A summary on labour market power literature

Seiro Ito

2023 年 11 月 20 日 16:04

(労働市場力研究に関する要約)

Seiro Ito

Abstract I summarize the various proposed methods of production approach to markups/markdowns/TFP estimation. I make distrinctions between FOCs of profit maximization and FOCs of cost minimization. The original method of Hall (1998) only uses cost minimization, so do De Loecker and Warzynski (2012). Once factor price markdown is considered, one needs the marginal revenue product of an input so it takes profit maximization FOC. GNR (2020) consider output price markups yet use profit maximization FOC, which can be their limitation.

Summary

- Generally speaking, concentration may be a misleading measure of market power.

- Monopolistic competition of heterogenous firms can result in a negative relationship between market power and concentration.

- A more solid definition of market power is given by output price markups (downstream) or input price markdowns (upstream).
- But using concentration can be OK in some markets: Studies on US/online labour markets show concentration is positively related to markdowns/supression of wages.

- So we can start with concentration, then proceed to wage markdowns.
- Many studies estimate markups/markdowns at the establishment level with financial statement data using the "production approach".

- It only assumes cost minimization of firms.

- Production approach looked very useful, but the criticism of nonidentification results by Gandhi, Navarro, and Rivers (2020).

- Literature has not come up with a method that responds to the criticism.
- For the time being, we can start with the Rubens' or DJ's method to estimate wage markdowns.

Intro

Following Hall (1988)'s industry level estimation, De Loecker and Warzynski (2012) developed the "production approach" to TFPR/markup estimation at the plant/firm level, by using only accounting data.

Their algorithm follows the tradition of inverting a factor demand function to obtain (to "proxy") productivity ω_{jt} (Olley and Pakes 1996; Levinsohn and Petrin 2003; synthesized by Ackerberg, Caves, and Frazer 2015).

There are debates over ACF-DLW method:

- Invertibility of factor demand, aka scalar unobservable ass. and monotinicity ass. (Bond et al. 2021; but see Appendix O.6.3 of Yeh, Macaluso, and Hershbein 2022)
- Nonidentification results when there is a flexible input (materials, energey) in gross output production (GNR)
- Nonidentification results when using deflated revenues in place of physcal quantity (Bond et al. 2021; De Loecker 2021)
- Endogeneity of inputs (or what the valid IVs are, Doraszelski and Jaumandreu 2021)
- Hicks neutral (factor non-augmenting) tech assumption (Demirer 2022; Raval 2023)

Market power and concentration measures

Output market power := Markup, or the difference between price and marginal cost.

Input market power := Markdown, or the difference between factor price and marginal revenue product.

Problem: We do not observe marginal X.

Old method: Structure-conduct-performance literature

Herfindahl-Hirschman index (HHI, a summary measure of market shares) was used as a proxy

Problem: Monopolistically competitive models with heterogenous firms can predict: concentration$\uparrow$ $\sim$ markup$\downarrow$

Current method

Use markups (and markdowns) by obtaining marginal X in some way

However, the old method may be valid in some markets. A positive relationship between market power and concentration:

Employment share and markdowns (Yeh, Macaluso, and Hershbein 2022)

Employment effects of minimum wages (if small, originally large markdown) and HHI (Azar et al. 2019)

Job application elasticity (if small, large power) and HHI (Azar, Marinescu, and Steinbaum 2019)

Earnings or wages (low) and HHI (Azar et al. 2020; Benmelech, Bergman, and Kim 2022)

Income inequality and HHI (Rinz 2022)

Slower wage growth and HHI (by hospital mergers) (Prager and Schmitt 2021)

DLW method on markup

The point: Marginal cost can be obtained from a cost minimization problem.

Markup formula

Production function with a measurement error ϵ (subscripts i, t are suppressed)

$$Y = Q^* e^{\epsilon}, \quad Q^* \leqslant F(L, K, M) e^{\omega}.$$

Cost minimization on variable input $V = L, M$:

$$\mathcal{L} = P^V V + rK + \lambda[Q^* - F(L, K, M)e^{\omega}].$$

FOC:

$$P^V - \lambda F_V e^{\omega} = 0 \quad \lambda = \frac{P^V}{F_V e^{\omega}}.$$

Note that λ is a measure of marginal cost. Markup $\mu = \frac{P}{MC}$ is[1]

$$\begin{aligned}
\mu = \frac{P}{\lambda} &= \frac{PF_V e^{\omega}}{P^V}, \\
&= \frac{PF_V e^{\omega} YV}{P^V V F e^{\epsilon+\omega}}, \\
&= \frac{PY}{P^V V} \frac{F_V V}{F} e^{-\epsilon}, \\
&= \frac{\theta^V}{S^V} e^{-\epsilon}, \quad S^V \equiv \frac{P^V V}{PY}, \ \theta^V \equiv \frac{F_V V}{F} = \frac{\partial F}{\partial V} \frac{V}{F}.
\end{aligned}$$

Once we back out ϵ, we get μ. With subscripts, firm specific markup $\mu_{it} = \frac{\theta^V_{it}}{S^V_{it}} e^{\epsilon_{it}}$ is affected by $\hat{\epsilon}_{it}$. Need to get rid of it.

Factor demand inversion

DLW suggests to invert the log factor demand $m = m(k, l, \mathbf{p}, \omega)$ to substitute with $\omega = h(k, l, m, \mathbf{p})$,[2] so log production is

$$y = f(k, l, m, \mathbf{p}) + h(k, l, m, \mathbf{p}) + \epsilon = \phi(k, l, m, \mathbf{p}) + \epsilon,$$

where

$$\phi(k, l, m, \mathbf{p}) = f(k, l, m, \mathbf{p}) + \omega.$$

Then DLW suggests to regress y on polynomials on $k, l, m, \mathbf{p}$ to get $\hat{\phi}$, and compute

$$\hat{\epsilon} = y - \hat{\phi}.$$

In inverting the factor demand, one needs:

1. Scalar unobservable assumption: There is only one latent factor ω in demand equation m.

2. Monotonicity assumption: The demand function is monotonic in ω.

Monotonicity ass. is not too strong. But scalar unobservable ass. is.

Other implicit assumptions:

3. There are no firm fixed effects α_i (or $\alpha_i = 0$).[3]
4. Output demand conditions which affect input demand differently by firms are all controlled by the observables.

DLW acknowledges the strong imposition of 4. Covariates $\mathbf{z}_{it}$ added to control for output market conditions are:

[1] Following the notation convention of the literature, I use S^V, not α^V as in this paper, for factor revenure share of V.

[2] If, for example,

$$m = a_1 k + a_2 l + a_3' \mathbf{p} + \omega,$$

then:

$$\omega = \underbrace{m - (a_1 k + a_2 l + \mathbf{a}_3' \mathbf{p})}_{\equiv h(k,l,m,\mathbf{p})}.$$

[3] But time-variant, partly random heterogenous productivity ω_{it} is incorporated. Not sure if this is a major drawback...

- DLW: Export status, lagged inputs in $h(k, l, m, \mathbf{p})$.[4]

- De Loecker et al. (2016): Location dummies, output prices, product dummies, market shares, input prices, export status, input tariffs, output tariffs (p.466).

- De Loecker (2011): Product information and import quota protection (assumed to be exogenous to individual firms).

Some can be endogenous to production but no discussions in DLW on IVs on firm state variables $\mathbf{z}_{it}$.

