From Imitation to Innovation: Where is all that Chinese R&D Going?

König, M., Storesletten, K., Song, Z., & Zilibotti, F. (2022).

March 7, 2026

Motivation

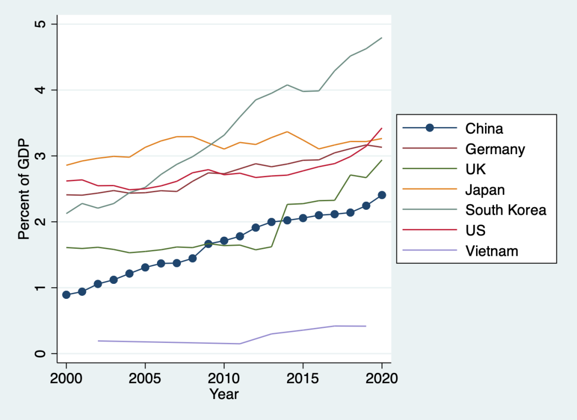
- Recently, the rapid economic growth in China has been accompanied by a boom in R&D expenditure and the government's growing emphasis on innovation
- R&D investments increased 1% of GDP in the 1990s to 2.4% of GDP by 2020.
- A common concern is that these investment decisions are distorted by policies and frictions that are pervasive in China.

$\Longrightarrow$ This paper constructs and structurally estimates an endogenous growth model to

- study the dynamic effects of misallocation on the investments firms make to improve their productivity growth
- quantify the contribution of these investments to China's aggregate growth

Motivation

Figure 1: Gross R&D Expenditure



Outline

Theory

- A dynamic Melitz-type (small open) economy populated by a unit measure of monopolistically competitive firms
- Produce differentiated goods using a CRS technology

$$Y_i(t) = A_i(t)K_i(t)^{\alpha}L_i(t)^{1-\alpha}$$

- Differentiated goods are combined into a homogeneous final good by a CES aggregator ($\eta > 1$)

$$Y = \left(\int_0^1 Y_i^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}}$$

- No entry and exit

Theory

- Firms are owned by overlapping generations of two-period–lived manager-entrepreneurs

$\Longrightarrow$ Break down the firm's problem into two steps

  - A static profit maximization
  - A dynamic (intertemporal investment) problem

Firm's static problem

- The young manager chooses $K_i(t)$ and $L_i(t)$ to maximize profits subject to its residual demand, TFP $A_i(t)$ and (persistent) output wedge $\tau_i(t) \in \{\tau_l, \tau_h\}$
- Firm i's current (period t) optimal profits:

$$\pi_i(t) = \left[\left(1 - \tau_i(t)\right) A_i(t) \right]^{\eta-1} \times \tilde{\Pi}(t) \quad (1)$$

$$\tilde{\Pi}(t) \equiv \left(\alpha^{\alpha} (1-\alpha)^{1-\alpha} (\eta - 1) \right)^{\eta-1} \eta^{-\eta} \times \frac{Y(t)}{\left(r^{\alpha} w(t)^{1-\alpha} \right)^{\eta-1}} \quad (2)$$

- The firm's value added

$$P_i(t) Y_i(t) = \eta \pi_i(t) \propto \left[A_i(t) \left(1 - \tau_i(t)\right) \right]^{\eta-1}$$

Firm's Dynamic Problem

- The young manager decides the strategy to improve the firm's TFP in the next period
- Each successful attempt to move up the productivity ladder results in a constant accrual of log TFP

$$\log(A_{i,t+1}) = \log(A_{i,t}) + \tilde{a}, \quad (\tilde{a} > 0)$$

- The ranking in the productivity ladder: $a \equiv \log(A)/\tilde{a}, \quad a \in \mathbb{N}^+$
- $\mathscr{A}$: the TFP distribution
- $\mathscr{A}_1, \mathscr{A}_2, \ldots$: the proportion of firms at each rung of the ladder
- $F_a = \sum_{j=1}^{a} \mathscr{A}_j$: the associated cumulative distribution

