The Size Distribution of Firms and Industrial Water Pollution
A Quantitative Analysis of China

Ji Qi [⋆] Xin Tang [†] Xican Xi [‡]

[⋆]Chinese Academy for Environmental Planning

[†]International Monetary Fund

[‡]Fudan University

October 6, 2019

The views expressed here are those of the authors and do not necessarily reflect the views of the CAEP or the IMF.

Motivation

- Rapid economic growth in China is paired with severe environmental consequences.
  - Is pollution an inevitable consequence of growth or exacerbated by policy distortions and/or market inefficiencies?
- We argue that misallocation across firms reduces output and amplifies industrial pollution by distorting firm size distribution.
- *Policy Implications*:
  - Reducing the distortions firms face increases output while decrease pollution.
  - Environmental regulations could amplify existing market distortions.
- We believe that the findings of our paper could be generalized to
  - air and solid waste pollution;
  - the agricultural sector;
  - other developing countries.

Intuition

- *Hypothesis*: Distortions $\rightarrow$ Size Distribution $\rightarrow$ Industrial Pollution.
  - Larger firms have lower pollution intensity because they use cleaner production and treatment technologies.
  - Distortions limit the expansion of productive firms in China.
  - Compared to an economy with fewer distortions (e.g., the U.S.), less productive but more polluting firms survive.
  - As a result, aggregate output is lower, but industrial pollution is higher.
- Technique effect dominates scale effect.
- *Significance*: Growth and environmental protection are not necessarily a trade-off.

Overview of Results

- Empirical Evidence:
  - Large firms have lower pollution intensity.
  - Correlated distortions limit the expansion of large firms.
- Model:
  - Two sectors (polluting and non-polluting);
  - Firms heterogeneous in productivity, pollution intensity and correlated distortions;
  - Endogenous treatment technology choice with imperfect regulations.
- Qualitative Results:
  - Correlated distortions discourage technology adoption.
  - When would technique effect dominates scale effect.
- Quantitative Exercises:
  - Removing distortions: $Y \uparrow 30\%$, $E \downarrow 20\%$.
  - Strengthening regulations: $Y =$, $E \downarrow 15\%$.

Related Literature

- Misallocation:
  - Restuccia and Rogerson (2008), Guner et al. (2008), Hsieh and Klenow (2009).
  - We consider both the output and pollution.
- Economic Growth and Environment:
  - Grossman and Krueger (1993, 1995), Copeland and Taylor (2004), Levinson (2009).
  - Cherniwchan et al. (2017), Shapiro and Walker (2018).
  - We focus on a theoretical interpretation behind the scale and technique effects.
- Technology Adoption in Developing Countries:
  - Parente and Prescott (1994, 1999), Bustos (2011), Acemoglu et al. (2012).
  - We study the role of size and misallocation.
- Industrial Pollution in China
  - Jiang et al. (2014), Kahn et al. (2015), Cai et al. (2016), Jia (2017).
  - We emphasize the role of economic factors.

Outline

- Empirical Evidence:
  - Data
  - Fact: Pollution Intensity
  - Fact: Employment Distribution
- The Model:
  - Measuring Distortions
  - Analytical Properties
- Quantitative Results:
  - Removing Distortions
  - Strengthening Regulations

1. Empirical Evidence

Data

Data Sources

- National General Survey of Pollution Sources (2007)
  - Firm output, emission and treatment equipment.
- China National Economic Census (2004)
  - Firm output, employment labor compensation, and capital stock.
- Statistics of U.S. Business (2004)
  - Firm size distribution in the U.S.