Algorithm (DLW 2012) on $y_{it} = f(\mathbf{x}_{it}, \beta) + \epsilon_{it} + \omega_{it}$ with a choice of f.

1. Get $\hat{\phi}_{it}$ by regressing y_{it} on polynomials in $k_{it}, l_{it}, m_{it}, \mathbf{p}'_{it} = \mathbf{x}'_{it}, \mathbf{p}'_{it}$.

2. Impute $\hat{\omega}_{it}(\beta) = \hat{\phi}_{it} - f(\mathbf{x}_{it}, \beta)$ and $\hat{\omega}_{it-1}(\beta) = \hat{\phi}_{it-1} - f(\mathbf{x}_{it-1}, \beta)$.

3. Assume $\omega_{it} = g(\omega_{it-1}) + \xi_{it}$ where g is a polynomial of choice in ω_{it-1}. Get $\hat{\xi}_{it}(\beta) = \hat{\omega}_{it}(\beta) - g\left\{\hat{\omega}_{it-1}(\beta)\right\}$.

4. Estimate β with GMM using $\varepsilon[\hat{\xi}_{it}(\beta)v_{it}] = 0$ where v_{it} is a flexible input chosen by firm i before observing ξ_{it}. Use additional moment conditions (e.g., capital as a predetermined input $v_{it} = k_{it}$) to gain efficiency.

5. Bootstrap to estimate covariance matrix: Set B as number of bootstrapping. Randomly sample firms with replacement for its entire lifespan. Do 1.-4. for B runs.

Comments on DLW algorithm:

- It takes at least 2 period panel data (vs. dynamic panel takes 3 periods).
- Scalar unobservable (= only 1 latent factor) is crucial.
- Materials costs: Cost of Goods and Services (COGS) in US manufacturing data. Some procured goods may be used for fixed investments or non-production purposes, so it is not a perfect measure.
- Formula is not new and inversion is done by OP. DLW (p.2444) notes their contribution is being able to recover firm-specific estimates of markup while providing consistent estimates of output elasiticities and allowing some inputs (capital, in some cases labour) to face adjustment costs.
- One of OP's motivation for factor demand inversion is "preserving the typically enormous cross-sectional variation for the sake of identification" (De Loecker and Syverson 2021).

Comparision with the dynamic panel (Blundell and Bond) estimator.

[4] DLW 2012 mention export status (p.2446), lagged variable inputs as IVs for current inputs (Appendix p.5), and is not explicit on what variables are used in $\mathbf{z}_{it}$.

Table 2—Model Estimates

	Ordinary least squares		Dynamic panel	
	Est.	SE	Est.	SE
Panel A. Production function				
Output elasticity of labor	0.563	0.082	0.532	0.147
Output elasticity of capital	0.569	0.066	0.630	0.105
Scale parameter	1.132	0.044	1.162	0.060
R^2	0.91		0.92	
Observations	1,130		849	
Panel B. Leaf price markdown				
Average	2.934	0.414	2.904	0.442
Median	2.134	0.066	2.126	0.079

Notes: Panel A reports the estimated output elasticities using both OLS and the dynamic panel estimator. Panel B contains the leaf markdown moment estimates. Standard errors are block-bootstrapped with 200 iterations.

- BB estimates tend to be noisier. See Table 2 from Rubens (2023).
- Because it double-differences out the confounds as it poses weaker assumptions on the correlation between unobservables and covariates $\varepsilon[\xi_{it} - \xi_{it-1} + (\epsilon_{it} - \rho\epsilon_{it-1}) - (\epsilon_{it} - \rho\epsilon_{it-1})|k_{it-1}, l_{it-1}] = 0$ (see 2.3.1 of Ackerberg 2023). This eliminates the need to invert a demand to net out ϵ_{it}, or of the scalar unobservable ass.
- Weaker statistical power leads to the use of additional moment conditions ("system GMM" estimator).[5]
- BB assumes AR(1) in ω_{it}, not polynomials (as under first-order Markov).
- BB's double-differencing eliminates the firm fixed effects α_i while DLW assumes $\alpha_i = 0$.

[5] But the economic validity of additional moment conditions can be ambiguous.

YMH method on markdown

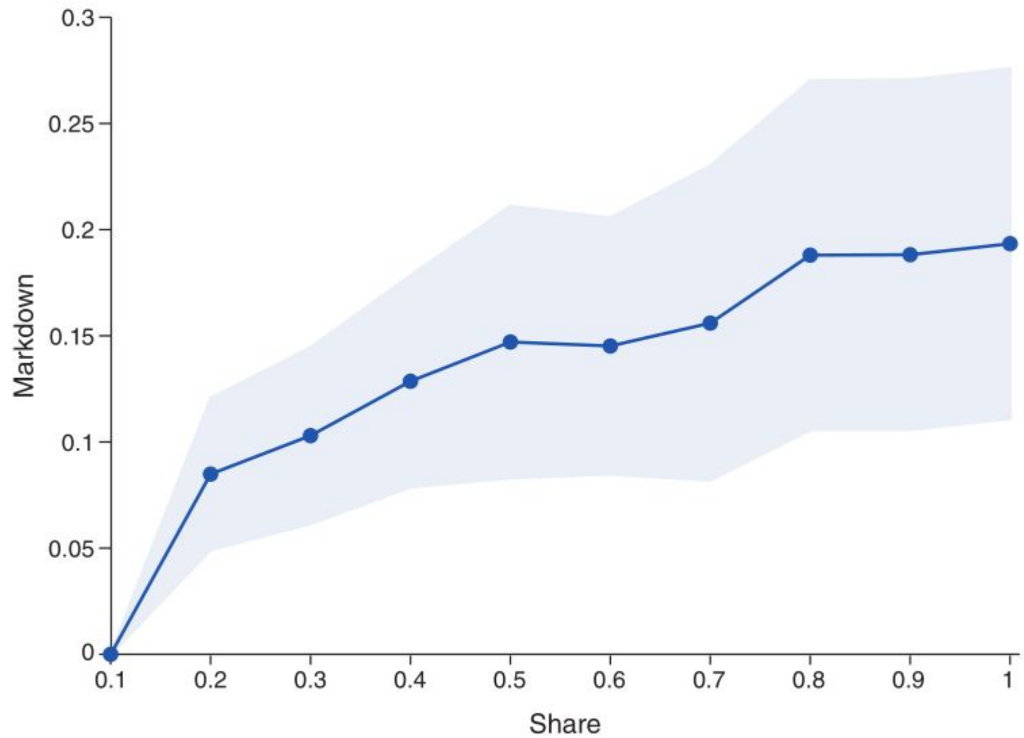


Figure 1. Markdown-Size Relationship

Notes: The figure shows point estimates and 95-percent confidence intervals of plant-specific markdowns on size (as measured by employment share) indicators, controlling for indicators for plant age and industry, as well as state and year fixed effects. The omitted group is the smallest size indicator, so coefficients reflect deviations relative to this baseline. The indicator labeled "0.1" is equal to unity for those plants with employment shares $s \in (0, 0.1]$. Other indicators are defined similarly. Standard errors are clustered at the industry level.

Source: Authors' own calculations from ASM/CM data in 1976–2014.