Firm's Dynamic Problem

- Firms can increase their TFP through:
  - **Imitation (costless):** Firm A_i meets A_j.
    - $A_j > A_i$: increases productivity with prob. $q > 0$
    - $A_j \leq A_i$: retains A_i

  $\implies$ Prob. of success: $q(1 - F_a)$
  - **Innovation (R&D cost):**
    - Succeed with prob. p drawn i.i.d from $G : [0, \bar{p}] \to [0, 1], \quad \bar{p} \leq 1$
    - If fail, can do passive imitation with probability $\delta q(1 - F_a) \geq 0$

  $\implies$ Total prob. of success: $p_i + (1 - p_i)\delta q(1 - F_a)$
- Firms pursue the strategy that yields the highest PV of profits (net of the R&D cost)

Equilibrium Dynamics With Costless Innovation

- No R&D cost and $\delta < 1$ (Why not $\delta \geq 1$?)
- Firm i chooses innovation $\iff$

$$p_i \geq Q(a, \tau; \mathscr{A}) \equiv \frac{q(1-\delta)(1-F_a)}{1-\delta q(1-F_a)} \tag{3}$$

- $\frac{\partial Q}{\partial a} < 0 \implies$ the proportion of innovating firms will be nondecreasing in the initial TFP
- the ex post TFP growth gap between innovating and imitating firms is increasing in the TFP level

Equilibrium Dynamics With Costless Innovation

- Decision function notations:

$$\chi^{im}(a,p,\tau;\mathscr{A}) = 1-\chi^{in}(a,p,\tau;\mathscr{A}) = \begin{cases} 1 \text{ if } p \leq Q(a,\tau;\mathscr{A}) \\ 0 \text{ if } p > Q(a,\tau;\mathscr{A}) \end{cases} \tag{4}$$

- The law of motion of TFP:

$$\begin{aligned}
&\mathscr{A}_a(t+1) - \mathscr{A}_a(t) \\
&= \int_0^{\bar{p}} \Big[\chi^{in}(a-1,p,\tau;\mathscr{A}) \times \Big(p+(1-p)\delta q\big(1-F_{a-1}(t)\big)\Big)\mathscr{A}_{a-1}(t) \\
&\quad + \chi^{im}(a-1,p,\tau;\mathscr{A})q\big(1-F_{a-1}(t)\big)\mathscr{A}_{a-1}(t) \\
&\quad - \chi^{in}(a,p,\tau;\mathscr{A}) \times \Big(p+(1-p)\delta q\big(1-F_a(t)\big)\Big)\mathscr{A}_a(t) \\
&\quad - \chi^{im}(a,p,\tau;\mathscr{A})q\big(1-F_a(t)\big)\mathscr{A}_a(t)\Big] dG(p)
\end{aligned} \tag{5}$$

Equilibrium Dynamics With Costless Innovation

PROPOSITION 1. Under som regularity conditions, there exists a traveling wave solution of the form $\mathscr{A}_a(t) = f(a - vt)$ with velocity $v = v(q, \delta, g(p)) > 0$, with left and right Pareto tails. For a given t, $\mathscr{A}_a$ is characterized as follows:

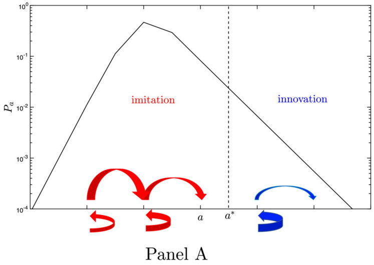
(i) For a sufficiently large, $\mathscr{A}_a(t) = O\left(e^{-\rho(a-vt)}\right)$, where the exponent ρ is the solution to the transcendental equation $\rho v = \hat{p}\left(e^{\rho} - 1\right)$.