Data

Measurement

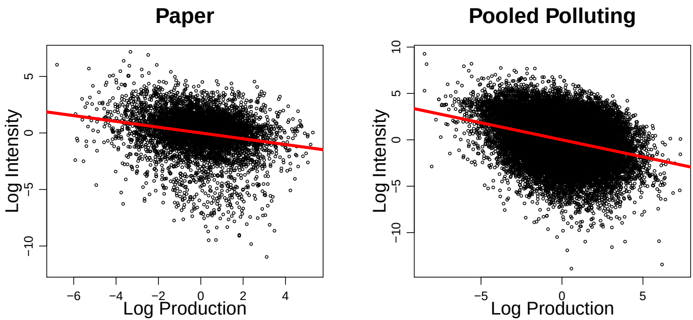
- We use Chemical Oxygen Demand (COD) as the main measure.
  - More than 70% firms in our sample have positive COD emission.
  - COD emission is highly correlated with those of other pollutants.
  - In the Paper industry, $\text{corr}(NH_4^+, COD) = 0.82$, and $\text{corr}(BOD, COD) = 0.94$.
- We focus on the 5 most polluting industries.
  - They account for 77% of total COD emission in China.
- We focus on the key sources.
  - For the polluting industries, they account for 90% of the emission and 80% of the production.

COD Definition | Pollutants

Firm Size and Pollution Intensity

Scatter Plots

- Emission intensity = emission/sales.



▸ Scatter by Industries

Firm Size and Pollution Intensity

Regression Results

- Regression Specification:

$$\log(\text{COD}_i) = \underset{(0.17)}{-3.75} + \underset{(0.01)}{0.63} \times \log(\text{Output}_i) + \mathbf{X}_i \gamma + \varepsilon_i,$$
$$N = 29,019, R^2 = 0.41.$$

- $\mathbf{X}_i$ controls for sectors, provinces and ownership rights.
  - Pollution intensity does not vary systematically with ownership rights.
- As output rises by 1%, emission intensity drops by $(1 - 0.63) = 0.37\%$.

Treatment Technologies and Firm Size

- Treatment efficiency: $1 -$ COD Emitted / COD Generated.
  - For the Paper industry:

Technology	Physical	Chemical	Biological
Adoption Rate	26%	34%	39%
Mean Efficiency	63%	75%	81%
Median Costs	100 (Normalized)	308	923
Median Sales	100 (Normalized)	219	500

- A linear probability model of biological equipment adoption on log-output:

$$y_i = \underset{(0.04)}{0.23} + \underset{(0.001)}{0.05} \times \log(\text{Output}_i) + \mathbf{X}_i \gamma + \epsilon_i.$$

▸ Examples ▸ Additional Tech Specs

Production Technologies and Firm Size

Regression Results

- Large firms also generate less pollution conditional on treatment technology.
- Regression Specifications:

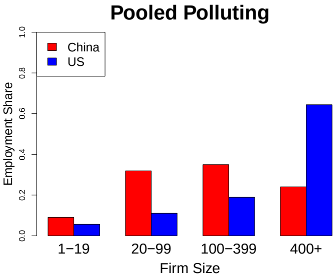
$$\begin{aligned}
\text{Physical:} \quad \log(\text{COD}_i) &= \underset{(0.40)}{-3.50} + \underset{(0.01)}{0.58} \times \log(\text{Output}_i) + \mathbf{X}_i\boldsymbol{\gamma} + \epsilon_i, \\
\text{Chemical:} \quad \log(\text{COD}_i) &= \underset{(0.49)}{-4.59} + \underset{(0.01)}{0.77} \times \log(\text{Output}_i) + \mathbf{X}_i\boldsymbol{\gamma} + \epsilon_i, \\
\text{Biological:} \quad \log(\text{COD}_i) &= \underset{(0.20)}{-4.37} + \underset{(0.01)}{0.67} \times \log(\text{Output}_i) + \mathbf{X}_i\boldsymbol{\gamma} + \epsilon_i.
\end{aligned}$$

Production Technologies and Firm Size

Possible Interpretations

- An example from the *Handbook of Emission Coefficients* by the *Chinese Academy of Sciences*:
  - Two technologies in paper pulp manufacturing use different inputs: bagasse and wood.
  - Bagasse is used mostly by firms with annual production of less than 100 k-tons with COD generation of 140–180 kg per ton.
  - Wood is used mostly by firms with annual production more than 100 k-tons with COD generation of 30–55 kg per ton.
- It could also be that more productive firms simply use less input (Bloom et al. 2010).
- We capture such decrease in intensity in a reduced-form way.

