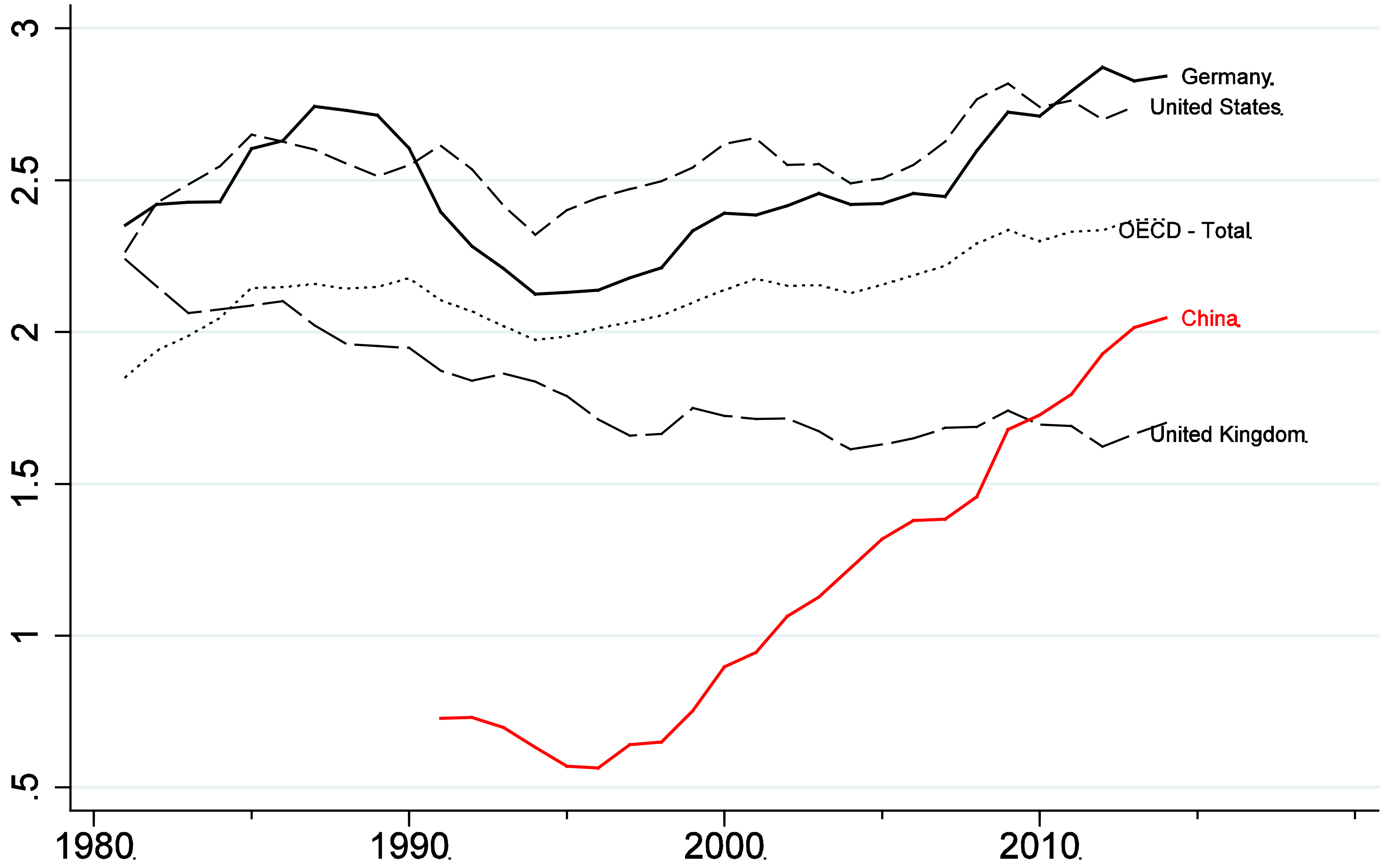
From imitation to innovation:
Where is all that Chinese R&D going?

Michael König (University of Zurich)
Zheng (Michael) Song (Chinese University of Hong Kong)
Kjetil Storesletten (University of Oslo)
Fabrizio Zilibotti (Yale University)

IMF / Atlanta Fed
May 19, 2017

Gross domestic expenditure on R&D (in % of GDP)



Source: OECD Science, Technology and Industry Outlook (latest available year)

R&D Misallocation

- Does R&D investment translate into productivity growth?
- Is China's allocation of R&D investment efficient?
  - E.g., SOE vs. DPE, connected firms, etc.
- Is R&D misallocation quantitatively important?
  - Policy distortions of R&D investments is likely a prime issue for China.
  - Proactive industrial policies, credit market frictions
  - … relevant also for R&D (cf. Schmitz 2016) and innovation?
- How does China compare with Taiwan (in earlier years)?