(ii) For a sufficiently small, $\mathscr{A}_a(t) = O\left(e^{\lambda(a-vt)}\right)$, where the exponent λ is the solution to the transcendental equation $\lambda v = q\left(1 - e^{-\lambda}\right)$.

$\Longrightarrow$ Basically, there exists a stationary equilibrium with the constant growth rate v, which is increasing in q, δ, and $\bar{p}$

Equilibrium Dynamics With Costless Innovation

Figure 2: Equilibrium dynamics with a stationary distribution



Equilibrium Dynamics With Costly Innovation

- Recall the discounted value of profits:

$$\pi_i(t) = \frac{1}{1+r} \times \left[(1-\tau_i(t))A_i(t)\right]^{\eta-1} \times \tilde{\Pi}(t)$$

$$\tilde{\Pi}(t) \equiv \left(\alpha^{\alpha}(1-\alpha)^{1-\alpha}(\eta-1)\right)^{\eta-1}\eta^{-\eta} \times \frac{Y(t)}{\left(r^{\alpha}w(t)^{1-\alpha}\right)^{\eta-1}}$$

$$w(t)^{1-\alpha} = \left(1-\frac{1}{\eta}\right)\frac{\alpha^{\alpha}(1-\alpha)^{1-\alpha}}{r^{\alpha}}\left(\int_0^1 (1-\tau_i)^{\eta-1}A_i(t)^{\eta-1}\,di\right)^{\frac{1}{\eta-1}}$$

- Given $A_i(t), \tau_i(t)$, profit $\pi_i(t)$ changes over time via two channels
  - \+ higher income $\implies$ higher demand for all varieties
  - \- the progress of competitors $\implies$ lower market share ($\eta > 1$)
- On equilibrium, wages, profits and aggregate output grow at the same rate

Equilibrium Dynamics With Costly Innovation

- The R&D cost for innovating firms

$$c_i(t) = \left(A_i(t)^{\theta}\bar{A}(t)^{1-\theta}\right)^{\eta-1} \times \bar{c} \times \tilde{\Pi}(t) \qquad (6)$$

where $\bar{c} > 0, \quad \theta \in [0,1], \quad \bar{A}(t) \equiv \int_0^1 A_i(t)di$

  - Larger profits, higher costs
  - The relative level of the firm TFP (but not its absolute value) matters for the R&D decision
  - R&D cost is increasing in $A_i(t)$, conditional on $\bar{A}(t)$
  - R&D cost follows the trend $\tilde{\Pi}(t)$

Equilibrium Dynamics With Costly Innovation

- Firm i chooses innovation $\iff p_i \geq Q(a,\tau;\mathscr{A})$

$$Q(a,\tau;\mathscr{A}) = \frac{q(1-\delta)(1-F_a)}{1-\delta q(1-F_a)} + \frac{e^{(1-\theta)(\eta-1)(\bar{a}-a)}}{\left(e^{\eta-1}-1\right)\mathbb{E}\left[(1-\tau')^{\eta-1} \mid \tau\right]} \frac{\bar{c}(1+r)(1+g)^{\eta-2}}{1-\delta q(1-F_a)} \tag{7}$$

  - g : the (endogenous) steady-state growth rate,
  - $a \equiv \log \bar{A}$
  - $\mathbb{E}[\tau' \mid \tau]$: the conditional expectation of next-period wedge.
- Comparison to the costless innovation
  - Imitation becomes more attractive $\implies$ the threshold Q is higher
  - Larger τ reduces future profit proportionally to *TFP* without affecting the R&D cost $\implies$ less innovation: $\frac{\partial Q}{\partial \tau} > 0$

