Unionization and the Dynamics of Firm Production Technology: A Supply-Side Analysis*

Man Jin[1], Shunan Zhao[1], Yulu Wang[2], and Fei Jia[3]

[1]Department of Economics, Oakland University, USA

[2]Department of Economics, University of Ottawa, Ottawa, ON, Canada

[3]Department of Economics, St. Louis University, USA

Abstract

This paper examines how unionization shapes firms' production technology and market behavior from a supply-side perspective. We develop a structural framework in which unionization affects productivity through both Hicks-neutral and labor-augmenting channels. To obtain causal estimates, we combine the structural model with a regression discontinuity design based on close union representation elections, exploiting the simple-majority threshold. Using matched National Labor Relations Board election records and Compustat data for U.S. manufacturing firms, we estimate a flexible translog production technology and recover firm-level productivity, output elasticities, markups, marginal costs, and wage markdowns. We find that unionization increases both labor-augmenting and Hicks-neutral productivity. An output decomposition shows that both components contribute to the increase in output, with a larger contribution from Hicks-neutral productivity. At the same time, unionization reduces markups and wage markdowns, suggesting lower product- and labor-market power. Unionization also lowers returns to scale and reduces the elasticity of substitution between capital and labor, indicating changes in the structure of production in addition to productivity. Overall, the results show that unionization affects multiple dimensions of firm production and market power that are not captured by conventional measures of overall productivity alone. These findings suggest that

*Emails: mjin@oakland.edu (Jin), shunanzhao@oakland.edu (Zhao), ywang45@uottawa.ca (Wang), fei.jia@slu.edu (Jia) .

labor policy may affect not only worker bargaining power but also firms' production technology and market power.

Keywords: Unionization; Labor-augmenting productivity; Hicks-neutral productivity; Production technology; Market power; Wage markdowns; Regression discontinuity

JEL Classification No.: C3, D24, O5, J5

1 Introduction

Labor unions have long been an important topic in labor economics. Over the past several decades, union membership in the United States has declined substantially. Between 1970 and 2000, the U.S. union membership rate fell by roughly 14 percentage points (Visser, 2006), with an even sharper decline in the private sector, from about 25% in the early 1970s to 9% by 2000 (Farber, 2005). Despite this long-run decline, union organizing has recently received renewed attention in the United States. Union election petitions filed with the National Labor Relations Board more than doubled between fiscal years 2021 and 2024.[1] These recent developments have renewed interest in understanding the economic consequences of unionization for workers and firms.

A central question is whether the gains to workers from unionization come at the expense of firm performance. On the one hand, unions raise wages and strengthen workers' bargaining power, potentially increasing production costs and reducing profits, investment, and employment. On the other hand, unions may improve productive efficiency by facilitating communication between workers and management, reducing worker turnover, and encouraging human-capital investment. Because production workers are directly involved in day-to-day operations, greater worker voice and improved workplace organization may also increase the effectiveness of labor and change how firms organize and combine their inputs. These competing mechanisms imply that the effect of unionization on firm productivity is theoretically ambiguous. More importantly, they suggest that unionization may affect not only the overall level of productivity, but also the way in which firms organize production and utilize their inputs.

A large literature has examined the effects of unionization on firm performance. Early studies

[1] See the report: `https://www.nlrb.gov/news-outreach/news-story/election-petitions-up-53-board-continues-to-reduce-case-processing-time-in`.

emphasize the potentially competing effects of unions on workplace efficiency. Freeman and Medoff (1984), for example, highlights the role of worker voice, lower turnover, and improved workplace organization in generating productivity gains, while also recognizing the potential costs associated with stronger bargaining power. Empirical evidence on productivity is mixed: Clark (1980) documents productivity gains among unionized establishments, whereas Addison and Hirsch (1989) concludes that average productivity effects are relatively modest and may depend on management responses and selection. Subsequent studies examine a broader range of outcomes, including investment (Fallick and Hassett, 1999), employment, output, productivity, and wages (DiNardo and Lee, 2004), firm value (Lee and Mas, 2012), innovation (Bradley et al., 2017), and worker composition and establishment survival (Frandsen, 2021). Evidence from other institutional settings similarly points to substantial heterogeneity in union productivity effects (Morikawa, 2010; Barth et al., 2020; Yang and Tsou, 2018).

While this literature provides important evidence on the overall consequences of unionization, knowing whether unionization raises or lowers productivity provides only part of the answer. Conventional measures such as labor productivity or total factor productivity summarize the net effect of unionization but provide limited information about the underlying production channels. In particular, they do not reveal whether unionization operates by increasing the effectiveness of labor, improving firm-wide technology or organization, or changing the way firms combine their inputs. Distinguishing among these channels is important because they imply different mechanisms through which collective bargaining affects firm performance.

To study these production-side responses, we build on the structural production-function literature. Standard proxy-variable approaches address the endogeneity of firms' input choices but typically represent unobserved productivity as a scalar Hicks-neutral component (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015; Gandhi et al., 2020). This scalar representation is not sufficient for our purpose because unionization may affect different dimensions of production asymmetrically. We therefore draw on more recent work that allows productivity and technological change to have factor-specific as well as factor-neutral components (Doraszelski and Jaumandreu, 2018; Zhang, 2019; Malikov et al., 2024).

Our paper takes a supply-side approach to the study of unionization by introducing multidimensional productivity into a flexible production framework. Specifically, we distinguish between

labor-augmenting and Hicks-neutral productivity. The former captures changes in the effectiveness of labor, while the latter captures broader firm-wide productivity changes. This distinction is particularly relevant in the union context because mechanisms such as worker voice, workplace organization, training, and human-capital investment may operate directly through labor effectiveness. At the same time, the flexible production framework allows us to examine whether unionization changes other dimensions of production technology, including output elasticities, returns to scale, and input substitution, as well as firms' product- and labor-market power.

Separating these production channels, however, does not by itself establish that the estimated differences are caused by unionization. Because unionization is not randomly assigned, differences between unionized and nonunionized firms may reflect preexisting firm characteristics rather than the effect of collective bargaining. To obtain causal variation, we combine the structural production model with a regression discontinuity design based on union representation elections. Specifically, we exploit the simple-majority rule in National Labor Relations Board elections, under which union representation is determined by whether the union receives more than 50% of the votes. Firms in which unions narrowly win or narrowly lose therefore provide quasi-experimental variation in union status. We incorporate this variation directly into the evolution of both labor-augmenting and Hicks-neutral productivity. Thus, the discontinuous change in union status around the 50% threshold identifies the effect of unionization, while the structural restrictions and firms' input choices identify the underlying production technology and productivity components.

Using matched National Labor Relations Board election records and Compustat data for U.S. manufacturing firms over the period 1980–2009, we find that unionization affects firms along several dimensions. First, unionization significantly increases both labor-augmenting and Hicks-neutral productivity. The implied overall output gain is approximately 10.3%, of which 2.9 percentage points operate through the labor-augmenting channel and 7.4 percentage points through the Hicks-neutral channel. Thus, both productivity components contribute to the increase in output, but the broader firm-wide productivity channel accounts for most of the gain. Second, unionization changes the structure of production: returns to scale decline, and capital and labor become less easily substitutable. Finally, unionization reduces both markups and wage markdowns, suggesting weaker firm market power in product and labor markets. Taken together, these findings indicate that unionization affects not only productive efficiency but also the organization of production and

the distribution of market power.

Our paper makes two main contributions. First, we contribute to the literature on unionization and firm productivity by moving beyond a single measure of overall productivity. Existing studies commonly examine outcomes such as output per worker or hour, total factor productivity, or other firm- and establishment-level performance measures (Clark, 1980; DiNardo and Lee, 2004). We instead distinguish between Hicks-neutral and labor-augmenting productivity, allowing us to examine whether unionization affects firm performance through changes in labor effectiveness or through broader firm-wide productivity. The same framework also allows us to study how unionization reshapes other dimensions of production technology and market behavior, including output elasticities, returns to scale, input substitution, markups, marginal costs, and wage markdowns.

Second, we contribute to the structural production-function literature by combining a multidimensional productivity framework with quasi-experimental variation in unionization. Building on recent approaches that distinguish Hicks-neutral from factor-augmenting productivity (Doraszelski and Jaumandreu, 2018; Zhang, 2019; Malikov et al., 2024), we incorporate unionization directly into the productivity evolution processes and identify its effects using the discontinuity generated by close union representation elections. This differs from a two-step approach in which productivity is first recovered without accounting for unionization and subsequently regressed on union status. In our framework, the potential effects of unionization are incorporated when productivity and production technology are jointly estimated. More broadly, the approach illustrates how quasi-experimental variation can be embedded within a structural production framework to study causal changes in multidimensional productivity while accommodating imperfect competition in both product and labor markets.

Beyond these contributions, our findings speak to the broader question of whether stronger worker bargaining power necessarily comes at the expense of firm efficiency. By examining productivity, production structure, and market power within the same framework, we show that these dimensions need not move in the same direction. In our setting, unionization is associated with productivity gains while simultaneously reducing firms' product- and labor-market power. This suggests that stronger worker bargaining power need not imply lower productive efficiency, and that the effects of labor-market institutions may extend beyond wages and employment to firms' production decisions and market power.

The remainder of the paper proceeds as follows. Section 2 presents the structural framework. Section 3 discusses the identification and estimation strategy. Section 4 describes the data and sample construction. Section 5 presents the empirical results, and Section 6 concludes.