Today’s presentation

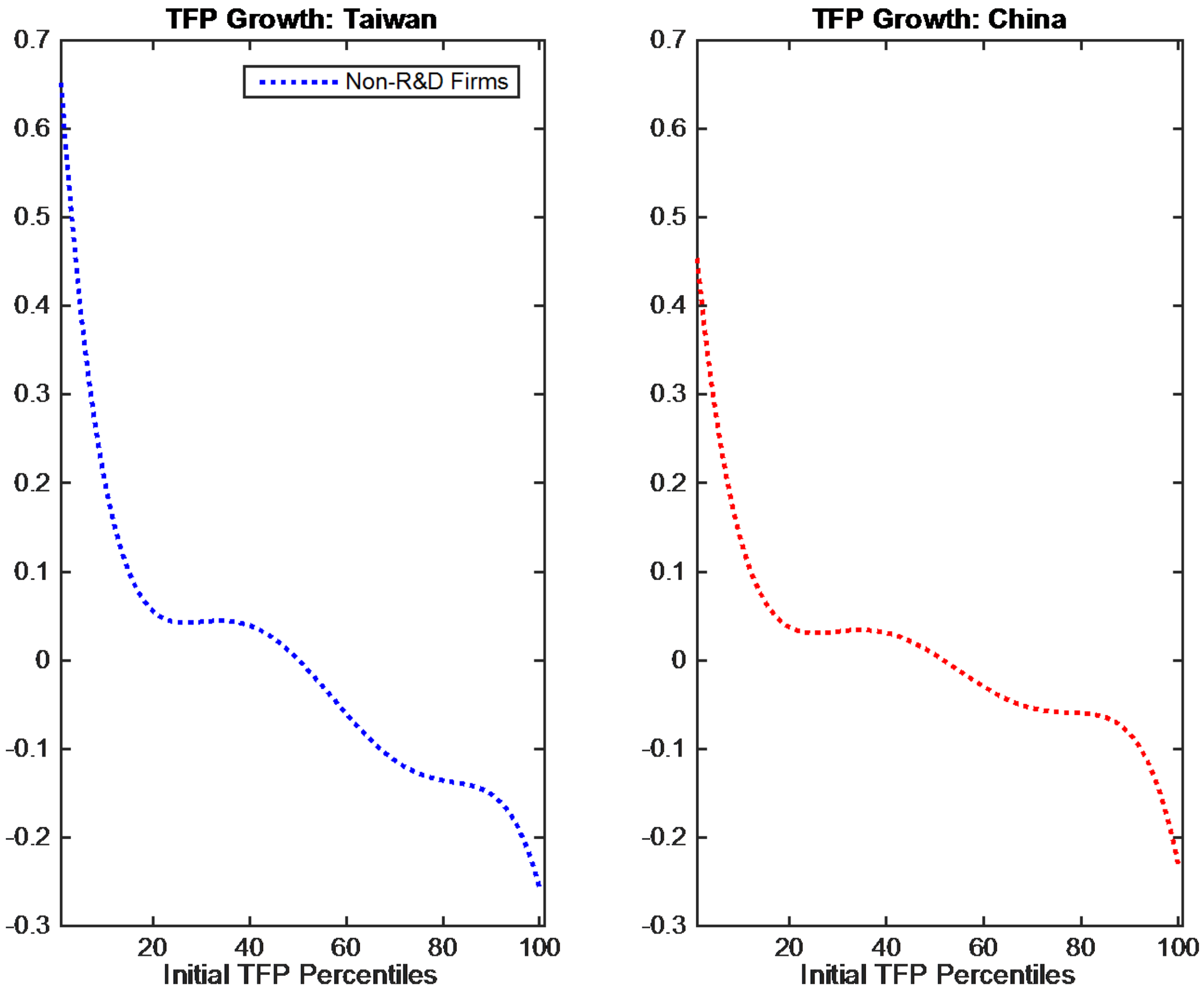
- Some facts on R&D from Chinese and Taiwanese firm-level data
  - Manufacturing, balanced panel, Taiwan: 1988-1993, China: 2001-2007

- A theoretical model

- Model estimation and policy counterfactuals

Stylized facts

1. Growth rates for non-R&D firms is falling in TFP
   - Roughly the same rate of decline in China and Taiwan
2. R&D firms grow faster than non-R&D firms
   - The gap is growing in the TFP level.
3. In Taiwan, larger growth difference R&D-vs-nonR&D than in China
   - Especially so for high TFP firms
4. R&D probability is increasing in TFP
   - More steeply so in Taiwan
5. Firm revenue ($P_i Y_i$) is positively correlated with R&D
   - Similar patterns in China and Taiwan



