A Friction–Robust Estimator of Capital–Labor Substitution

Lieyuan Yang* Yutong Zhong†‡

November 2, 2026

Abstract

The elasticity of substitution between capital and labor is important for long-run economic growth. Conventional estimators infer this elasticity from firms' input responses to relative factor prices. Capital frictions can distort these responses and bias estimates even with exogenous instruments. We develop an estimation framework that is robust to these frictions by exploiting firms' optimal labor choices to jointly identify capital–labor substitution and the demand elasticity. Monte Carlo experiments show that our estimators recover the true parameters in economies with capital frictions and policy distortions. We apply the framework to China's 2004 Industrial Census. Compared with conventional methods, our estimator yields greater capital–labor substitutability and lower markups and produces more stable estimates across firm groups with different financial conditions.

Keywords: Elasticity of substitution, capital frictions, financial constraints, labor share, markup estimation.

JEL codes: D24, E22, E23, E25

*Renmin University of China, No. 59, Zhongguancun Street, Haidian District, Beijing. Email: lieyuanyang@ruc.edu.cn.

†University of Minnesota, 1925 Fourth Street South, Minneapolis, MN 55455. Email: zhong462@umn.edu.

‡Acknowledgment: We thank Anmol Bhandari, Hannes Malmberg, and Kjetil Storesletten for their valuable feedback during presentations as we developed the project. We also thank Adam Oppenheimer, Kurt See, and Yu Zhu for the insights gained during discussions of the project.

1 Introduction

The elasticity of substitution between capital and labor is a fundamental production parameter with important implications for long-run economic growth, employment, the labor share, and social inequality (Uzawa, 1961; Karabarbounis and Neiman, 2014; Piketty, 2014; Grossman et al., 2017). However, empirical estimates differ substantially. Many micro-level estimates are below one, whereas some macro-level estimates exceed one (Karabarbounis and Neiman, 2014). Oberfield and Raval (2021) address this micro–macro discrepancy by deriving an aggregate elasticity that incorporates both capital–labor substitution within plants and reallocation across plants, governed by substitution across goods varieties. Their aggregate estimate remains below one.

Most studies estimate plant-level capital–labor substitution from firms' input responses to relative factor price changes, abstracting from capital frictions. Adjustment costs and financial constraints, however, can distort these responses, leading conventional estimators to attribute the resulting input responses to technological substitutability. Capital frictions can also affect markups inferred from the revenue–cost ratio and, consequently, estimates of substitution across goods varieties. When the shadow cost of capital exceeds its measured user cost, measured input costs understate effective production costs. The revenue–cost ratio therefore overstates the markup and understates the implied demand elasticity. Together, these channels can bias estimates of aggregate capital–labor substitution constructed using the framework of Oberfield and Raval (2021).

To address these concerns, we develop an estimation framework that accommodates capital frictions. We characterize the bias in conventional estimates and derive alternative moment conditions from firms' optimal labor choices, CES production technology, and product demand. These moments jointly identify the elasticity of substitution between capital and labor and the demand elasticity without imposing a frictionless capital optimality condition.

This paper starts by showing that firms' relative input responses to relative factor prices do not generally identify the elasticity of substitution between capital and labor in the presence of capital frictions. In a parsimonious firm dynamic model with capital adjustment costs, we represent the resulting deviations from the frictionless capital op-

timality condition by a wedge between the markup-adjusted marginal revenue product of capital and its frictionless user cost. This wedge responds endogenously to factor-price shocks. For example, a decline in investment-good prices lowers the user cost of capital and encourages investment, but adjustment costs can limit capital expansion, increasing the wedge. Consequently, an instrument exogenous to technology can be correlated with the omitted capital wedge, biasing conventional instrumental variable (IV) estimates of capital–labor substitution.

Within the same framework, we develop an alternative identification strategy that exploits firms' optimal labor choices, conditional on their realized capital stocks. Under CES production, the relationship between the labor share and capital intensity identifies capital–labor substitution when the markup is known. Holding technology fixed, greater capital intensity raises the labor share when capital and labor are gross complements: additional capital makes labor a tighter production bottleneck and thereby raises the labor share. When capital and labor are gross substitutes, greater capital intensity instead reduces labor's relative contribution to output and lowers its share. To identify the markup jointly with substitution, we combine labor optimality with product demand to derive a revenue equation in employment and the labor share. Conditional on technology and the labor share, output rises proportionally with employment. Under monopolistic competition, firms must lower their prices to sell the additional output, so revenue rises less than proportionally. The magnitude of this revenue response identifies the demand elasticity. With valid instruments and sufficient identifying variation, the two relationships jointly identify both parameters. The main novelty of our framework is that we rely only on firms' labor optimal choices without imposing a frictionless capital optimality condition. This makes our estimators robust to frictions that distort firms' capital choice.

To assess how capital frictions affect conventional estimation and to evaluate the performance of our proposed estimators, we conduct a series of Monte Carlo experiments. We simulate a heterogeneous-firm dynamic model with factor-biased technological change, convex and fixed capital adjustment costs, convex financing costs, a collateral constraint, and distortionary taxation. We introduce two exogenous policy shocks that lower the cost of capital relative to labor: a payroll tax increase and the

elimination of the value-added tax (VAT) on equipment purchases. The latter is motivated by China's 2009 VAT reform. In the frictionless benchmark, both conventional and proposed estimators recover the true parameters. When capital frictions are introduced, the markup inferred from the revenue–cost ratio and the conventional IV estimates of capital–labor substitution are biased, whereas our proposed estimators remain close to the true parameters. We also explore the sensitivity of our estimators to measurement error.

To compare our estimators with conventional methods in an empirical setting, we apply our framework to China's 2004 Industrial Census (CIC), using local industry employment shares in 1990 as instruments. The conventional IV estimate of capital–labor substitution is 0.51, compared with our generalized method of moments (GMM) estimate of 0.67. Our estimated markup is 1.26, substantially below the revenue–cost estimate of 2.87, implying a higher demand elasticity. These differences suggest that conventional methods that abstract from capital frictions may understate both capital–labor substitution and substitution across goods varieties.

Our framework explicitly abstracts from labor market frictions and assumes that firms can adjust labor freely. However, labor market power, firing costs, and search and matching frictions can distort the labor optimality condition and potentially bias our estimates. To examine the sensitivity of our results to labor market conditions, we control for local employment concentration, measured by the Herfindahl–Hirschman index, the employment share of state-owned enterprises (SOE), and each firm's share of local employment. Our GMM estimates remain largely unchanged after including these controls.

To assess whether our estimator is less sensitive to capital frictions than conventional methods, we compare estimates across firms grouped by age, ownership, and leverage, which serve as proxies for access to external finance. Conventional IV yields lower elasticities for young firms, non-SOEs, and low-leverage firms—groups that may face more limited access to external finance. By contrast, our GMM estimator delivers much more similar estimates across these groups. These patterns are consistent with our argument that conventional estimates are sensitive to capital frictions, whereas our proposed estimator is more robust to these frictions.

Related Literature. Our paper contributes to several strands of literature. First, we contribute to the estimation of capital–labor substitution, pioneered by Arrow et al. (1961). Recent studies use microdata and exogenous variation in relative factor prices to identify this technological parameter. Raval (2019) exploits local wage variation, using Bartik-style shift-share instruments and local amenities as IV to estimate plant-level substitution elasticities. Other studies exploit policy-induced changes in factor costs, including minimum-wage increases (Harasztosi and Lindner, 2019), U.S. bonus depreciation (Curtis et al., 2023), and China's VAT reform (Meng, Zhang and Zhang, 2024). The latter two studies use structural factor-demand models to distinguish substitution from changes in production scale. Most micro-level estimates are below one, implying that capital and labor are gross complements.

In contrast, Karabarbounis and Neiman (2014) provide macro evidence that capital and labor are gross substitutes. Using country-level data, they estimate the elasticity from the relationship between long-run changes in relative investment-good prices and labor shares. Countries experiencing larger declines in investment-good prices also experienced larger declines in labor shares, yielding an estimated elasticity of approximately 1.25. Their findings contrast with the predominantly below-one estimates from microdata.

Oberfield and Raval (2021) address this micro–macro discrepancy by developing a framework that links plant-level estimates to aggregate substitution. Their framework combines substitution within plants with reallocation across plants, governed by substitution across goods varieties and heterogeneity in capital intensities. Using U.S. manufacturing data, they estimate plant-level elasticities of 0.3–0.5 and aggregate elasticities of 0.5–0.7. Thus, incorporating reallocation raises the estimated elasticity, but their aggregate estimates remain below one.

These micro-level price-based approaches infer technological substitution from firms' input responses under assumptions that abstract from capital adjustment frictions. At the plant level, adjustment costs, fixed investment costs, and irreversibility shape investment decisions and capital dynamics (Cooper and Haltiwanger, 2006; Khan and Thomas, 2008; Asker, Collard-Wexler and De Loecker, 2014). Bloom, Bond and Van Reenen (2007) show that uncertainty can reduce investment responsiveness in the presence of

partial irreversibility. Borrowing constraints, limited access to external finance, and collateral requirements can also prevent otherwise profitable investments (Kiyotaki and Moore, 1997; Clementi and Hopenhayn, 2006; Chaney, Sraer and Thesmar, 2012). Consequently, firms' capital–labor responses reflect both technological substitutability and constraints on capital adjustment.

Across plants, a large literature documents capital misallocation following Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Asker, Collard-Wexler and De Loecker (2014) show that productivity volatility and capital adjustment costs can explain a substantial fraction of the observed capital misallocation pattern. Midrigan and Xu (2014) and Gopinath et al. (2017) examine how financial frictions shape capital allocation. These studies show how constraints on firm expansion can affect capital reallocation across producers.

Moreover, policies used to identify capital–labor substitution may themselves change capital frictions. Chen et al. (2023) show that China's 2009 VAT reform reduced not only the purchase cost of equipment but also partial irreversibility by narrowing the gap between the purchase and resale prices of capital. By making capital easier to adjust, the reform increased investment largely through more frequent investment spikes. Similarly, Zwick and Mahon (2017) find stronger investment responses to bonus depreciation when the policy generates immediate cash-flow benefits, consistent with financial frictions shaping firms' responses. Thus, these reforms change both factor costs and the frictions governing capital adjustment, complicating the identification of technological substitution.

Our work contributes to these literatures by developing a novel estimation framework that jointly identifies the elasticity of substitution between capital and labor and the demand elasticity while allowing for capital frictions. The key innovation is to exploit firms' optimal labor choices conditional on realized capital stocks, without imposing a frictionless capital optimality condition. To identify the demand elasticity jointly with capital–labor substitution, we additionally use a revenue relationship implied by CES production and product demand. Under valid exclusion restrictions and appropriate identification conditions, the resulting moments identify both elasticities.

Moreover, our paper broadly relates to the macroeconomic and policy implications

of capital–labor substitution. The contrasting estimates discussed above imply different explanations for declining labor shares: Karabarbounis and Neiman (2014) emphasize falling investment-good prices, whereas Oberfield and Raval (2021) attribute the decline in U.S. manufacturing primarily to biased technical change. Beyond factor shares, Grossman et al. (2017) show how endogenous schooling can reconcile capital-augmenting technological progress with balanced growth when capital and labor are gross complements. Krusell et al. (2000) examine how capital–skill complementarity affects the skill premium and wage inequality, while Acemoglu, Manera and Restrepo (2020) study how the relative taxation of capital and labor influences automation and employment. These studies illustrate the importance of credible substitution estimates, and of accounting for reallocation when translating plant-level estimates into aggregate elasticities, for understanding macroeconomic outcomes and evaluating policy.

Finally, although markup estimation is not our main focus, our paper also relates to the literature on estimating markups using production-side information (De Loecker and Warzynski, 2012; Bond et al., 2021; Raval, 2023).

The remainder of the paper is organized as follows. In Section 2, we provide a parsimonious firm dynamic model with capital adjustment costs, where we characterize the bias of conventional estimators and propose our identification strategy. In Section 3, we perform several Monte Carlo experiments and compare our estimator with conventional methods in different model settings. In Section 4, we apply our method to the Chinese Industrial Census and conduct a series of robustness checks. Section 5 concludes.

2 A Framework for Identifying Capital–Labor Substitution with Capital Friction

In Oberfield and Raval (2021), the aggregate elasticity of substitution between capital and labor is given by

$$\sigma^{agg} = \chi\sigma + (1-\chi)\varepsilon$$

where σ is the plant-level elasticity of substitution between capital and labor to capture how firms replace labor by capital when the relative input price changes, and ε is the

elasticity of substitution between goods varieties to capture capital reallocation towards capital-intensive plants. They infer σ by exploiting the relationship between the input cost ratio and the relative price and infer ε through the average revenue-cost ratio.

In this section, we propose a parsimonious firm dynamic model in the presence of nonlinear capital adjustment costs. The aim is twofold. First, we show that the estimation of σ and ε is biased in Oberfield and Raval (2021) when there are capital frictions. We are able to characterize the bias under specific assumptions. Second, we provide our friction-robust estimator for both parameters.