2 Model

Consider the following production technology for firm i in period t through which the firm converts its inputs into a single output:

$$Q_{it} = F(K_{it}, e^{\varphi_{it}} L_{it}, M_{it})\ e^{\omega_{it}+\eta_{it}}, \tag{2.1}$$

where Q_{it} is the firm's output quantity, and K_{it}, L_{it}, and M_{it} are capital, labor, and materials, respectively; φ_{it} is labor-augmenting productivity and ω_{it} is the conventional Hicks-neutral productivity in the literature; η_{it} is a pure noise, such as a measurement error in output; $F(\cdot)$ is a parametric production function, as specified below.

In the production function formulation (2.1), there are two productivity terms $(\varphi_{it}, \omega_{it})$. While ω_{it} captures the neutral technological changes that affect the marginal products of all inputs in the same proportion, φ_{it} accommodates non-neutral labor "augmenting" effects of technological changes and allows the marginal productivity of input to change at different rates. Compared with capital and materials, we specifically model the labor-augmenting productivity since a large number of previous studies have documented the potential effects of unionization on employees and labor productivity.

We use a translog specification for $F(\cdot)$ due to its flexibility. Specifically, define the effective labor as $\widetilde{L}_{it} \equiv e^{\varphi_{it}} L_{it}$. Then, we have

$$\begin{aligned}\ln F(K_{it}, \widetilde{L}_{it}, M_{it}) = \beta_k k_{it} + \beta_l \widetilde{l}_{it} + \beta_m m_{it} + \beta_{kl} k_{it} \widetilde{l}_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} \widetilde{l}_{it} m_{it} + \\ \tfrac{1}{2}\beta_{kk} k_{it}^2 + \tfrac{1}{2}\beta_{ll} \widetilde{l}_{it}^2 + \tfrac{1}{2}\beta_{mm} m_{it}^2,\end{aligned} \tag{2.2}$$

where, throughout the paper, lower-case letters denote logs of their upper-case counterparts.

The productivity components are controlled first-order autoregressive processes with the fol-

lowing representations:

$$\varphi_{it} = \alpha_{\varphi,1}\varphi_{it-1} + \phi_{\varphi}(u_i) + \alpha_{\varphi,2}d_{it'} + \epsilon_{\varphi,it}, \tag{2.3}$$

$$\omega_{it} = \alpha_{\omega,1}\omega_{it-1} + \phi_{\omega}(u_i) + \alpha_{\omega,2}d_{it'} + \epsilon_{\omega,it}, \tag{2.4}$$

where $\phi_{\varphi}(\cdot)$ and $\phi_{\omega}(\cdot)$ are flexible unknown functions; u_i denotes employees' support for unionization in firm i, which we later measure using the share of votes in favor of unionization in the representation election; $d_{it'}$ is the firm's unionization status dummy in period t'; and $\epsilon_{\varphi,it}$ and $\epsilon_{\omega,it}$ are zero-mean productivity innovations. We let $t' \leq t$, allowing for the possibility that the effects of unionization on productivity take time to materialize. Processes (2.3) and (2.4) do not include intercepts, as these cannot be separately identified from $\phi_{\varphi}(\cdot)$ and $\phi_{\omega}(\cdot)$.

3 Identification and Estimation

Define $\mathcal{I}_{it}$ as firm i's information set in period t, based on which the firm makes its production decisions. Within the structural framework of production function and productivity estimation, we have the following identification assumptions.

Assumption A *(a) Firm's production technologies can be specified by Equations* (2.1) *and* (2.2)*, and the pure noise* η_{it} *is mean independent of variables that belong to* $\mathcal{I}_{it}$*.*

(b) Productivity terms φ_{it} *and* ω_{it} *follow first-order autoregressive processes specified in Equations* (2.3) *and* (2.4)*. They are known for firm* i *when it makes its production decisions in period* t*, i.e.,* $(\varphi_{it}, \omega_{it}) \in \mathcal{I}_{it}$*. The productivity innovation terms in period* t*, denoted by* $\epsilon_{\varphi,it}$ *and* $\epsilon_{\omega,it}$*, are mean independent of variables that belong to* $\mathcal{I}_{it-1}$*.*

(c) In the production function (2.1)*,* φ_{it} *has a zero mean, i.e.,* $\mathbb{E}(\varphi_{it}) = 0$*.*

The above assumptions are about firms' production technologies. Compared with previous studies in the literature of non-neutral productivity (Doraszelski and Jaumandreu, 2018; Demirer, 2022; Malikov et al., 2024), Assumption A(a) specifies a flexible relationship between production inputs and output by not only using a translog (log-quadratic) production function but also allowing for three productivity terms. For Assumption A(b), although productivity terms $(\varphi_{it}, \omega_{it}) \in \mathcal{I}_{it}$

are known to firm i in period t, they are not observed by econometricians. Also, according to the following assumption regarding firms' behavior, they are correlated with the optimal input usage (K_{it}, L_{it}, M_{it}), and therefore lead to endogeneity issues if not being appropriately controlled for.

The mean independence of productivity innovation terms from lagged variables in the information set $\mathcal{I}_{it-1}$ is standard in the literature on the structural identification of production functions. This assumption is slightly weaker than the conventional Markov assumption on productivity processes. Under the Markov assumption, the productivity innovation terms $\epsilon_{\varphi,it}$ and $\epsilon_{\omega,it}$ are independent of all variables in $\mathcal{I}_{it-1}$, including $d_{it'}$ for $t' < t$. As a result, the causal effects of unionization on productivity can be directly estimated under the Markov assumption. In our setting, however, such an assumption may not be appropriate when $t' = t$ and both u_i and $d_{it'}$ are included in the productivity processes. A more suitable identification strategy, which directly exploits the data-generating process of $d_{it'}$, is the regression discontinuity design, as outlined in the following section.

Assumption (c) is an identification assumption on the location of the φ_{it} distribution. A similar assumption has also been used by Jaumandreu and Mullens (2024), and it helps to better separate the latent labor-augmented productivity from the unknown coefficients in the production function.

Assumption B *The input K_{it} is quasi-fixed, with its optimal value determined dynamically by the firm i in period $t-1$, whereby $K_{it} \in \mathcal{I}_{it}$. In each period t, firm i minimizes its expected variable cost conditional on state variables in $\mathcal{I}_{it}$ by choosing inputs $(L_{it}, M_{it})'$.*

The above assumptions are about firms' behaviors. The input timing Assumption B is widely-used in the literature and has been adopted by previous studies such as Levinsohn and Petrin (2003), Doraszelski and Jaumandreu (2018), Zhang (2019), and Malikov et al. (2024). In our context, we assume that even for unionized workers, hiring and firing decisions can be finalized within one period, which corresponds to one year in our data. We consider this assumption reasonable. Nevertheless, if it were violated to some extent, it would introduce a discrepancy in the first-order condition with respect to labor in the firm's static profit maximization problem. We return to this point in the next section. In addition, we do not impose specific assumptions on the output market conditions beyond assuming that firms are cost minimizers, which is standard in production theory. This modeling choice accommodates a wide range of competitive environments, including perfect

competition, monopolistic competition, and more general market structures.

3.1 Two-stage Identification Strategy

According to Assumption B, we can write the static cost minimization problem of firm i in period t as

$$\min_{M_{it},L_{it}} \quad P_{it}^M M_{it} + P_{it}^L L_{it} + \lambda_{it} \left\{ Y^0 - F(K_{it}, \widetilde{L}_{it}, M_{it})\ e^{\omega_{it}} \Lambda \right\}, \tag{3.1}$$

where P_{it}^L and P_{it}^M denote the prices of labor and material inputs, respectively; λ_{it} is the Lagrangian multiplier, which economically represents the marginal cost of production; $\widetilde{L}_{it} \equiv e^{\varphi_{it}} L$ is the effective labor; and $\Lambda^* \equiv \mathbb{E}(e^{\eta_{it}} \mid \mathcal{I}_{it}) = \mathbb{E}(e^{\eta_{it}})$, and it is introduced into the cost-minimization problem since η_{it} is unobserved by firm i at the time it makes its production decisions in period t.

We have the following two first-order conditions based on the above problem:

$$P_{it}^M = \lambda_{it} \frac{\partial F(K_{it}, \widetilde{L}_{it}, M_{it})}{\partial M} e^{\omega_{it}} \Lambda \quad \text{and} \quad (1+\tau) P_{it}^L = \lambda_{it} \frac{\partial F(K_{it}, \widetilde{L}_{it}, M_{it})}{\partial \widetilde{L}} e^{\omega_{it} + \varphi_{it}} \Lambda,$$

where we assume an expected wage markdown $\tau \geq 0$, defined as the average percentage gap between the value of the marginal product of labor and the wage. Similar to the output market, we do not impose a specific structure on labor market competition, but we allow the wage markdown to potentially depend on union presence. Arranging the first-order conditions, we have the following share equations

$$S_{it}^M = \frac{\lambda_{it}}{P_{it}} \frac{\Lambda}{e^{\eta_{it}}} \Big(\beta_m + \beta_{km} k_{it} + \beta_{mm} m_{it} + \beta_{lm} (l_{it} + \varphi_{it}) \Big), \tag{3.2}$$

$$(1+\tau) S_{it}^L = \frac{\lambda_{it}}{P_{it}} \frac{\Lambda}{e^{\eta_{it}}} \Big(\beta_l + \beta_{kl} k_{it} + \beta_{lm} m_{it} + \beta_{ll} (l_{it} + \varphi_{it}) \Big), \tag{3.3}$$

where S_{it}^L and S_{it}^M denote the labor and material input shares in output, respectively, i.e., $S_{it}^X \equiv \frac{P_{it}^X X_{it}}{P_{it} Q_{it}}$ for $X \in \{L, M\}$, and the right-hand side corresponds to the output elasticities with respect to labor and materials, computed from the translog specification in (2.2).