Firm Size Distributions



▸ By Industry

Accounting Exercise

- Question: How will average pollution intensity change if the size distribution is mechanically changed to that in the U.S.?
  - Production level is fixed.
  - The relation between pollution intensity and firm size in China is used.
  - It captures the *Technique Effect* while holding the *Scale Effect* constant.
- Answer: Average pollution intensity would drop by 33%.

Robustness

2. The Model

Why Are We *Not* Done Yet?

- The accounting exercise has several limitations:
  - Only the polluting sector is considered.
  - Aggregate output is fixed.
  - Firm size distribution is mechanically changed to that in the U.S.
  - The firm size and pollution intensity relationship is taken as invariant.
- To better evaluate the environmental consequences of distortions to firm size, we need a model:
  - features two sectors;
  - where distortions affect both aggregate output and pollution;
  - which reveals why firm size is distorted and how it will be changed;
  - and which features treatment technology choice.

Overview of the Model

- A two-sector model with perfect substitution between products.
- Each sector features a Lucas (1978) span-of-control economy.
- Firms differ in productivity and pollution intensity.
- Firms can use treatment equipment to reduce emissions.
  - The cost of treatment equipment is independent of firm size.
  - There are penalties for not using clean technology.
- Correlated distortions that increase with firm's productivity (Restuccia and Rogerson, 2008).
  - Domestic local protectionism, size-dependent tax treatments, etc.
- Steady state analysis.

Household

- A representative household with a continuum of members:

$$\sum_{t=0}^{\infty} \beta^t U(C_t).$$

- Each member is endowed with z units of sector-specific managerial talent.
  - The CDFs are $G^c(z)$ and $G^d(z)$ with a common support $Z \triangleq [0, \overline{z}]$;
  - z is fixed once drawn;
  - A fraction μ of members possess talent in the polluting sector.
- Members can only become entrepreneurs in their designated sector, but are free to enter both sectors as worker.
- The household owns all firms and capital in the economy.

Firms

Non-polluting Sector

- The production function is the same in both sectors:

$$y = F(z, k, l) = z^{1-\gamma}(k^{\alpha} l^{1-\alpha})^{\gamma},$$

  where $0 < \gamma < 1$ is the span-of-control parameter.

- Firms in the non-polluting sector do not discharge pollutants and are not subject to environmental regulations.

Firms

Polluting Sector: Clean Technology and Regulations

- Treatment Technologies:
  - Emissions e are by-product of production:

$$\log\left(\frac{e}{y}\right) = \psi_0^j + \psi_1^j \log y, \quad j = 0, 1,$$

    where $j = 1$ indicates clean technology and $\psi_1^j < 0$.

  - Clean technology requires fixed installation costs k_E.
- Regulations:
  - If a firm using dirty technology is inspected, a fraction ξ of its total profits $\pi(z)$ is confiscated.
  - Confiscated profits are rebated back to household as lump-sum transfers.

Which Distortion Matters?

Identification Assumption

- Let τ_{z_i}, τ_{k_i} and τ_{l_i} be the wedges firm i faces on the product, capital and labor market, the problem of firm i is

$$\pi_i = \max_{k_i, l_i} \left\{ (1-\tau_{z_i}) z_i^{1-\gamma} (k_i^{\alpha} l_i^{1-\alpha})^{\gamma} - (1+\tau_{k_i}) R k_i - (1+\tau_{l_i}) W l_i \right\}.$$

- We use the variations in average factor products to measure these wedges (Hsieh and Klenow, 2009).

- Assuming $\tau_{l_i} = 0$:

$$\frac{y_i}{l_i} = \frac{W}{(1-\alpha)\gamma(1-\tau_{z_i})} \propto \frac{1}{1-\tau_{z_i}},$$
$$\frac{k_i}{l_i} = \frac{\alpha}{1-\alpha} \cdot \frac{W}{(1+\tau_{k_i})R} \propto \frac{1}{1+\tau_{k_i}}.$$

Which Distortion Matters?