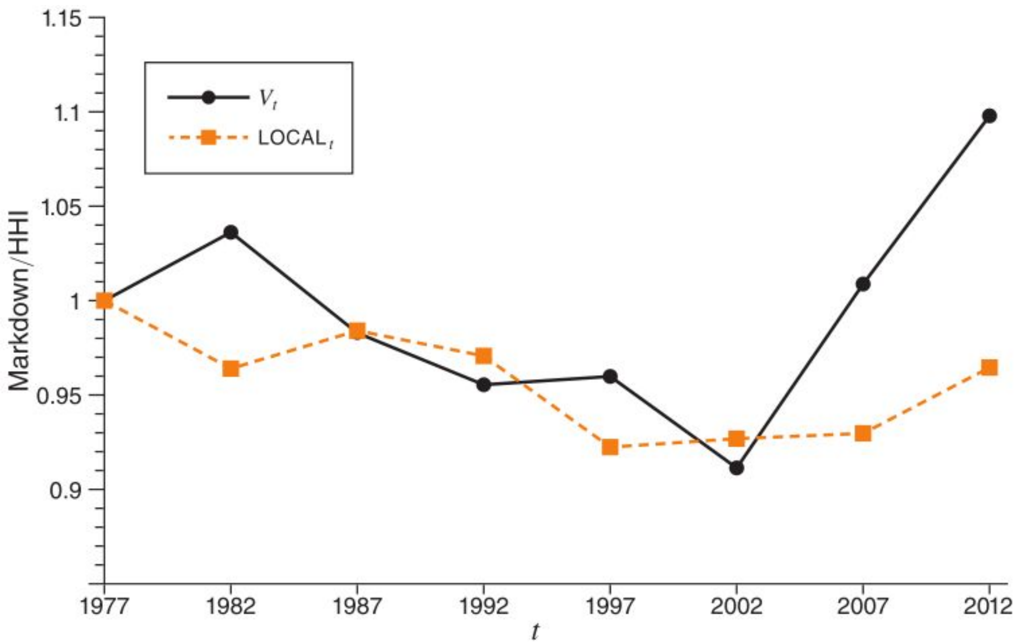


FIGURE 6. AGGREGATE MARKDOWN AND LOCAL CONCENTRATION, 1977–2012

Notes: The solid black line shows the time series for the aggregate markdown as in (14), and the dashed orange line shows the time series of local employment concentration as in (18). Both are normalized to their initial respective values in 1977.

Source: Authors' own calculations from quinquennial CM data from 1977–2012.

Yeh, Macaluso, and Hershbein (2022) showed markdown can be estimated with a formula similar to DLW.

US all industry aggregate markdown decreased (1977-2002) but rose after 2002. Markdown is increasing in the local concentration measure (HHI).

Profit maximization (with a single input)

$$\max_{L} \quad R(L) - w(L)L$$

FOC

$$R'(L) = w'(L)L + w(L) = \left(\frac{w'(L)L}{w(L)} + 1\right) w(L) = \left(\varepsilon_s^{-1} + 1\right) w(L)$$

where ε_s^{-1} is an inverse elasticity of input L supply (or input price elasticity of input supply function). This gives the markdown ψ^L:[6]

$$\psi^L \equiv \frac{R'(L)}{w(L)} = \varepsilon_s^{-1} + 1.$$

(Variable) Cost minimization (with a single input)

$$\mathcal{L} = w(L)L + \lambda[Q^* - F(L; \mathbf{V}, \omega)].$$

FOC

$$w'(L)L + w(L) = \lambda F_L \quad \Leftrightarrow \quad \underbrace{\frac{w'(L)L}{w(L)} + 1}_{=\varepsilon_s^{-1}+1} = \frac{\lambda F_L}{w(L)}.$$

[6]How much more revenue is generated with an additional input cost. > 1 under monopsony/oligoposny.

Then

$$\psi^L = \varepsilon_s^{-1} + 1 = \frac{\lambda}{P}\frac{F_L L}{Y}\frac{PY}{w(L)L} = \mu^{-1}\frac{\theta^L}{S^L}.$$

Or, with any other flexible input $V = K, M$ and markup μ^V derived from V, input price markdown for L is given by:[7]

$$\psi^L = \frac{\left(\frac{\theta^L}{S^L}\right)}{\left(\frac{\theta^V}{S^V}\right)}.$$

For implementation, Yeh et al. (2022) use the DLW algorithm.

Data

Census of Manufactures

Establishment level data of all manufactures in 1976-2014, collected in years ending with 2 and 7.
Revenues and inputs (capital, labor, materials, energy) in monetary units.

Census Bureau

NBER-CES Manufacturing Database

Annual industry-level data from 1958-2018 on output, employment, payroll and other input costs, investment, capital stocks, TFP, and various industry-specific price indexes.

National Bureau of Economic Research

Annual Survey of Manufactures

Establishment level data of all manufactures in 1976-2014, rotating panel collected in every year.
Large plants are sampled with near certainty, small plants are sampled less frequently.

Census Bureau

Rubens' method on markup and markdown

Nonsubstitutability of materials

Production with Leontief in (raw) materials (tobacco leaves)[8] [9]

$$Q^* \leqslant \min\left\{\beta^M M, H(L,K)e^{\omega}\right\}$$

M is more like a constraint that firm has to keep up with when changing other inputs… Cost minimization of planned production (free of ϵ):

$$\mathcal{L} = W^M(M)M + W^L(L)L + \lambda\left[Q^* - \min\left\{\beta^M M, H(L,K)e^{\omega}\right\}\right].$$

FOC

$$W^{M\prime}(M)M\frac{dM}{dL} + W^M(M)\frac{dM}{dL} + W^{L\prime}(L)L + W^L(L) = \lambda H_L e^{\omega}.$$

Rewriting

$$\underbrace{\left[\frac{W^{M\prime}(M)}{W^M}M + 1\right]}_{\equiv\psi^M} W^M\frac{dM}{dL} + \underbrace{\left[\frac{W^{L\prime}(L)}{W^L}L + 1\right]}_{\equiv\psi^L} W^L = \lambda\underbrace{H_L e^{\omega}}_{\frac{\partial Q^*}{\partial L}}$$

[7] This formula has an extra benefit that it cancels out bias terms of using deflated revenues in place of quantity. See Proposition 2 of Appendix.

[8] ϵ is missing, but it will not be a problem, because, in the end the paper does not invert a factor demand. But for comparability, I add it in below. To be exact with the paper, assume $\epsilon = 0$.

[9] DLW also consider Leontief in materials to which they call the "value added production function" as opposed to "gross output production function" in which they treat materials as substitutable. However, even with VA production function, they subtract the materials as if it is a preprocess of cost minimization and do not consider how materials affect FOCs.