One concern regarding Equation (3.3) is whether the wage markdown and labor-augmenting productivity can be separately identified. In particular, a change in τ could be mistakenly attributed

to a change in φ_{it}, since both are unobserved. However, the presence of the material share equation helps resolve this concern: a change in φ_{it} simultaneously affects both Equations (3.2) and (3.3), whereas τ only influences Equation (3.3). However, to facilitate the subsequent derivation and estimation procedure, we impose the following simplifying assumption.

Assumption C *In the translog specification of the production function in* (2.2)*,* $\beta_{ll} = 0$*.*

We emphasize the following points regarding this assumption: (1) With this restriction, it becomes transparent that τ and φ_{it} can be separately identified in the share equation system above; (2) Without this restriction, our identification strategy still applies but requires solving a substantially more complex nonlinear optimization problem; (3) This assumption is less restrictive than the conventional homothetic separability assumption often imposed in production functions with non-neutral technologies, which would require $\beta_{kl} = \beta_{km} = 0$ and $\beta_{ll} = \beta_{mm} = -\beta_{lm}$.

In addition, union power affects the above share equality indirectly through multiple channels. It influences φ_{it} via the union support u_i and the unionization dummy $d_{it'}$, alters input shares through its impact on firm production, and also affects the marginal cost by changing productivity. We will return to this discussion in the results section.

Under Assumption C, we can take the ratio of Equations (3.2) and (3.3) and solve for φ_{it} to construct a proxy for the unobserved labor-augmenting productivity:

$$\varphi_{it} = -b_1 - b_2 k_{it} - b_3 m_{it} - l_{it} + b_4 S^r_{it} + b_5 k_{it} S^r_{it} + b_6 m_{it} S^r_{it} \tag{3.4}$$

where $b_1 \equiv \frac{\beta_m}{\beta_{lm}}$, $b_2 \equiv \frac{\beta_{km}}{\beta_{lm}}$, $b_3 \equiv \frac{\beta_{mm}}{\beta_{lm}}$, $b_4 \equiv \frac{\beta_l}{\beta_{lm}(1+\tau)}$, $b_5 \equiv \frac{\beta_{kl}}{\beta_{lm}(1+\tau)}$, $b_6 \equiv \frac{1}{1+\tau}$, and $S^r_{it} \equiv S^M_{it}/S^L_{it}$. Note that the φ proxy in (3.4) is not uniquely defined in terms of unknown parameters $(\beta_l, \beta_m, \beta_{km}, \beta_{kl}, \beta_{lm}, \tau)$ since coefficients b_1 to b_5 are ratios of the β parameters. Once estimates of b_j, $j = 1, \ldots, 6$, are obtained, the β coefficients can be recovered up to a scale parameter β_{lm} by exploring their relationships.

Next, we discuss a two-stage procedure for the identification of the production technology. The two-stage structure is needed because labor-augmenting and Hicks-neutral productivity are both unobserved and cannot be separately identified from the production function alone. We therefore first use the input-share conditions to obtain a proxy for labor-augmenting productivity and identify

the associated parameter combinations. Conditional on this component, the remaining production-function parameters and Hicks-neutral productivity can then be identified in the second stage.

Stage One. To estimate b_1 through b_6, we exploit the evolution process of φ together with its proxy defined in Equation (3.4). This process follows a dynamic panel model with unobserved productivity. By substituting φ_{it} and its lagged value with their respective proxies from (3.4), we obtain

$$\begin{aligned} l_{it} =& b_1^* + \alpha_{\varphi,1} l_{it-1} - b_2 \left(k_{it} - \alpha_{\varphi,1} k_{it-1}\right) - b_3 \left(m_{it} - \alpha_{\varphi,1} m_{it-1}\right) + b_4 \left(S_{it}^r - \alpha_{\varphi,1} S_{it-1}^r\right) + \\ & b_5 \left(k_{it} S_{it}^r - \alpha_{\varphi,1} k_{it-1} S_{it-1}^r\right) + b_6 \left(m_{it} S_{it}^r - \alpha_{\varphi,1} m_{it-1} S_{it-1}^r\right) - \phi_\varphi(u_i) - \alpha_{\varphi,2} d_{it'} - \epsilon_{\varphi,it}, \end{aligned} \tag{3.5}$$

where $b_1^* \equiv -(1 - \alpha_{\varphi,1}) b_1$. This transformation yields a new dynamic panel model that no longer contains unobserved productivity terms.

To identify Equation (3.5), we next discuss the data-generating processes of its right-hand-side variables. A firm's unionization decision is determined by the outcome of its union representation election:

$$d_{it'} = \begin{cases} 1 & \text{if } u_i > u^0 \text{ and } t' > t_i^0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.6}$$

where t_i^0 denotes the election period for firm i, and $u^0 = 50\%$ is the threshold for winning the election under the simple majority rule. Since $d_{it'}$ is a deterministic function of u_i (and period t'), the coefficient $\alpha_{\varphi,2}$, which captures the marginal effect of unionization on the labor-augmented productivity, can be consistently estimated from Equations (3.5) following the idea of the sharp regression discontinuity design.

However, in order to estimate the remaining parameters in the equation above, we need to discuss the identification assumptions imposed in our structural model. According to Assumptions A(b) and B, S_{it}^r and m_{it} are correlated with $\epsilon_{\varphi,it}$ and are therefore endogenous. In contrast, their lagged values S_{it-1}^r and m_{it-1}, together with $(k_{it}, k_{it-1}, l_{it-1})$, are exogenous.

Regarding u_i, we have so far remained agnostic about its econometric properties. Is u_i correlated with production inputs (K_{it}, L_{it}, M_{it})? Likely so, since inputs are chosen based on firm productivity terms, which themselves vary with u_i if $\phi_\varphi(\cdot)$ and $\phi_\omega(\cdot)$ are not constant. Consequently, if u_i is

correlated with the productivity innovation term in Equation (3.5), the b coefficients and $\alpha_{\varphi,1}$ would be estimated with bias. To obtain consistent estimation, we impose the following identification assumption:

Assumption D *(a) The employees' support for unionization u_i is independent of the productivity innovations $\epsilon_{\varphi,it}$ and $\epsilon_{\omega,it}$.*

(b) For some i and t, $k_{it} \neq \alpha_{\varphi,1} k_{it-1}$, $m_{it} \neq \alpha_{\varphi,1} m_{it-1}$, $S^r_{it} \neq \alpha_{\varphi,1} S^r_{it-1}$, $k_{it} S^r_{it} \neq \alpha_{\varphi,1} k_{it-1} S^r_{it-1}$, and $m_{it} S^r_{it} \neq \alpha_{\varphi,1} m_{it-1} S^r_{it-1}$.

(c) Define $\phi_{it} = \left(\beta_l \widetilde{l}_{it} + \beta_m m_{it} + \beta_{kl} k_{it} \widetilde{l}_{it} + \beta_{km} k_{it} m_{it} + \beta_{lm} \widetilde{l}_{it} m_{it} + \frac{1}{2}\beta_{mm} m^2\right) / \beta_{lm}$. For some i and t, $k_{it} \neq \alpha_{\omega,1} k_{it-1}$, $k^2_{it} \neq \alpha_{\omega,1} k^2_{it-1}$, and $\phi_{it} \neq \alpha_{\omega,1} \phi_{it-1}$.

To understand Assumption D(a), from an econometric perspective, the components of u_i in the productivity evolution processes (2.3) and (2.4) are analogous to fixed effects in panel models. In such models, the error terms are conventionally assumed to have zero mean conditional on the independent variables and the fixed effects. From an economic perspective, although employees' support for unionization may be correlated with overall firm productivity, as described in their evolution processes (2.3) and (2.4), it is unclear whether it is correlated with the productivity innovation terms. Nevertheless, we adopt Assumption D as a technical identification condition in order to structurally estimate the production function parameters.

For Assumption D(b), it ensures that the components inside the parentheses are nonzero. Assumption D(c) serves the same role, but for the second-stage estimation discussed below. These assumptions are mild, as the production variables are determined by firms' optimization behavior and are unlikely to be perfectly dependent over time.

Under Assumptions A to D, we can estimate Equation (3.5) to identify the unknown parameters $(\alpha_{\varphi,1}, \alpha_{\varphi,2}, b^*_1, b_2, \ldots, b_6)$ and function $\phi_\varphi(\cdot)$ based on the following moment conditions:

$$\mathbb{E}\Big(\epsilon_{\varphi,it} \;\Big|\; 1, S^r_{it-1}, k_{it}, k_{it-1}, l_{it-1}, m_{it-1}, u_i, d_{it'}\Big) = 0.$$

Note that once the b parameters are estimated, the unobserved productivity φ_{it} can be identified from (3.4), so can the effective labor input $\widetilde{l}_{it}$. The wage markdown τ is then recovered from $\frac{1-b_6}{b_6}$, and the β coefficient ratios $\left(\frac{\beta_m}{\beta_{lm}}, \frac{\beta_{km}}{\beta_{lm}}, \frac{\beta_{mm}}{\beta_{lm}}, \frac{\beta_l}{\beta_{lm}}, \frac{\beta_{kl}}{\beta_{lm}}\right)$ are recovered from $\left(\frac{b^*_1}{-(1-\alpha_{\varphi,1})},\ b_2,\ b_3,\ \frac{b_4}{b_6},\ \frac{b_5}{b_6}\right)$.

