Estimating Factor Price Markdowns using Production Models

Michael Rubens* Yingjie Wu Mingzhi (Jimmy) Xu

February 11, 2025

Abstract

Factor price markdowns are a key object of interest when studying monopsony power. This paper evaluates the performance of production-function-based estimators of markdowns, which have been used increasingly often in the literature. Using Monte Carlo simulations, we assess these estimators across different data-generating processes. We also examine the key assumptions underlying this approach and explore the methodological challenges of relaxing them, including deviations from Hicks neutrality, the presence of non-substitutable inputs, and alternative labor market structures.

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 It combines new results with a discussion of existing published and unpublished work, as cited throughout the paper. Mingzhi Xu thanks financial support from the Natural Science Foundation of China (Grant No. 72322007).

1 Introduction

There is an increased interest in the monopsony power of firms on their labor and other factor markets. Production functions are increasingly used to estimate 'factor price markdowns', which are a key object of interest when studying monopsony power.[1] However, little is known about the consistency of these estimators. To fill this gap, we conduct Monte Carlo simulations in which we generate oligopsonistic labor market equilibria. Using these simulated data, we apply production function-based estimators to recover factor price markdowns. Comparing these estimates to the true markdown distribution enables us to assess how well these 'cost-based' estimators succeed at recovering true markdowns under a variety of data-generating processes.

We start the paper by discussing existing production approaches to markdown estimation, which extend the markup estimation approach of Hall (1988) and De Loecker and Warzynski (2012) to allow for endogenous factor prices (Morlacco, 2017; Brooks, Kaboski, Li, & Qian, 2021; Yeh, Hershbein, & Macaluso, 2022; Mertens, 2022; Rubens, 2023; Delabastita & Rubens, in press; Mertens & Schoefer, 2024). These methods share a common identification strategy: they assume that at least one variable input has an exogenous price and normalize the markup expressions derived from the first-order conditions of cost minimization for all other inputs relative to this exogenous-price input.[2] This enables the estimation of markdowns for all other variable inputs.

We evaluate this class of markdown estimators using simulated data generated from 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.[3] On the labor demand side, we assume firms minimize costs under a Cobb-Douglas production function with two variable inputs: labor and materials. Given the oligopsonistic structure of the labor market, 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 during 10 years, and solve for labor market equilibrium in each of these markets. This process is repeated for 200 random draws. We find

[1] For an excellent survey of the literature on markups and markdowns, see Syverson (2024).

[2] This identification approach was proposed in Appendix D of De Loecker, Goldberg, Khandelwal, and Pavcnik (2016).

[3] Unlike Card et al. (2018), which assumes monopsonistic competition, we allow for granular employers, as in Berger, Herkenhoff, and Mongey (2022) and Azar, Berry, and Marinescu (2022).

that under the Hicks-neutral data-generating process (DGP), the production-based markdown estimator delivers consistent and precise 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 detail by Rubens, Wu, and Xu (2024). To test the robustness of this assumption, we relax it by introducing random coefficients in the production function, allowing for unobserved technological heterogeneity across firms—a realistic concern in many applications. We then re-run our Monte Carlo simulation, keeping the labor supply side unchanged but replacing the Hicks-neutral production function on the labor demand side. The results show a stark contrast: unlike in the Hicks-neutral case, both the estimated production function coefficients and wage markdowns exhibit significant bias. The root of this bias lies in the inability to separately identify latent technological heterogeneity and wage markdowns using the first-order conditions of cost minimization.

We test an alternative production function estimator, proposed by Rubens et al. (2024), which extends the approach of Doraszelski and Jaumandreu (2018) to account for imperfect competition in factor markets. 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, and find that the estimator of Rubens et al. (2024) remains unbiased but is less efficient than the standard Hicks-neutral markdown estimator. Despite this efficiency loss, it still estimates the true production function coefficients with minimal bias.

We end the paper by discussing how some of the other commonly made assumptions have been relaxed in the literature. Key extensions include allowing for non-substitutable inputs in production, alternative labor market conduct beyond Nash-Bertrand competition, differentiated goods and inputs, adjustment frictions, and scenarios where no factor prices can reasonably be assumed to be competitive. These considerations highlight the broader challenges in markdown estimation and reveal opportunities for refining existing methods in more complex economic settings.

The remainder of this paper is structured as follows. In Section 2, we discuss cost-side markdown estimators when the production function is Hicks-neutral. In Section 3, we introduce technological heterogeneity in the production function, which relaxes the assumption of Hicks-neutrality. Section 4 discusses further extensions, and Section 5 concludes.