2.1 Environment

Competitive final-good producers assemble a continuum of intermediate inputs $i \in [0, 1]$ into a final good Y_t at a price P_t according to:

$$Y_t = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}},$$

where y_{it} denotes the quantity of input i and $\varepsilon > 1$ is the elasticity of substitution between input varieties. Final-good producers purchase intermediate inputs at prices p_{it}. Cost minimization implies that the demand function for input variety i is

$$y_{it} = \left(\frac{p_{it}}{P_t} \right)^{-\varepsilon} Y_t. \tag{1}$$

Intermediate producers are monopolistically competitive and charge a constant markup $\mu = \varepsilon/(\varepsilon - 1)$. The final good is used for consumption and investment. We normalize the price of the final consumption good to one, $P_t = 1$, and let ξ_t denote the effective price of investment goods in units of the final good.

For convenience, we omit the subscript i and consider a representative intermediate producer. Intermediate output is produced using capital k_t and labor n_t according to the CES production function:

$$y_t = F(k_t, n_t) = \left[\alpha (A_{k,t} k_t)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(A_{n,t} n_t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{2}$$

where σ is the elasticity of substitution between capital and labor, and α is a distri-

bution parameter. The terms $A_{k,t}$ and $A_{n,t}$ denote the capital-augmenting and labor-augmenting technology, respectively. Capital and labor are gross substitutes when $\sigma > 1$ and gross complements when $\sigma < 1$. When $\sigma \to 1$, the production function is Cobb-Douglas production function, $F(k_t, n_t) = (A_{kt}k_t)^{\alpha}(A_{nt}n_t)^{1-\alpha}$.

We next state the assumptions on input timing and capital adjustment costs used in the analysis.

Assumption 1 (Timing of input choice)**.** *Firms accumulate capital according to*

$$k_t = \kappa(k_{t-1}, x_{t-1}),$$

where investment x_{t-1} is chosen in period $t-1$. Labor input n_t is chosen at t and is freely adjustable.

In each period, an intermediate producer chooses labor n_t, next-period capital k_{t+1}, price p_t, and dividends d_t to maximize the present value of dividends. Current capital k_t is predetermined by past investment, whereas labor n_t is chosen flexibly within the period. Our framework allows for frictions in capital adjustment but assumes that the firm's static labor optimality condition is undistorted. A time-varying labor wedge correlated with the policy shock would enter the estimating equation and invalidate the proposed identification strategy.

Assumption 2 (Capital adjustment frictions)**.** *A firm incurs a cost $\Phi(k_t, k_{t+1})$ when adjusting its capital stock from k_t to k_{t+1}.*

Given its initial capital stock k_0, the intermediate producer chooses sequences of dividends, prices, labor, and future capital to maximize the present value of dividends:

$$\max_{\{d_t, p_t, n_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \frac{d_t}{\prod_{s=1}^{t}(1+r_s)}$$

subject to the budget constraint:

$$d_t + \xi_t x_t + \Phi(k_t, k_{t+1}) = p_t y_t - w_t n_t,$$

the capital accumulation equation $k_{t+1} = x_t + (1-\delta)k_t$, the demand function (1), and

the production function (2). r_s is the risk-free interest rate in period s.

Define the frictionless rental rate of capital as:

$$R_{t+1} = (1 + r_{t+1})\xi_t - \xi_{t+1}(1 - \delta).$$

The optimality condition for labor and the Euler equation for capital are, respectively,

$$\frac{p_t F_n(k_t, n_t)}{\mu} = w_t, \tag{3}$$

and

$$\begin{aligned} \frac{p_{t+1} F_k(k_{t+1}, n_{t+1})}{\mu} &= (1 + r_{t+1})[\xi_t + \Phi_2(k_t, k_{t+1})] - \xi_{t+1}(1 - \delta) + \Phi_1(k_{t+1}, k_{t+2}) \\ &= R_{t+1}(1 + \tau_{t+1}). \end{aligned} \tag{4}$$

where $(1 + \tau)$ is the wedge between the firm's effective capital cost and risk free rental rate R, following the wedge representation in Hsieh and Klenow (2009). Here, the wedge is generated by the firm's dynamic adjustment costs.

2.2 Input Response to Relative Price Changes

We first characterize how capital frictions affect input responses to relative factor prices. Define the output elasticities of labor and capital as

$$e_n = \frac{F_n n}{y}, \qquad e_k = \frac{F_k k}{y}, \qquad e_n + e_k = 1,$$

where the last equality follows from constant returns to scale. In the local comparative-static analysis, we hold factor-augmenting technology and the demand shifter $P_t^{\varepsilon} Y_t$ fixed. We present the main results here and leave the derivations to Appendix A.

Consider local responses around a common initial allocation satisfying the frictionless capital optimality condition, $pF_k/\mu = R$. Let $d \log \tilde{k}$ denote the capital response when this condition continues to hold following the shock. The frictionless capital response satisfies

$$d \log \tilde{k} = e_n(\sigma - \varepsilon) d \log w - (\sigma e_n + \varepsilon e_k) d \log R. \tag{5}$$

This expression reflects both substitution and scale effects. A wage increase encourages substitution toward capital, contributing $e_n \sigma \, d\log w$ to the capital response. It also raises production costs and reduces demand, contributing $-e_n \varepsilon \, d\log w$. When $\sigma < \varepsilon$, the scale effect dominates and frictionless capital declines; when $\sigma > \varepsilon$, the substitution effect dominates and frictionless capital increases. By contrast, an increase in the user cost of capital reduces capital through both channels: firms substitute toward labor, while higher production costs reduce demand and production.

The change in the capital–labor ratio can be decomposed as:

$$d\log \frac{k}{n} = \underbrace{\sigma \, d\log \frac{w}{R}}_{\text{a. technological substitution}} + \underbrace{\frac{d\log k - d\log \tilde{k}}{e_n + (\varepsilon/\sigma) e_k}}_{\text{b. deviation from the frictionless capital response}} . \tag{6}$$

The first term captures the response implied by technological substitution. The second captures the additional response arising from deviations of actual capital adjustment from the frictionless benchmark. This decomposition relates the resulting adjustment to the bias in conventional estimation.

To illustrate the direction of bias when frictions dampen capital adjustment, suppose actual and frictionless capital responses have the same sign, with $|d\log k| < |d\log \tilde{k}|$. We consider two shocks: an increase in wages and a decrease in the user cost of capital. Both increase w/R, but they generate different benchmark capital responses and can therefore produce biases in opposite directions. Holding R fixed, a wage increase reduces frictionless capital when $\sigma < \varepsilon$. Under the assumed attenuation of capital adjustment, $d\log \tilde{k} < d\log k < 0$. Term b is therefore positive: the capital–labor ratio rises more than in the frictionless benchmark, generating an upward bias in the conventional elasticity estimate. By contrast, holding wages fixed, a decline in the user cost of capital increases frictionless capital through both substitution and scale effects. Under the same attenuation assumption, $d\log \tilde{k} > d\log k > 0$. Term b is therefore negative: the capital–labor ratio rises less than in the frictionless benchmark, generating a downward bias in the conventional elasticity estimate.

We now characterize the conventional IV estimator. Let Z denote a policy instrument that shifts the relative factor price w/R. We impose the following exclusion and relevance conditions.

Assumption 3 (Exclusion restriction)**.** *The instrument is uncorrelated with factor-augmenting technology:*

$$\operatorname{cov}(\log A_k, Z) = \operatorname{cov}(\log A_n, Z) = 0.$$

Assumption 4 (Relevance condition)**.** *The instrument predicts relative factor prices:*

$$\operatorname{cov}\left(\log\left(\frac{w}{R}\right), Z\right) \neq 0.$$

The technology exclusion condition does not restrict the response of the capital wedge to the instrument. By changing relative factor prices, the policy shock alters firms' desired capital stocks and investment decisions. The resulting changes in the capital adjustment path affect current and future marginal adjustment costs, $\Phi_2(k_t, k_{t+1})$ and $\Phi_1(k_{t+1}, k_{t+2})$, and can therefore change the capital wedge. If the policy also changes the frictionless user cost R, it additionally changes the denominator of $1+\tau$. By construction, we have

$$d\log(1+\tau) = -\frac{d\log k - d\log \tilde{k}}{\sigma e_n + \varepsilon e_k}.$$

Combining the capital and labor optimality conditions (3) and (4), we have

$$\log\left(\frac{k}{n}\right) = \sigma \log\left(\frac{w}{R}\right) - \sigma\log(1+\tau) + (\sigma-1)\log\left(\frac{A_k}{A_n}\right) + \sigma\log\frac{\alpha}{1-\alpha}.$$

Therefore, the IV estimator from a regression of the log capital–labor ratio on the log relative factor price, using Z as the instrument, satisfies

$$\operatorname{plim}\hat{\sigma}^p = \frac{\operatorname{cov}\left(\log\left(\frac{k}{n}\right), Z\right)}{\operatorname{cov}\left(\log\left(\frac{w}{R}\right), Z\right)} = \sigma \underbrace{-\sigma\frac{\operatorname{cov}(\log(1+\tau), Z)}{\operatorname{cov}\left(\log\left(\frac{w}{R}\right), Z\right)}}_{\text{bias}}.$$

Consequently, the conventional IV estimator is inconsistent whenever the instrument is correlated with the capital wedge. Exogeneity with respect to technology alone is

therefore insufficient for identification in the presence of capital frictions.

2.3 Friction-Robust Estimator

Now, we propose our friction-robust estimator. Define the labor share as

$$s = \frac{wn}{py}.$$

The labor optimality condition (3) implies:

$$\mu s = \frac{\mu w n}{py} = \frac{(1-\alpha)(A_n n)^{\frac{\sigma-1}{\sigma}}}{\alpha(A_k k)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(A_n n)^{\frac{\sigma-1}{\sigma}}}. \tag{7}$$

Because the production function exhibits constant returns to scale, $1-\mu s$ is the corresponding markup-adjusted capital share. It follows that

$$\frac{\mu s}{1-\mu s} = \frac{1-\alpha}{\alpha}\left(\frac{A_n n}{A_k k}\right)^{\frac{\sigma-1}{\sigma}}.$$

Taking log of each side gives

$$\log\frac{\mu s}{1-\mu s} = \frac{1-\sigma}{\sigma}\log\left(\frac{k}{n}\right) + \frac{\sigma-1}{\sigma}\log\left(\frac{A_n}{A_k}\right) + \log\frac{1-\alpha}{\alpha}. \tag{8}$$

Equation (8) relates the markup-adjusted labor-share odds, $\log[\mu s/(1-\mu s)]$, to the log capital–labor ratio. The capital wedge does not enter this relationship because its derivation uses the flexible-labor optimality condition and the production technology, but not the capital Euler equation. Holding factor-biased technology fixed, a 1 percent increase in the capital–labor ratio changes the labor-share odds by $(1-\sigma)/\sigma$ percent. The sign and magnitude of this response identify the elasticity of substitution: an increase in capital intensity raises the relative share of labor compensation when capital and labor are gross complements, but lowers it when they are gross substitutes. Therefore, when the markup is known, Equation (8) can identify the elasticity of capital–labor substitution with a valid instrument for capital-labor ratio. We refer to this equation as the *labor-share moment*.

A potential limitation of the IV estimator arises when the first stage is weak or zero:

$\text{cov}\left(\log \frac{k}{n}, Z\right) \approx 0$. One possibility is that σ is close to zero. As $\sigma \to 0$, the production function converges to the Leontief form $F(k, n) = \min\{A_k k, A_n n\}$. Capital and labor are then used in fixed proportions,

$$\frac{k}{n} = \frac{A_n}{A_k},$$

so the capital–labor ratio does not respond to relative factor prices. Consequently, the covariance between Z and the change in the capital–labor ratio converges to zero as $\sigma \to 0$.

More generally, under the technology exclusion restriction,

$$\text{cov}\left(\log \frac{k}{n}, Z\right) = \sigma \left[\text{cov}\left(\log \frac{w}{R}, Z\right) - \text{cov}\left(\log(1+\tau), Z\right)\right]$$

A weak first stage can therefore arise because: (i) $\sigma \to 0$, (ii) the policy generates little variation in observed relative factor prices; or (iii) capital adjustment frictions largely offset the policy-induced relative-price variation. These cases cannot generally be distinguished from the first stage alone. We therefore impose a modified relevance condition requiring the policy to generate nonzero variation in the effective relative factor price $w/((1+\tau)R)$.

Assumption 4a (Modified relevance condition)**.** *After accounting for the endogenous response of the capital wedge, the exogenous shock Z generates nonzero variation in the effective relative factor price:*

$$\text{cov}\left(\log\left(\frac{w}{(1+\tau)R}\right), Z\right) \neq 0.$$

Or equivalently, under Assumption 3 and for $\sigma > 0$,

$$\text{cov}\left(\log\left(\frac{k}{n}\right), Z\right) \neq 0.$$

Proposition 1 presents our friction–robust estimator of the elasticity of substitution between capital and labor when the markup is known. Rather than using variation in observed relative factor prices as the endogenous regressor, we regress the markup-

adjusted log labor-share odds on the log capital–labor ratio and instrument the latter with the exogenous policy shock. Capital adjustment frictions affect the first-stage response of the capital–labor ratio to the policy shock, but they do not enter the labor-share equation as an omitted determinant. Consequently, the same first-stage response appears in both the numerator and denominator of the IV estimand and cancels from their ratio. Whether the policy shock changes the capital–labor ratio through relative factor prices or through the endogenous capital wedge is therefore irrelevant for identification. As long as the shock generates a nonzero first stage and is uncorrelated with factor-augmenting technical change, the estimator consistently identifies the elasticity of substitution.