Stage Two. To estimate the remaining unknown parameters, we write the production function as follows:

$$\begin{aligned} y_{it} &= \beta_{lm}\left(\tfrac{\beta_l}{\beta_{lm}}\widetilde{l}_{it} + \tfrac{\beta_m}{\beta_{lm}}m_{it} + \tfrac{\beta_{kl}}{\beta_{lm}}k_{it}\widetilde{l}_{it} + \tfrac{\beta_{km}}{\beta_{lm}}k_{it}m_{it} + \widetilde{l}_{it}m_{it} + \tfrac{1}{2}\tfrac{\beta_{mm}}{\beta_{lm}}m_{it}^2\right) + \beta_k k_{it} + \tfrac{1}{2}\beta_{kk}k_{it}^2 + \omega_{it} + \eta_{it}, \\ &= \beta_{lm}\phi_{it} + \beta_k k_{it} + \tfrac{1}{2}\beta_{kk}k_{it}^2 + \omega_{it} + \eta_{it}, \end{aligned}$$

where ϕ_{it} is defined in Assumption D(c) and identified in the first stage. We can solve for ω_{it} from the above equation:

$$\omega_{it} = y_{it} - \beta_{lm}\phi_{it} - \beta_k k_{it} - \tfrac{1}{2}\beta_{kk}k_{it}^2 - \eta_{it}, \tag{3.7}$$

which serves as a proxy for the unobserved Hicks-neutral productivity. Then, plugging the proxy and its lagged value into the law of motion (2.4), we obtain the following dynamic panel model:

$$\begin{aligned} y_{it} =& \alpha_{\omega,1}y_{it-1} + \beta_{lm}\left(\phi_{it} - \alpha_{\omega,1}\phi_{it-1}\right) + \beta_k\left(k_{it} - \alpha_{\omega,1}k_{it-1}\right) + \\ & \tfrac{1}{2}\beta_{kk}\left(k_{it}^2 - \alpha_{\omega,1}k_{it-1}^2\right) + \phi_\omega(u_i) + \alpha_{\omega,2}d_{it'} + \epsilon^*_{\omega,it}, \end{aligned} \tag{3.8}$$

where $\epsilon^*_{\omega,it} \equiv \epsilon_{\omega,it} + \eta_{it} - \alpha_{\omega,1}\eta_{it-1}$.

Under Assumptions A to D, we can identify unknown parameters $(\alpha_{\omega,1}, \alpha_{\omega,2}, \beta_{lm}, \beta_k, \beta_{kk})$ and function $\phi_\omega(\cdot)$ from the following moment condition:

$$\mathbb{E}\left(\epsilon^*_{it} \mid y_{it-1}, \phi_{it-1}, k_{it}, k_{it-1}, u_i, d_{it'}\right) = 0.$$

Note that ϕ_{it-1} is identified after the first stage and can be treated as data in the above moment. With the estimated β_{lm}, $(\beta_l, \beta_m, \beta_{km}, \beta_{kl}, \beta_{mm})$ can be recovered from their ratios in the first stage.

3.2 Identification of Productivity

In the first stage, once the b parameters are estimated, the labor-augmenting productivity φ_{it} can be identified from Equation (3.4). Likewise, the factor-neutral productivity ω_{it} can be identified from Equation (3.7), provided that the coefficients $(\beta_{lm}, \beta_k, \beta_{kk})$ and the error term η_{it} are also estimated. To estimate η_{it}, we follow the approach of Ackerberg et al. (2015), replacing φ_{it} and ω_{it} with their implicit functional forms and estimating a reduced-form nonparametric regression model.

In particular, Equation (3.4) implies that φ_{it} can be expressed as a function of $(k_{it}, m_{it}, l_{it}, S^r_{it})$:

$$\varphi_{it} = g_\varphi(k_{it}, m_{it}, l_{it}, S^r_{it}).$$

The first-order condition for labor input in the cost minimization problem (3.1) implies that ω_{it} is a function of $(k_{it}, l_{it}, m_{it}, p^L_{it}, \varphi_{it}, \lambda_{it})$. If we assume that λ_{it} is an (unknown) function of firm characteristics and location factors, denoted by z_{it}, then ω_{it} can be implicitly expressed as

$$\omega_{it} = g_\omega(k_{it}, l_{it}, m_{it}, p^L_{it}, S^r_{it}, z_{it}),$$

where φ_{it} is substituted by its implicit function $g_\varphi(k_{it}, m_{it}, l_{it}, S^r_{it})$ and subsequently incorporated into $g_\omega(\cdot)$.

By replacing φ_{it} and ω_{it} with their respective implicit functions, we can rewrite the production function in a non-structural form:

$$y_{it} = \Psi(k_{it}, l_{it}, m_{it}, p^L_{it}, S^r_{it}, z_{it}) + \eta_{it}, \tag{3.9}$$

where $\Psi(\cdot) = \ln F(\cdot; g_\varphi(\cdot)) + g_\omega(\cdot)$. Estimating the reduced-form regression model (3.9) with standard nonparametric methods yields the estimate of η_{it} as the regression residual.

3.3 Estimation Details and Inference

We now discuss the estimation of the production-function parameters, the productivity processes, the wage markdown, and the unionization effects. To facilitate inference and improve estimation efficiency, we estimate the two stages described above jointly within a GMM framework.

Let

$$\boldsymbol{\beta} = (\beta_k, \beta_l, \beta_m, \beta_{kl}, \beta_{km}, \beta_{lm}, \beta_{kk}, \beta_{mm})' \tag{3.10}$$

collect the production-function parameters, where $\beta_{ll} = 0$ is imposed throughout estimation, and define

$$\boldsymbol{\vartheta} = (\boldsymbol{\beta}', \tau)'. \tag{3.11}$$

Using the labor-augmenting productivity proxy derived in Equation (3.4), we write

$$q_{it}(\boldsymbol{\vartheta}) = -\, b_1 - b_2 k_{it} - b_3 m_{it} - l_{it} + b_4 S^r_{it} + b_5 k_{it} S^r_{it} + b_6 m_{it} S^r_{it} \tag{3.12}$$

where $b_1 \equiv \frac{\beta_m}{\beta_{lm}}$, $b_2 \equiv \frac{\beta_{km}}{\beta_{lm}}$, $b_3 \equiv \frac{\beta_{mm}}{\beta_{lm}}$, $b_4 \equiv \frac{\beta_l}{\beta_{lm}(1+\tau)}$, $b_5 \equiv \frac{\beta_{kl}}{\beta_{lm}(1+\tau)}$, $b_6 \equiv \frac{1}{1+\tau}$, and $S^r_{it} \equiv S^M_{it}/S^L_{it}$. Because the absolute level of labor-augmenting productivity is not separately identified from the level of effective labor in the production function, a location normalization is required. We therefore normalize labor-augmenting productivity to have zero sample mean:

$$\widetilde{\varphi}_{it}(\boldsymbol{\vartheta}) = q_{it}(\boldsymbol{\vartheta}) - \overline{q}(\boldsymbol{\vartheta}), \tag{3.13}$$

where $\overline{q}(\boldsymbol{\vartheta}) = \frac{1}{N}\sum_{i,t} q_{it}(\boldsymbol{\vartheta})$. Effective labor in logs is therefore

$$\widetilde{l}_{it}(\boldsymbol{\vartheta}) = l_{it} + \widetilde{\varphi}_{it}(\boldsymbol{\vartheta}). \tag{3.14}$$

For implementation, we approximate the unknown functions of union support in the two productivity processes using cubic polynomials:

$$\phi_\varphi(u_i) = \gamma_{\varphi,1} u_i + \gamma_{\varphi,2} u_i^2 + \gamma_{\varphi,3} u_i^3, \tag{3.15}$$

$$\phi_\omega(u_i) = \gamma_{\omega,1} u_i + \gamma_{\omega,2} u_i^2 + \gamma_{\omega,3} u_i^3. \tag{3.16}$$

As defined above, the unionization indicator is

$$d_{it} = \mathbf{1}\{u_i > u^0\}\mathbf{1}\{t > t_i^0\}, \qquad u^0 = 0.5, \tag{3.17}$$

where t_i^0 denotes the year of the union representation election. Thus, in the baseline specification, the effect of unionization on productivity begins in the first period following the election year.