ψ^M is inverse material supply elasticity + 1 and ψ^L is inverse labor supply elasticity + 1. So

$$\psi^M W^M \frac{\frac{dM}{dL}}{\frac{\partial Q^*}{\partial L}} + \psi^L W^L \frac{1}{\frac{\partial Q^*}{\partial L}} = \lambda.$$

In Leontief, any increase in M increases Q^* at the same rate, or at $\frac{Q^*}{M}$:

$$\frac{\partial Q^*}{\partial L}\frac{dL}{dM} = \frac{Q^*}{M} = \frac{Ye^{-\epsilon}}{M}.$$

This gives

$$\psi^M W^M \frac{1}{\frac{Ye^{-\epsilon}}{M}} + \psi^L W^L \frac{1}{\frac{\partial Ye^{-\epsilon}}{\partial L}} = \lambda.$$

Usual substitution gives

$$\begin{aligned}\lambda &= P\psi^M \frac{W^M M}{PYe^{-\epsilon}} + P\psi^L \frac{W^L L}{PY}\frac{1}{\frac{\partial Ye^{-\epsilon}}{\partial L}}\frac{1}{\frac{L}{Y}}, \\ &= P\psi^M S^M e^{\epsilon} + P\psi^L S^L \frac{1}{\theta^L} e^{\epsilon},\end{aligned}$$

or[10]

$$\mu = \left[S^M \psi^M + S^L \psi^L \frac{1}{\theta^L}\right]^{-1} e^{-\epsilon}. \tag{4a}$$

Note: There is no input demand inversion involved up to here.

- To be concrete, DLW's formula does not necessarily imply its implementation (algorithm with inversion). Neither does YMH's or Rubens'.

But we need to deal with ϵ.

Identification strategy "choices"

Design choice

Yeh et al. (2022) estimated markups and markdowns simultaneously.

Rubens (2023) showed this is no longer possible (without further assumptions) once material is nonsubstitutable.

- 1 equation, 6 variables
- 3 unknowns: μ, ψ^L, ψ^M
- 3 knowns (1 estimable: θ^L, 2 observables: S^L, S^M)

Need to come up with a structure that gives values to 2 unknowns.

- $\psi^L = 1$, $\mu_i = \mu = 1$ and exogenously set by wholesalers, estimate ψ^M (main text).
  - Other values of μ in Appendix.
- $\psi^L = 1$, ψ^M is given by a nested logit (over farmer occupations) framework (with many accompanying assumptions) as a function of estimated parameters of occupation share equation, estimate μ (Appendix A, C.1).
  - Noisy estimates.

[10](4a) nests DLW (substitutable $S^M = 0$ and input market competition $\psi^V = 1$), Morlacco (2017, substitutable $S^M = 0$ and input market non-competition $\psi^V > 1$), and De Loecker and Scott (2022, nonsubstitutable $S^M > 0$ and input market competition $\psi^V = 1$) as special cases.

- Others are possible.

Why are $\psi^L = 1$, $\mu_i = \mu$ plausible?

- Wage rates did not change before and after consolidation.

- Wholesaler (state monopoly) unilaterally sets P.

Once we assume uniform exogenous markup and $\psi^L = 1$, we have:

$$\begin{aligned}\mu &= \frac{\beta^L}{S_{it}^L + \beta^L S_{it}^M \psi_{it}^M}, \\ (S_{it}^L + \beta^L S_{it}^M \psi_{it}^M)\mu &= \beta^L, \\ \psi_{it}^M &= \frac{\beta^L - \mu S_{it}^L}{\mu \beta^L S_{it}^M}, \\ &= \frac{1}{S_{it}^M}\left(\frac{1}{\mu} - \frac{S_{it}^L}{\beta^L}\right).\end{aligned} \tag{8}$$

Moment condition choice

How do we deal with ϵ?

Proxy variable approach eliminates ϵ in the moment condition by inverting a factor demand to substitute for ω.

Dynamic panel estimator leaves ϵ in the moment condition.

Rubens chooses the latter (proxy variable approach is in Appendix), because:

- Input demand (and its inversion) must control for (markups and) markdowns. This requires an additional structure or additional restrictions on their distributions. This limits the data from expressing heterogenous markups/markdowns.
- With additional parametric restrictions on productivity dynamics in dynamic panel estimator, heterogenous markdowns are obtained. Since markdown estimation is the main objective, this is better suited to the purpose.[11]

Estimation steps

Main results set $\epsilon = 0$ and use dynamic panel estimator.

Robustness check results invert a factor demand to purge ϵ.

One could have retained ϵ and estimated dynamic panel, too.

Estimation without ϵ

1. Assume a production function with quality-unadjusted inputs $\tilde{l}, \tilde{k}$ and quality-unadjusted revenue with a quality adjustment function $a\left(p, w^L\right)$:

$$q = \delta\left(\tilde{l}, \tilde{k}, \beta\right) + a\left(p, w^L\right) + \omega$$

2. Incorporate a consolidation dummy in productivity dynamics and assume AR(1) in redisdual productivity:

$$\omega_{it} = \beta^{z'} \mathbf{z}_{it} + \tilde{\omega}_{it}, \quad \tilde{\omega}_{it} = \rho \tilde{\omega}_{it-1} + \nu_{it}$$

3. By ρ-differencing, form moment conditions $\varepsilon[\hat{\nu}_{it}(\tilde{l}_{it-1}, \tilde{k}_{it}, \tilde{k}_{it-1})] = 0$, etc. and estimate β.

[11] I do not fully understand this argument.

Estimation with ϵ

Appendix C.1. deals with $\epsilon \neq 0$. It assumes leaf as flexible but nonsubstitutable inputs and inverts its demand (termed as "control function approach").[12]

- The "CF Approach" uses Leontief: leaf price per case = leaf costs/output in cases. Use leaf price per case w^M in the leaf demand inversion to control for differences in input price (input quality).
- Input demand is given by $m(p, w^L, w^M, \tilde{\mathbf{x}}, \beta, \mu, \psi^M, \omega)$ in Appendix E[13] and substitute $\mu = \mu(p, \gamma^P, \mathbf{s})$ with $\mathbf{s}$ is a vector of various market shares and γ^P is price elasticity of output demand. γ^P is assumed to be unique and to be subsumed in the intercept term (a strong ass.). ψ^M is derived as $\psi^M(w^M, \mathbf{s})$. Everything is sort of first-order Taylor approximation.

$$
\begin{aligned}
q &= \delta\left(\tilde{\mathbf{x}}, \beta\right) + a\left(p, w^L\right) + \omega + \epsilon, \\
&= \delta\left(\tilde{\mathbf{x}}, \beta\right) + a\left(p, w^L\right) + h\left(p, w^L, w^M, \tilde{\mathbf{x}}, \beta, \gamma^P, \mathbf{s}\right) + \epsilon, \\
&= \phi\left(\tilde{\mathbf{x}}, \beta, p, w^L, w^M, \mathbf{s}\right) + \epsilon.
\end{aligned}
$$

- Productivity is backed out with the use of quality adjustment function:

$$
\omega = \hat{\phi} - \delta\left(\tilde{\mathbf{x}}, \beta\right) - a\left(p, w^L\right).
$$

- Results do not change much (Table A4).

  > "(T)he production approach to markup measurement does not hinge on a particular approach to estimate the production function. Rather it is an approach that delivers markups (and potentially marginal costs) for each individual producer (and time period), by exploiting standard cost minimization of a variable input in production."
  >
  > — De Loecker, 2020

GNR criticism and the proposed solution

Gandhi, Navarro, and Rivers (2020) showed DLW algorithm does not uniquely identify the production function parameters. This is a fatal blow to the "production approach".

Criticism: Nonidentification of elasticity

In the absence of time-series variations in relative prices $\frac{P_t}{P_t^V}$, IV estimation (GMM) using lagged flexible inputs as instruments for current flexible inputs does not give unique values for output elasticity θ^K, θ^L on k, l.

Even with time-series variations in relative prices, IV estimates suffer from weak identification. This is also shown in Monte Carlo by Kasahara and Sugita (2023).