Data

- Using the China National Economic Census, we find:

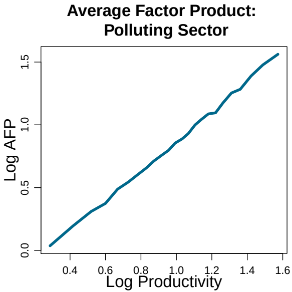
- The firm's problem becomes:

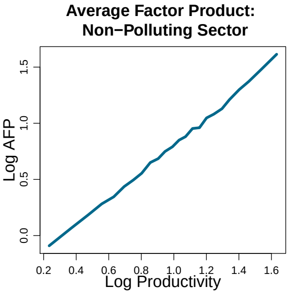
$$\pi_i = \max_{k_i, l_i} \left\{ (1 - \tau_{z_i}) z_i^{1-\gamma} (k_i^{\alpha} l_i^{1-\alpha})^{\gamma} - Rk_i - Wl_i \right\}.$$

- Tax revenues from τ_z is rebated to household as lump-sum transfers.

Correlated Distortions

Polluting and Non-polluting Sectors





- Assume the correlated distortions are the same in both sectors:

$$\tau_z = \max\left\{0, 1 - \phi_0 z^{\phi_1}\right\},$$

where τ_z is progressive when $\phi_1 < 0$.

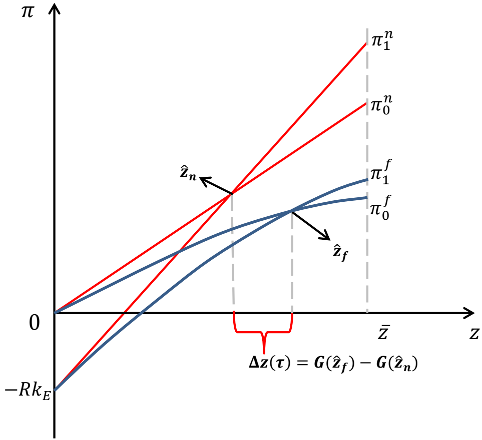
Stationary Equilibrium

- In the equilibrium, there are three thresholds:
  - individuals choose to be manager iff $z > z_c(z_d)$ in each sector,
  - firms in the polluting sector install clean technology iff $z > z_E$.
- In the equilibrium,
  - Household and firms optimize.
  - Labor, capital and product markets clear.
  - Aggregate emission:

$$E = \mu \left[\underbrace{\int_{z_d}^{z_E} e_0(y_z) dG(z)}_{\text{Dirty}} + \underbrace{\int_{z_E}^{\overline{z}} e_1(y_z) dG(z)}_{\text{Clean}} \right].$$

Optimization

Correlated Distortions and Technology Adoption



Algebra

Correlated Distortions and Aggregate Pollution

- Consider a simplified one-sector model with no occupational and technology adoption choice.
- There exists a threshold

$$\hat{\phi}_1 = \frac{\psi_1(1-\gamma)}{1-\gamma(1+\psi_1)},$$

  such that E is decreasing in $|\phi_1|$ when $|\phi_1| < |\hat{\phi}_1|$.
- Intuition:
  - Technique effect dominates when the progressivity of the distortions is low.
  - $|\hat{\phi}_1|$ is increasing in $|\psi_1|$.

3. Quantitative Analysis

Calibration

Calibration Strategy

- *Our Objective*: how firm size distribution affects aggregate pollution.
- Crucial to match the Chinese economy on:
  - firm size distribution;
  - relative distortions faced by large versus small firms;
  - pollution intensity of differently sized firms.
- τ_z and $G(z)$ jointly determine the firm size distribution.
- *Identification*: Let $\theta_l = y/l$, then

$$\left(\frac{z_i}{z_j}\right)^{\phi_1} = \frac{\theta_{l,j}}{\theta_{l,i}},$$
$$l^{1-\gamma} = \Phi(1-\tau_z)z^{1-\gamma}.$$

Calibration

Parameterizations: Exogenously Calibrated

Parameter		Value	Target
Production	δ	0.1000	Depreciation Rate
	α	0.5376	Capital Share 0.5
	μ	0.20	Fraction of firms in the polluting sector
Treatment	ψ_0^0	-3.4144	Physical Intensity-output Elasticity
	ψ_1^0	-0.3636	
	ψ_0^1	-4.3748	Biological Intensity-output Elasticity
	ψ_1^1	-0.3288	