Proposition 1 (Identification with labor-share moment)**.** *Under Assumptions 1, 2, 3, and 4a, regressing the markup-adjusted log labor-share odds on the log capital–labor ratio, instrumented by the exogenous policy shock, delivers a consistent estimator for* $(1-\sigma)/\sigma$*, which can be used to recover the elasticity of substitution between capital and labor. That is:*

$$\text{plim}\hat{\rho} = \frac{\text{cov}\left(\log \frac{\mu s}{1-\mu s}, Z\right)}{\text{cov}\left(\log \frac{k}{n}, Z\right)} = \frac{1-\sigma}{\sigma},$$

Then,

$$\hat{\sigma} = \frac{1}{1+\hat{\rho}}.$$

The friction–robust estimator also addresses a practical limitation of the price estimator: actual firm-specific rental rates of capital are rarely observed. Several policies generate plausibly exogenous changes in the user cost of capital, including China's 2009 value-added tax reform studied by Chen et al. (2023) and Meng, Zhang and Zhang (2024) and the bonus depreciation studied by Zwick and Mahon (2017) and Curtis et al. (2023). Because actual rental rates are unavailable, researchers typically use the theoretically implied effect of the policy on the rental rate as the endogenous regressor. This theoretical effect, however, may differ from the actual change experienced by firms, causing the price estimator to be biased. By contrast, our estimator does not require quantifying the policy's effect on rental rates. It requires only that the policy generate a nonzero first-stage response in the capital–labor ratio.

An advantage of this strategy is that it accommodates any positive elasticity of substitution, including unity. However, it requires knowledge of the demand elasticity ε. Identifying ε often requires product-price data, which are typically unavailable in industrial surveys and censuses. We therefore develop an alternative strategy that jointly identifies ε and σ. Equation (7) also implies

$$\mu s = (1-\alpha) A_n^{\frac{\sigma-1}{\sigma}} \left(\frac{y}{n}\right)^{\frac{1-\sigma}{\sigma}}. \tag{9}$$

Combining Equation (9) with the demand function (1) and taking logarithms yields, for $\sigma \neq 1$,

$$\log py = \frac{1}{\mu}\log A_n + \frac{\sigma}{\mu(\sigma-1)}\log(1-\alpha) + \frac{1}{\mu}\log n - \frac{\sigma}{\mu(\sigma-1)}\log(\mu s) \tag{10}$$

where we normalize the price index P to unity.

Equation (10) relates plant-level revenue to labor input and the markup-adjusted labor share. Holding technology and the labor share constant, an increase in labor represents an expansion of production. Output rises proportionally, but revenue rises less because the plant must lower its price to sell the additional output. The revenue response therefore identifies the markup. Conditional on the markup, the relationship between revenue and the labor share identifies capital–labor substitution. Two valid instruments that provide sufficiently independent variation in labor and the labor share can therefore identify both coefficients. We refer to this relation as the *revenue moment*. However, this strategy excludes $\sigma = 1$, for which Equation (9) implies $\mu s = 1 - \alpha$. Proposition 2 summarizes the result.

Proposition 2 (Identification with revenue moment)**.** *Under Assumptions 1, 2, 3, and 4a, suppose that $\sigma \neq 1$, and the instrument–regressor covariance matrix has full column rank. An IV regression of revenue $\log py$ on labor $\log n$ and labor share $\log s$, with two instruments Z_1, Z_2, consistently estimates the coefficients $1/\mu$ and $-\sigma/[\mu(\sigma-1)]$. These coefficients identify the markup μ and the capital–labor elasticity of substitution σ. The moment condition is*

$$\mathbb{E}\left[\log py - \frac{1}{\mu}\log n + \frac{\sigma}{\mu(\sigma-1)}\log s \,\middle|\, Z_1, Z_2\right] = 0$$

with the constant terms omitted.

Finding two valid instruments can be difficult. Equation (8) provides an additional moment condition, allowing the same instrument to be used in both equations. Both moments require the instrument to be uncorrelated with factor-biased technology. Under this exclusion restriction and the rank condition in Assumption 4b, the two moments jointly identify μ and σ with a single instrument through the generalized method of moments (GMM). Proposition 3 summarizes this approach.

Assumption 4b (Rank condition for joint identification)**.** *Let $\theta = (\mu, \sigma)'$, and define*

$$g(\theta) = \begin{pmatrix} \mathbb{E}\left[Z\left\{\log \frac{\mu s}{1-\mu s} - \frac{1-\sigma}{\sigma}\log \frac{k}{n}\right\}\right] \\ \mathbb{E}\left[Z\left\{\log(py) - \frac{1}{\mu}\log n + \frac{\sigma}{\mu(\sigma-1)}\log s\right\}\right] \end{pmatrix}.$$

The moment function is continuously differentiable in a neighborhood of the true parameter vector θ_0, and

$$\operatorname{rank} G(\theta_0) = 2, \qquad G(\theta_0) = \left.\frac{\partial g(\theta)}{\partial \theta'}\right|_{\theta=\theta_0}.$$

Proposition 3 (Joint identification with two moments)**.** *Under Assumptions 1, 2, 3, and 4b, suppose that the true parameter vector $\theta_0 = (\mu_0, \sigma_0)'$ lies in the interior of the admissible parameter space, with $\sigma_0 \neq 1$. Then the labor-share and revenue moments locally identify the markup μ_0 and the capital–labor elasticity of substitution σ_0 using a single instrument Z.*

2.4 Discussion

In this subsection, we discuss the properties and limitations of our methods and relate them to the existing literature. First, a key limitation is the assumption that firms can adjust labor freely at the observed wage, as is standard in neoclassical growth models. Labor market power, hiring and firing costs, and search and recruitment frictions can invalidate the frictionless labor optimality condition in Equation (3) (Bentolila and Bertola, 1990; Cooper, Haltiwanger and Willis, 2007; Bloom, 2009; Berger, Herkenhoff and Mongey, 2022). Our identification strategies therefore require labor to be free of

these frictions, or additional restrictions on the resulting labor wedges. Two approaches may help address this concern. One is to derive alternative moment conditions using more flexible intermediate inputs, subject to additional identification conditions. The other is to control for observable labor market conditions and examine the sensitivity of our estimates to differences in exposure to labor market frictions.

Second, estimators using the labor-share moment can be sensitive to measurement error because this moment involves a nonlinear transformation of the observed labor share. Even independent, mean-zero errors in the log variables can generate transformed errors whose conditional means depend on the true labor share. If the instrument shifts the true labor share, the transformed errors may therefore be correlated with the instrument.[1] By contrast, the revenue moment is linear in the observed log variables. Additive measurement errors preserve its moment conditions provided that they remain orthogonal to the instruments in the estimation sample. This distinction creates a trade-off: using the labor-share moment allows joint identification with a single instrument under the stated conditions, whereas using the revenue moment alone avoids this nonlinear transformation but requires instruments that provide sufficiently independent variation in employment and the labor share.

The labor-share moment also requires that each firm's markup-adjusted labor share satisfies $\mu s < 1$. Although this restriction holds in our model, observed labor shares can exceed one. With a common markup, retaining all observations therefore imposes the upper bound $\mu < 1/\max_i s_i$. Allowing for higher markups requires excluding observations with sufficiently high labor shares, creating a trade-off between the admissible markup range and the sample size. We examine the sensitivity of our estimator to this issue in Monte Carlo experiments with measurement error.

[1] To illustrate this mechanism, suppose that only the wage bill contains measurement error. Consistent with our simulation design, let $\log(wn) = \log(wn)^* + \iota$, where an asterisk denotes the true value. Then $s = s^* e^{\iota}$, and Equation (8) implies

$$\log\left(\frac{\mu s}{1-\mu s}\right) = \frac{1-\sigma}{\sigma}\log\frac{k}{n} + \frac{\sigma-1}{\sigma}\log\frac{A_n}{A_k} + \log\frac{1-\alpha}{\alpha} + \underbrace{\iota - \log\left(1 - \frac{\mu s^*}{1-\mu s^*}(e^{\iota}-1)\right)}_{\text{measurement-error term}}.$$

Even when ι is independent and mean zero, the conditional mean of the transformed error generally depends on the true labor share. It can therefore be correlated with the instrument if the instrument shifts the true labor share.

In addition, the revenue moment in Equation (10) shares with De Loecker and Warzynski (2012) the idea of using production-side information to identify markups. They first estimate production function parameters using a control function approach and then recover markups as the ratio of a flexible input's output elasticity to its expenditure share in revenue. Our approach instead jointly identifies production and demand parameters in a single IV regression. Concerns about the production approach include the difficulty of recovering physical output elasticities from revenue data and its commonly imposed assumption of Hicks-neutral productivity (Bond et al., 2021; Raval, 2023). These concerns are less directly applicable to our approach because our revenue moment explicitly incorporates the demand function and allows for factor-biased technological change. Our approach does, however, impose CES restrictions on production and demand, whereas their markup formula does not require these functional forms.

3 Monte Carlo Experiments

To compare the performance of conventional methods and our proposed estimator, we conduct several Monte Carlo experiments using a dynamic firm model with uncertainty, taxation, and capital frictions. We briefly describe the economic environment here and provide further details in Appendix B. The firm faces the demand function in Equation (1) and produces using the CES technology in Equation (2). Each firm enters a period with capital k, debt b, and capital- and labor-augmenting technologies A_k, A_n, which follow AR(1) processes. The firm chooses labor n, price p, investment x, next-period debt b', and dividends d.

Capital adjustment involves convex and fixed costs,

$$\mathcal{C}(k, x) = \frac{\psi x^2}{2k} + \chi \mathbb{1}\{x \neq 0\}.$$

Firms also face convex financing costs,

$$\mathcal{B}(b', k') = \frac{\phi b'^2}{2k'},$$

and a collateral constraint $b' \leq \theta k'$ that limits borrowing to a fraction θ of next-period capital.

Investment is subject to a value-added tax (VAT), $\mathcal{T}^v(x) = \nu \max\{x, 0\}$, where ν is the VAT rate. Firms also pay a payroll tax at rate τ_n, initially set to zero, and a corporate income tax at rate τ_c. Payroll-inclusive labor expenditures, capital adjustment costs, and interest payments are deductible, whereas financing costs are not. We represent depreciation allowances by their present value, so that net corporate income tax payments are

$$\mathcal{T}^c = \tau_c \left[py - (1 + \tau_n)wn - \mathcal{C}(k, x) - rb - \mathcal{D}\left(x + \mathcal{T}^v(x)\right) \right],$$

where $\mathcal{D} = \omega + (1 - \omega)\mathcal{D}^{\text{ord}}$. Here, ω is the immediately deductible share of investment expenditure, and $\mathcal{D}^{\text{ord}}$ is the present value of ordinary depreciation deductions per unit of investment. The tax basis includes VAT paid on investment. We allow full tax-loss refunds and apply tax recapture when capital is sold.

Given these features, the firm's recursive problem is

$$\begin{aligned} v(k, b, A_k, A_n) &= \max_{p,n,x,d,b'} \{d + \beta \mathbb{E} v(k', b', A_k', A_n')\} \\ \text{s.t.} \quad d + x + (1 + r)b + \mathcal{T}^v(x) + \mathcal{T}^c + \mathcal{C}(k, x) + \mathcal{B}(b', k') &= py - (1 + \tau_n)wn + b' \\ b' &\leq \theta k' \\ k' &= (1 - \delta)k + x. \end{aligned}$$

3.1 Baseline Experiments

We consider three data-generating processes. The first, DGP1, is a benchmark economy without uncertainty and frictions. We specify deterministic technology paths, assume perfect foresight over these paths, and set $\nu = \psi = \chi = \phi = \tau_c = 0$ and $\theta = +\infty$. By allowing firms to anticipate future technology and adjust investment accordingly, this setting approximates the static benchmark underlying Oberfield and Raval (2021). We then unexpectedly raise the payroll tax rate from zero to 5% for 50% of firms randomly and use treatment assignment as an instrument. Capital remains predetermined when the shock occurs.

The second data-generating process, DGP2, introduces uncertainty, capital adjustment costs, financing costs, collateral constraints, and taxes. Parameter values are reported in Table B1. Comparing DGP2 with DGP1 allows us to assess how these distortions affect the estimators. We implement the same unanticipated payroll tax increase from zero to 5% for 50% of firms randomly and use treatment assignment as an instrument.