Solution: Use price-taking profit maximization FOC to pin down material demand

But, we are trying to estimate markdowns and markups. Why a price taker?

What shall we do?

- Some papers (e.g., Kasahara and Sugita 2023) are working on the solution.

[12] Control function approach, proxy variable approach, production approach…all of them seem to refer to the same thing.

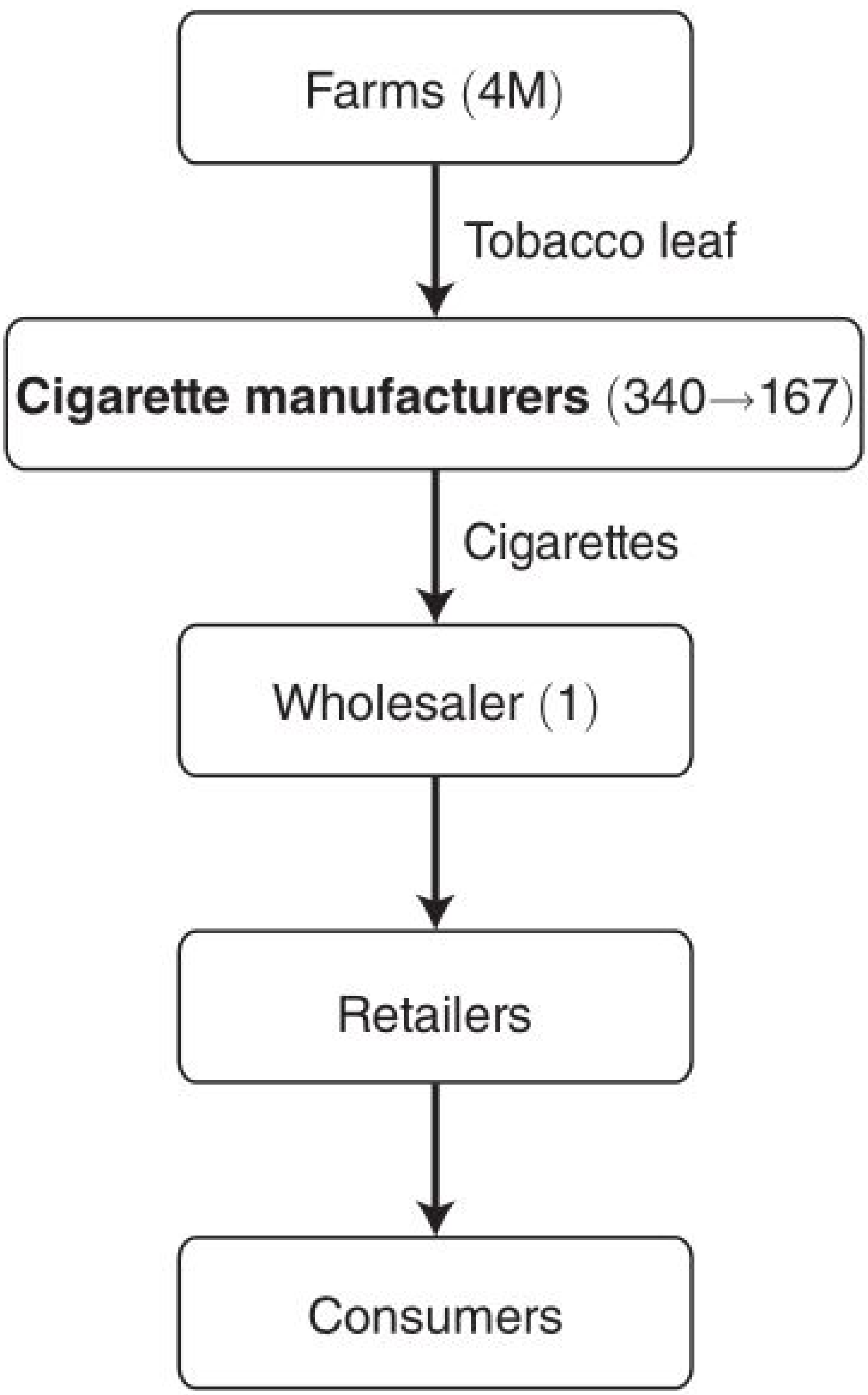
[13] Just derive $Q^*(\cdot)$ from cost minimzation FOCs and form $\frac{Q^*}{\beta^M}$ where β^M is the quantity per case.

- A parametric solution under a Cobb-Douglas production had been proposed (Doraszelski and Jaumandreu 2013).

- We can first do concentration, then proceed to Rubens' or DJ's method, and wait to see if any non-parametric (non-functional form dependent) alternative comes up.

Empirical background

Panel A. Value chain



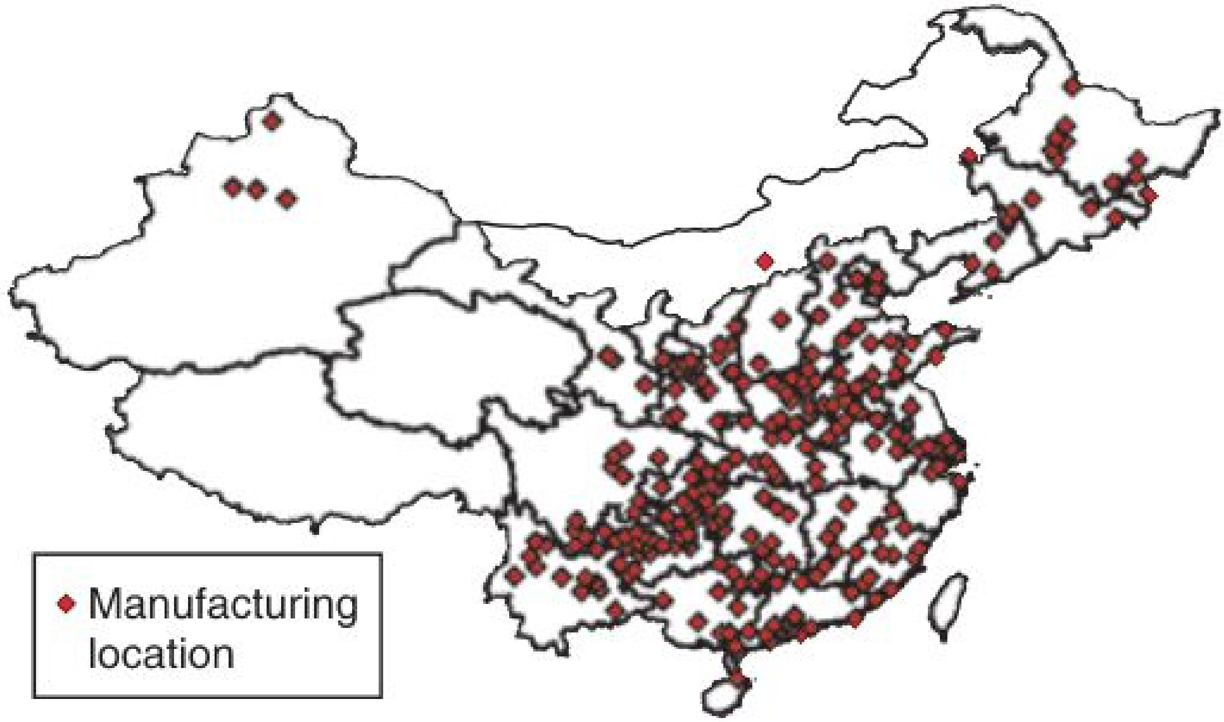
Farmers

- 4 million tobacco farms in China (2003), mostly small scale around .3-.4 ha (FAO 2003).
- Planting is annual, not perennial.
- Farmers choose which purchasing stations or intermediaries (both under particular manufacturers) within county borders to sell the leaves.
- State Tobacco Monopoly Administration (STMA) annually sets leaf prices by quality grades, but manufacturers can flexibly change the effective prices through redefining grades and influencing local STMA management.
- Tobacco farming got less profitable by year, but switching costs and local polical pressures for tax base prevented farmer conversions. Farmers lose cultivation rights once they move to elsewhere, so they are more tied to the location.