2 Markdown Estimation under Hicks Neutrality

2.1 Primitives

In this section, we test the canonical ‘production approach’ to markdown estimation, which relies on Hicks-neutrality. The analysis begins by outlining the model’s core primitives: the production function and the factor supply model.

Production Function

Let firms be indexed by f and time periods by t. We assume 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 $H(.)$. The residual Ω_{ft} captures Hicks-neutral productivity variation, whereas the production function coefficients β are assumed to be common across firms. This implies that the productivity residual is a scalar.

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

We start by highlighting three assumptions. As a convention throughout the paper, we list the assumptions that are used to estimate the production function and markdowns, while those used solely for data simulation—without directly impacting estimation—are noted but not formally stated.

assumption Production includes an unobserved scalar term, Ω_{ft}, which captures firm-specific productivity shocks.

Assumption 1 imposes Hicks neutrality, as it rules out unobserved heterogeneity in the production coefficients β. In Section 3, we relax this assumption and analyze its implications.

***Assumption* 1** *The production function $H(.)$ is assumed to be twice continuously differentiable.*

Assumption 1 rules out perfect complementarities between inputs, as seen in a Leontief production function. We relax this assumption in Section 4.1.

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

Assumption 2 treats both the output and inputs as undifferentiated. In Section **??**, we explore how this assumption can be relaxed.

For the simulations, we impose a simple Cobb-Douglas production function, which in logs yields Equation (2). However, any production function satisfying Assumptions 1 and 1 could be used. The output elasticities of labor and materials are denoted by β^l and β^m, respectively.

$$q_{ft} = \beta^l l_{ft} + \beta^m m_{ft} + \omega_{ft} \tag{2}$$

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

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

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

Finally, we assume that both labor and materials are variable, static inputs, meaning they adjust freely each period without frictions and fully depreciate by the end of each period.[4]

Labor Supply

We model labor supply using a discrete choice framework with oligopsonistic competition, following Berry (1994) and Card et al. (2018), to simulate an environment where markdowns vary across firms and wages are set strategically. Firms pay per-unit wages W^l_{ft} 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. Firms are assumed to pay uniform wages and cannot discriminate among homogeneous workers. We assume that a worker's utility from working at firm f depends on the offered wage W_{ft}, an unobserved amenity ξ_{ft}, and an idiosyncratic type-I extreme value error term υ_{ift}, as specified in Equation (3).

$$U_{ift} = \underbrace{\gamma W_{ft} + \xi_{ft}}_{\equiv \delta_{ft}} + \upsilon_{ift} \tag{3}$$

We define mean utility as δ_{ft} and, following standard practice, normalize the utility of

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

the outside option to zero: $U_{i0t} = 0$. Applying the logit model, the labor market share, $s_{ft} = \frac{L_{ft}}{\sum_{g \in \mathcal{F}_t} L_{gt}}$, is given by:

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

Let $\overline{L}$ denote the total labor force. The labor supply function, $H(.)$, is then given by:

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

We define the inverse residual supply elasticities of labor and materials as $\psi^l_{ft} - 1$ and $\psi^m_{ft} - 1$, respectively, so that:

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

Under the logit labor supply framework, the inverse residual labor supply elasticity faced by firm f, $(\psi^l_{ft} - 1)$, is given by:

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

2.2 Behavioral Assumptions

Firms choose inputs each period to minimize current variable costs. Let λ_{ft} denote marginal cost, so that the cost minimization problem is given by Equation (7):

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

As demonstrated in De Loecker et al. (2016), the markup of the final goods price P_{ft} over marginal cost, defined as $\mu^p_{ft} \equiv (P_{ft} - \lambda_{ft})/\lambda_{ft}$, is equal to Equation (8):

$$\mu^p_{ft} = \frac{\beta^j_{ft}}{\alpha^j_{ft} \psi^j_{ft}} - 1 \quad \forall j \in \{l, m\} \tag{8}$$

where α^j_{ft} denotes the expenditure on input j as a share of gross revenues of firm f in year t, such that $\alpha^l_{ft} \equiv W^l_{ft}L_{ft}/P_{ft}Q_{ft}$ and $\alpha^m_{ft} \equiv W^m_{ft}L_{ft}/P_{ft}Q_{ft}$. Following Morlacco (2017), Brooks et al. (2021), and Yeh et al. (2022), the inverse supply elasticity of labor can be expressed relative to that of materials by scaling the ratio of input expenditures with the corresponding output elasticities of both inputs:

$$\psi^l_{ft} = \frac{\beta^l}{\beta^m}\frac{\alpha^m_{ft}}{\alpha^l_{ft}}\psi^m_{ft} \tag{9}$$

The wage markdown, which measures the gap between wages and the marginal revenue product of labor ($MRPL_{ft}$), is defined as $\mu^w_{ft} \equiv (MRPL_{ft} - W_{ft})/MRPL_{ft}$. This markdown can be expressed as a function of the inverse labor supply elasticity:

$$\mu^w_{ft} = \frac{\psi^l_{ft} - 1}{\psi^l_{ft}} \tag{10}$$

The less elastic the labor supply curve, the greater a firm's ability to exert monopsony power and depress wages.

