptions in this class of estimators, and we test this class of estimators using Monte Carlo simulations for oligopsonistic labor markets in which firms compete in wages in a static Nash-Bertrand equilibrium. We find that when production is Hicks-neutral, existing production-approach markdown estimators recover markdowns consistently. This implies that it is possible to estimate wage markdowns without having to specify and estimate a labor supply model, and while remaining agnostic about the underlying model of labor market conduct.

However, we find that allowing for unobserved technological heterogeneity in production leads to severely biased estimates of factor price markdowns using the production approaches that rely on Hicks neutrality. When implementing the estimation procedure suggested by Rubens et al. (2024), which is designed to allow for departures from Hicks neutrality, we find that the production-function coefficients and heterogeneity can be estimated consistently in the presence of imperfect labor market competition.

Finally, we discuss approaches in the literature that have extended production approaches to markdown estimation to relax other assumptions, such as allowing for nonsubstitutable inputs, different types of labor market conduct, and multiproduct production.

References

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

Avignon, R., & Guigue, E. (2022). *Markups and markdowns in the french dairy market* (Tech. Rep.). Technical report, mimeo.

Azar, J., Berry, S., & Marinescu, I. (2022). *Estimating Labor Market Power* (Working Paper No. 30356). National Bureau of Economic Research.

Azzam, A., Jaumandreu, J., & Lopez, R. (2025). Input and output market power with non-neutral productivity: Livestock and labor in us meatpacking.

Berger, D., Herkenhoff, K., & Mongey, S. (2022). Labor Market Power. *American Economic Review*, *112*(4), 1147–93.

Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics*, 242–262.

Blundell, R., & Bond, S. (2000). GMM Estimation with Persistent Panel Data: an Application to Production Functions. *Econometric reviews*, *19*(3), 321–340.

Brooks, W. J., Kaboski, J. P., Li, Y. A., & Qian, W. (2021). Exploitation of Labor? Classical Monopsony Power and Labor's Share. *Journal of Development Economics*, *150*, 102627.

Caldwell, S., Haegele, I., & Heining, J. (2025). *Bargaining and inequality in the labor market* (Tech. Rep.). National Bureau of Economic Research.

Card, D., Cardoso, A. R., Heinig, J., & Kline, P. (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, *36*(1), 13-70.

Chan, M., Mattana, E., Salgado, S., & Xu, M. (2023). Wage Setting and Passthrough: The Role of Market Power, Production Technology and Adjustment Costs. *Production Technology and Adjustment Costs (August 17, 2023)*.

Delabastita, V., & Rubens, M. (2025).
Journal of Political Economy, *6*(133), 1796–1839.

De Loecker, J., Goldberg, P., Khandelwal, A., & Pavcnik, N. (2016). Prices, Markups and Trade Reform. *Econometrica*, *84*(2), 445-510.

De Loecker, J., & Scott, P. T. (2022). *Markup Estimation using Production and Demand Data: An Application to the US Brewing Industry* (Mimeo).

De Loecker, J., & Warzynski, F. (2012). Markups and Firm-Level Export Status. *The American Economic Review*, *102*(6), 2437-2471.

Demirer, M. (2019). *Production Function Estimation with Factor-Augmenting Technology:*

An Application to Markups.

Dhyne, E., Petrin, A., Smeets, V., & Warzynski, F. (2022). *Theory for extending single-product production function estimation to multi-product settings* (Tech. Rep.). National Bureau of Economic Research.

Dobbelaere, S., & Mairesse, J. (2013). Panel data estimates of the production function and product and labor market imperfections. *Journal of Applied Econometrics*, *28*(1), 1–46.

Doraszelski, U., & Jaumandreu, J. (2018). Measuring the Bias of Technological Change. *Journal of Political Economy*, *126*(3), 1027-1084.

Hahn, N. (2024). Who is in the driver's seat? markups, markdowns, and profit sharing in the car industry. *Markups, Markdowns, and Profit Sharing in the Car Industry (July 08, 2024). ZEW-Centre for European Economic Research Discussion Paper*(24-047).

Hall, R. E. (1988). The relation between price and marginal cost in US industry. *Journal of Political Economy*, *96*(5), 921–947.