The third data-generating process, DGP3, retains the frictional environment and payroll tax shock of DGP2 and adds an unanticipated elimination of the 17% VAT on equipment purchases. Motivated by China's 2009 VAT reform, this experiment reduces both the effective purchase price of investment goods and partial irreversibility, since VAT is paid on equipment purchases but not refunded upon resale (see Liu and Mao (2019); Chen et al. (2023); Meng, Zhang and Zhang (2024); Cao et al. (2026)). The VAT and payroll tax treatments are independently and randomly assigned for 50% of firms each, and their indicators serve as instruments for evaluating the joint identification strategy in Proposition 2. Table B2 reports summary statistics for selected key variables. DGP3 captures several features of firms' investment and financing behavior observed in the data. Approximately 50% of observations do not have positive investment, while 4% do not have positive net debt.

In each DGP, the economy is initially in its pre-reform stationary equilibrium, with baseline observations drawn from periods -5 through -1. An unanticipated reform occurs in period 0, and we solve the subsequent transition over 100 periods. Appendix B defines the stationary and dynamic equilibria. All estimators use long-run changes between the pre-reform window and periods 80 through 99, by which time the economy has approximately converged to its post-reform stationary equilibrium. We report estimates based on firms' short-run responses in periods 1–20 in Tables B3 and B4. Conventional methods exhibit larger biases because capital frictions have a greater effect on short-run input adjustments.

In each DGP, we first estimate the parameters without measurement error. We then add increasing levels of i.i.d. measurement errors to log value added ($\log(py)$), log wage bill ($\log((1+\tau_n)wn)$), log capital ($\log k$), and log employment ($\log n$). The error variances are calibrated to increase the variance of each observed log variable by 5% and

10%.[2] Labor shares and input ratios are reconstructed from the noisy observations. This exercise assesses the sensitivity of conventional and proposed estimators to measurement error.

Table 1 reports estimates of the demand elasticity ε and capital–labor substitution elasticity σ, with Panels A, B, and C corresponding to DGP1, DGP2, and DGP3. Columns (1)-(3) report the estimates of the demand elasticity ε. Column (1) infers ε from the pre-reform ratio of aggregate revenue to aggregate cost. The frictionless rental rate is imputed as:

$$R = (1+\nu)\frac{1-\tau_c\mathcal{D}}{1-\tau_c}(r+\delta).$$

Column (2) reports IV estimates using the revenue moment, denoted by R Moment, with the two policy shocks as instruments. Column (3) reports GMM estimates using both the labor-share and revenue moments with a single policy shock as the instrument. Columns (4)–(7) report estimates of the elasticity of substitution between capital and labor, σ. Column (4) reports conventional estimates obtained by regressing the log capital-to-labor expenditure ratio on the $\log((1+\tau_n)w/R)$. Column (5) reports IV estimates using the labor-share moment, denoted by L Moment, with a single policy shock as the instrument and the true value of ε assumed to be known. Columns (6) and (7) report the estimates of σ corresponding to the joint estimation procedures in Columns (2) and (3), respectively. We also compute firm-clustered standard errors within each replication, using the delta method for nonlinear transformations, and report their average across replications.

We set the true values of the capital–labor substitution elasticity to 0.8 and the demand elasticity to 5. Panel A presents the frictionless benchmark. Without measurement error, both the conventional and proposed estimators recover values close to the true parameters. The revenue–cost method and joint GMM yield demand elasticities around 5, while conventional IV, labor-share IV with known ε, and joint GMM yield capital–labor substitution elasticities around 0.8. The revenue–cost method has

[2]Under our measurement-error design, a 10% increase in the variance of the log variables translates into an approximately 21%–100% increase in their variance in levels, depending on the variable and DGP. Overall, the increase is largest in DGP1, whereas DGP2 exhibits the smallest increase in variance. Splitting the results by variable, the increase ranges from 35.5% in DGP2 to 93.2% in DGP1 for value added, from 33.2% to 91.0% for labor, from 21.2% to 100.7% for capital, and from 33.1% to 90.9% for labor compensation.

a smaller reported standard error than GMM for the demand elasticity, whereas the reported standard errors for capital–labor substitution are similar across estimators.

Panel B shows how capital frictions affect estimation under the payroll tax shock. The conventional revenue–cost method yields a demand elasticity of 4.65, below the true value of 5. Because the frictionless rental rate omits the shadow cost associated with capital frictions, measured factor costs understate effective production costs. The resulting revenue–cost ratio thereby overstates the markup and understates the demand elasticity. By contrast, our GMM estimator yields a demand elasticity of 5.09, close to the true value.

For capital–labor substitution, conventional IV produces an upward-biased estimate of 0.86. Higher payroll taxes raise labor costs, increasing marginal costs and reducing output. In the frictionless benchmark, the contraction in production scale dominates substitution toward capital, so both capital and labor decline. Capital frictions, however, dampen the decline in capital, causing the capital–labor ratio to rise more than technological substitution alone would imply. Conventional IV therefore overestimates σ, consistent with the mechanism in Section 2.2. By contrast, the labor-share IV and GMM estimators both recover 0.8.

Panel C introduces a VAT reform alongside the payroll tax reform in the frictional economy, providing a second instrument. Column (1) shows that the revenue–cost method continues to underestimate the demand elasticity. Column (4) reports the conventional IV estimate using VAT treatment assignment as the instrument. The estimate is now downward biased, at about 0.63 compared with the true value of 0.8. Eliminating VAT reduces the user cost of capital and encourages capital accumulation. However, adjustment costs and financing constraints limit this expansion relative to the frictionless response. The capital–labor ratio consequently rises less than technological substitution alone would imply, generating downward bias.

For our proposed estimators, Column (5) uses the labor-share moment with VAT treatment assignment as the instrument and the demand elasticity fixed at its true value. Columns (2) and (6) jointly estimate both elasticities using the revenue moment with both policy assignments as instruments. Columns (3) and (7) jointly estimate both elasticities using GMM with both moments and VAT treatment assignment as

the single instrument. All three approaches yield capital–labor substitution elasticities close to 0.8, while both joint estimation procedures yield demand elasticities close to 5.

We then add increasing levels of i.i.d. measurement error to our simulated data. Measurement error affects the estimators differently. Conventional IV estimates of capital–labor substitution change little across the reported error levels, although their bias from capital frictions remains. By contrast, the labor-share IV and joint GMM estimators overestimate capital–labor substitution when measurement error is introduced. For example, at the 10% measurement-error level in Panel C, both estimates rise to about 0.9, compared with the true value of 0.8. Joint GMM estimates of the demand elasticity are also sensitive to measurement error: the direction of the deviation varies across DGPs, and the reported standard errors increase substantially in Panels A and B. The conventional revenue–cost estimates are more stable in the frictionless benchmark but decline further below the true demand elasticity in the frictional economies.

These results are consistent with the measurement-error properties discussed in Section 2.4. The nonlinear labor-share transformation can invalidate the corresponding moment conditions even when errors in the underlying log variables are independent and mean zero. By contrast, the revenue-moment IV estimator in Panel C yields capital–labor substitution estimates close to 0.8 and demand elasticity estimates close to 5 across the reported error levels, although its reported standard errors increase.

Overall, both conventional and proposed estimators recover the true parameters in the frictionless benchmark without measurement error. Capital frictions bias the conventional estimates, while measurement error introduces bias into the proposed estimators that use the nonlinear labor-share moment. The revenue-moment estimator performs well in the presence of both capital frictions and the measurement errors considered in these experiments.

3.2 Correlated Policy Shocks and Capital Irreversibility

We conduct two additional experiments to examine how conventional estimators perform when policy shocks are correlated or when a policy directly changes capital adjustment frictions.

DGP4 retains the frictional environment of DGP3 but allows the payroll tax and

Table 1: Monte Carlo Experiments

	$\varepsilon = 5$: substitution between goods varieties			$\sigma = 0.8$: substitution between capital and labor			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Meas. error	Conventional revenue-cost	R Moment 2 IVs	2 Moments GMM+ 1 IV	Conventional IV	L Moment known ε + 1 IV	R Moment 2 IVs	2 Moments GMM+ 1 IV
A. DGP1: frictionless economy with payroll tax shock							
0%	5.000	—	5.034	0.796	0.796	—	0.795
	(0.000)		(0.415)	(0.023)	(0.023)		(0.021)
5%	5.008	—	6.925	0.793	0.939	—	0.941
	(0.025)		(19.564)	(0.041)	(0.108)		(0.078)
10%	5.015	—	7.012	0.792	0.949	—	0.949
	(0.037)		(5.362)	(0.054)	(0.125)		(0.091)
B. DGP2: Frictional economy with payroll tax shock							
0%	4.648	—	5.086	0.857	0.796	—	0.797
	(0.007)		(0.449)	(0.060)	(0.021)		(0.020)
5%	4.548	—	5.966	0.856	0.870	—	0.909
	(0.009)		(5.502)	(0.063)	(0.072)		(0.073)
10%	4.454	—	6.122	0.856	0.911	—	0.927
	(0.011)		(2.097)	(0.066)	(0.091)		(0.070)
C. DGP3: Frictional economy with VAT reform and payroll tax shock							
0%	4.577	5.083	5.031	0.626	0.799	0.797	0.800
	(0.006)	(0.359)	(0.254)	(0.018)	(0.009)	(0.023)	(0.012)
5%	4.483	5.144	4.654	0.627	0.871	0.800	0.897
	(0.009)	(0.647)	(0.316)	(0.018)	(0.028)	(0.039)	(0.027)
10%	4.394	5.196	4.420	0.627	0.905	0.801	0.916
	(0.011)	(0.885)	(0.516)	(0.019)	(0.035)	(0.050)	(0.035)

Note: L and R moments denote the labor-share and revenue moments, respectively. Each cell reports the Monte Carlo mean across replications; parentheses report the average estimator-reported, firm-clustered standard error across these replications. At measurement-error level x, each variable log-error variance equals $x/100$ times the log variance of that variable in the estimation sample.

VAT treatments to be correlated. Half of the firms receive each treatment, and 32.5% receive both. Consequently, 65% of firms receiving the payroll tax shock also receive the VAT reform. This design captures the possibility that an instrument used to identify one policy's effect also picks up the effects of another policy.

DGP5 implements the VAT reform in an otherwise frictionless economy, isolating the effects of the investment tax from other capital frictions. Firms pay VAT when purchasing equipment but receive no refund when selling used equipment. The tax therefore creates a gap between purchase and resale prices, discouraging equipment resales following adverse productivity shocks. We eliminate the tax for 50% of firms randomly, lowering their purchase price of capital and reducing partial irreversibility.

Table 2 reports the results of the additional experiments. Panel A shows how correlated policy shocks affect conventional IV estimation. Without measurement error, the conventional estimate of capital–labor substitution falls from 0.86 under the isolated payroll tax shock in Panel B of Table 1 to 0.64 when the payroll tax and VAT treatments are correlated. Thus, the correlation reverses the direction of conventional IV bias relative to the true elasticity of 0.8. The payroll tax instrument now captures both higher labor costs and the effects of the VAT reform on capital costs and adjustment incentives. Conventional IV attributes the resulting input responses to changes in relative factor prices without accounting for the endogenous capital wedge. Figure B1 confirms that the estimate declines as treatment correlation increases, falling below 0.8 at a correlation of approximately 0.1.

By contrast, all three proposed strategies recover the capital–labor elasticity closely in the absence of measurement error. Their validity does not require the instrument to operate exclusively through a particular factor price. Instead, we require that the instrument generate variation in firms' input choices while satisfying the exclusion restriction. With measurement error, estimates based on the labor-share moment exhibit bias, whereas the revenue-moment estimates remain close to the true parameters.

Panel B examines the VAT reform in an otherwise frictionless economy. We assume that VAT reform theoretically reduces the user cost of capital by 15.7% following Chen et al. (2023). The conventional IV estimate of capital–labor substitution is 0.19, implying a larger downward bias compared to 0.63 in Panel C of Table 1. Eliminating VAT

not only lowers the purchase price of investment goods but also reduces partial irreversibility. Firms may therefore sell more capital following adverse productivity shocks, offsetting some of the investment response to lower capital costs. The conventional revenue–cost method also understates the demand elasticity, yielding 4.81 instead of 5, because the imputed rental rate does not capture the effective capital costs associated with partial irreversibility. By contrast, our labor-share IV estimator recovers a capital–labor substitution elasticity of 0.80. Joint GMM yields estimates of 0.79 for capital–labor substitution and 5.02 for demand elasticity. All are very close to the true values.

Together, these experiments show that conventional IV estimates can be biased when instruments capture multiple factor-cost shocks or when policies change capital adjustment frictions. Our approach accommodates both channels without requiring their effects on the capital wedge to be quantified.

4 Empirical Application

4.1 Data

Our main data source is the 2004 Chinese Industrial Census (CIC), which provides manufacturing establishment data comparable to that used by Oberfield and Raval (2021) for the United States. Brandt, Kambourov and Storesletten (2026) provide a detailed description of the dataset. The CIC covers the entire manufacturing sector and reports gross output, value added, gross and net capital stocks, employment, and the wage bill.