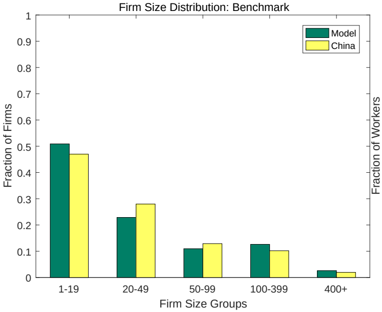
Calibration

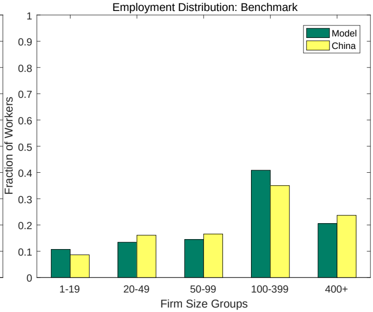
Parameterizations: Endogenously Calibrated

Parameter		Value	Target
Treatment	k_E	4.60	Clean Firms: Fixed Cost/Output Ratio 2.5%
	ξ	0.23	Clean Technology Adoption Rate 57%
Preference	β	0.8750	Capital-output Ratio 1.65
Distortions	ϕ_0	1.15	Average Value Added Tax 13%
	ϕ_1	-0.03	θ_l–z Elasticity
Production	γ	0.9300	Size Distributions
Productivity	μ	-2.4567	Size Distributions
	σ	4.0020	
	z'_{max}	24855	
	g_{max}	0.00048	

Model Fit

- Size and Employment Distributions:





Two Experiments

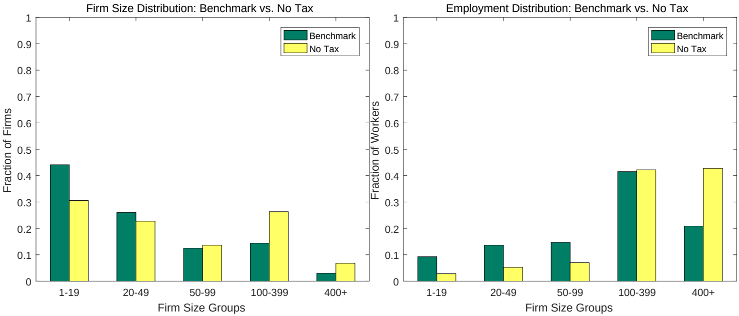
Results

- Two Exercises:
  - (i) Removing all distortions.
  - (ii) Increase regulation ξ such that the fraction of firms using clean technology is the same 85% as in (i).
- None of the exercises has large compositional effect, hence we focus on the polluting sector.

Statistic	Benchmark	(i)	(ii)
Aggregate Output	100.0	131.2	99.0
Capital	100.0	163.1	99.0
Number of Firms	100.0	44.1	89.9
Mean Size	60	139	66
Aggregate Pollution	100.0	76.7	85.8
Average Intensity	100.0	58.5	86.7
Clean Share	57.8	85.6	85.1

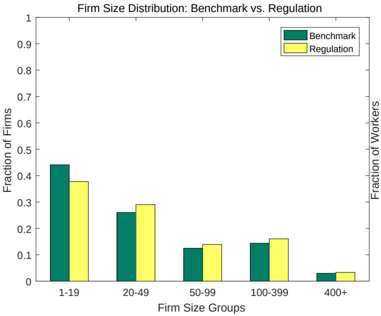
Mechanism

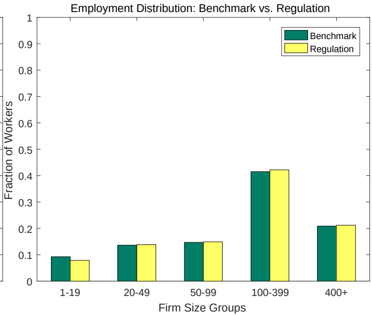
No Distortions



Mechanism

Stronger Regulations





Mechanism

Misallocation

- Extensive and Intensive Margin:
  - Extensive margin: selection of more productive firms.
  - Intensive margin: production distribution among active firms.