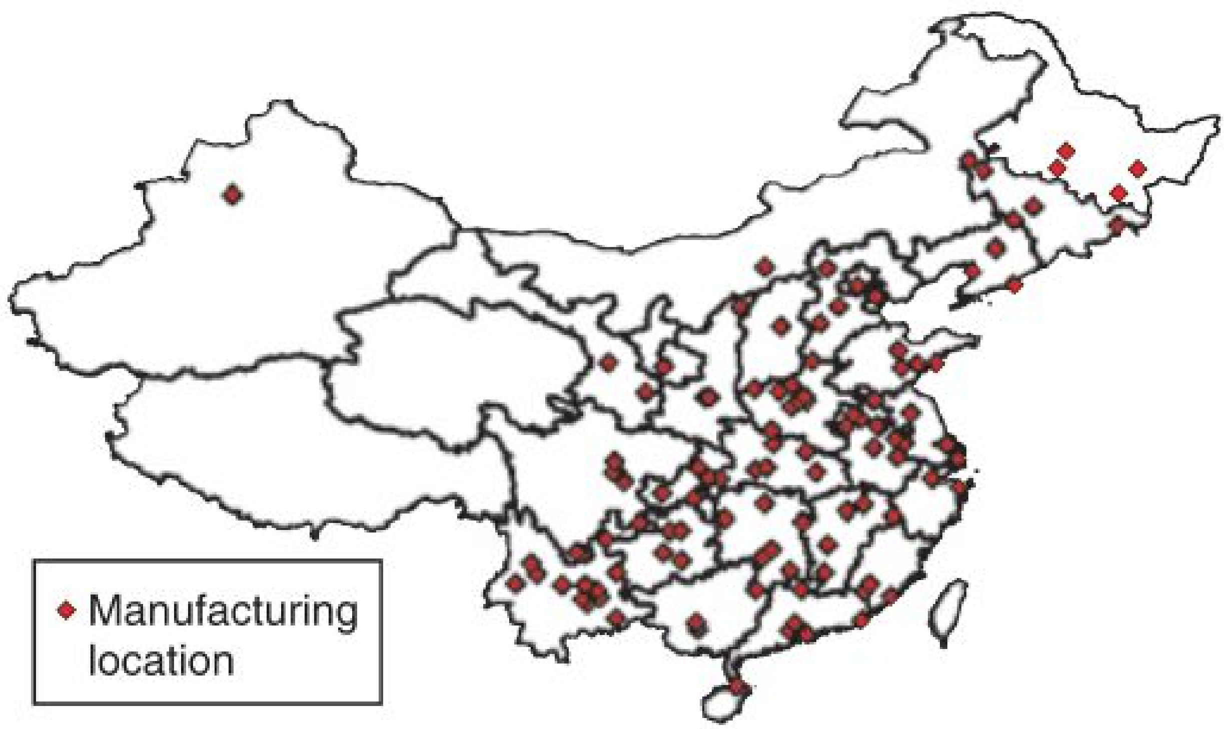
Cigarrete manufacturers

- Variable input costs consist of intermediate inputs (90%, of which 2/3 is tobacco leaves, 1/3 is paper) and labour (10%).
- Operate as separate enterprises although formally belong to Chinese National Tobacco Corporation (CNTC).
- Sell to a (monopsonistic) wholesaler controlled by Chinese National Tobacco Trade Corporation (CNTTC) and State STMA.
- Wholesaler unilaterally sets cigarette factory gate prices and sells across the country.
- Exports count 1% and imports count .2% of total industry revenue in 2019.

Panel B. Manufacturing locations in 1999



Panel C. Manufacturing locations in 2006



Consolidation policy

May 2002, STMA ordered:

- Close down of all (98) SOEs producing less than 100K cigarette cases per year.[14]
- Encouragement to 99 SOEs producing below 300K cases to be mergered with larger manufactures.[15]
- Number of cigarette manufacturers: 340 (1999) → 167 (2006)

Data

Annual survey of industrial firms

Establishment level data of all cigarette manufactures in 1999-2006, non-SOEs with sales greater than 5 million RMB and all SOEs, 470 firms with 2025 observations.
Production and cost data in monetary units.
Including 1132 observations of 257 firms that come with quantity information.

National Bureau of Statistics

2000 Census of Population

County level demographic information.

Harvard Dataverse

Weather data

County level weather

Chinese Meteorological Agency

Brand information data

Brand level cigarette characteristics

O'Connor et al. (2010)

Product information data

Product level characteristics. Quality grades, subsidy.