We measure py using value added and wn using the wage bill. We construct the real capital stock following Brandt, Kambourov and Storesletten (2026). Appendix C.1 provides details on data and variable construction. We exclude observations with negative or missing values for key variables. To limit the influence of outliers and measurement error, we follow Oberfield and Raval (2021) and exclude observations in the top or bottom 1% of the value-added-to-cost ratio distribution. As a robustness check, Appendix C.2 reports estimates obtained by instead excluding observations in the top or bottom 1% and 5% of each key variable's distribution. Our main estimates

Table 2: Additional Monte Carlo Experiments

	$\varepsilon = 5$: substitution between goods varieties			$\sigma = 0.8$: substitution between capital and labor			
Meas. error	(1) Conventional revenue-cost	(2) R Moment 2 IVs	(3) 2 Moments GMM + 1 IV	(4) Conventional IV	(5) L Moment known ε + 1 IV	(6) R Moment 2 IVs	(7) 2 Moments GMM + 1 IV
A. DGP4: Frictional economy with correlated payroll tax shock and VAT reform							
0%	4.578	5.104	5.039	0.641	0.799	0.798	0.800
	(0.006)	(0.419)	(0.309)	(0.016)	(0.008)	(0.020)	(0.012)
5%	4.484	5.172	4.546	0.641	0.872	0.801	0.895
	(0.009)	(0.759)	(0.349)	(0.017)	(0.026)	(0.034)	(0.024)
10%	4.395	5.247	4.320	0.641	0.904	0.803	0.915
	(0.011)	(1.056)	(0.489)	(0.018)	(0.031)	(0.044)	(0.032)
B. DGP5: Frictionless economy with VAT reform							
0%	4.811	—	5.020	0.188	0.795	—	0.785
	(0.003)		(0.688)	(0.009)	(0.030)		(0.031)
5%	4.782	—	5.250	0.190	0.883	—	0.907
	(0.020)		(57.494)	(0.013)	(0.100)		(0.093)
10%	4.753	—	5.260	0.190	0.929	—	0.951
	(0.029)		(8.471)	(0.016)	(0.128)		(0.129)

Note: L and R moments denote the labor-share and revenue moments, respectively. Each cell reports the Monte Carlo mean across replications; parentheses report the average estimator-reported, firm-clustered standard error across these replications. At measurement-error level x, each variable log-error variance equals $x/100$ times the log variance of that variable in the estimation sample. The correlation between treatment assignments in Panel A is 0.3.

are robust to this alternative sample restriction. For GMM estimation, we restrict the markup to $[1, 2]$ and retain observations with labor shares below 50%, ensuring that $\mu s < 1$ throughout the admissible parameter range. This restriction excludes approximately 38% of firms. These firms are predominantly small, accounting for only 18% of sales and 17% of value added in the sample before the labor-share restriction. The final estimation sample contains 591,528 firms. Table C5 reports summary statistics of key variables.

City-level wages come from the *China City Statistical Yearbook*. Following the approach of Oberfield and Raval (2021), we construct a residual local wage measure by regressing log city-level wages on the share of the population enrolled in higher education, the male population share, and the tertiary-sector share. We use the residual from this regression as our measure of local wages.

4.2 Instrument

Oberfield and Raval (2021) exploit the idea that nationwide industry expansion raises labor demand in areas initially more exposed to that industry, improving workers' outside options and increasing local wages. They construct an instrument by interacting initial local industry employment shares with national industry employment growth. Following this approach, we use China's 1990 Population Census to calculate city-level employment shares at the 3-digit industry level.

National industry employment growth between 1990 and 2004 at the three-digit level is not available in our data. We therefore use the initial employment shares directly as instruments rather than combining them with national growth rates into a single shift-share instrument. This approach follows the identification argument based on exogenous shares discussed by Goldsmith-Pinkham, Sorkin and Swift (2020) and Borusyak, Hull and Jaravel (2025): national industry growth rates determine how the shares are combined, while identification can rest on the exogeneity of the underlying shares.

Specifically, our instruments are the 1990 employment shares of nonmanufacturing industries in each city. The identifying assumption is that, conditional on other controls, these initial shares are uncorrelated with the factor-augmenting technology terms

entering the estimating equations. Let z_c denote the column vector of these employment shares in city c. The specification uses 46 separate nonmanufacturing employment shares in 248 city clusters.

4.3 Identification

Following Oberfield and Raval (2021), we estimate the conventional expenditure-ratio specification:

$$\log \frac{Rk_{ijc}}{wn_{ijc}} = \rho_0 + \rho_1 \log w_c + X_{ijc} + \epsilon_{ijc}$$

where plant i operates in industry j and city c, and $\log w_c$ denotes the residualized log city wage. The wage bill wn_{ijc} accounts for differences in worker skills across plants when labor is measured in efficiency units. The controls X_{ijc} include four-digit industry and plant age fixed effects. We instrument the residualized log wage with the initial industry employment shares Z_c. Under the conventional identifying assumptions, $\rho_1 = \sigma - 1$, and the estimated elasticity is $1 + \hat{\rho}_1$. We use standard two-stage least squares and cluster standard errors by city.

For our proposed estimator, define the labor-share and revenue expressions as

$$u_{1ijc}(\theta) = \log \frac{\mu s_{ijc}}{1 - \mu s_{ijc}} - \frac{1-\sigma}{\sigma} \log \left(\frac{k_{ijc}}{n_{ijc}} \right),$$
$$u_{2ijc}(\theta) = \log py_{ijc} - \frac{1}{\mu} \log n_{ijc} + \frac{\sigma}{\mu(\sigma - 1)} \log s_{ijc}.$$

where $\theta = (\varepsilon, \sigma)'$. At each candidate parameter vector, we remove the intercept, industry fixed effects, and plant age controls from each expression. We impose the parameter restrictions:

$$\Theta = \left\{ (\varepsilon, \sigma) : \varepsilon > 1, \quad \sigma > 0, \quad \sigma \neq 1, \quad \frac{\varepsilon}{\varepsilon - 1} s_{ijc} < 1 \text{ for every retained observation} \right\}.$$

The plant-level moment vector is

$$g_{ijc}(\theta) = \begin{pmatrix} z_c u_{1ijc}(\theta) \\ z_c u_{2ijc}(\theta) \end{pmatrix},$$

and its sample average is

$$g_N(\theta) = \frac{1}{N} \sum_{ijc} g_{ijc}(\theta).$$

Identification requires the population moments to equal zero at the true parameters and to provide sufficiently independent variation to identify both parameters.

The GMM estimator is

$$\hat{\theta} = \arg\min_{\theta \in \Theta} g_N(\theta)' W_N g_N(\theta),$$

The first step uses a block-diagonal weighting matrix constructed from the instruments. The second step uses the generalized inverse of the estimated city-clustered covariance matrix of the stacked moment conditions, evaluated at the first-step estimates. This weighting matrix remains fixed during the second-step optimization. Standard errors are clustered by city.

Our GMM specification includes the same four-digit industry and firm age fixed effects as the conventional specification. An additional identification concern is that initial local industry composition may be correlated with demand shifters entering the revenue moment. In the baseline specification, we assume that demand shifters vary only across four-digit industries and are therefore absorbed by industry fixed effects. More generally, identification permits demand heterogeneity across regions, provided that the remaining demand component is orthogonal to the instruments conditional on the included controls.

We use both the labor-share and revenue moments in our baseline GMM specification. Although the revenue moment alone can identify both parameters with suitable instruments, the resulting estimates are sensitive to the labor-share sample restriction in our application. We therefore also use the labor-share moment to incorporate information from the relationship between labor shares and capital intensity for identification.

4.4 Main Results

Table 3 reports the conventional estimates and our proposed estimates, denoted by YZ. Column (1) reports an OLS estimate of the capital–labor substitution elasticity of 0.68. Instrumenting local wages with the vector of initial industry employment shares

reduces the estimate to 0.51, close to the plant-level estimates in Oberfield and Raval (2021). The first-stage cluster-robust F-statistic is 54, indicating that the instruments jointly predict local wages. We also infer the markup from the ratio of aggregate revenue to measured factor costs. The resulting markup is 2.87, implying an elasticity of substitution between goods varieties, ε, of 1.54.

Column (5) reports our joint GMM estimates of σ and ε. The estimated plant-level capital–labor substitution elasticity is 0.67, compared with the conventional IV estimate of 0.51. The difference is statistically significant at the 5% level. Both estimates imply that capital and labor are gross complements, but our estimate indicates weaker complementarity. Under the maintained assumptions, our estimator identifies technological substitution without imposing a frictionless capital optimality condition. The estimated markup is 1.26, implying a demand elasticity of 4.76. This markup is substantially lower than the revenue–cost estimate. The discrepancy highlights the sensitivity of revenue–cost markup estimates to whether imputed capital rental expenditures capture the effective cost of capital services.

To assess the sensitivity of our GMM estimates to the instrument set, we conduct several robustness checks. Dropping each employment share individually yields capital–labor substitution estimates between 0.64 and 0.67, compared with the baseline estimate of 0.65 using all 46 initial employment shares. Excluding agriculture-related shares yields a similar estimate of 0.65. Aggregating the instruments into 12 broad sectors slightly raises the estimate to 0.73 and increases its standard error from 0.01 to 0.05. Overall, the substitution estimates remain broadly similar across these instrument sets. The individual-share deletion results also suggest that the baseline estimate is not driven by any single industry share.

The difference between the GMM and conventional IV estimates suggests a downward bias in the latter under our maintained assumptions. Our Monte Carlo experiments provide a possible explanation. An isolated payroll tax shock produces an upward bias, whereas correlating this shock with the VAT reform reverses the bias. Initial industry employment shares may similarly capture changes in both labor and capital costs.[3] Nationwide industry expansion raises local labor demand and wages, but it may

[3]Provincial equipment price indices show substantial regional dispersion. Between 2000 and 2005, equipment prices rose by 7.4% in Xinjiang, 5.0% in Inner Mongolia, and 4.3% in Gansu, but fell

also increase equipment demand and investment-good prices where equipment supply is inelastic. The conventional estimate can therefore reflect both channels rather than the wage response alone. Our proposed estimator accommodates these channels because it does not require the instrument to affect wages exclusively.

Table 3: Estimates of Capital-Labor Substitution and Demand Elasticities

	Conventional				YZ		
	OLS	IV	OLS	IV	GMM	GMM	GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
σ	0.679	0.510	0.752	0.581	0.670	0.654	0.650
	(0.061)	(0.070)	(0.069)	(0.080)	(0.012)	(0.010)	(0.014)
ε			1.536		4.760	4.771	3.664
			(0.016)		(0.599)	(0.541)	(0.359)
Labor market controls	No	No	Yes	Yes	No	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	No
Firm age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province × Industry	No	No	No	No	No	No	Yes

Note: All specifications include firm age fixed effects. Columns (1)–(6) include four-digit industry fixed effects, while column (7) includes province × four-digit industry fixed effects. Labor market controls include the industry × city employment Herfindahl–Hirschman index, the firm's own employment share, and the SOE employment share. Standard errors are clustered by city and reported in parentheses. Standard errors for nonlinear transformations of the GMM estimates are computed using the delta method. Conventional markups are calculated as aggregate revenue divided by aggregate measured factor costs, with the implied demand elasticity given by $\varepsilon = \mu/(\mu - 1)$. YZ jointly estimates μ and σ using GMM. All specifications use 591,528 firms.

4.5 Threats to Identification

We examine several potential threats to identification and assess the sensitivity of our estimates to these concerns. First, our GMM estimation restricts the markup to $1 < \mu \leq 2$. To ensure that $\mu s < 1$ throughout the parameter space, we exclude observations with labor shares of 50% or higher. Conditional on labor output elasticities, this restriction may disproportionately exclude low-markup firms, shifting the retained sample toward higher-markup firms. This selection mechanism would tend to work against our finding

by 18.4% in Beijing, 16.0% in Fujian, and 15.1% in Guangdong. Initial industrial composition may capture such differences through industry-specific equipment demand and exposure to different types of capital goods.

of a lower markup and a higher demand elasticity than the revenue–cost estimates. Consistent with this interpretation, in the robustness check, we lower the markup upper bound and retain more high-labor-share firms. This process further reduces our markup estimate and raises the implied demand elasticity. These results suggest that excluding high-labor-share firms is unlikely to be the main driver of our lower markup estimate.

Second, our framework assumes that firms can adjust labor freely within each period. Labor market power and search and matching frictions can distort the labor optimality condition and bias our estimates. Four-digit industry fixed effects absorb common industry-level components of the estimating equations, including those associated with labor market frictions. To assess sensitivity to differences in local labor market conditions within industries, we additionally include the following controls. Motivated by Berger, Herkenhoff and Mongey (2022), we construct an employment Herfindahl–Hirschman index (HHI) within each city–industry labor market and include each plant's share of employment in that market. These variables capture local employer concentration and the plant's potential influence on the local labor market.[4] We also control for the city-level employment share of SOEs. As Brandt, Kambourov and Storesletten (2026) show, the size of the state sector is positively correlated with entry barriers for non-state firms; the presence of the state sector may affect the local competitive environment and employment opportunities.