***Assumption* 4** *The residual supply of intermediate inputs is perfectly elastic:* $\psi^m_{ft} = 1$.

Assumption 4 implies that intermediate input prices are exogenous to individual firms. As shown in Equation (8), this assumption enables the point identification of the wage markdown, rather than merely its value relative to the material price markdown.

Solving the cost minimization problem in Equation (7) yields the labor demand function for the Cobb-Douglas case, where factor prices are denoted as W^m_{ft} and W^l_{ft}:

$$L_{ft} = \Big[\frac{\beta^l}{W^l_{ft}\psi^l_{ft}}\big(\frac{\beta^m\Omega_{ft}}{W^m_{ft}}\big)^{\frac{\beta^m}{1-\beta^m}}\Omega_{ft}\Big]^{\frac{1-\beta^m}{1-\beta^l-\beta^m}} \tag{11}$$

The corresponding optimal demand for intermediate inputs is:

$$M_{ft} = \Big(\frac{\beta^m L^{\beta^l}_{ft}\Omega_{ft}}{W^m_{ft}}\Big)^{\frac{1}{1-\beta^m}}$$

These expressions capture how input choices respond to factor prices and the firm's productivity.

2.3 Identification and Estimation

Given Assumption 3 (input variability) and the AR(1) process for productivity in Equation (2), the production function can be estimated using a dynamic panel approach. Taking ρ-differences, as in Blundell and Bond (2000), the productivity shock can be written as:

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

Similarly to Ackerberg, Caves, and Frazer (2015), assuming that labor and materials are both variable inputs, we construct the following moment conditions for lags $r = 1$ to $r = T - 1$, where T represents the panel length. As in Ackerberg et al. (2015), identification relies on the assumption that variable inputs—materials and labor in our case—are chosen after the firm observes the productivity shock e_{ft}.

$$\mathbb{E}\Big[e_{ft}(\rho, \beta^l, \beta^m) | \begin{pmatrix} L_{ft-r} \\ M_{ft-r} \end{pmatrix} \Big]_{r=1}^{T-1} = 0 \tag{12}$$

We estimate the production function coefficients (β^l, β^m) using these moment conditions with two time lags. The resulting estimates, $(\hat{\beta}^l, \hat{\beta}^m)$, are then used to compute the wage markdown ψ^l_{ft} via Equation (9). This approach allows us to estimate the inverse residual labor supply elasticity directly from the production function, without the need to estimate the labor supply parameters γ and ξ_{ft}.[5]

A dynamic panel approach to identify the production function, and the associated assumption of the AR(1) productivity transition, is not strictly needed. One could instead identify the production function using more widely used productivity inversion techniques, provided that imperfect labor market competition is taken into account, as discussed in Ackerberg and De Loecker (2021).

2.4 Monte Carlo Simulation

Parametrization

We simulate a dataset of 50 independent labor markets that each contain 5 firms, which are observed during 10 years. Hence, the simulated dataset contains 250 firms that are observed during 10 times each ($N = 2500$). We parametrize the true output elasticities of labor and

[5]While we specify the logit labor supply model to generate the simulated data, the labor supply curve itself does not need to be estimated when applying Equation (9).

materials 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 $\omega_{f1} \sim \mathcal{N}(1, 0.01)$ and let the productivity shocks be $\mathcal{N}(0, 0.01)$ distributed. The serial correlation in productivity is set at $\rho = 0.6$, yielding a steady-state log productivity distribution with a mean of $1/4$ and a standard deviation of $1/3$. The total labor market size is normalized to one.

Solving for Equilibrium

We perform a Monte Carlo simulation with 200 independent draws. In each iteration, we numerically solve for equilibrium wages and market shares by ensuring that, at every firm in every year, labor demand (11) equals labor supply (4) and that labor markets clear in the aggregate.

Using the simulated dataset, we estimate the production function parameters ρ, β^l, and β^m based on the moment conditions in Equation (12). We then use these estimates in Equation (9) to compute the inverse residual labor supply elasticities ψ^l_{ft} for all firms across all years.