We first consider the labor-augmenting productivity process. Let

$$\boldsymbol{\pi}_\varphi = (\alpha_{\varphi,1}, \alpha_{\varphi,2}, \gamma_{\varphi,1}, \gamma_{\varphi,2}, \gamma_{\varphi,3})'. \tag{3.18}$$

Substituting the proxy in Equation (3.12) into the productivity process gives

$$\xi_{\varphi,it}(\boldsymbol{\vartheta}, \boldsymbol{\pi}_\varphi) = \; q_{it}(\boldsymbol{\vartheta}) - \alpha_{\varphi,1} q_{it-1}(\boldsymbol{\vartheta}) - \alpha_{\varphi,2} d_{it} - \gamma_{\varphi,1} u_i - \gamma_{\varphi,2} u_i^2 - \gamma_{\varphi,3} u_i^3.$$

Under the identification assumptions discussed above, we use the instrument vector

$$Z_{\varphi,it} = \Big(1,\; k_{it},\; k_{it-1},\; m_{it-1},\; S^r_{it-1},\; k_{it-1}S^r_{it-1},\; m_{it-1}S^r_{it-1}, d_{it},\; u_i,\; u_i^2,\; u_i^3\Big)'.$$

The associated moment conditions are

$$\mathbb{E}\left[Z_{\varphi,it}\xi_{\varphi,it}(\boldsymbol{\vartheta}_0, \boldsymbol{\pi}_{\varphi,0})\right] = 0. \tag{3.19}$$

We next turn to the Hicks-neutral productivity process. Given $\boldsymbol{\vartheta}$ and the corresponding measure of effective labor, define

$$\begin{aligned} q^\omega_{it}(\boldsymbol{\vartheta}) = y_{it} &- \beta_k k_{it} - \beta_l \widetilde{l}_{it}(\boldsymbol{\vartheta}) - \beta_m m_{it} - \beta_{kl} k_{it} \widetilde{l}_{it}(\boldsymbol{\vartheta}) \\ &- \beta_{km} k_{it} m_{it} - \beta_{lm} \widetilde{l}_{it}(\boldsymbol{\vartheta}) m_{it} - \frac{1}{2}\beta_{kk} k_{it}^2 - \frac{1}{2}\beta_{mm} m_{it}^2. \end{aligned} \tag{3.20}$$

At this stage, q^ω_{it} contains both Hicks-neutral productivity and the transitory output disturbance. Let

$$\boldsymbol{\pi}_\omega = (\alpha_{\omega,0}, \alpha_{\omega,1}, \alpha_{\omega,2}, \gamma_{\omega,1}, \gamma_{\omega,2}, \gamma_{\omega,3})'. \tag{3.21}$$

The corresponding estimating residual is

$$\xi_{\omega,it}(\boldsymbol{\vartheta}, \boldsymbol{\pi}_\omega) = \; q^\omega_{it}(\boldsymbol{\vartheta}) - \alpha_{\omega,0} - \alpha_{\omega,1} q^\omega_{it-1}(\boldsymbol{\vartheta}) - \alpha_{\omega,2} d_{it} - \gamma_{\omega,1} u_i - \gamma_{\omega,2} u_i^2 - \gamma_{\omega,3} u_i^3.$$

The composite residual in Equation (3.22) contains the innovation to Hicks-neutral productivity together with the current and lagged transitory output disturbances. Using the timing assumptions discussed in the identification section, we use the instrument vector

$$Z_{\omega,it} = \Big(1,\; k_{it},\; k_{it-1},\; l_{it-1},\; m_{it-1},\; k_{it-1}l_{it-1},\; k_{it-1}m_{it-1},\; l_{it-1}m_{it-1}, k_{it}^2,\; k_{it-1}^2,\; m_{it-1}^2,\; d_{it},\; u_i,\; u_i^2,\; u_i^3\Big)'.$$

The corresponding moment conditions are

$$\mathbb{E}\left[Z_{\omega,it}\xi_{\omega,it}(\boldsymbol{\vartheta}_0, \boldsymbol{\pi}_{\omega,0})\right] = 0. \tag{3.22}$$

The two sets of moment conditions share the production-function and wage-markdown parameters contained in $\boldsymbol{\vartheta}$. Rather than estimating the two sets of moments separately, we therefore estimate them jointly. Define the full parameter vector as

$$\boldsymbol{\theta} = \left(\boldsymbol{\vartheta}', \boldsymbol{\pi}_{\varphi}', \boldsymbol{\pi}_{\omega}'\right)'. \tag{3.23}$$

For each observation, we collect the two sets of moment conditions in

$$g_{it}(\boldsymbol{\theta}) = \begin{bmatrix} Z_{\varphi,it}\xi_{\varphi,it}(\boldsymbol{\vartheta}, \boldsymbol{\pi}_{\varphi}) \\ Z_{\omega,it}\xi_{\omega,it}(\boldsymbol{\vartheta}, \boldsymbol{\pi}_{\omega}) \end{bmatrix}, \qquad \mathbb{E}[g_{it}(\boldsymbol{\theta}_0)] = 0. \tag{3.24}$$

The corresponding sample moment vector is

$$\overline{g}_N(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i,t} g_{it}(\boldsymbol{\theta}), \tag{3.25}$$

where N denotes the number of firm-year observations used in estimation.

The joint GMM estimator is

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} \overline{g}_N(\boldsymbol{\theta})'\widehat{W}_N\overline{g}_N(\boldsymbol{\theta}), \tag{3.26}$$

where $\widehat{W}_N$ is a positive-definite weighting matrix. We implement a two-step GMM estimator. After obtaining a preliminary consistent estimate using an initial weighting matrix, we estimate the long-run covariance matrix of the stacked moments, $\widehat{\Omega}_N$, and use the optimal second-step weighting matrix

$$\widehat{W}_N = \widehat{\Omega}_N^{-1}. \tag{3.27}$$

Because the two sets of moment conditions are estimated jointly, $\widehat{\Omega}_N$ also incorporates the covariance between the labor-augmenting and Hicks-neutral productivity equations.

For numerical implementation, the production-function parameters are initialized using estimates from a conventional translog regression. The restriction $\beta_{ll} = 0$ is imposed throughout estimation.

Once the joint GMM estimates are obtained, the remaining step is to separate Hicks-neutral productivity from the transitory output disturbance. As discussed in Section 3.2, we recover transitory output disturbance from the reduced-form representation of the production function. Let

$$X_{it} = \left(k_{it}, l_{it}, m_{it}, \ln P_{it}^{L}, S_{it}^{r}, z_{it}\right)', \tag{3.28}$$

where z_{it} contains observed firm characteristics. We approximate the unknown reduced-form function $\Psi(\cdot)$ using a third-order polynomial and estimate

$$y_{it} = \Psi_3(X_{it}) + \eta_{it}. \tag{3.29}$$

The regression residual provides an estimate of the transitory output disturbance, $\widehat{\eta}_{it}$. We then estimate

$$\widehat{\Lambda} = \frac{1}{N}\sum_{i,t} \exp(\widehat{\eta}_{it}), \tag{3.30}$$

and recover Hicks-neutral productivity from the production function as

$$\widehat{\omega}_{it} = y_{it} - \ln F\left(K_{it}, \exp\{\widehat{\widetilde{\varphi}}_{it}\}L_{it}, M_{it}; \widehat{\boldsymbol{\beta}}\right) - \widehat{\eta}_{it}. \tag{3.31}$$

Labor-augmenting productivity is recovered using the normalized proxy,

$$\widehat{\varphi}_{it} = q_{it}(\widehat{\boldsymbol{\vartheta}}) - \frac{1}{N}\sum_{i,t} q_{it}(\widehat{\boldsymbol{\vartheta}}). \tag{3.32}$$

Finally, inference is based on the heteroskedasticity- and autocorrelation-consistent covariance matrix of the joint GMM estimator. Let

$$G_N = \left.\frac{\partial \overline{g}_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}. \tag{3.33}$$

Under the standard regularity conditions for GMM,

$$\sqrt{N}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right) \xrightarrow{d} N\left(0, V_\theta\right), \tag{3.34}$$

where, under optimal weighting,

$$V_\theta = \left(G'\Omega^{-1}G\right)^{-1}. \tag{3.35}$$

Standard errors are computed using the corresponding sample HAC covariance matrix.

For nonlinear functions of the estimated parameters, including output elasticities, returns to scale, and the output decomposition, we obtain standard errors using the delta method. For a scalar function $h(\boldsymbol{\theta})$,

$$\widehat{\text{Var}}\left[h(\widehat{\boldsymbol{\theta}})\right] = \nabla_\theta h(\widehat{\boldsymbol{\theta}})'\widehat{V}_\theta \nabla_\theta h(\widehat{\boldsymbol{\theta}}). \tag{3.36}$$

For the reported mean output elasticities and returns to scale, these functions are evaluated at the sample means of the production inputs and normalized labor-augmenting productivity.

4 Data

Following Lee and Mas (2012), our sample data come from two sources: election information from the National Labor Relations Board (NLRB) and firm production information from the Compustat North America Dataset.

To obtain union information, we use the dataset from the NLRB, which publicly reports representation election vote details. This dataset includes the number of voters in favor of the union, the name of the firm, the city and state of the election, a two-digit industry code for the company, the month the NLRB opened and closed the election, and the number of eligible voters. We use the 1962–2009 NLRB election data to identify firms with union representation elections and determine the outcomes of these elections.[2]

The firm-level production data used in this study are obtained from Compustat North America through WRDS. Compustat provides detailed financial information for publicly traded firms in the United States and Canada, including balance-sheet and income-statement information. For this paper, we focus on firms in the manufacturing sector.

[2]The 1962–2009 NLRB election dataset was cleaned by Henry Farber and is available in Sojourner (2013).

The merging process of the NLRB and Compustat data is complex because the NLRB file provides only the company name and the location of the election, while firm identifiers that can be directly linked to Compustat are unavailable. We therefore follow the matching procedure in Lee and Mas (2012) to link the two datasets. For firms with multiple election records, we retain the earliest election in the matched sample.

Our final sample includes U.S. manufacturing firms from 1980 to 2009 with more than five employees. Output (Y) is defined as the value of net sales. Labor (L) is measured by the number of employees, and capital (K) is defined as the value of property, plant, and equipment, less depreciation. Materials (M) are constructed as the cost of goods sold plus administrative and selling expenses, less depreciation and wage expenditures, where wage expenditures are calculated as the number of workers multiplied by the average industry wage, following Keller and Yeaple (2009). We use total assets as a proxy for firm size and the logarithm of total assets, denoted by Z, in the empirical model. We obtain the price indices from the NBER-CES Manufacturing Industry Database,[3] the World Bank, and the Bureau of Labor Statistics. For the union election data, we use information on the number of eligible voters, the total votes cast, the number of votes in favor of union representation, and the number of votes against union representation. Table 1 reports the summary statistics for the main variables used in the analysis.