Columns (3), (4), and (6) report estimates that include these controls. The conventional OLS and IV estimates increase from 0.68 to 0.75 and from 0.51 to 0.58, respectively. By contrast, our GMM estimate changes only slightly, from 0.67 to 0.65. The estimated demand elasticity is also nearly unchanged, increasing from 4.760 to 4.771. Thus, controlling for observed labor market conditions leaves our GMM estimates largely unchanged, while narrowing the difference between the conventional IV and GMM estimates of capital–labor substitution. These results suggest that our main findings are not driven by the included labor market characteristics, although these controls cannot fully account for distortions to firms' labor choices.

Finally, the baseline specification assumes that demand shifters vary only across in-

[4]We also use the wage-bill HHI and each firm's share of the local wage bill as proxies for employer concentration and firms' wage-setting power, respectively. Our estimates are robust to this alternative specification.

dustries. If initial local industry composition is correlated with regional demand shifters, the revenue moment may be invalid. To assess sensitivity to this concern, we include province $\times$ four-digit industry fixed effects, allowing demand shifters to vary across province–industry pairs. These fixed effects also absorb additive province–industry components associated with labor market frictions. City fixed effects would absorb all variation in our city-level instruments and therefore cannot be included. Column (7) reports the results. The estimated capital–labor substitution elasticity remains close to 0.65. The estimated demand elasticity declines from 4.8 to 3.7, corresponding to an increase in the markup from 1.27 to 1.38. Thus, the estimated capital–labor substitution elasticity is stable after controlling for province–industry demand shifters. The demand elasticity is more sensitive to these controls but remains substantially above the conventional estimate.

4.6 Financial Constraints and Estimation Results

The main contribution of our framework is to identify the elasticities using firms' optimal labor choice without imposing a frictionless capital optimality condition. Our estimators therefore accommodate capital frictions that can bias conventional estimates. In this subsection, we examine whether estimates across firms with different financial conditions are consistent with this distinction. Because financial constraints are not directly observed, we group firms by age, ownership, and leverage and estimate capital–labor substitution separately for each group. All estimates use the same baseline specifications, controlling for four-digit industry fixed effects, plant age fixed effects, and the labor market conditions described above. Table 4 reports the main results.

Columns (1) and (2) compare old and young firms. Firm age serves as a proxy for financing conditions, as older firms generally face fewer obstacles to external finance (Hadlock and Pierce, 2010; Cloyne et al., 2023). We classify firms at or below the median age within their two-digit industry as young and the remaining firms as old. The conventional IV estimate is 0.81 for old firms and 0.39 for young firms. This difference is consistent with limited access to external finance constraining younger firms' input adjustments. By contrast, the GMM estimates are much closer: 0.64 for old firms and 0.69 for young firms. This pattern is consistent with our framework's

prediction that conventional IV estimates are more sensitive to capital frictions than our GMM estimates.

Columns (3) and (4) compare SOEs and non-SOEs. In China, SOEs generally enjoy better access to credit, whereas private firms rely more heavily on internal finance (Cull and Xu, 2005; Guariglia, Liu and Song, 2011; Cong et al., 2019). The conventional IV estimate is 0.63 for SOEs and 0.57 for non-SOEs. Our GMM estimator instead yields nearly identical elasticities of approximately 0.65 for both groups. The larger difference between conventional IV and GMM among non-SOEs is consistent with greater estimation bias among firms facing more restricted access to credit.

Columns (5) and (6) compare firms with different leverage ratios, measured as total liabilities divided by total assets. Leverage has an ambiguous relationship with financial constraints. Higher leverage can indicate greater debt burdens and limited additional borrowing capacity, but it can also reflect better access to external finance (Faulkender and Petersen, 2006; Leary, 2009). This second interpretation is relevant in China. In Song, Storesletten and Zilibotti (2011), entrepreneurial firms face limited access to credit and finance investment through entrepreneurs' accumulated savings, whereas SOEs have better access to external borrowing. Low leverage may therefore reflect restricted financing access rather than ample unused borrowing capacity. The conventional IV estimate is 0.67 for high-leverage firms and 0.52 for low-leverage firms, compared with GMM estimates of 0.67 and 0.62, respectively. Under this credit-access interpretation, the larger discrepancy among low-leverage firms is consistent with conventional estimation being more affected by financial constraints.

Overall, conventional estimates vary substantially across firms with different financing conditions, whereas our GMM estimates are more stable across groups. Moreover, conventional IV produces lower estimates relative to GMM among young firms, non-SOEs, and low-leverage firms—groups potentially facing greater obstacles to external finance. These patterns support our argument that conventional estimates are sensitive to financial constraints, while our proposed estimator is more robust to the frictions governing capital adjustment.

Table 4: Substitution Elasticity by Firm Characteristics

	Plant Age		Ownership		Liabilities / Assets	
	Old	Young	SOE	Non-SOE	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
Conventional	0.806	0.385	0.632	0.572	0.671	0.516
	(0.074)	(0.093)	(0.106)	(0.081)	(0.078)	(0.097)
YZ GMM	0.638	0.687	0.657	0.653	0.670	0.618
	(0.011)	(0.013)	(0.007)	(0.010)	(0.010)	(0.010)
Observations	273,744	317,784	10,419	581,106	295,463	296,015

Note: All specifications include four-digit industry and firm age fixed effects. Labor market controls include the industry × city employment Herfindahl–Hirschman index, the firm's own employment share, and the SOE employment share. Conventional estimates use IV, while YZ jointly estimates the markup and the capital–labor substitution elasticity using GMM. Standard errors are clustered by city and reported in parentheses.

5 Conclusions

This paper develops a framework for estimating capital–labor substitution that is robust to capital frictions. Conventional estimators infer substitution from firms' input responses to relative factor prices. These responses also reflect constraints on capital adjustment and can therefore generate substantial estimation bias. Our main contribution is to derive moment conditions from firms' optimal labor choices conditional on realized capital stocks. Together with CES production and product demand, these conditions jointly identify capital–labor substitution and markups without imposing a frictionless capital optimality condition. Our framework is therefore more general than those widely applied in the previous literature.

Monte Carlo experiments show that conventional estimators become biased with capital frictions and policy distortions, whereas our proposed estimators recover values close to the true parameters. In our empirical application, our estimator yields higher capital–labor substitution and demand elasticities than conventional methods. It also delivers more similar capital–labor substitution estimates across firm groups with different financial conditions. These patterns are consistent with our argument that the proposed estimator is more robust to capital frictions.

Our findings highlight the importance of distinguishing technological substitution from input adjustments shaped by capital frictions. A small response of capital in-

tensity to lower capital costs need not imply limited technological substitutability; it may instead reflect constraints on capital adjustment. Conventional elasticity estimates can therefore lead researchers to attribute sluggish investment to production technology rather than capital frictions. This distinction matters for evaluating investment policies: policies that lower capital costs and those that relax financing constraints operate through different channels, even when both encourage capital accumulation. Our framework provides a way to estimate technological substitution while allowing capital adjustment to remain constrained.

We also recognize that these gains come with limitations. The nonlinear transformation in the labor-share moment makes estimators using this moment sensitive to measurement error. The requirement that markup-adjusted labor shares remain below one also excludes some observations and creates a trade-off between the admissible markup range and sample coverage. Moreover, identification relies on an undistorted labor optimality condition, which may fail in the presence of labor market power or hiring and firing costs. Future research could explicitly account for measurement error and labor market distortions and extend the framework to allow for heterogeneous markups.

References

Acemoglu, Daron, Andrea Manera, and Pascual Restrepo. 2020. "Does the US tax code favor automation?" National Bureau of Economic Research.

Arrow, Kenneth J, Hollis B Chenery, Bagicha S Minhas, and Robert M Solow. 1961. "Capital-labor substitution and economic efficiency." *The review of Economics and Statistics*, 225–250.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic inputs and resource (mis) allocation." *Journal of Political Economy*, 122(5): 1013–1063.

Bentolila, Samuel, and Giuseppe Bertola. 1990. "Firing costs and labour demand: how bad is eurosclerosis?" *The Review of Economic Studies*, 57(3): 381–402.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor market power." *American Economic Review*, 112(4): 1147–1193.

Bloom, Nicholas. 2009. "The impact of uncertainty shocks." *econometrica*, 77(3): 623–685.

Bloom, Nick, Stephen Bond, and John Van Reenen. 2007. "Uncertainty and investment dynamics." *The review of economic studies*, 74(2): 391–415.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. 2021. "Some unpleasant markup arithmetic: Production function elasticities and their estimation

from production data." *Journal of monetary economics*, 121: 1–14.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2025. "A practical guide to shift-share instruments." *Journal of Economic Perspectives*, 39(1): 181–204.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten. 2026. "Barriers to entry and regional economic growth in China." *Review of Economic Studies*, 93(1): 286–326.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing." *Journal of development economics*, 97(2): 339–351.

Brandt, Loren, Thomas G Rawski, et al. 2008. *China's great economic transformation.* Vol. 9, Cambridge university press, Cambridge.

Cao, Jing, Guangrong Ma, Jie Mao, and Haoru Yin. 2026. "Who benefits from investment tax incentives? Evidence from China's VAT reform." *Journal of Economic Behavior & Organization*, 245: 107532.

Chaney, Thomas, David Sraer, and David Thesmar. 2012. "The collateral channel: How real estate shocks affect corporate investment." *American Economic Review*, 102(6): 2381–2409.

Chen, Zhao, Xian Jiang, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu. 2023. "Tax policy and lumpy investment behaviour: Evidence from China's VAT reform." *The Review of Economic Studies*, 90(2): 634–674.

Clementi, Gian Luca, and Hugo A Hopenhayn. 2006. "A theory of financing constraints and firm dynamics." *The Quarterly Journal of Economics*, 121(1): 229–265.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico. 2023. "Monetary policy, corporate finance, and investment." *Journal of the European Economic Association*, 21(6): 2586–2634.

Cong, Lin William, Haoyu Gao, Jacopo Ponticelli, and Xiaoguang Yang. 2019. "Credit allocation under economic stimulus: Evidence from China." *The review of financial studies*, 32(9): 3412–3460.

Cooper, Russell, John Haltiwanger, and Jonathan L Willis. 2007. "Search frictions: Matching aggregate and establishment observations." *Journal of Monetary Economics*, 54: 56–78.

Cooper, Russell W, and John C Haltiwanger. 2006. "On the nature of capital adjustment costs." *The Review of Economic Studies*, 73(3): 611–633.

Cull, Robert, and Lixin Colin Xu. 2005. "Institutions, ownership, and finance: the determinants of profit reinvestment among Chinese firms." *Journal of financial economics*, 77(1): 117–146.

Curtis, E Mark, Daniel G Garrett, Eric Ohrn, Kevin A Roberts, and Juan-Carlos Suarez-Serrato. 2023. "Capital Investment and Labor Demand: Evidence from 21st Century Stimulus Policy." Working Paper.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and firm-level export status." *American economic review*, 102(6): 2437–2471.

Faulkender, Michael, and Mitchell A Petersen. 2006. "Does the source of capital

affect capital structure?" *The Review of Financial Studies*, 19(1): 45–79.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review*, 110(8): 2586–2624.

Gopinath, Gita, Şebnem Kalemli-Özcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. 2017. "Capital allocation and productivity in South Europe." *The Quarterly Journal of Economics*, 132(4): 1915–1967.

Grossman, Gene M, Elhanan Helpman, Ezra Oberfield, and Thomas Sampson. 2017. "Balanced growth despite Uzawa." *American Economic Review*, 107(4): 1293–1312.

Guariglia, Alessandra, Xiaoxuan Liu, and Lina Song. 2011. "Internal finance and growth: Microeconometric evidence on Chinese firms." *Journal of Development economics*, 96(1): 79–94.

Hadlock, Charles J, and Joshua R Pierce. 2010. "New evidence on measuring financial constraints: Moving beyond the KZ index." *The review of financial studies*, 23(5): 1909–1940.

Harasztosi, Peter, and Attila Lindner. 2019. "Who pays for the minimum wage?" *American Economic Review*, 109(8): 2693–2727.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly journal of economics*, 124(4): 1403–1448.

Hsieh, Chang-Tai, and Zheng Michael Song. 2015. "Grasp the large, let go of the small: The transformation of the state sector in China." National Bureau of Economic Research.

Karabarbounis, Loukas, and Brent Neiman. 2014. "The global decline of the labor share." *The Quarterly journal of economics*, 129(1): 61–103.

Khan, Aubhik, and Julia K Thomas. 2008. "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics." *Econometrica*, 76(2): 395–436.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit cycles." *Journal of political economy*, 105(2): 211–248.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante. 2000. "Capital-skill complementarity and inequality: A macroeconomic analysis." *Econometrica*, 68(5): 1029–1053.

Leary, Mark T. 2009. "Bank loan supply, lender choice, and corporate capital structure." *The Journal of Finance*, 64(3): 1143–1185.

Liu, Yongzheng, and Jie Mao. 2019. "How do tax incentives affect investment and productivity? Firm-level evidence from China." *American Economic Journal: Economic Policy*, 11(3): 261–291.

Meng, Lingsheng, Yifan Zhang, and Yunbin Zhang. 2024. "Estimating capital-labor substitution in China: Evidence from firm-level data." *Available at SSRN 5037049*.