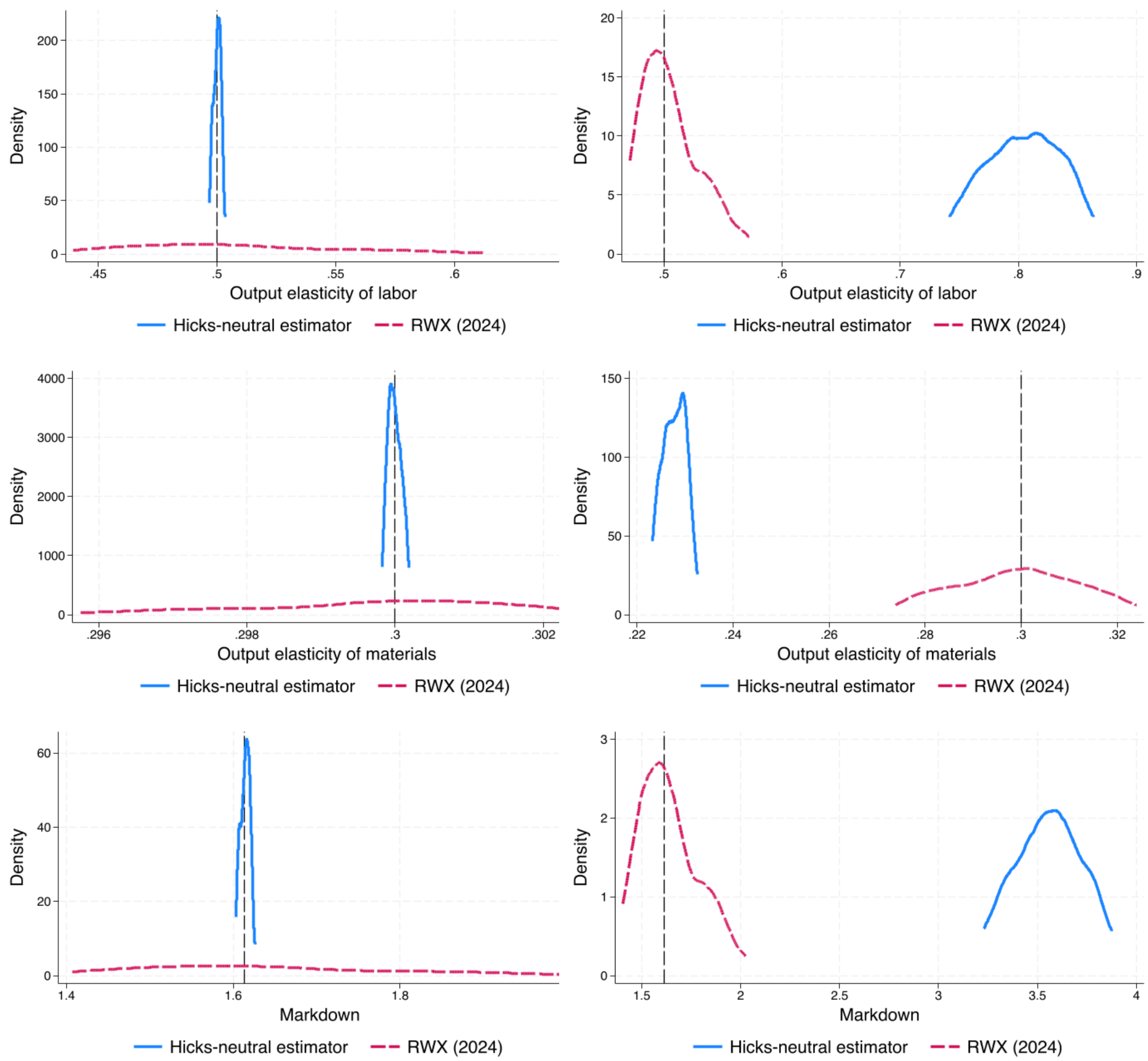
Results

The distribution of the estimated parameters is shown by the solid blue lines in Panel (a) of Figure 1. The results indicate that the Hicks-neutral production function estimator yields consistent and precise estimates of the output elasticities of labor and materials. As reported in Panel (a) of Table 1, the estimated output elasticities closely match their true values of 0.5 and 0.3, with remarkably low standard deviations across bootstrap iterations—0.003 for labor and below 0.001 for materials. This confirms that the production function remains identifiable even in the presence of imperfect labor market competition. Furthermore, the estimator provides a reliable estimate of the wage markdown ψ^l_{ft}, which has a true value of 1.614 and is estimated at 1.615, with a standard deviation of just 0.009 across draws.