Table 1: Data Summary Statistics

Variable	Mean	1st Q.	Median	3rd Q.	S.D.
Sales (Y)	489.43	30.78	122.59	453.25	991.97
Capital (K)	134.98	5.02	23.39	101.23	342.68
Labor (L)	3.34	0.25	0.97	3.21	6.86
Materials (M)	321.71	20.92	81.26	299.77	660.26
Total assets	476.49	25.12	98.23	388.96	1,140.61
Votes for union	20.90	0	0	12	73.25
Total votes	50.29	0	0	32	177.25
Vote percent	45.93	29.23	42.86	59.57	23.06

Notes: Sales, capital, materials, and total assets are measured in millions of current U.S. dollars. Total assets are used as a proxy for firm size, while their logarithm is used as Z in the empirical model. Labor is measured in thousands of employees. Votes for union representation and total votes are counts, while vote percent denotes the percentage of votes cast in favor of union representation.

[3]The NBER-CES dataset is available at `https://www.nber.org/research/data/nber-ces-manufacturing-industry-database`.

5 Results

This section presents the empirical results from applying the proposed model and estimation procedure to the matched NLRB–Compustat sample. We first report the estimated production-function parameters, productivity dynamics, and wage markdown. We then examine how unionization affects labor-augmenting and Hicks-neutral productivity and quantify their contributions to the resulting change in output. Next, we study the effects of unionization on other features of firm production, including returns to scale, input substitutability, markups, and wage markdowns.

5.1 Production Technology

Table 2 reports the estimated mean output elasticities of capital, labor, and materials, together with returns to scale. All estimates are statistically significant at the 1% level. Materials have the largest output elasticity, with an estimated mean of 0.568, followed by labor at 0.267 and capital at 0.143, highlighting the material-intensive nature of production in our sample. The estimated returns to scale are 0.979, indicating that the production technology is close to constant returns to scale on average.

Table 2: Estimates of Output Elasticities

	e_k	e_l	e_m	RTS
Mean	0.143^{***}	0.267^{***}	0.568^{***}	0.979^{***}
s.e.	(0.020)	(0.019)	(0.040)	(0.007)

Note: e_k, e_l, and e_m denote the mean output elasticities of capital, labor, and materials, respectively. RTS is the return to scale.

We next turn to the evolution of the productivity components. Figures 1 and 2 plot the annual averages of labor-augmenting and Hicks-neutral productivity, normalized relative to the first year, together with their 95% confidence intervals. Overall, labor-augmenting productivity shows a clear upward trend over the sample period. Before 2000, it increases gradually with some short-run fluctuations. After 2001, the increase becomes more pronounced and remains at relatively high levels throughout the second half of the 2000s.

In contrast, Hicks-neutral productivity exhibits a broadly declining trend over the sample period. It falls through much of the 1980s and early 1990s, followed by a partial recovery around the early

2000s. After 2002, it declines again and remains relatively low through the end of the sample period. Taken together, the two figures show distinct dynamics in the two productivity components: labor-augmenting productivity generally increases over time, whereas Hicks-neutral productivity generally declines.

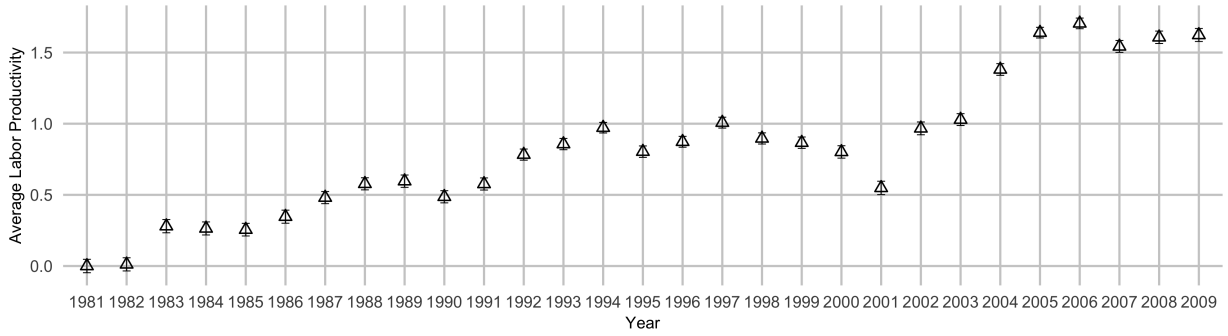


Figure 1: Labor-augmenting productivity with 95% CI

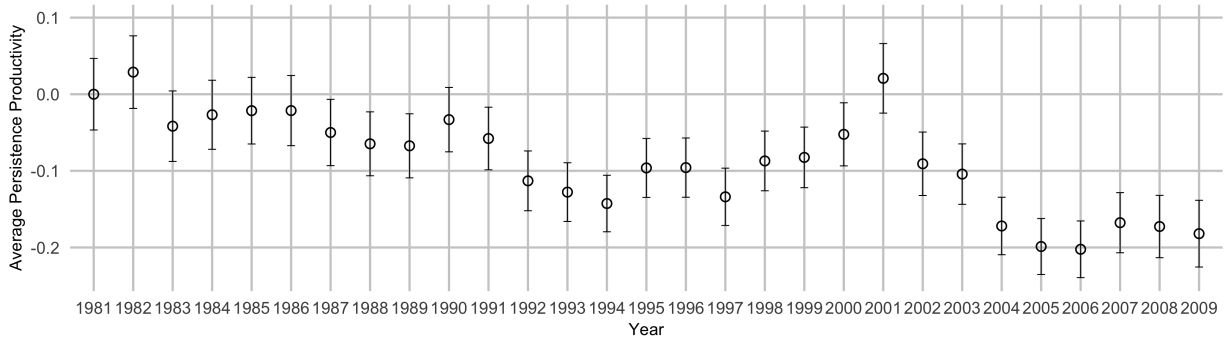


Figure 2: HN productivity with 95% CI

We now turn to the effects of unionization on firms' production technology and market behavior. Rather than focusing on a single measure of firm productivity, our framework allows us to examine several dimensions of firms' supply-side responses to unionization.

We first consider the effects on productivity. The coefficients $\alpha_{\varphi,2}$ and $\alpha_{\omega,2}$ capture the effects of unionization on labor-augmenting and Hicks-neutral productivity, respectively. Since both productivity components enter the production function in logs, these coefficients can be interpreted as semi-elasticities with respect to unionization. Specifically, φ_{it} and ω_{it} would change by $100 \times \alpha_{\varphi,2}\%$ and $100 \times \alpha_{\omega,2}\%$, respectively, if firm i becomes unionized in period t'. Table 3 reports estimates of 0.109 for $\alpha_{\varphi,2}$ and 0.074 for $\alpha_{\omega,2}$, both statistically significant at the 1% level. Thus, unionization

is associated with an approximately 10.9% increase in labor-augmenting productivity and a 7.4% increase in Hicks-neutral productivity.

These results indicate that unionization improves both the effectiveness of labor and broader firm-level productivity. One possible explanation is that unionization may operate partly through human-capital and workplace mechanisms. Higher wages, lower worker turnover, and greater incentives for firms and workers to invest in firm-specific skills may improve worker effort and labor effectiveness. At the same time, greater worker voice and more stable employment relationships may improve workplace organization and production practices, contributing to gains in Hicks-neutral productivity.

Table 3: Unionization Parameters

	$\alpha_{\varphi,2}$	$\alpha_{\omega,2}$
Point Estimates	0.109^{***}	0.074^{***}
s.e.	(0.037)	(0.011)

Note: *p<0.1; **p<0.05; ***p<0.01

However, the two productivity components affect firm output through different channels. Labor-augmenting productivity affects output through effective labor, whereas Hicks-neutral productivity enters the production function directly. Consequently, the output response to unionization can be decomposed into the following two components:

$$\frac{\Delta y_{it}}{\Delta d_{it'}} = \frac{\partial \ln F(\cdot)}{\partial \widetilde{l}_{it}} \frac{\Delta \varphi_{it}}{\Delta d_{it'}} + \frac{\Delta \omega_{it}}{\Delta d_{it'}} = \varepsilon_{it}^{L} \cdot \alpha_{\varphi,2} + \alpha_{\omega,2}, \tag{5.1}$$

where ε_{it}^{L} denotes the output elasticity of labor. The first term represents the factor-specific channel through which unionization affects output, while the second term represents the factor-neutral channel. The expression in Equation (5.1) is also a semi-elasticity, implying that output changes by $100 \times \frac{\Delta y_{it}}{\Delta d_{it'}}\%$ when firm i becomes unionized.

The results show that the estimated labor-augmenting productivity effect implies an approximately 2.9% increase in output. The Hicks-neutral productivity channel contributes an additional 7.4%, resulting in an overall output increase of approximately 10.3% following unionization. Thus, both productivity components contribute positively to the increase in output, with a larger contribution coming from the Hicks-neutral productivity channel.