Equilibrium Dynamics With Costly Innovation

The law of motion of the TFP distribution:

$$
\begin{aligned}
&A_a(t+1) - A_a(t) \\
&= \sum_{j \in \{l,h\}} \omega_{\tau_j}(t) \int_0^P \Big[\chi^{\text{in}}(a-1, p, \tau_j; A)\big(p + (1-p)\delta q(1 - F_{a-1}(t))\big) A_{a-1}(t) \\
&\quad + \chi^{\text{in}}(a-1, p, \tau_j; A) q(1 - F_{a-1}(t)) A_{a-1}(t) \\
&\quad - \chi^{\text{in}}(a, p, \tau_j; A)\big(p + (1-p)\delta q(1 - F_a(t))\big) A_a(t) \\
&\quad - \chi^{\text{in}}(a, p, \tau_j; A) q(1 - F_a(t)) A_a(t) \Big] \, dG(p)
\end{aligned} \tag{8}
$$

- ω_{τ_l} and ω_{τ_h}: the proportion of low- and high-wedge firms

$$\omega_{\tau_l}(t+1) - \omega_{\tau_l}(t) = (1-\rho_h)(1-\omega_{\tau_l}(t)) - (1-\rho_l)\omega_{\tau_l}(t),$$

Predictions of the theory

1. ceteris paribus, the proportion of firms engaged in R&D is increasing in TFP;
2. ceteris paribus, firms with higher wedges are less likely to engage in R&D. Then, conditional on TFP, larger firms are more likely to engage in R&D
3. expected TFP growth is falling in current TFP, especially for nonR&D firms
4. the gap in average TFP growth between R&D firms and nonR&D firms is increasing in TFP.

Data

- the Annual (firm-level) Survey of Industries conducted by China's National Bureau of Statistics, $1998 - 2007$ and $2011 - 2013$:
  - A balanced panel for all manufacturing firms, $2007 - 2012$
  - Data for R&D expenditure at the firm level are 2007
  - Innovators $=$ positive R&D expenditure
- Plant level data collected by Taiwan's Ministry of Economic Affairs, $1999 - 2004$.

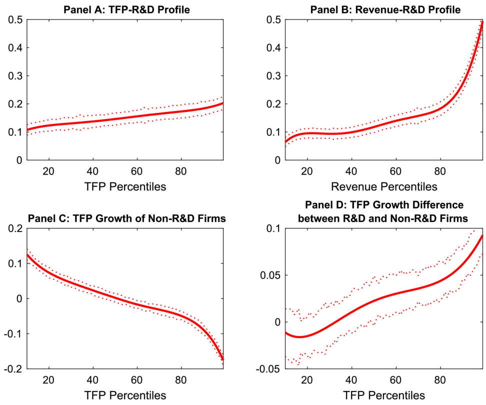
Summary Statistics

Table 1: Summary Statistics

Year	Number of Firms	Number of R&D Firms	Median Value Added (Million USD)	Mean Value Added (Million USD)	Median R&D Intensity (%)	Aggregate R&D Intensity (%)
			Balanced Panel of Chinese Firms			
2007	123,368	18,140	1.48	5.81	1.73	1.86
2012	123,368	N.A.	3.33	11.45	N.A.	N.A.
			Private Chinese Firms in the Balanced Panel			
2007	117,983	15,828	1.43	4.67	1.65	1.54
2012	117,771	N.A.	3.26	9.57	N.A.	N.A.
			Balanced Panel of Taiwanese Firms			
1999	11,229	1,487	0.16	2.91	8.50	3.14
2004	11,229	1,144	0.17	4.78	6.42	2.93