It is worth noting that the assumptions imposed on the labor supply model and on conduct were only necessary to simulate the dataset, but were not used for production function estimation: 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:**



Notes: Panel (a) shows the distribution of the production function estimates when assuming a Hicks-neutral DGP. The solid blue lines report the estimates under the Hicks-neutral estimator, the dashed red lines report the estimates using Rubens et al. (2024). Panel (b) visualizes these estimators for the Factor-Biased DGP, in which there is latent heterogeneity about the output elasticities.

3 Introducing Unobserved Technological Heterogeneity

3.1 Extended Model

We now revisit the identification strategy from Section 2, relaxing Assumption 1, which imposed Hicks neutrality. This assumption underpins most production-function-based markdown estimators, including those in Morlacco (2017), Brooks et al. (2021), Yeh et al. (2022), Mertens (2022), Rubens (2023), Delabastita and Rubens (in press), and Mertens and Schoefer (2024). By lifting this restriction, we examine the implications for markdown estimation and assess the robustness of existing approaches.

Instead of the Cobb-Douglas production function with constant output elasticities from Equation (2), we introduce unobserved random coefficients β^l_{ft} and β^m_{ft}, as specified in Equation (13). These random coefficients can arise due to various reasons, such as latent heterogeneity in production technologies or differences in capital intensity.

$$q_{ft} = \beta^l_{ft} l_{ft} + \beta^m_{ft} m_{ft} + \omega_{ft} \tag{13}$$

s We let the output elasticities of labor and materials be distributed around the same values β^l and β^m as before, with idiosyncratic error terms ϵ^l_{ft} and ϵ^m_{ft}:

$$\begin{cases} \beta^l_{ft} &= \beta^l + \epsilon^l_{ft} \\ \beta^m_{ft} &= \beta^m + \epsilon^m_{ft} \end{cases}$$

Equation (13) provides a straightforward way to introduce unobserved heterogeneity while preserving the analytical tractability of the Cobb-Douglas production function. For a more flexible approach, Rubens et al. (2024) estimate a Constant Elasticity of Substitution (CES) production function under imperfect labor market competition and apply it empirically to the Chinese nonferrous metals industry.

3.2 Identification Challenge

We reformulate the markdown estimator from Equation (9) to account for heterogeneous output elasticities. As shown in Equation (14), identifying the wage markdown requires accurately estimating the random coefficients β^m_{ft} and β^l_{ft}.

$$\psi^l_{ft} = \frac{\beta^l_{ft}}{\beta^m_{ft}} \frac{\alpha^m_{ft}}{\alpha^l_{ft}} \psi^m_{ft} \tag{14}$$

While existing studies, such as Doraszelski and Jaumandreu (2018) and Demirer (2019), have developed methods for estimating production functions with non-scalar unobservables, these approaches assume perfect factor market competition, imposing $\psi^l_{ft} = \psi^m_{ft} = 1$. In contrast, Rubens et al. (2024) introduce an estimator that accommodates both non-scalar unobservables and imperfect factor market competition. Below, we outline this estimation procedure within the framework of our simplified production model.

3.3 Estimation

Rubens et al. (2024) propose a joint estimation of the labor supply curve and the production function. Building on the discrete choice labor supply model introduced earlier, the labor supply equation to be estimated is:

$$s_{ft} - s^0_{it} = \gamma \ln(W_{ft}) + \xi_{ft} \tag{15}$$

Under the Nash-Bertrand conduct assumption, the markdown ψ^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}:

$$\hat{\psi}^l_{ft} = 1 + \frac{1}{\hat{\gamma}(1 - s_{ft})} \tag{16}$$

From the first-order conditions, the output elasticity of labor can be expressed 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.

$$\hat{\beta}^l_{ft} = \frac{\hat{\psi}^l_{ft} \alpha^l_{ft} \beta^m}{\alpha^m_{ft}} \tag{17}$$

Substituting this expression for the output elasticity of labor into the production function yields Equation (18), 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}.

$$q_{ft} = \beta^m \underbrace{[\frac{\hat{\psi}^l_{ft}\alpha^l_{ft}l_{ft}}{\alpha^m_{ft}} + m_{ft}]}_{a_{ft}} + \omega_{ft} \tag{18}$$

Applying 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 for estimating the parameters (β^m, ρ) are given by:

$$\mathbb{E}\Big[e_{ft}(\rho, \beta^m)| \begin{pmatrix} L_{ft-r} \\ M_{ft-r} \end{pmatrix} \Big]_{r=1}^{T-1} = 0 \tag{19}$$

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} s can be recovered using Equation (17), which is now a function of data and estimated parameters.