Economy	QU_1	QU_2	QU_3	QU_4	QU_5
Benchmark	2.79	4.26	7.45	16.81	68.70
Case (i)	1.53	2.89	6.46	18.25	70.87
Case (ii)	3.02	4.63	8.01	17.79	66.55

- Regulations worsen resource allocation on the intensive margin by amplifying the effects of distortions.
- Core mechanics:
  - General equilibrium feedback through wages.
  - The interaction between selection and correlated distortions.

Monopolistic Competition

Mechanism

The Progressiveness in Distortions

- Experiment (i') imposes a uniform distortion to all the firms that collects the same amount of tax revenue.

Statistic	Benchmark	(i')	(i)
Aggregate Output	100.0	108.2	131.2
Capital	100.0	110.9	163.1
Number of Firms	100.0	44.1	44.1
Mean Size	60	139	139
Aggregate Pollution	100.0	70.0	76.7
Average Intensity	100.0	64.8	58.5
Clean Share	57.4	73.3	85.6

- The progressiveness of the distortions is particularly damaging to the environment.

Compositional Effect

Setting

- Consider when the products from the two sectors have a constant elasticity of substitution ρ:

$$Y = \left[\varphi\left(Y_d\right)^{\frac{\rho-1}{\rho}} + (1-\varphi)\left(Y_c\right)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}.$$

- The price of the final good Y is normalized to one:

$$\left[\varphi^{\rho} p_d^{1-\rho} + (1-\varphi)^{\rho} p_c^{1-\rho}\right]^{\frac{1}{\rho-1}} = 1.$$

- Two Exercises:
  (i) Removing all distortions.
  (ii) Removing only the distortions in the polluting sector.

Compositional Effect

Results

	Polluting			Non-polluting		
Statistics	Benchmark	(i)	(ii)	Benchmark	(i)	(ii)
Physical Output	100.00	129.64	134.63	100.00	129.90	98.03
Price	100.00	99.97	85.00	100.00	100.00	102.34
Revenue	100.00	129.60	114.42	100.00	129.90	100.32
# of Firms	100.00	42.63	46.44	100.00	41.88	95.38
Mean Size	64.31	152.21	176.17	51.16	123.50	49.38
Pollution	100	74.69	80.05			
Intensity	100	57.81	59.47			
Clean Share	56.18	83.80	73.10			

Conclusions

- This paper proposes a new explanation for the industrial pollution problem in China:
  - a negative relationship between firm size and the pollution intensity;
  - expansion of productive firms is limited by distortions.
- We build a model consistent with the empirical regularities.
- Quantitative results show that distortions reduce output by 30% and increase pollution by 20%.
- Regulations improves resource allocation on the extensive margin, but worsens the intensive margin by amplifying distortions.

Thank You!

Appendix

Chemical Oxygen Demand Definition

- In environmental chemistry, the chemical oxygen demand (COD) test is commonly used to indirectly measure the amount of organic compounds in water. Most applications of COD determine the amount of organic pollutants found in surface water (e.g. lakes and rivers) or waste water, making COD a useful measure of water quality. It is expressed in milligrams per liter (mg/L) also referred to as ppm (parts per million), which indicates the mass of oxygen consumed per liter of solution.

- COD is an indirect measure that indicates the presence of contaminants that will eventually cause oxygen loss.

▸ Data Source

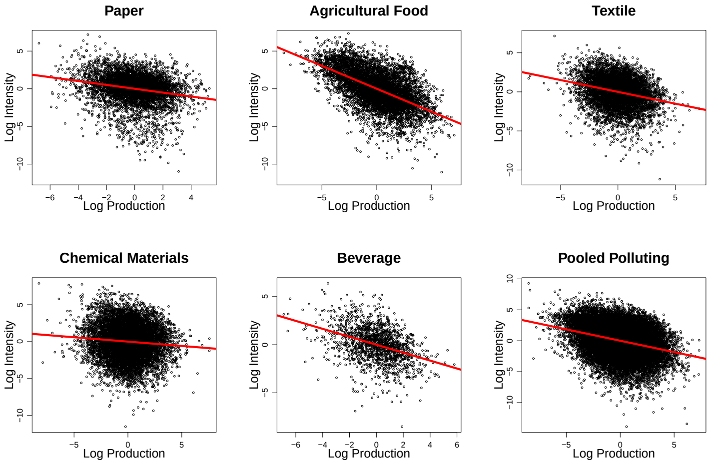
Major Pollutant Discharge

	Waste	COD	Petro	NH_4^+	BOD	CN	Cr^{6+}	Phenol	As	Cr
Key	76.2	73.2	31.4	25.2	17.5	4.90	4.86	2.42	2.27	2.01
Reg	35.2	28.3	7.91	6.49	2.56	0.13	N/A	0.04	0.07	N/A