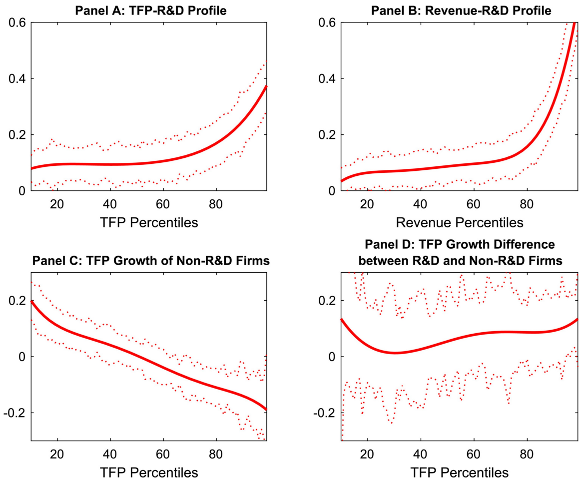
Empirical results

Figure 3: Chinese firms in the balanced panel 2007–2012



Empirical results

Figure 4: Taiwanese firms in the balanced panel 1999–2004



Outline

Structural estimation

- Calibrate α (equal to the measured industry-specific labor income share), $\eta = 5$, θ, $\tilde{a}$ and estimate the rests
- Target two sets of moments
    - Economics: 16 (four intervals (10th-49th, 50th-79th, 80th-94th, and the top five) of the distribution of TFP and size in each of the four panels)
    - Measurement error: 1 (the empirical variance of m.e. in TFP $\hat{\upsilon}_{\mu a}$)
- Use Simulated Method of Moments (SMM)

Measurement Error

- Classical m.e. of revenues and inputs
  - Generate an attenuation bias in the relationships between the propensity to engage in R&D and both TFP and size, and
  - Exaggerate TFP convergence
- Identification:
  - the firm-specific wedges τ_i are constant during the $2007-2012$ period

  $\Longrightarrow$ the time series variation ($2007-2012$) in value added and input measures at the firm level can be used to infer the extent of m.e.
  - The paper shows how to back out (empirical) $\hat{\upsilon}_{\mu a}$ from data

Distribution of output wedges

- The estimated output wedges and TFP are correlated in the data

$\Longrightarrow$ Assume

$$-\ln(1-\tau_i) = b \cdot (a_{it} - a_t) + \varepsilon_{it}^{\tau} \quad , \quad \varepsilon_{it}^{\tau} \sim \mathcal{N}\left(0, \text{var}(\varepsilon_{it}^{\tau})\right)$$

- Estimate $b = 0.779$ and $\text{var}(\varepsilon_{it}^{\tau} = 0.042$ (correcting for m.e.)

The Technology of R&D

- Cost function: estimate $\theta \approx 0.25$ by targeting the R&D cost-to-value added ratio: the top five percent and the 10th–49th percentile..

$$\frac{\mathbb{E}[\psi_j \mid a_j]}{\mathbb{E}[\psi_i \mid a_i]} = \exp\left(\frac{1-\eta}{\eta}(1-\theta+b)(a_i - a_j)\right)$$

- Step size: $\tilde{a} \approx 0.78$ implies an average cost of innovation of 3.7% of the industrial value added (about twice the ratio in the Chinese data)
- Productivity of innovation p: drawn from an i.i.d uniform probability distribution with support $[0, \bar{p}]$

Outline

Parsimonious Model

Estimate five parameters:

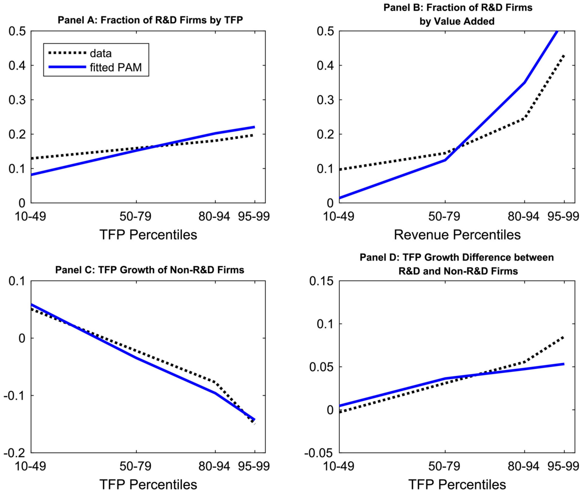
- $q \Longleftarrow$ the TFP convergence rate across imitating firms (Panel C) conditional on m.e.
- $\delta \Longleftarrow$ the TFP convergence rate across both imitating and innovating firms (Panels C and D).
- $\bar{p} \Longleftarrow$ the extent to which, conditional on TFP, innovating firms grow faster than imitating firms (Panel D).
- $\bar{c} \Longleftarrow$ the total share of innovating firms, given the other parameters
- $\upsilon_{\mu a} \Longleftarrow$ flatten the schedules in Panels A and B, and steepen the schedule in Panel C.