Lamadon, T., Mogstad, M., & Setzler, B. (2022). Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review*, *112*(1), 169-212.

Mertens, M. (2019). *Micro-Mechanisms Behind Declining Labour Shares: Market Power, Production Processes, and Global Competition* (Tech. Rep.).

Mertens, M., & Schoefer, B. (2024). *From labor to intermediates: Firm growth, input substitution, and monopsony* (Tech. Rep.). National Bureau of Economic Research.

Morlacco, M. (2017). *Market Power in Input Markets: Theory and Evidence From French Manufacturing* (Tech. Rep.).

Orr, S. (2022). Within-firm productivity dispersion: Estimates and implications. *Journal of Political Economy*, *130*(11), 2771–2828.

Rubens, M. (2023). Market Structure, Oligopsony Power, and Productivity. *American Economic Review*, *113*(9), 2382-2410.

Rubens, M. (2024). *Labor market power and factor-biased technology adoption.*

Rubens, M., Wu, Y., & Xu, M. (2024). *Augmenting or exploiting labor?*

Syverson, C. (2024). *Markups and Markdowns* (Tech. Rep.). National Bureau of Economic Research.

Treuren, L. (2022). *Wage markups and buyer power in intermediate input markets.* KU Leuven, Department of Economics.

Valmari, N. (2023). Estimating production functions of multiproduct firms. *Review of*

Economic Studies, *90*(6), 3315–3342.

Yeh, C., Hershbein, B., & Macaluso, C. (2022). Monopsony in the U.S. Labor Market. *American Economic Review*, *112*(7).

Zheng, H. (2024). Death by market power and the production-safety tradeoff in the coal mining industry.

A Online Appendix

A.1 Derivation of First-Order Conditions

Denote expected output as $Q^* = H(.)\Omega$. We omit firm and time subscripts. First, we take the first derivative of the cost-minimization problem (7) w.r.t. materials, which gives

$$W^m = \lambda \frac{\partial Q^*}{\partial M}$$

Substituting $\lambda = \frac{P}{\mu+1}$ into this equation and working out the first derivative results in the FOC from the main text:

$$W^m = \frac{P}{\mu + 1} \beta^m L^{\beta^l} M^{\beta^m - 1} \Omega$$

For labor, we need to take the first derivative w.r.t. the wage on both sides of Equation (7):

$$\frac{\partial (W^l L)}{\partial W^l} = \lambda \frac{\partial Q^*}{\partial L} \frac{\partial L}{\partial W^l}$$

Working out the derivative on the left-hand side and dividing both sides by $\frac{\partial L}{\partial W^l}$ results in

$$\frac{L}{\frac{\partial L}{\partial W^l}} + W^l = \lambda \frac{\partial Q^*}{\partial L}$$

Factoring out W^l and rearranging terms, we obtain

$$W^l \left(\frac{1}{\frac{\partial L}{\partial W^l} \frac{W^l}{L}} + 1 \right) = \lambda \frac{\partial Q^*}{\partial L}$$

Finally, working out the derivative on the right-hand side and using the ψ^l notation results in the FOC from the main text:

$$W^l (1 + \psi^l) = \frac{P}{\mu + 1} \beta^l L^{\beta^l - 1} M^{\beta^m} \Omega$$

A.2 Derivation of Labor Supply Elasticity

Using the logit formula, the labor market share of firm f is

$$s_f = \frac{\exp(\delta_f)}{\sum_k \exp(\delta_k)}$$

with $\delta_f \equiv \gamma \ln(w_f) + \xi_f$

Taking the first derivative w.r.t., the wage gives

$$\begin{aligned}\frac{\partial s_f}{\partial w_f} &= \frac{\sum_k \exp(\delta_k)\exp(\delta_f)\frac{\gamma}{w_f} - (\exp(\delta_f))^2\frac{\gamma}{w_f}}{(\sum_k \exp(\delta_k))^2} \\ &= s_f(1-s_f)\frac{\gamma}{w_f}\end{aligned}$$

Multiplying by w_f and dividing by s_f obtains the labor supply elasticity that was used in the main text:

$$\frac{\partial s_f}{\partial w_f}\frac{w_f}{s_f} = \gamma(1-s_f)$$