3.4 Monte Carlo Simulations

Parametrization and Estimation

To demonstrate the potential bias in the markdown estimates when the Hicks-neutral assumption is imposed, while the DGP includes heterogeneity in output elasticity, we estimate the production function twice. First, we "naively" estimate the production function assuming the DGP is Hicks-neutral, using the moment conditions in Equation (12), and estimate the markdown using the cost-side markdown estimator from Equation (9).

Second, we estimate the production function using the estimation procedure from Rubens et al. (2024) that was outlined above. We start by estimating Equation (15). 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 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 the labor demand shifter in hand, we estimate the labor supply curve (15) using 2SLS. Using the estimated parameter $\hat{\gamma}$ and the observed labor market share s_{ft}, we compute the wage markdowns based on Equation (16). We then substitute the markdown estimate ψ^l_{ft} into Equation (18) and form the moment conditions in Equation (19) to estimate the production function parameters β^m and ρ. Finally, we recover the full distribution of the output elasticities α^l_{ft} and α^m_{ft} using Equation (18).

Results under the Factor-Biased Data Generating Process

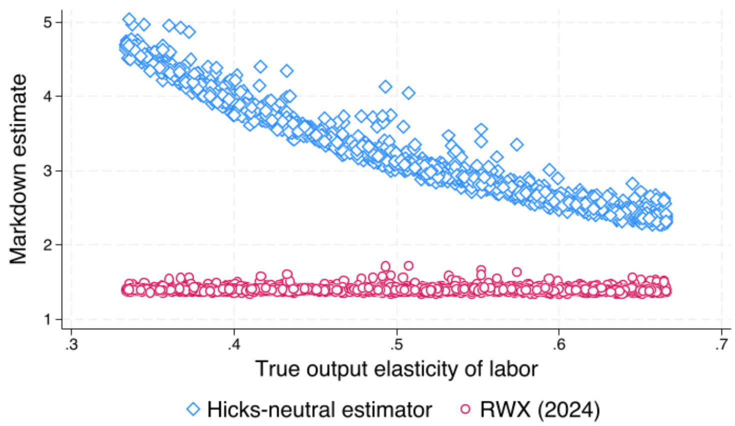
We visualize the production function estimates for the DGP with random coefficients in the production function in panel (b) of Figure 1. The solid blue lines in Figure 1 report the estimates using the Hicks-neutral production function estimator that assumes homogeneous output elasticities. It is evident that the Hicks-neutral estimator performs poorly in estimating the production function coefficients: the labor coefficient is estimated at 0.8, 60% higher than the true value, while the materials coefficient is estimated at 0.23, 25% below the true value. As a result, the Hicks-neutral model estimates the inverse labor supply elasticity at 3.559 on average—three times higher than the true average value of 1.613. This leads the econometrician to conclude that wages are marked down by 72% relative to the marginal revenue product of labor, whereas the actual markdown is only 38%.[6]

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

The red dashed lines in panel (b) of Figure 1 show the estimates using the method from Rubens et al. (2024) for the random coefficients DGP. The markdown is estimated with a small negative bias, likely due to the small-sample properties of the instrumental variables

[6]The wage markdown is calculated as $1 - 1/3.559$, which is approximately 72%.

Figure 2: Markdowns and the Output Elasticity of Labor



Notes: The blue diamonds report firm-level markdown estimates compared with these firms' output elasticities of labor in the first iteration of the Monte Carlo simulation, when using the Hicks-neutral estimator. The red circles show the corresponding markdown estimates when using the estimator of Rubens et al. (2024).

estimator for labor supply, but it remains close to the true value of 1.613. As for the production function coefficients, our estimator provides consistent estimates of the output elasticity. This demonstrates that the production function can be accurately estimated even with random coefficients and imperfect labor market competition, though it must be estimated jointly with the labor supply curve.

Results under the Hicks-Neutral Data Generating Process

How does the estimator from Rubens et al. (2024) perform when there is no unobserved heterogeneity in the output elasticities? Panel (a) of Table 1 shows that the output elasticity of labor is still estimated reasonably close to the true value, at 0.516, which reflects a small upward bias of 3.2%, while the materials elasticity is consistently estimated. The standard errors of these estimates—0.072 for labor and 0.003 for materials—are notably higher than those from the Hicks-neutral estimator, but remain relatively precise. The full distribution of output elasticity and markdown estimates is shown by the red lines in Panel (a) of Figure 1.

Assuming Exogenous Input Prices

Finally, we re-estimate the production function under both DGPs using the method of Rubens et al. (2024), but assume exogenous input prices. This effectively corresponds to the estimator of Doraszelski and Jaumandreu (2018). 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.