Results

Table III. Estimation, Chinese firm balanced panel 2007–2012

	(1) PAM	(2) FLM	(3) IPM	(4) FRM	(5) PAM	(6) IPM
					Higher R&D cutoff	
Imitation prob. q	0.175	0.271	0.361	0.275	0.294	0.546
	(0.031)	(0.019)	(0.019)	(0.051)	(0.058)	(0.052)
Second chance δ	0.008	0.020	0.001	0.019	0.141	0.001
	(0.011)	(0.021)	(0.027)	(0.024)	(0.106)	(0.080)
Innov. prod. $\bar{p}$	0.096	0.114	0.113	0.237	0.107	0.111
	(0.008)	(0.006)	(0.006)	(0.016)	(0.013)	(0.010)
Innov. cost $\bar{c}$	1.627	3.374	2.318	10.486	3.601	9.393
	(0.136)	(0.174)	(0.177)	(1.363)	(0.448)	(2.403)
Std. dev. m.e. $\sigma_{\mu a}$	0.549	0.472	0.431	0.459	0.476	0.391
	(0.014)	(0.008)	(0.005)	(0.025)	(0.022)	(0.011)
Std. dev. innov. subs. σ_c		1.243	1.092	0.011		1.969
		(0.038)	(0.037)	(0.036)		(0.213)
Policy inter. c_a			1.888	-10.46		2.499
			(0.159)	(1.042)		(0.356)
Fake share Y				0.099		
				(0.005)		
J-statistic	1.518	0.507	0.368	0.362	2.690	0.516

Figure 5: Parsimonious model (PAM).



Heterogeneous Innovation Costs

Three cases:

FLM: innovation wedges $\bar{c}_i$ are i.i.d. across firms

$$c_i(t) \propto \left[\bar{c} - \exp\left(\bar{\xi}_i(t) - \frac{\sigma_c}{2}\right) + 1\right] \cdot \left(A_i(t)^{\theta} A(t)^{1-\theta}\right)^{\eta-1},$$

where $\bar{\xi}_i \sim \mathcal{N}(0, \sigma_c^2)$.

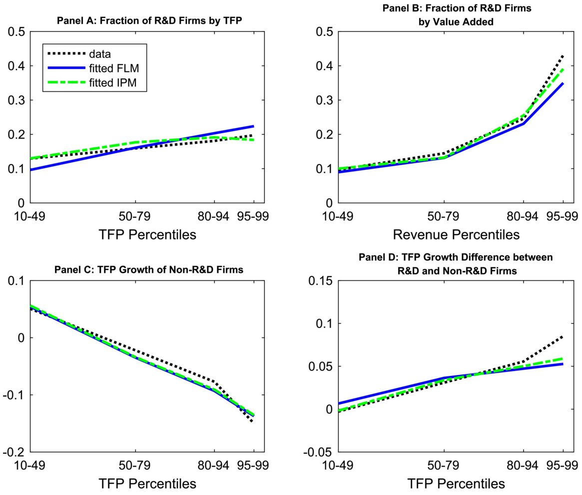
IPM: $\bar{c}_i$ to be correlated with observable firm characteristics

$$c_i(t) \propto \left[\bar{c} - \exp\left(\bar{\xi}_i(t) - \frac{\sigma_c}{2}\right) + 1 + c_a\left(G(a_i) - \frac{1}{2}\right)\right] \left(A_i(t)^{\theta} \bar{A}(t)^{1-\theta}\right)^{\eta-1},$$