Midrigan, Virgiliu, and Daniel Yi Xu. 2014. "Finance and misallocation: Evidence from plant-level data." *American economic review*, 104(2): 422–458.

Oberfield, Ezra, and Devesh Raval. 2021. "Micro data and macro technology."

Econometrica, 89(2): 703–732.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century.* Harvard University Press.

Raval, Devesh. 2023. "Testing the production approach to markup estimation." *Review of Economic Studies*, 90(5): 2592–2611.

Raval, Devesh R. 2019. "The micro elasticity of substitution and non-neutral technology." *The RAND Journal of Economics*, 50(1): 147–167.

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic dynamics*, 11(4): 707–720.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing like china." *American economic review*, 101(1): 196–233.

Uzawa, Hirofumi. 1961. "Neutral inventions and the stability of growth equilibrium." *The Review of Economic Studies*, 28(2): 117–124.

Zwick, Eric, and James Mahon. 2017. "Tax policy and heterogeneous investment behavior." *American Economic Review*, 107(1): 217–248.

A Details of the Framework in Section 2

The CES production function implies the marginal product to labor as:

$$F_n = (1-\alpha)A_n^{\frac{\sigma-1}{\sigma}} y^{1/\sigma} n^{-1/\sigma} \quad \Rightarrow \quad d\log F_n = \frac{1}{\sigma}(d\log y - d\log n). \tag{A1}$$

Differentiating the production function gives

$$d\log y = e_k d\log k + e_n d\log n, \tag{A2}$$

where $e_k + e_n = 1$ by constant returns to scale. The demand function implies

$$d\log p = -\frac{1}{\varepsilon} d\log y. \tag{A3}$$

The labor optimality condition (3) implies:

$$d\log w = d\log p + d\log F_n. \tag{A4}$$

Combining these expressions yields

$$\begin{aligned} d\log w &= -\frac{1}{\varepsilon} d\log y + \frac{1}{\sigma}(d\log y - d\log n) \\ &= \left(\frac{1}{\sigma} - \frac{1}{\varepsilon}\right)(e_k d\log k + e_n d\log n) - \frac{1}{\sigma} d\log n \\ &= e_k \left(\frac{1}{\sigma} - \frac{1}{\varepsilon}\right) d\log k - \left(\frac{e_k}{\sigma} + \frac{e_n}{\varepsilon}\right) d\log n. \end{aligned}$$

The first equality is to replace $d\log p$ and $d\log F_n$ in Equation (A4) using Equations (A3) and (A1). The second equality is to replace $d\log y$ using Equation (A2). The third equality follows from the fact that $e_k + e_n = 1$. Rearranging yields

$$d\log \frac{k}{n} = \frac{\varepsilon\, d\log w + d\log k}{e_n + (\varepsilon/\sigma) e_k}. \tag{A5}$$

From the Euler equation (4):

$$\frac{(1+\tau)R}{w} = \frac{F_k}{F_n} = \frac{\alpha}{1-\alpha}\left(\frac{A_k}{A_n}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{k}{n}\right)^{-\frac{1}{\sigma}}$$

This gives the input response to the relative price change in the frictionless economy:

$$d\log \frac{k}{n} = \sigma[d\log w - d\log R - d\log(1+\tau)], \tag{A6}$$

Plugging this equation into Equation (A5) gives

$$\begin{aligned} d\log k &= \left(e_n + \frac{\varepsilon}{\sigma} e_k\right) d\log \frac{k}{n} - \varepsilon d\log w \\ &= \left(e_n + \frac{\varepsilon}{\sigma} e_k\right) \sigma[d\log w - d\log R - d\log(1+\tau)] - \varepsilon d\log w \\ &= e_n(\sigma - \varepsilon)\, d\log w - (e_n\sigma + \varepsilon e_k)\, d\log R - (e_n\sigma + \varepsilon e_k)\, d\log(1+\tau) \end{aligned} \tag{A7}$$

In the frictionless benchmark, $\tau = 0$ remains fixed. Evaluating the response at the common initial allocation therefore gives

$$d\log \tilde{k} = e_n(\sigma - \varepsilon) d\log w - (\sigma e_n + \varepsilon e_k) d\log R.$$

which is Equation (5) in the main text. Subtracting this expression from Equation (A7)

yields

$$d\log k - d\log\tilde{k} = -(\sigma e_n + \varepsilon e_k)d\log(1+\tau),$$

or equivalently,

$$-\sigma d\log(1+\tau) = \frac{d\log k - d\log\tilde{k}}{e_n + (\varepsilon/\sigma)e_k}.$$

Substituting back into Equation (A6) establishes

$$d\log\frac{k}{n} = \sigma\, d\log\frac{w}{R} + \frac{d\log k - d\log\tilde{k}}{e_n + (\varepsilon/\sigma)e_k},$$

which is Equation (6) in the main text.

Propositions 1 and 2 follow directly from Equations (8) and (10). We provide details of Proposition 3 here. Under regularity conditions permitting differentiation under the expectation, the derivatives of the first moment are

$$\frac{\partial g_1(\theta)}{\partial\mu} = \mathbb{E}\left[\frac{Z}{\mu(1-\mu s)}\right], \qquad \frac{\partial g_1(\theta)}{\partial\sigma} = \frac{\mathbb{E}[Z\log\frac{k}{n}]}{\sigma^2}.$$

The derivatives of the second moment are

$$\frac{\partial g_2(\theta)}{\partial\mu} = \frac{\mathbb{E}[Z\log n]}{\mu^2} - \frac{\sigma\mathbb{E}[Z\log s]}{\mu^2(\sigma-1)}, \qquad \frac{\partial g_2(\theta)}{\partial\sigma} = -\frac{\mathbb{E}[Z\log s]}{\mu(\sigma-1)^2}.$$

Therefore,

$$G(\theta) = \begin{pmatrix} \mathbb{E}\left[\frac{Z}{\mu(1-\mu s)}\right] & \frac{\mathbb{E}[Z\log\frac{k}{n}]}{\sigma^2} \\ \frac{\mathbb{E}[Z\log n]}{\mu^2} - \frac{\sigma\mathbb{E}[Z\log s]}{\mu^2(\sigma-1)} & -\frac{\mathbb{E}[Z\log s]}{\mu(\sigma-1)^2} \end{pmatrix}.$$

By Assumption 4b, $G(\theta_0)$ is nonsingular. The inverse function theorem implies that g is one-to-one in a neighborhood of θ_0. Consequently, θ_0 is the unique solution to $g(\theta) = 0$ in that neighborhood, establishing local identification. The rank condition is equivalent to

$$\frac{\mathbb{E}\left[\frac{Z}{\mu_0(1-\mu_0 s)}\right]\mathbb{E}[Z\log s]}{(\sigma_0-1)^2} + \frac{\mathbb{E}[Z\log(py)]\mathbb{E}\left[Z\log\frac{k}{n}\right]}{\sigma_0^2} \neq 0.$$

B Monte Carlo Setup

Capital- and labor-augmenting technologies follow AR(1) processes:

$$\log A_k' = \rho_k \log A_k + \epsilon_k', \quad \epsilon_k' \sim \mathcal{N}(0, \sigma_{\epsilon_k}^2),$$
$$\log A_n' = \rho_n \log A_n + \epsilon_n', \quad \epsilon_n' \sim \mathcal{N}(0, \sigma_{\epsilon_n}^2).$$

Both processes have zero intercepts. We discretize each process using the Rouwenhorst method with 11 states.

We first define the stationary equilibrium before the reform and then the transition following the reform.

Definition 1 (Stationary general equilibrium)**.** *Given the structural parameters, the policy vector (τ, ν, ω) and the constant world interest rate r, a stationary equilibrium is a wage w, final output Y, a value function v, optimal decision rules (g_n, g_k, g_b), an invariant firm distribution $\Omega(k, b, A_k, A_n)$, and implied individual and aggregate quantities such that: (i) firms maximize discounted dividends subject to their constraints; (ii) varieties obey CES demand and optimal pricing; (iii) the labor market clears $\int g_n d\Omega = 1$.; (iv) the final-good price index equals one and aggregation is consistent; and (v) Ω is invariant under firm policies and factor-augmenting technology transitions.*

Definition 2 (Dynamic equilibrium after a policy reform)**.** *Given an initial distribution Ω_T, an announced policy path $\{\tau_{it}, \nu_{it}, \omega_{it}\}_{t \geq T}$, and the fixed interest rate r, a dynamic equilibrium consists of sequences $\{w_t, Y_t, V_t, g_{n,t}, g_{k,t}, g_{b,t}, \Omega_t\}_{t \geq T}$ satisfying individual optimality given anticipated future aggregate variables and tax policies, labor clearing, final-good consistency, distribution transitions, and the government/resource identities at every date.*

Table B1 reports the baseline parameter values. Parameters are set to zero or otherwise modified where indicated in the description of each DGP. We solve the firm's problem using 160 capital grid points and nine leverage grid points. The grid used to propagate the firm distribution includes an additional positive-debt point.

We initialize firms from the pre-reform invariant distribution and solve the transition for 100 periods following the reform. Each simulated economy contains 100,000

firms. The reform occurs unexpectedly in period 0. The long-run estimates compare observations from periods −5 through −1 with those from periods 80 through 99. Market-clearing conditions are imposed on the model-implied firm distributions rather than separately on each finite simulated panel.

Table B1: Parameterization

Parameter		Value
Technology		
σ	Capital-labor substitution	0.8
α	CES capital weight	0.33
δ	Depreciation factor	0.08
ρ_k, ρ_n	Persistence of technology process	0.85, 0.85
$\sigma_{\epsilon_k}, \sigma_{\epsilon_n}$	Standard deviation of technology innovation	0.3, 0.3
Preference		
ε	Demand elasticity	5
β	Discount rate	0.95
Policy		
τ_c	Corporate income tax rate	25%
ν	VAT rate	17%
τ_n	Payroll tax rate	0
Frictions		
χ	Fixed cost of capital adjustment	0
ψ	Quadratic capital adjustment cost	0.45
ϕ	Quadratic borrowing cost	0.02
θ	Collateral constraint	0.12
Other parameters		
r	Interest rate	0.054
ω	Bonus depreciation rate	0
$\mathcal{D}^{\text{ord}}$	Present value of ordinary depreciation deductions	0.6

B.1 Additional Results

Table B2 reports summary statistics in DGP3.

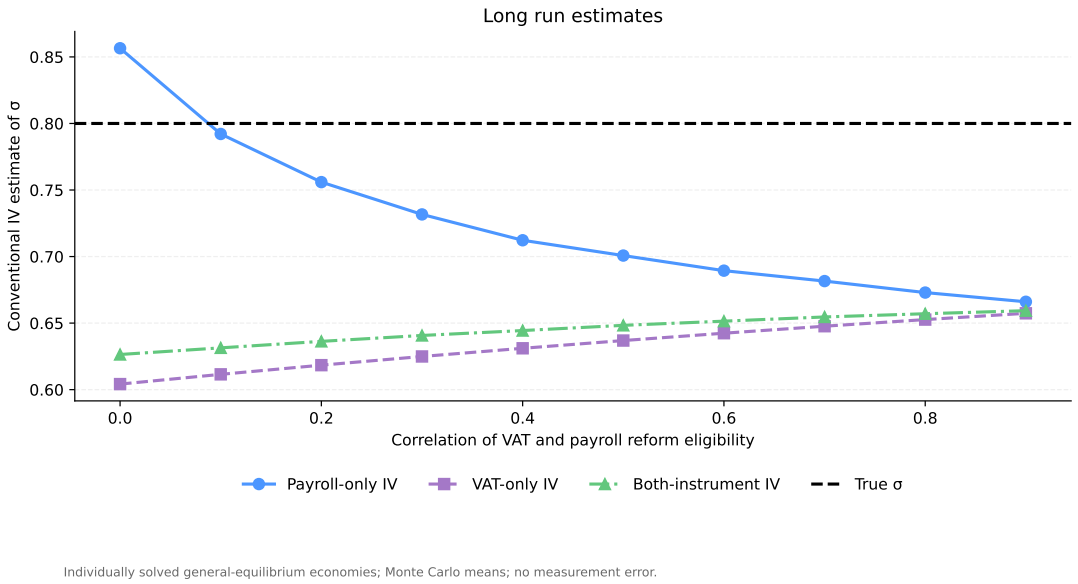
Figure B1 plots conventional IV estimates of capital–labor substitution against the correlation between VAT and payroll tax treatment assignments. The estimate using the payroll tax instrument decreases as the correlation rises, falling below the true elasticity of 0.8 at a correlation of 0.1. Estimates using the VAT instrument or both instruments remain below the true value throughout.