Table 4: Output Decomposition

	$\frac{\Delta y_{it}}{\Delta d_{it'}}$	$\varepsilon_{it}^{L} \cdot \alpha_{\varphi,2}$	$\alpha_{\omega,2}$
Mean	0.103^{***}	0.029^{***}	0.074^{***}
s.e.	(0.005)	(0.01)	(0.011)

Note: *p<0.1; **p<0.05; ***p<0.01

Second, the above output decomposition also suggests that unionization may affect other aspects of firms' production technologies beyond productivity. A common measure of production technologies embedded in the production function is the returns to scale (RTS, denoted by ν), which is defined as the summation of input elasticities, i.e., $\nu_{it} = \varepsilon_{it}^{K} + \varepsilon_{it}^{L} + \varepsilon_{it}^{M}$. To examine the effect of unionization on returns to scale, we project ν_{it} on union support u_i and unionization status $d_{it'}$:

$$\nu_{it} = \phi_{\nu}(u_i) + \alpha_{\nu} d_{it'} + \epsilon_{\nu,it}, \tag{5.2}$$

where $\phi_{\nu}(\cdot)$ is a flexible function of union support, specified similarly to $\phi_{\varphi}(\cdot)$ and $\phi_{\omega}(\cdot)$, and α_{ν} captures the marginal effect of unionization. Under the translog specification, ν_{it} depends on both the production-function parameters and firms' input choices. In particular, it can be written as

$$\nu_{it} = \sum_{x \in \{k,l,m\}} \beta_x + \sum_{x,x' \in \{k,l,m\}} \beta_{xx'} x_{it} + \bar{\beta}_l \alpha_{\varphi,1} \varphi_{it-1} + \bar{\beta}_l \phi_{\varphi}(u_i) + \bar{\beta}_l \alpha_{\varphi,2} d_{it'} + \bar{\beta}_l \epsilon_{\varphi,it},$$

where $\bar{\beta}_l \equiv \sum_{x \in \{k,l,m\}} \beta_{lx}$. We emphasize that α_{ν} differs from $\bar{\beta}_l \alpha_{\varphi,2}$. The latter captures the conditional effect of unionization on returns to scale through labor-augmenting productivity while holding firms' input usage fixed, whereas α_{ν} also reflects firms' adjustments in input usage following unionization. We therefore focus on the estimate of α_{ν} in our analysis.

Table 5 reports the estimated effect of unionization on returns to scale. The coefficient α_{ν} is -0.041 and is statistically significant at the 1% level, indicating that unionization is associated with lower returns to scale in our sample. This result is distinct from the positive productivity effects documented above. While unionization increases the level of productivity, it also changes how output responds to proportional increases in all production inputs. In particular, the decline in returns to scale implies that, after unionization, a proportional increase in all inputs is associated with a smaller proportional increase in output.

Table 5: Unionization on RTS Parameters

	α_ν	$\phi_{\nu,1}$	$\phi_{\nu,2}$	$\phi_{\nu,3}$
Point Estimates	-0.041***	-0.444***	0.977***	-0.543***
s.e.	(0.001)	(0.0236)	(0.078)	(0.056)

Note: $\phi_{\nu,1}$, $\phi_{\nu,2}$, and $\phi_{\nu,3}$ are the estimated coefficients for u, u^2, and u^3, respectively. *p<0.1; **p<0.05; ***p<0.01

Third, we examine whether unionization affects firms' product-market power by recovering firm-level markups, defined as the ratio of output price to marginal cost. Following De Loecker and Warzynski (2012), markups are backed out from the first-order condition of the material input under the assumption of cost minimization. This allows us to quantify and compare the markup differences between unionized and non-unionized firms. Concretely, we have the firm-specific markup from

$$\mu_{it} \equiv \frac{P_{it}}{\lambda_{it}} = \frac{\varepsilon_{it}^M}{S_{it}^M} \frac{\Lambda}{e^{\eta_{it}}}. \tag{5.3}$$

By definition, Λ is estimated as $\widehat{\Lambda} = \frac{1}{N} \sum_{i,t} \exp(\widehat{\eta}_{it})$, where N is the total number of observations and $\widehat{\eta}_{it}$ is obtained from Section 3.2. Even if all firms face the same demand function—implying that unionization does not directly affect product demand—systematic differences in markups may still arise between unionized and non-unionized firms due to heterogeneity in marginal costs, which are shaped by their production technologies. To capture these differences, we project the estimated firm-level markups on union support u_i and unionization status $d_{it'}$ using the following regression model:

$$\mu_{it} = \phi_\mu(u_i) + \alpha_\mu d_{it'} + \epsilon_{\mu,it}, \tag{5.4}$$

where α_μ captures the marginal effect of unionization on markups.

Table 6 reports the estimated effect of unionization on firm markups. The coefficient α_μ is -0.043 and is statistically significant at the 5% level, indicating that unionization is associated with lower markups in our sample. Since the markup measures the ratio of output price to marginal cost, this result suggests that unionization reduces firms' product-market power. One possible explanation is that unionization raises firms' marginal costs, for example through higher labor costs, while firms may be unable to fully pass these costs on to output prices. We examine this possibility below by decomposing the estimated markup effect into its output-price and marginal-

cost components.

Table 6: Effect of Unionization on Markups

	α_μ	$\phi_{\mu,1}$	$\phi_{\mu,2}$	$\phi_{\mu,3}$
Point Estimates	-0.043***	-0.161***	0.564***	-0.332***
s.e.	(0.007)	(0.035)	(0.118)	(0.084)

Note: $\phi_{\mu,1}$, $\phi_{\mu,2}$, and $\phi_{\mu,3}$ are the estimated coefficients for u, u^2, and u^3, respectively. *p<0.1; **p<0.05; ***p<0.01

Since $\mu_{it} = \frac{P_{it}}{\lambda_{it}}$, the effect of unionization on markups can be decomposed as $\frac{\Delta\mu_{it}}{\Delta d_{it'}} = \frac{1}{\lambda_{it}}\frac{\Delta P_{it}}{\Delta d_{it'}} - \frac{\mu_{it}}{\lambda_{it}}\frac{\Delta\lambda_{it}}{\Delta d_{it'}}$. After estimating the production function, we can recover firms' marginal costs from the first-order condition as $\lambda_{it} = (1+\tau)P_{it}^L\left(\frac{\partial F(K_{it},\widetilde{L}_{it},M_{it})}{\partial\widetilde{L}}e^{\omega_{it}+\varphi_{it}}\Lambda\right)^{-1}$. Therefore, we can estimate the marginal cost responses to unionization ($\frac{\Delta\lambda_{it}}{\Delta d_{it'}}$) and assess whether market demand differs between unionized and non-unionized firms, i.e., whether $\frac{\Delta P_{it}}{\Delta d_{it'}} = 0$.

Table 7 reports the estimated effect of unionization on marginal costs. The coefficient α_λ is 0.027 and is statistically significant at the 1% level, indicating that unionization is associated with higher marginal costs. Table 8 then decomposes the estimated markup effect into its output-price and marginal-cost components. The overall markup effect is -0.043, while the output-price and marginal-cost components are 0.089 and 0.132, respectively. Thus, both output prices and marginal costs increase following unionization, but marginal costs increase by more. As a result, markups decline. One possible explanation is that unionization raises firms' production costs, particularly through higher labor costs, while firms are able to pass only part of these additional costs on to consumers through higher output prices. The decomposition is consistent with this partial cost pass-through: the positive output-price response offsets some, but not all, of the increase in marginal costs.

Table 7: Effect of Unionization on Marginal Costs

	α_λ	$\phi_{\lambda,1}$	$\phi_{\lambda,2}$	$\phi_{\lambda,3}$
Point Estimates	0.027**	-0.170***	-0.095	0.091
s.e.	(0.011)	(0.062)	(0.205)	(0.147)

Note: $\phi_{\lambda,1}$, $\phi_{\lambda,2}$, and $\phi_{\lambda,3}$ are the estimated coefficients for u, u^2, and u^3, respectively. *p<0.1; **p<0.05; ***p<0.01

Table 8: Unionization on MC Decomposition

	$\frac{\Delta\mu_{it}}{\Delta d_{it'}}$	$\frac{1}{\lambda_{it}}\frac{\Delta P_{it}}{\Delta d_{it'}}$	$\frac{\mu_{it}}{\lambda_{it}}\frac{\Delta\lambda_{it}}{\Delta d_{it'}}$
Mean	-0.043	0.089	0.132

Note: $\frac{1}{\lambda_{it}}\frac{\Delta P_{it}}{\Delta d_{it'}}$ is market demand decomposition.
*p<0.1; **p<0.05; ***p<0.01

Fourth, we examine whether unionization reduces firms' monopsony power in the labor market through the wage markdown. In Equation 3.3, τ represents the average wage markdown across all firms. To examine its heterogeneity between unionized and non-unionized firms, we note that taking the ratio between Equations 3.3 and 3.2 yields $(1+\tau) = \frac{\varepsilon_{it}^{L} S_{it}^{M}}{\varepsilon_{it}^{M} S_{it}^{L}}$. We then project the right-hand side on the unionization dummy as

$$\frac{\varepsilon_{it}^{L} S_{it}^{M}}{\varepsilon_{it}^{M} S_{it}^{L}} = \phi_{\tau}(u_i) + \alpha_{\tau} d_{it'} + \epsilon_{\tau,it}, \tag{5.5}$$

where the effect of unionization on the wage markdown is captured by α_{τ}.

Table 9 reports the estimated effect of unionization on the wage markdown. The coefficient α_{τ} is -0.123 and is statistically significant at the 1% level, indicating that unionization substantially reduces the wage markdown. Since a larger wage markdown reflects a larger gap between workers' marginal revenue product and their wages, the negative estimate suggests that unionization reduces firms' labor-market power and allows workers to capture a larger share of their marginal contribution to production. Taken together, the markup and markdown results indicate that unionization reduces firms' market power in both product and labor markets.