$c_a > 0$: industrial policy favors low-TFP firms

FRM: a share Υ of firms can collect subsidies by just *claiming* to invest in R&D

Figure 6: Flexible (FLM) and Industrial Policy (IPM) models



Outline

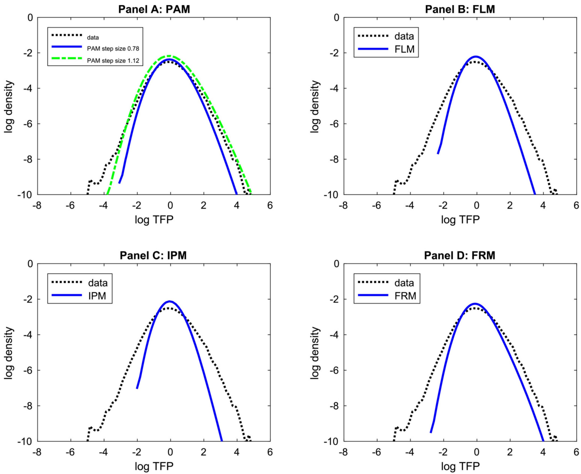
Indirect inference

Table IV Indirect inference, balanced panel of Chinese firms, 2007-2012

	(1) Data	(2) PAM	(3) FLM	(4) IPM	(5) FRM
Panel A: Dependent Variable: R&D decision in 2007					
log(TFP)	0.368	0.712	0.287	0.298	0.322
	(0.0284)	(0.0145)	(0.0167)	(0.0181)	(0.0167)
wedge	-0.410	-0.824	-0.274	-0.327	-0.357
	(0.0357)	(0.0177)	(0.0202)	(0.0214)	(0.0203)
Panel B: Dependent Variable: TFP Growth					
log(TFP)	-0.062	-0.094	-0.104	-0.112	-0.081
	(0.0035)	(0.0002)	(0.0003)	(0.0003)	(0.0004)
$R\&D_d$	0.036	0.034	0.033	0.028	0.053
	(0.0042)	(0.0005)	(0.0005)	(0.0005)	(0.0008)

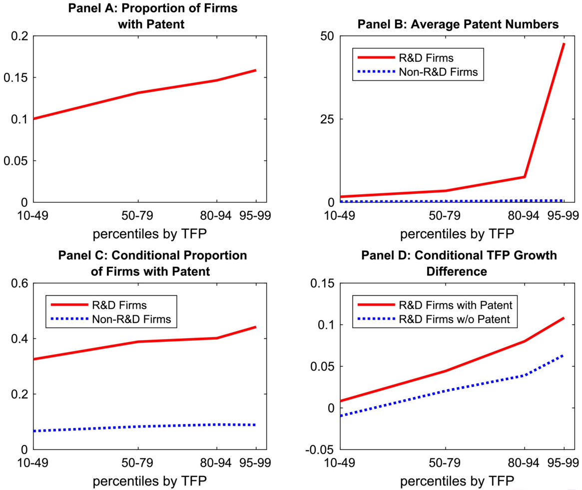
TFP Distribution

Figure 7: TFP distribution



Patents

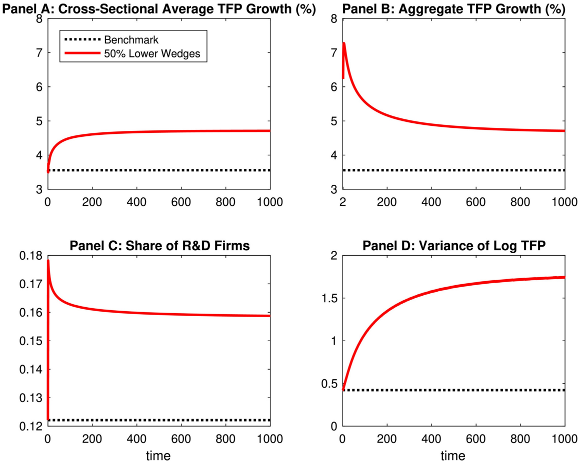
Figure 8: Patents of Chinese firms, balanced panel 2007–2012