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	<0.001	0.515	0.072
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 = 1.614	1.615	0.009	0.000	.	1.669	0.252
$corr(\beta^l, \psi^l)$		.	.	.	.	0.012	0.996
(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	0.805	0.048	0.503	0.001	0.512	0.042
sd(β^l)	true = 0.096	0.000	.	0.050	<0.001	0.098	0.008
β^m	true = 0.3	0.228	0.004	0.497	0.001	0.299	0.019
ψ^l	true = 1.613	3.559	0.252	0.000	.	1.669	0.246
$corr(\beta^l, \psi^l)$		0.000	.	.	.	-0.111	0.033

Notes: This Table reports the results of the Monte-Carlo simulations, which are 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 features unobserved random coefficients in production instead.

4 Further Extensions

4.1 Non-substitutable Inputs

Rubens (2023) relaxes Assumption 1 by allowing labor and materials to be perfect complements, while still permitting labor to be substitutable with other inputs, such as capital (K). The revised production function, shown in Equation (20), is used in Rubens (2023) to study cigarette production in China. For inputs like tobacco leaves, and many other intermediate goods, assuming perfect complementarity with labor is more realistic than assuming substitutability.

$$Q_{ft} = \min\{\beta^l l_{ft} \beta^k k_{ft}; \beta^m m_{ft}\}\Omega_{ft} \tag{20}$$

As demonstrated in Rubens (2023), which incorporates imperfect factor market competition into the model of De Loecker and Scott (2022), the markup takes on a new form. This reflects the fact that marginal costs are additive in both labor and materials:

$$\mu_{ft} = (\frac{\alpha^l_{ft}}{\beta^l}\psi^l_{ft} + \alpha^m_{ft})^{-1} \tag{21}$$

Unlike the models in Sections 2 and 3, the first-order conditions for labor and materials are no longer linearly independent. Instead, there is a single first-order condition that incorporates both input prices and supply elasticities. This reduction in the number of first-order conditions arises because firms no longer choose labor and materials separately; the choice of one input determines the quantity of the other. This poses an identification challenge for the input price markdowns, as the two first-order conditions can no longer be divided to express the markdown in terms 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 one over which monopsony power is exerted (in this case, labor), the markdown on the substitutable input can still be identified by solving for the energy first-order condition and the markup expression (21). However, this approach doesn't apply when firms exert monopsony power over a non-substitutable input, such as materials in Equation (20). In this scenario, the markup expression becomes:

$$\mu_{ft} = (\frac{\alpha^l_{ft}}{\beta^l}\psi^l_{ft} + \alpha^m_{ft}\psi^m_{ft})^{-1} \tag{22}$$

Even if additional variable inputs that substitute for labor are introduced, leading to more first-order conditions, it does not enable expressing the inverse intermediate input supply elasticity ψ^m_{ft} as a function of output elasticities and data. This is because materials are perfect complements to any of these additional variable inputs. In this case, one must either estimate or impose a markup, or estimate the factor supply elasticity, as outlined in Rubens (2023). This identification strategy has been applied in various industries, including Chinese tobacco manufacturing in Rubens (2023), German car manufacturing in Hahn (2024), French dairy production in Avignon and Guigue (2022), and Chinese coal mining in Zheng (2024).

4.2 Labor Market Conduct

The labor market simulations in Sections 2 and 3 assume firms compete oligopsonistically under Nash-Bertrand conduct. This framework encompasses monopsonistic competition as a special case, where firms become atomistic and labor market shares approach zero. In this subsection, we explore alternative forms of labor market competition beyond oligopsonistic and monopsonistic models.

Collusion

One possibility is that firms collude in their input markets, coordinating wage or employment decisions instead of making them independently. Delabastita and Rubens (in press) examines markdown estimation under the assumption of potential collusion. They demonstrate that, even with firms colluding, the wage markdown can still be estimated using the production approach, provided Hicks neutrality holds. Further, by combining labor supply model estimation with production estimates, they identify labor market conduct and find that their collusion estimates match 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 posting wages (Caldwell, Haegele, & Heining, 2025). This bargaining can occur individually or collectively through labor unions. Rubens (2024) explores cost-side markdown estimation in the context of bargained wages, with an empirical application focused on Illinois coal operators negotiating wages with miner unions. A key methodological challenge arises: to identify the bargaining parameters, an estimate of the marginal revenue product of labor is needed, which requires estimating the production function. However, the bargaining parameters themselves must be known to estimate the production function. To resolve this, Rubens (2024) employs a fixed-point estimator, where production function estimation is embedded in a loop that iteratively guesses the bargaining parameters. As explained in Rubens (2024), this procedure converges quickly to a stable set of estimates for both bargaining power and production function coefficients.