Table 9: Effect of Unionization on Wage Markdowns

	α_{τ}	$\phi_{\tau,1}$	$\phi_{\tau,2}$	$\phi_{\tau,3}$
Point Estimates	-0.123***	0.130	0.785	-0.415
s.e.	(0.0282)	(0.151)	(0.502)	(0.360)

Note: $\phi_{\tau,1}$, $\phi_{\tau,2}$ and $\phi_{\tau,3}$ are the estimated coefficients for u, u^2, and u^3, respectively. *p<0.1; **p<0.05; ***p<0.01

Finally, unionization may alter firms' decisions regarding input composition in the production process. Higher labor costs may create incentives for firms to substitute capital for labor, while employment protection and collective bargaining may work in the opposite direction. To examine

these potentially offsetting effects, we first estimate the following reduced-form regression:

$$\frac{L}{K} = \phi_l(u_i) + \alpha_l d_{it'} + \epsilon_{l,it}, \tag{5.6}$$

where α_l captures the effect of unionization on the labor-to-capital ratio. Table 10 reports the estimated effect of unionization on the labor-to-capital ratio. The coefficient α_l is -0.002 and is not statistically significant, providing little evidence that unionization leads to a systematic change in the relative use of labor and capital.

Table 10: Effect of Unionization on the Labor-to-Capital Ratio

	α_l	$\phi_{l,1}$	$\phi_{l,2}$	$\phi_{l,3}$
Point Estimates	-0.002	-0.129**	0.279***	-0.175**
s.e.	(0.002)	(0.012)	(0.039)	(0.028)

Note: $\phi_{l,1}$, $\phi_{l,2}$, and $\phi_{l,3}$ are the estimated coefficients for u, u^2, and u^3, respectively. *p<0.1; **p<0.05; ***p<0.01

The labor-to-capital ratio, however, reflects both firms' input choices and the underlying substitution possibilities embedded in the production technology. We therefore also examine the elasticity of substitution (ES) between capital and labor implied by the estimated production function. Specifically, we use the Allen-Uzawa elasticity of substitution, a commonly used measure in empirical work. For a production function $F(X_1, \ldots, X_J)$, the Allen-Uzawa ES between inputs X_p and X_q is defined as

$$\text{ES}_{pq} = \frac{\sum_s X_s F_s}{X_p X_q} \frac{\text{C}_{pq}}{|\overline{\text{H}}|}, \tag{5.7}$$

where $F_s = \frac{\partial}{\partial X_s} F(\cdot)$ is the marginal product of X_s, $\overline{\text{H}}$ is the bordered Hessian matrix, and C_{pq} is the cofactor of $\overline{\text{H}}$ corresponding to X_p and X_q. A negative value indicates that the two inputs are gross complements, whereas a positive value indicates gross substitutes. The measure is symmetric, so the elasticity of substitution between K and L is identical to that between L and K. Moreover, since the ES is scale-invariant, the ES of any input with L is identical to that with the productivity-adjusted input $\widetilde{L}$.

We then project the estimated elasticity of substitution between capital and labor, ES_{kl}, on

union support and unionization status:

$$\text{ES}_{kl} = \phi_e(u_i) + \alpha_e d_{it'} + \epsilon_{e,it}, \tag{5.8}$$

where α_e captures the effect of unionization on the substitutability between capital and labor.

Table 11 reports the estimated effect of unionization on the elasticity of substitution between capital and labor. The coefficient α_e is -0.376 and is statistically significant at the 1% level, indicating that unionization reduces the substitutability between capital and labor. In other words, capital and labor become less easily substituted for one another following unionization. The contrast with the insignificant effect on the labor-to-capital ratio suggests that unionization changes the underlying substitution possibilities in the production technology even though firms do not exhibit a large change in their observed input mix.

Table 11: Effect of Unionization on the Elasticity of Substitution

	α_e	$\phi_{e,1}$	$\phi_{e,2}$	$\phi_{e,3}$
Point Estimates	-0.376***	-2.388***	6.484***	-4.296**
s.e.	(0.132)	(0.712)	(2.367)	(1.697)

Note: $\phi_{e,1}$, $\phi_{e,2}$, and $\phi_{e,3}$ are the estimated coefficients for u, u^2, and u^3, respectively. *p<0.1; **p<0.05; ***p<0.01

Overall, the results show that unionization affects firms along several dimensions of production and market behavior. Unionization increases both labor-augmenting and Hicks-neutral productivity, while also changing returns to scale and the substitutability between capital and labor. At the same time, it reduces both product-market markups and labor-market wage markdowns. These findings suggest that the effects of unionization cannot be characterized solely as a transfer of rents from firms to workers. In our setting, stronger worker bargaining power is accompanied by productivity gains and weaker firm market power, although firms also face higher marginal costs and adjust their production technology in response.

From a policy perspective, these results suggest that labor-market institutions may have effects beyond wages and employment. Policies that strengthen workers' bargaining power may also affect firms' production decisions, their ability to exercise market power, and the extent to which cost increases are passed through to consumers. More broadly, evaluating labor policy solely through

its effects on labor costs may miss important adjustments in productivity, production technology, and market power.

6 Conclusion

This paper develops a structural framework to study how unionization affects firms' production technology and market behavior from a supply-side perspective. By distinguishing between labor-augmenting and Hicks-neutral productivity within a flexible translog production function, we move beyond conventional approaches that summarize the effect of unionization using a single measure of overall productivity. We further combine the structural production framework with the regression discontinuity generated by close union representation elections, allowing the effects of unionization to enter directly into the evolution of multidimensional productivity. Using matched NLRB election records and Compustat data for U.S. manufacturing firms, the framework also allows us to examine changes in returns to scale, input substitutability, product-market markups, marginal costs, and labor-market wage markdowns within a common setting. In this way, the paper contributes to the unionization literature by separating the channels through which collective bargaining affects firm productivity and production, and to the production-function literature by incorporating quasi-experimental variation directly into a multidimensional productivity framework.

Empirically, we find that unionization increases both labor-augmenting and Hicks-neutral productivity. The implied increase in output is approximately 10.3%, with both productivity components contributing positively to the overall effect. Unionization also changes the structure of production: returns to scale decline, and capital and labor become less easily substitutable, although the observed labor-to-capital ratio does not change significantly. At the same time, unionization reduces both product-market markups and labor-market wage markdowns. The markup decomposition shows that both marginal costs and output prices increase following unionization, but marginal costs increase by more, resulting in lower markups. Taken together, these results show that stronger worker bargaining power can coexist with productivity gains while also reshaping firms' production technology and reducing their market power.

From a policy perspective, our findings suggest that the effects of labor-market institutions should be evaluated along a broader set of margins than wages and employment alone. Stronger

worker bargaining power can affect not only labor costs, but also firm productivity, production technology, product-market power, and the pass-through of costs to consumers. In our setting, the productivity gains associated with unionization coexist with lower wage markdowns and lower product-market markups, suggesting that stronger worker bargaining power need not come at the expense of productive efficiency. At the same time, the increase in marginal costs and the corresponding rise in output prices indicate that part of the adjustment is borne by consumers, while the remaining increase in costs is absorbed through lower firm markups. These results highlight that the incidence of unionization is shared across workers, firms, and consumers, and that labor policy can have implications for both labor-market outcomes and product-market performance.

References

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

Addison, J. T. and Hirsch, B. T. (1989). Union effects on productivity, profits, and growth: has the long run arrived? *Journal of Labor Economics*, 7(1):72–105.

Barth, E., Bryson, A., and Dale-Olsen, H. (2020). Union density effects on productivity and wages. *The Economic Journal*, 130(631):1898–1936.

Bradley, D., Kim, I., and Tian, X. (2017). Do unions affect innovation? *Management Science*, 63(7):2251–2271.

Clark, K. B. (1980). Unionization and productivity: Micro-econometric evidence. *The Quarterly Journal of Economics*, 95(4):613–639.

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

Demirer, M. (2022). Production function estimation with factor-augmenting technology: An application to markups. *MIT Working Paper*.

DiNardo, J. and Lee, D. S. (2004). Economic impacts of new unionization on private sector employers: 1984–2001. *The quarterly journal of economics*, 119(4):1383–1441.

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

Fallick, B. C. and Hassett, K. A. (1999). Investment and union certification. *Journal of labor Economics*, 17(3):570–582.

Farber, H. S. (2005). Union membership in the united states: The divergence between the public and private sectors. Technical report.

Frandsen, B. R. (2021). The surprising impacts of unionization: Evidence from matched employer-employee data. *Journal of Labor Economics*, 39(4):861–894.

Freeman, R. B. and Medoff, J. L. (1984). What do unions do. *Industrial & Labor Relations Review*, 38:244.

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

Jaumandreu, J. and Mullens, I. (2024). Technology and us manufacturing employment (lessons from the estimation of a microeconomic multiproductivity production function). *Boston University*.

Keller, W. and Yeaple, S. R. (2009). Multinational enterprises, international trade, and productivity growth: firm-level evidence from the united states. *The review of economics and statistics*, 91(4):821–831.

Lee, D. S. and Mas, A. (2012). Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999. *The Quarterly Journal of Economics*, 127(1):333–378.

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

Malikov, E., Zhao, S., and Zhang, J. (2024). A system approach to structural identification of production functions with multi-dimensional productivity. *Advances in Econometrics*. forthcoming.

Morikawa, M. (2010). Labor unions and productivity: An empirical analysis using japanese firm-level data. *Labour Economics*, 17(6):1030–1037.

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

Sojourner, A. (2013). Effects of unionization on workplace safety: Regression discontinuity evidence from osha's enforcement data. *The Department of Labor 2013-2014 Scholars Program*, page 230.

Visser, J. (2006). Union membership statistics in 24 countries. *Monthly Lab. Rev.*, 129:38.

Yang, C.-H. and Tsou, M.-W. (2018). Labour unions and firm productivity: Evidence from china. *The Manchester School*, 86(6):699–721.

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