Table B2: Summary Statistics of Key Variables and Moments in DGP3: Frictional economy with VAT and payroll-tax reforms

	Pre-reform steady state			$t = 0$			$t = 10$			Post-reform steady state		
Variable	Mean	SD	$P_{90} - P_{10}$	Mean	SD	$P_{90} - P_{10}$	Mean	SD	$P_{90} - P_{10}$	Mean	SD	$P_{90} - P_{10}$
Investment rate I/k	0.096	0.186	0.419	0.121	0.217	0.544	0.099	0.204	0.514	0.100	0.204	0.515
Fraction of positive investment	0.496	—	—	0.590	—	—	0.571	—	—	0.573	—	—
Log capital $\log k$	0.388	0.825	2.121	0.388	0.825	2.121	0.426	0.880	2.269	0.423	0.882	2.269
Log labor $\log n$	-0.588	1.084	2.788	-0.590	1.086	2.797	-0.633	1.125	2.899	-0.634	1.126	2.899
Log revenue $\log(py)$	-0.160	1.126	2.896	-0.161	1.126	2.900	-0.171	1.164	2.995	-0.172	1.165	2.998
Worker compensation wn	0.880	1.283	1.856	0.859	1.254	1.818	0.894	1.376	1.915	0.894	1.379	1.915
Total labor cost $(1+\tau)wn$	0.880	1.283	1.856	0.880	1.283	1.859	0.914	1.404	1.957	0.914	1.407	1.957
Labor-cost share $(1+\tau)wn/(py)$	0.575	0.040	0.104	0.575	0.040	0.105	0.578	0.040	0.103	0.578	0.040	0.103
Log capital/labor $\log(k/n)$	0.976	0.515	1.311	0.977	0.517	1.318	1.059	0.507	1.286	1.058	0.506	1.286
Leverage b'/k'	0.105	0.054	0.045	0.105	0.052	0.045	0.098	0.074	0.060	0.098	0.075	0.060
Fraction of borrowing	0.961	—	—	0.962	—	—	0.944	—	—	0.944	—	—

Note: Each column summarizes the cross-sectional population at that date. SD is the population standard deviation. $P_{90} - P_{10}$ is the 90–10 percentile range; percentiles use the inverse population CDF. Worker compensation is the firm's wage bill wn. Total labor cost is the firm's full employer wage bill $(1+\tau)wn$, including employer payroll tax, not the wage per worker. The leverage is next-period net debt divided by next-period capital, b'/k'. The impact date $t = 0$ uses inherited pre-reform capital and the new policy. The $t = 10$ distribution uses ten transitions under policies $0, \ldots, 9$, and date-10 outcomes. Both steady states use their solved invariant distributions, not averages over panel dates.

Figure B1: Conventional IV Estimates under Correlated Policy Treatments



Individually solved general-equilibrium economies; Monte Carlo means; no measurement error.

Our baseline Monte Carlo estimates use long-run responses to policy shocks. Since empirical data often cover only a limited period after a reform, we also estimate the parameters using observations from the first 20 post-reform periods. Tables B3 and B4 report these shorter-horizon results.

C Details on Empirical Application

C.1 Data and Variable Construction

Our main data source is the 2004 Chinese Industrial Census (CIC). The dataset contains approximately 1.38 million firms, including 1.26 million manufacturing firms. Approximately 254,000 firms are classified as above-scale enterprises. In 2004, this category included all state-owned industrial enterprises and non-state industrial enterprises with annual sales of at least RMB 5 million. The census reports firm location, industry, year of establishment, and ownership, as well as gross output, value added, sales, intermediate inputs, employment, wages, fixed assets, total assets, depreciation, and liabilities. In the original dataset, median employment is 20 workers and median annual sales are RMB 1.9 million.

We follow Brandt, Kambourov and Storesletten (2026) in constructing our key vari-

Table B3: Monte Carlo Experiments, Periods 1-20

	$\varepsilon = 5$: substitution between goods varieties			$\sigma = 0.8$: substitution between capital and labor			
Meas. error	(1) Conventional revenue-cost	(2) R Moment 2 IVs	(3) 2 Moments GMM+ 1 IV	(4) Conventional IV	(5) L Moment known ε + 1 IV	(6) R Moment 2 IVs	(7) 2 Moments GMM+ 1 IV
A. DGP1: frictionless economy with payroll tax shock							
0%	5.000	—	5.025	0.795	0.796	—	0.795
	(0.000)		(0.357)	(0.020)	(0.017)		(0.016)
5%	4.996	—	7.292	0.795	0.923	—	0.939
	(0.050)		(25.005)	(0.039)	(0.096)		(0.073)
10%	4.997	—	7.633	0.795	0.958	—	0.952
	(0.073)		(7.165)	(0.052)	(0.118)		(0.088)
B. DGP2: Frictional economy with payroll tax shock							
0%	4.645	—	5.070	1.129	0.798	—	0.797
	(0.012)		(0.482)	(0.059)	(0.013)		(0.013)
5%	4.544	—	6.082	1.129	0.871	—	0.903
	(0.017)		(1.417)	(0.063)	(0.054)		(0.040)
10%	4.449	—	8.221	1.129	0.904	—	0.928
	(0.021)		(20.559)	(0.067)	(0.066)		(0.050)
C. DGP3: Frictional economy with VAT reform and payroll tax shock							
0%	4.645	5.048	5.046	0.537	0.799	0.799	0.800
	(0.012)	(0.326)	(0.260)	(0.017)	(0.008)	(0.020)	(0.013)
5%	4.544	5.112	4.666	0.537	0.871	0.801	0.893
	(0.018)	(0.682)	(0.356)	(0.018)	(0.031)	(0.038)	(0.033)
10%	4.450	5.182	4.578	0.537	0.905	0.802	0.916
	(0.021)	(0.977)	(0.907)	(0.019)	(0.039)	(0.050)	(0.043)

Note: L and R moments denote the labor-share and revenue moments, respectively. Each cell reports the Monte Carlo mean across successful replications; parentheses report the average estimator-reported, firm-clustered standard error across these replications. At measurement-error level x, each primitive log-error variance equals $x/100$ times the pooled clean-log variance of that primitive in the estimation sample.

Table B4: Additional Monte Carlo Experiments, Periods 1-20

	$\varepsilon = 5$: substitution between goods varieties			$\sigma = 0.8$: substitution between capital and labor			
Meas. error	(1) Conventional revenue-cost	(2) R Moment 2 IVs	(3) 2 Moments GMM + 1 IV	(4) Conventional IV	(5) L Moment known ε + 1 IV	(6) R Moment 2 IVs	(7) 2 Moments GMM + 1 IV
A. DGP4: Frictional economy with correlated payroll tax shock and VAT reform							
0%	4.645	5.099	5.057	0.573	0.799	0.799	0.800
	(0.012)	(0.391)	(0.316)	(0.016)	(0.007)	(0.017)	(0.012)
5%	4.544	5.189	4.530	0.573	0.871	0.801	0.894
	(0.018)	(0.830)	(0.389)	(0.017)	(0.028)	(0.034)	(0.028)
10%	4.449	5.294	4.412	0.573	0.903	0.802	0.915
	(0.021)	(1.215)	(0.751)	(0.017)	(0.034)	(0.044)	(0.037)
B. DGP5: Frictionless economy with VAT reform							
0%	4.564	—	5.381	0.154	0.799	—	0.776
	(0.006)		(0.674)	(0.008)	(0.030)		(0.030)
5%	4.529	—	4.869	0.155	0.886	—	0.888
	(0.030)		(2.165)	(0.013)	(0.120)		(0.104)
10%	4.498	—	5.933	0.155	0.925	—	0.957
	(0.043)		(44.791)	(0.016)	(0.157)		(0.161)

Note: L and R moments denote the labor-share and revenue moments, respectively. Each cell reports the Monte Carlo mean across successful replications; parentheses report the average estimator-reported, firm-clustered standard error across these replications. At measurement-error level x, each primitive log-error variance equals $x/100$ times the pooled clean-log variance of that primitive in the estimation sample. The correlation between treatment assignments in Panel A is 0.3.

ables. We measure py using value added. Reported value added and intermediate inputs are primarily available for above-scale firms and state-owned enterprises (SOEs); only 1.1% of below-scale firms report these variables. We therefore calculate a value-added-to-gross-output ratio for each two-digit industry–ownership cell using firms with available data. Assuming that this ratio is common within each cell, we construct value added as gross output multiplied by the corresponding ratio. We apply this procedure to all firms, including those with reported value added.

We measure k using the real capital stock. Starting from firms' reported fixed assets at original purchase prices in 2004, we reconstruct initial capital and subsequent nominal investment flows using establishment years and capital growth rates for the corresponding province–industry cells. We deflate initial capital and reconstructed investment flows using the investment price index provided by Brandt, Rawski et al. (2008) and accumulate them through the perpetual-inventory method, assuming an annual depreciation rate of 9%. This method is also used in Brandt, Van Biesebroeck and Zhang (2012) and Hsieh and Song (2015).

We measure labor input n using the annual average number of employees. Annual labor compensation wn is the sum of reported wages and employee benefits. When these components are unavailable, we use total labor remuneration reported in the abbreviated census form. We construct the labor share as the ratio of labor compensation to value added, $s = wn/py$.

We classify firms as state-owned if the census identifies them as being under absolute or relative state control. Firms in all other control categories are classified as non-state-owned.

We retain only manufacturing firms and remove duplicate observations. After constructing the key variables py, k, n, and wn,, we exclude observations with missing or negative values for any of these variables. In the baseline analysis, we follow Oberfield and Raval (2021) and exclude observations in the top or bottom 1% of the value-added-to-cost ratio distribution. As a robustness check, Appendix C.2 reports estimates obtained by instead excluding observations in the top or bottom 1% and 5% of each key variable's distribution. Our main estimates are robust to this alternative sample restriction. For GMM estimation, we restrict the markup to $\mu \in [1, 2]$. To ensure that

the labor-share moment satisfies $\mu s < 1$ throughout this range, we exclude observations with labor shares of 50% or higher. The final estimation sample contains 591,528 firms. Table C5 reports summary statistics of key variables for the full census sample and the estimation sample.

Table C5: Summary Statistics

	Full Census			**Analysis Sample**		
Observations	**N=1,258,471**			**N = 591,528**		
Variable	**Mean**	**SD**	**Median**	**Mean**	**SD**	**Median**
Employment	64.06	341.88	20.00	64.17	410.61	20.00
Firm age	6.44	8.04	4.00	6.44	6.72	4.00
Sales	15.04	269.14	1.90	22.51	364.35	3.12
Gross output	15.40	262.50	2.00	23.08	353.61	3.26
Value added	4.07	69.51	0.52	6.14	95.05	0.86
Labor compensation	0.86	10.10	0.16	0.90	13.16	0.17
Real capital stock	5.67	135.07	0.43	7.59	187.24	0.54
Total assets	14.38	229.23	1.26	19.11	306.63	1.60
Labor share	1.152	69.289	0.368	0.243	0.122	0.230
SOE (%)	2.97	—	—	1.76	—	—
Above-scale firms (%)	20.21	—	—	24.47	—	—

C.2 Additional Results

Table C6: Capital–Labor Substitution and Demand Elasticities under Alternative Trimming Rules

	Conventional				YZ		
	OLS	IV	OLS	IV	GMM	GMM	GMM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. 1% trimming of four core variables							
σ	0.677	0.517	0.738	0.571	0.619	0.610	0.612
	(0.057)	(0.067)	(0.066)	(0.079)	(0.011)	(0.009)	(0.013)
ε			1.480		5.541	5.175	3.255
			(0.011)		(0.592)	(0.469)	(0.181)
Observations				610,673			
B. 5% trimming of four core variables							
σ	0.684	0.519	0.739	0.560	0.620	0.606	0.606
	(0.053)	(0.061)	(0.063)	(0.075)	(0.013)	(0.010)	(0.014)
ε			1.492		4.120	4.006	3.068
			(0.011)		(0.316)	(0.301)	(0.167)
Observations				527,062			
Labor market controls	No	No	Yes	Yes	No	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	No
Firm age fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province × Industry	No	No	No	No	No	No	Yes

Note: Within each two-digit industry, we exclude observations outside the 1st–99th percentile range in Panel A and the 5th–95th percentile range in Panel B for any of four variables: value added, employment, real capital, and labor compensation. All other specifications are the same as those in Table 3.

Table C7: Substitution Elasticity by Firm Characteristics under Alternative Trimming Rules

	Plant Age		Ownership		Liabilities / Assets	
	Old	Young	SOE	Non-SOE	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)
A. 1% trimming of four core variables						
Conventional	0.747	0.418	0.354	0.570	0.650	0.543
	(0.072)	(0.090)	(0.102)	(0.080)	(0.082)	(0.086)
YZ GMM	0.591	0.637	0.620	0.609	0.634	0.582
	(0.011)	(0.009)	(0.010)	(0.009)	(0.007)	(0.011)
Observations	278,275	332,398	11,401	599,267	305,289	305,324
B. 5% trimming of four core variables						
Conventional	0.700	0.437	0.349	0.562	0.634	0.554
	(0.068)	(0.086)	(0.099)	(0.075)	(0.081)	(0.078)
YZ GMM	0.585	0.641	0.597	0.607	0.640	0.570
	(0.012)	(0.011)	(0.014)	(0.010)	(0.008)	(0.011)
Observations	244,228	282,834	7,142	519,917	263,396	263,618

Note: Panels use the same four-core-variable trimming as Table C6. All other specifications are the same as those in Table 4.