Outline

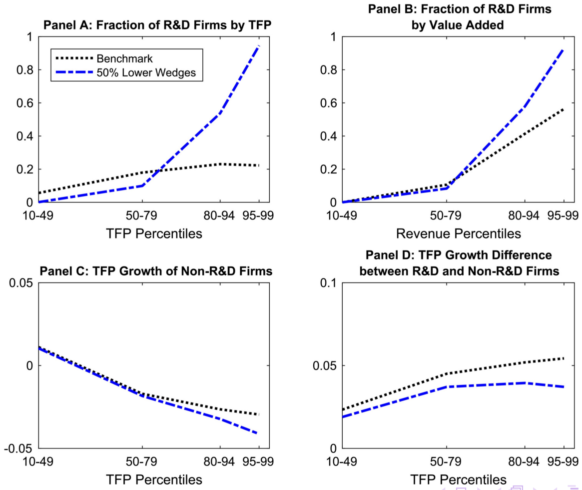
Reducing misallocation

Figure 9: Transition after lowering wedges



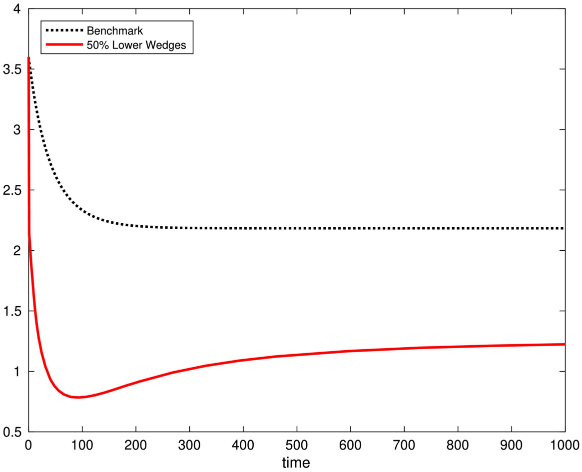
Reducing misallocation

Figure 10: Steady-state moments in the counterfactual model



International spillovers

Figure 11: TFP Gap during transition with international spillovers



The innovation technology

Table 2: Counterfactuals, parsimonious model

	(1) PAM estim. model	(2) 50% lower output wedges	(3) Taiwan's q	(4) Taiwan's $\bar{p}$ and $\bar{c}$	(5) Taiwan's $\bar{p}$, $\bar{c}$ and q	(6) Increase $\bar{c}$ so share R&D firms is 5%	(7) Decrease $\bar{c}$ so share R&D firms is 20%	(8) All firms do R&D
Fraction of R&D firms (%)	12.2	15.8	10.7	14.1	12.2	5	20	100
Steady-state TFP growth (%)	3.56	4.70	4.49	4.92	6.03	2.41	4.42	3.80

Conclusion

- Construct and structurally estimate a theory of TFP growth driven by innovation and technology diffusion through random interactions
- Estimate the model to earn new insights about the nature and effects of the R&D expenditure boom in China in recent years.
- R&D investments appear to have contributed significantly to the productivity growth of China.
- the return to productivity of R&D investments is lower in China than in Taiwan.
- Pervasive output wedges often induce the wrong firms (and, conversely, deter the right ones) from investing in R&D, thereby reducing the productivity of R&D investments.

Discussion

- Indirect approach to infer wedges $\Longrightarrow$ debate and dispute + limited policy implications (Bergquist et al. (2026))
- Abstract from entry and exit (an important margin)
- Under-explore in firm dynamics and the role of an intensive margin of R&D (the focus on innovation literature).

References I

Bergquist, Lauren F, Danial Lashkari, and Eric Verhoogen (2026), "Wedges: A microeconomic perspective on misallocation." Technical report, National Bureau of Economic Research.