National Bureau of Statistics

FAOSTAT

Agricultural prices

FAO

UN COMTRADE

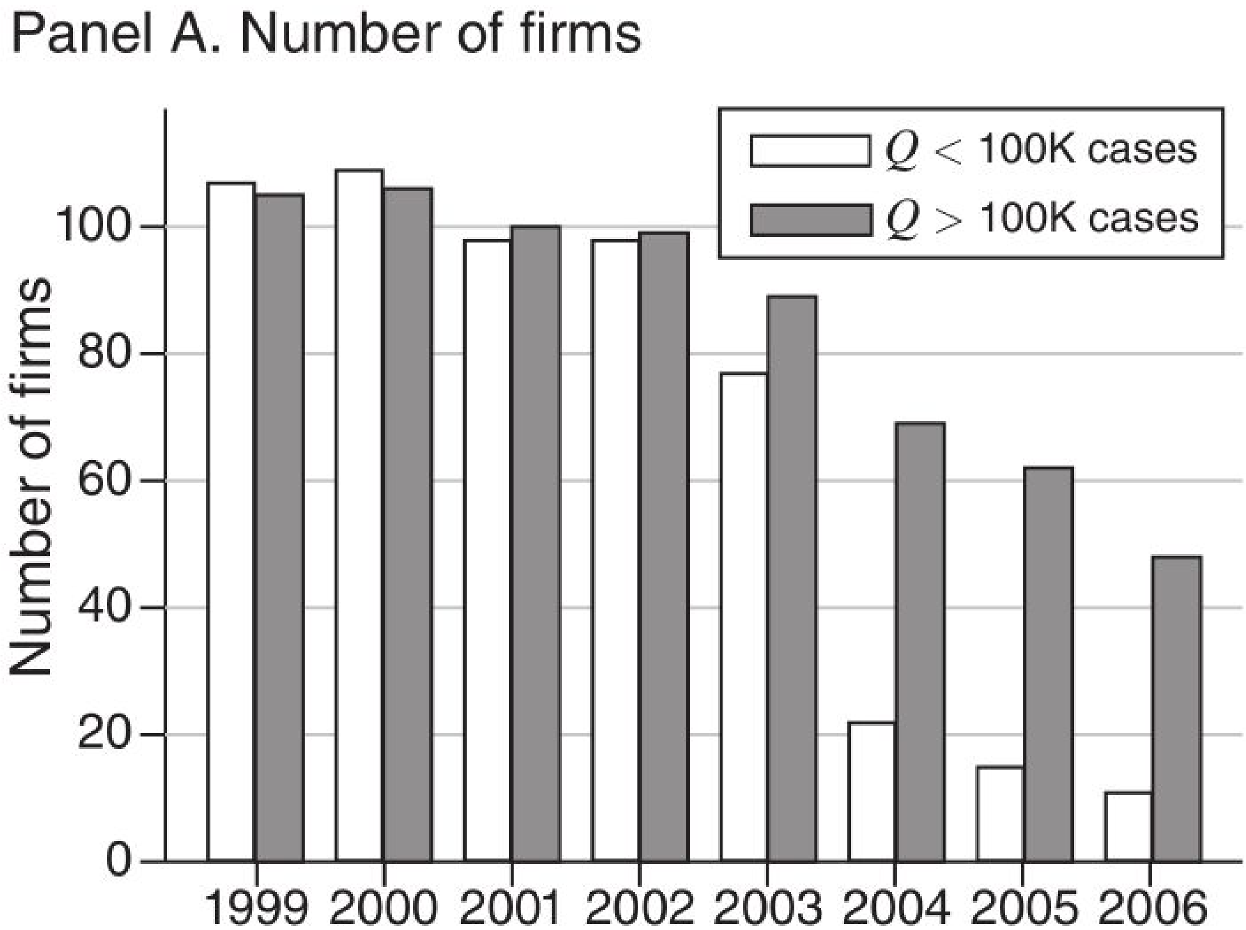
Aggregate trade flows

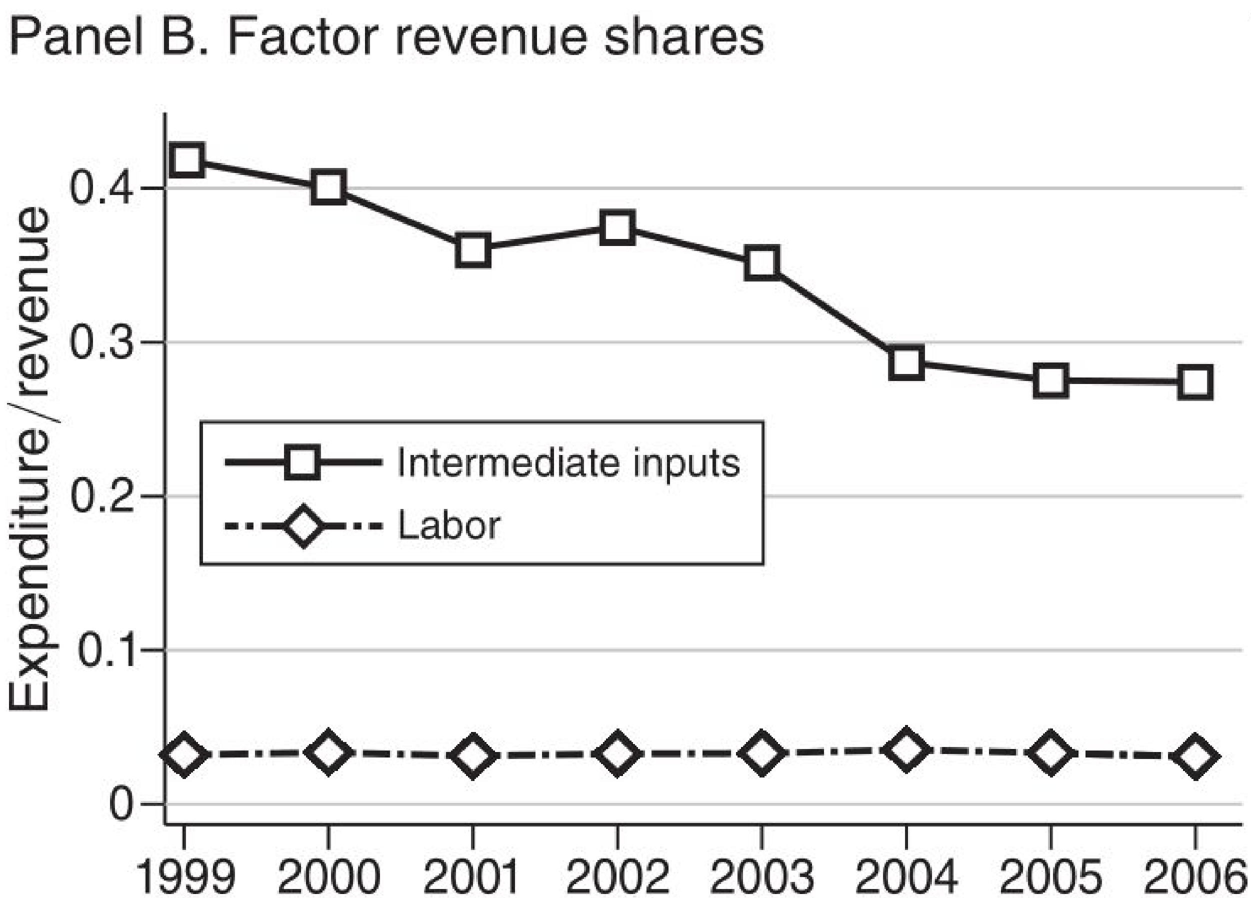
UN

[14] 3 noncomplier firms survived by 2006.
[15] 51 noncomplier firms survived by 2006.

Results

Cost shares





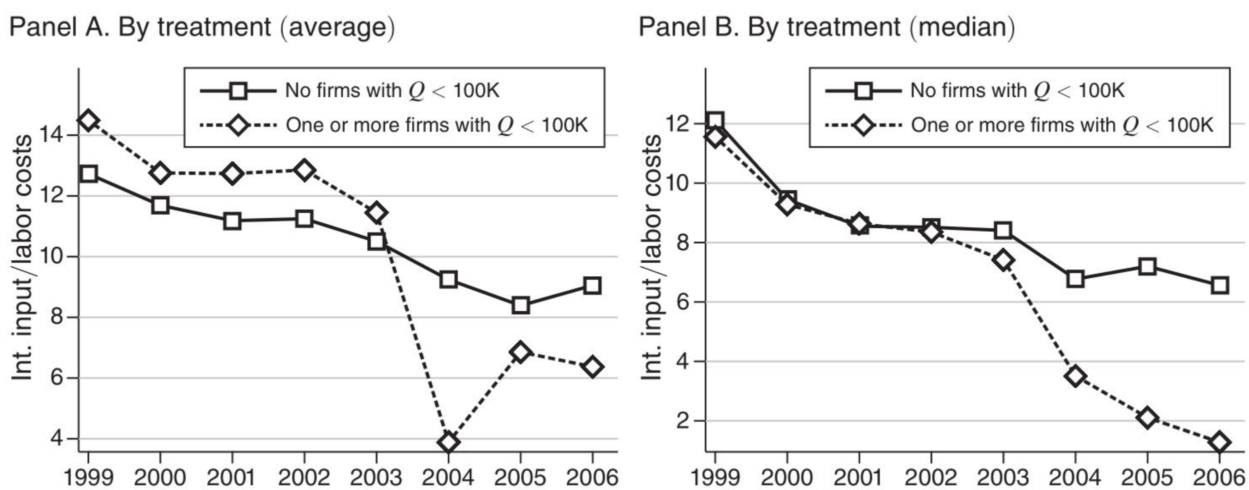


FIGURE 3. RELATIVE COST SHARES AND CONSOLIDATION

Note: Panels A and B compare the average and median ratio of labor expenditure over intermediate input expenditure over time between the consolidation treatment and control groups.