† Data Source: National General Survey of Pollution Sources.
Acronyms:
Waste: Wastewater; COD: Chemical Oxygen Demand; Petro: Petrochemicals; NH_4^+: Ammonian
BOD: Biochemical Oxygen Demand; CN: Cyanidium; Cr^{6+}: Hexavalent Chromium
Phenol: Volatile Phenols; As: Arsenium; Cr: Chromium.

Data Sources

Firm Size and Pollution Intensity



Treatment Technologies

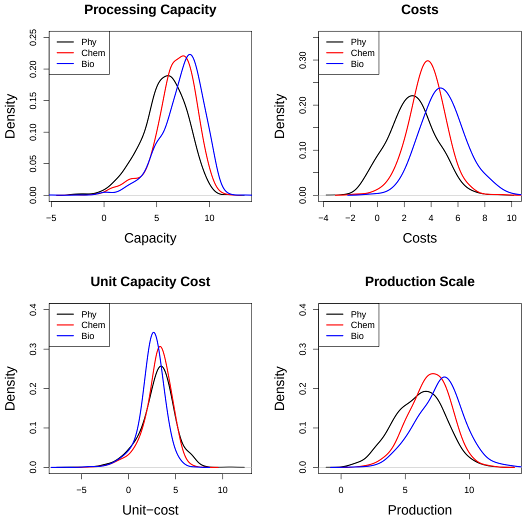
Efficiency of Technologies

- Examples of the treatment technologies:
  - Physical: Filtering, Centrifuging, Precipitation Separation, etc.
  - Chemical: Oxidation-reduction, neutralization, etc.
  - Biological: Aerobic Biological Treatment, Activated Sludge Process, etc.
- In the U.S. wastewater treatment system,
  - Physical technology: preliminary and primary treatments.
  - Chemical and biological technologies: secondary and tertiary treatments.

Treatment Technology

Treatment Technologies

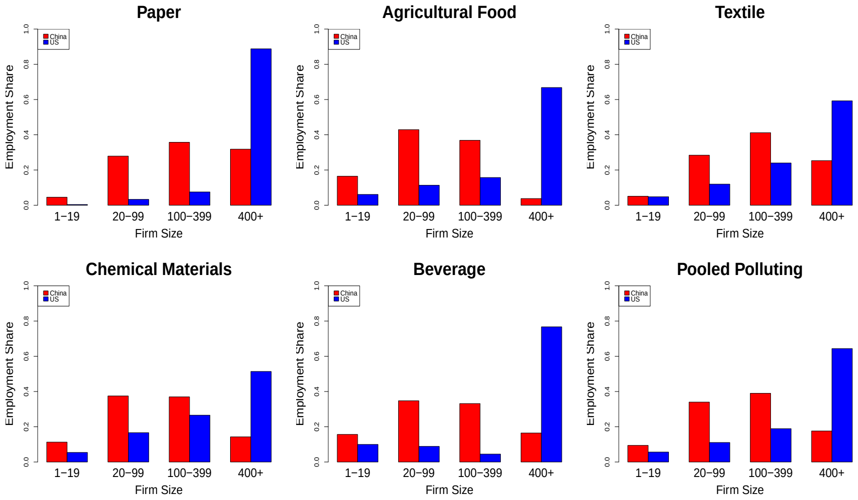
Costs, Capacity and Production Scale



Treatment Technology

Firm Size Distribution

By Industry



▸ Polluting and All Manufacturing

Robustness of the Accounting Exercise

- We need to estimate the employment–production relationship of Chinese firms because
  - the U.S data contain only information on number of employees;
  - the pollution census has information on production, but not on number of employees.
- China National Economic Census data are used.
- Results:

Method	Paper	Agri	Tex	Chem	Bever	Pooled	Reduc
Non-parametric	39.8%	60.7%	81.6%	102.5%	103.7%	63.5%	28.2%
Piecewise-linear	34.8%	69.4%	93.5%	180.1%	N/A	75.4%	19.0%
Parametric	43.5%	61.1%	97.5%	101.2%	89.0%	67.0%	25.5%

Accounting Exercise

Firm's Optimization

Setting

- Non-polluting sector:

$$\pi^c(z) = \max_{k,l} \left\{ (1-\tau_z) z^{1-\gamma} (k^\alpha l^{1-\alpha})^\gamma - Wl - Rk \right\}.$$

- Polluting sector:
  - Technology Adoption:

$$\pi^d(z) = \max_{i \in \{0,1\}} \left\{ \pi_0^d(z), \pi_1^d(z) \right\}.$$

  - Profits when using dirty technology

$$\pi_0^d(z) = \max_{k,l} \left\{ (1-\xi) \left[(1-\tau_z)\, z^{1-\gamma} (k^\alpha l^{1-\alpha})^\gamma - Wl - Rk \right] \right\}.$$

  - Profits when using clean technology

$$\pi_1^d(z) = \max_{k,l} \left\{ (1-\tau_z)\, z^{1-\gamma} (k^\alpha l^{1-\alpha})^\gamma - Wl - R(k + k_E) \right\}.$$

▸ Equilibrium

Firm's Optimization

Technology Adoption Decision

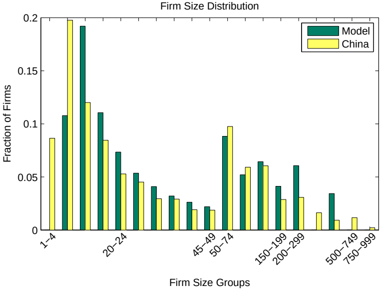
- Difference in potential profits using different technologies:

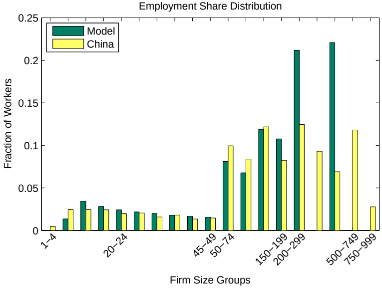
$$\pi_1^d(z) - \pi_0^d(z) = \frac{\xi(1-\gamma)}{(1-\alpha)\gamma} Wl^d(z) - Rk_E.$$

- There exists a threshold z^E where firms switch technology.
  - z_E decreases with ξ and increases with k_E.
  - z_E increases with τ_z.

Graphical

Model Fit Benchmark Full Groups





Model Fit

Monopolistic Competition

Setting

- The crucial elements of our analysis:
  - a non-degenerate distribution of firms;
  - profits be a function of productivity.
- Melitz (2003) setting:
  - CES preference:

$$U = \left[\int_{\omega \in \Omega} q(\omega)^{\frac{\rho-1}{\rho}} d\omega \right]^{\frac{\rho}{\rho-1}}.$$

  - Linear production function:

$$y = F(z, l) = zl.$$

  - Monopolistic competition:

$$\max_{p(z), q(z)} \left\{ p(z) q(z) - Wl \right\}.$$

Perfect Substitutes

Monopolistic Competition

Equivalence between Lucas (1978) and Melitz (2003)

- The equivalence between Lucas (1978) and Melitz (2003):
  - Intensive margin

$$l(z) = \left(\frac{\gamma}{W}\right)^{\frac{1}{1-\gamma}} z, \qquad l(z) = Q\left(\frac{W\rho}{\rho-1}\right)^{-\rho} z^{\rho-1}.$$

  - Intensive margin

$$\pi(z) = \left[\left(\frac{\gamma}{W}\right)^{\frac{\gamma}{1-\gamma}} - W\left(\frac{\gamma}{W}\right)^{\frac{1}{1-\gamma}}\right] z$$

$$\pi(z) = \frac{Q}{\rho}\left(\frac{\rho-1}{\rho}\right)^{\rho-1} z^{\rho-1}.$$

- The equivalence of diminishing returns to scale and to utility.

Perfect Substitutes