4.3 Differentiation and Multi-Product Firms

Product and Input Differentiation

Assumption 2 assumes both goods and inputs are homogeneous, a strong assumption in many contexts. While vertical product differentiation can be incorporated using a price control in the production function (De Loecker et al., 2016), most goods are also horizontally differentiated. Hahn (2024) addresses this challenge by estimating a hedonic price model for car manufacturers, which incorporates car characteristics alongside the production function. This model is then used to estimate markdowns and analyze bargaining between car manufacturers and parts suppliers. A different challenge arises when inputs, rather than products, are differentiated. Lamadon, Mogstad, and Setzler (2022) addresses this by accounting for heterogeneous worker quality using matched employer-employee data.

Multi-Product Firms

Estimating production functions for multi-product firms is challenging, even with perfectly competitive input markets, as inputs are typically not disaggregated at the product level in the data. Several approaches have been proposed to address this issue (De Loecker et al., 2016; Orr, 2022; Dhyne, Petrin, Smeets, & Warzynski, 2022; Valmari, 2023), all without assuming imperfect factor market competition. In contrast, Avignon and Guigue (2022) develop a model that incorporates both imperfect factor market competition and multi-product firms. Their approach leverages engineering data to allocate input costs across different products. They apply this model to estimate factor price markdowns and goods price markups in the French dairy industry.

Assumption 4 assumes that both materials and labor are variable, static inputs. However, in many real-world applications, a subset of these inputs may face adjustment frictions, such as hiring or firing costs. While these frictions do not hinder production function estimation—since timing assumptions can be easily adapted—they complicate markdown identification using the production function approach. This is because the markup and markdown expressions (8) and (10) are derived from solving a static cost minimization problem. Adjustment frictions introduce additional wedges between marginal revenue products and input prices, which are unrelated to monopsony power. One way to separately identify adjustment costs from monopsony distortions is to jointly estimate both a labor supply model and a production model, as demonstrated by Chan, Mattana, Salgado, and Xu (2024) using Danish data.

4.4 No Competitive Input Market

Finally, Assumption 3 assumes that intermediate input prices are exogenous to firms. This is a common assumption in the literature (Morlacco, 2017; Brooks et al., 2021; Yeh et al., 2022; Delabastita & Rubens, in press), and it is essential for point-identifying the markdown when only using the production function, as shown in Equation (10). If all input markets are imperfectly competitive, meaning no input price is exogenous, there are two possible solutions. First, one could impose a model of imperfect competition and estimate a factor supply curve for one of the inputs, as done in Section 3, allowing the markdown of the remaining inputs to be identified using the production approach. Alternatively, Treuren (2022) suggests estimating a revenue production function, in contrast to the quantity production functions used in this paper, to identify wage markdowns without assuming competitive material markets. While the advantage of allowing endogenous material prices is clear, using a revenue production function comes with the tradeoff of imposing homogeneous demand elasticities across firms, limiting the range of imperfect competition models that can be applied to the product market. As with any assumption, the balance between imposing additional restrictions on product market competition while relaxing those in input market competition depends on the specific empirical application and industry context.

5 Conclusions

In this article, we review 'production approaches' for estimating factor price markdowns. We discuss the commonly made assumptions in this class of estimators and 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 'cost-side' 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. By implementing the estimation procedure proposed by Rubens et al. (2024), which accommodates departures from Hicks neutrality, we demonstrate that both production function coefficients and heterogeneity can be consistently estimated in the presence of imperfect labor market competition. Finally, we discuss approaches in the literature that have extended cost-side markdown estimation to relax other assumptions, such as allowing for nonsubstitutable inputs, different types of labor market conduct, and multi-product production.

References

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

Ackerberg, D. A., & De Loecker, J. (2021). *Production function identification under imperfect competition* (Tech. Rep.). Mimeo, University of Leuven.

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.

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. (2024). Dynamic Wage Setting: The Role of Monopsony Power and Adjustment Costs.

Delabastita, V., & Rubens, M. (in press). Colluding Against Workers. *Journal of Political Economy*.

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, Working Paper 30784.

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. (2022). Micro-mechanisms behind declining labor shares: Rising market power and changing modes of production. *International Journal of Industrial Organization*, *81*, 102808.

Mertens, M., & Schoefer, B. (2024). *From labor to intermediates: Firm growth, input substitution, and monopsony* (Tech. Rep.). National Bureau of Economic Research, Working Paper 33172.

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.