Cost share of the leaves decreased relative to labor among the treated (below 100K cases) relative to the controls (above 100K case firms).

Markdowns and productivity

TABLE 3—CONSOLIDATION TREATMENT EFFECTS

	log(*Markdown*)		log(*Productivity*)			
	Est.	SE	Est.	SE		
Panel A. Markdown and productivity						
Treatment $\times$ $\mathbf{1}\{Year \geq 2002\}$	0.315	0.103	−0.055	0.083		
R^2	0.72		0.88			
Observations	1,123		1,132			
	log(*Agg. TFP*)		log(*Avg. TFP*)		Reallocation	
	Est.	SE	Est.	SE	Est.	SE
Panel B. Allocative efficiency						
Treatment $\times$ $\mathbf{1}\{Year \geq 2002\}$	−0.544	0.166	−0.084	0.135	−0.460	0.106
R^2	0.65		0.33		0.77	
Observations	221		221		221	
	log(*Agg. output*)		log(*Avg. output*)		Reallocation	
	Est.	SE	Est.	SE	Est.	SE
Panel C. Output						
Treatment $\times$ $\mathbf{1}\{Year \geq 2002\}$	−0.485	0.171	0.220	0.154	−0.704	0.090
R^2	0.65		0.48		0.85	
Observations	221		221		221	

Notes: Panel A reports the estimated treatment effects from equation (1) with the logarithms of the markdown ratio and productivity as the dependent variables. Controls are firm fixed effects and a linear time trend. Panel B estimates the effects of the consolidation on log aggregate productivity, weighted by labor usage; log unweighted average productivity; and a reallocation term, all at the province-year level. Panel C reports the effects of the consolidation on log total province-level output, log average output, and the difference between these two. All standard errors are block-bootstrapped with 200 iterations.

Markdown increased (37%).

Manufacturers reduced outputs (38%), thereby reducing leaf demand and prices.

Manufacturers' provincial aggregate productivity[16] decreased (42 %), mostly due to leaf (mis)reallocation to low productivity manufacturers (37 pp), less so to individual TFP growth slowdown.

When markdown is modeled and estimated as farmers' buyer choices (Table A1, A2), markup becomes the estimate. Estimated markup is larger which indicates cigarette price manipulations by manufacturers (with nonmarket power on the wholesaler).

- Markup estimates are noisy.[17]

Robustness checks

- Endogenous ψ^M with farmer choice
- Endogenous firm exit
- Leaf substitutablity in production

[16] Using number of workers as weights. Note the difference with the simple average.
[17] I think this choice could have been author's first choice if the markups and markdowns are more precisely estimated.

- Labor augmeting tech
- Cost share approach
- Leaf market definition
- Leaf definition changes, only large firms, different markup calibration

Ackerberg, Daniel A., Kevin Caves, and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators." *Econometrica* 83 (6): 2411–51. https://doi.org/10.3982/ECTA13408.

Azar, José, Emiliano Huet-Vaughn, Ioana Marinescu, Bledi Taska, and Till Von Wachter. 2019. "Minimum Wage Employment Effects and Labor Market Concentration." *NBER Working Paper Series*. Vol. 26101. National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w26101/w26101.pdf.

Azar, José, Ioana Marinescu, and Marshall Steinbaum. 2019. "Measuring Labor Market Power Two Ways." *AEA Papers and Proceedings* 109 (May): 317–21. https://doi.org/10.1257/pandp.20191068.

Azar, José, Ioana Marinescu, Marshall Steinbaum, and Bledi Taska. 2020. "Concentration in US Labor Markets: Evidence from Online Vacancy Data." *Labour Economics* 66: 101886. https://doi.org/10.1016/j.labeco.2020.101886.

Benmelech, Efraim, Nittai K. Bergman, and Hyunseob Kim. 2022. "Strong Employers and Weak Employees." *Journal of Human Resources* 57 (S): S200–250. https://doi.org/10.3368/jhr.monopsony.0119-10007R1.

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. https://doi.org/10.1016/j.jmoneco.2021.05.004.

De Loecker, Jan. 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity." *Econometrica* 79 (5): 1407–51. https://doi.org/10.3982/ECTA7617.

———. 2021. "Comment on (Un)pleasant ... By Bond Et Al (2020)." *Journal of Monetary Economics* 121: 15–18. https://doi.org/10.1016/j.jmoneco.2021.04.009.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik. 2016. "Prices, Markups, and Trade Reform." *Econometrica* 84 (2): 445–510. https://doi.org/10.3982/ECTA11042.

De Loecker, Jan, and Chad Syverson. 2021. "An Industrial Organization Perspective on Productivity." In *Handbook of Industrial Organization*, 4:141–223. 1. Elsevier.

De Loecker, Jan, and Frederic Warzynski. 2012. "Markups and Firm-Level Export Status." *American Economic Review* 102 (6): 2437–71. https://doi.org/10.1257/aer.102.6.2437.

Demirer, Mert. 2022. "Production Function Estimation with Factor-Augmenting Technology: An Application to Markups." https://demirermert.github.io/Papers/Demirer_production_function%202.pdf.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2013. "R& D and productivity: Estimating endogenous productivity." *The Review of Economic Studies* 80 (4): 1338–83. https://doi.org/10.1093/restud/rdt011.

———. 2021. "Reexamining the de Loecker & Warzynski (2012) Method for Estimating Markups."

Gandhi, Amit, Salvador Navarro, and David A. Rivers. 2020. "On the Identification of Gross Output Production Functions." *Journal of Political Economy* 128 (8): 2973–3016. https://doi.org/10.1086/707736.

Hall, Robert E. 1988. "The Relation Between Price and Marginal Cost in US Industry." *Journal of Political Economy* 96 (5): 921–47.

Levinsohn, James, and Amil Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." *Review of Economic Studies* 70 (2): 317–41.

Olley, G. Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64 (6): 1263–97.

Prager, Elena, and Matt Schmitt. 2021. "Employer Consolidation and Wages: Evidence from Hospitals." *American Economic Review* 111 (2): 397–427. https://doi.org/10.1257/aer.20190690.

Raval, Devesh. 2023. "Testing the production approach to markup estimation." *The Review of Economic Studies* 90 (5): 2592–2611. https://doi.org/10.1093/restud/rdad002.

Rinz, Kevin. 2022. "Labor Market Concentration, Earnings, and Inequality." *Journal of Human Resources* 57 (S): S251–83. https://doi.org/10.3368/jhr.monopsony.0219-10025R1.

Rubens, Michael. 2023. "Market Structure, Oligopsony Power, and Productivity." *American Economic*

Review 113 (9): 2382–2410. https://doi.org/10.1257/aer.20210383.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein. 2022. "Monopsony in the US Labor Market." *American Economic Review* 112 (7): 2099–2138. https://doi.org/10.1257/aer.20200025.