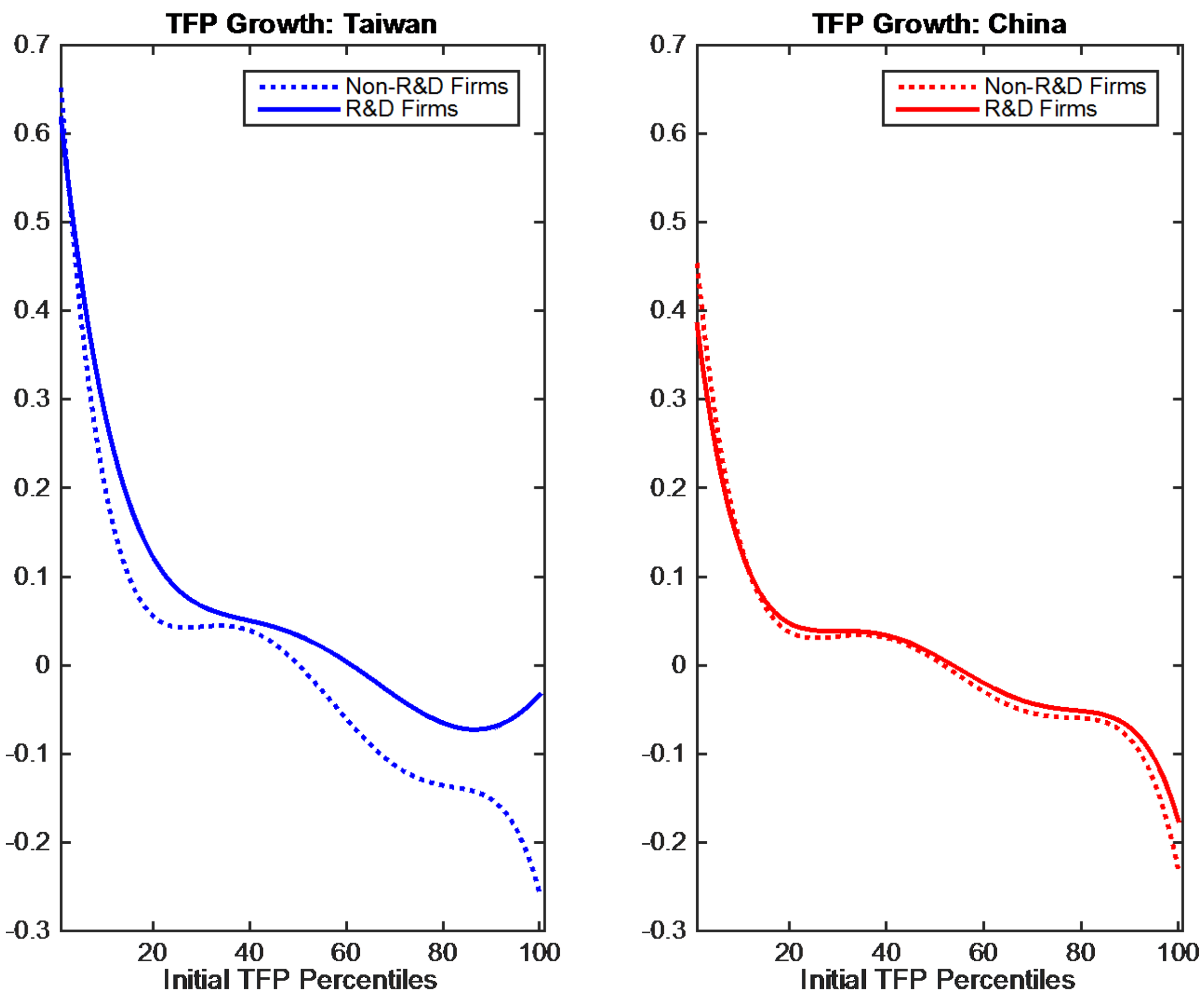
TFP Growth: Taiwan
Non-R&D Firms
Initial TFP Percentiles
TFP Growth: China
Initial TFP Percentiles

Stylized facts

1. Growth rates for non-R&D firms is falling in TFP
   - Roughly the same rate of decline in China and Taiwan
2. R&D firms grow faster than non-R&D firms
   - The gap is growing in the TFP level
3. In Taiwan, larger growth difference R&D-vs-nonR&D than in China
   - Especially so for high TFP firms
4. R&D probability is increasing in TFP
   - More steeply so in Taiwan
5. Firm revenue ($P_i Y_i$) is positively correlated with R&D
   - Similar patterns in China and Taiwan

Stylized facts

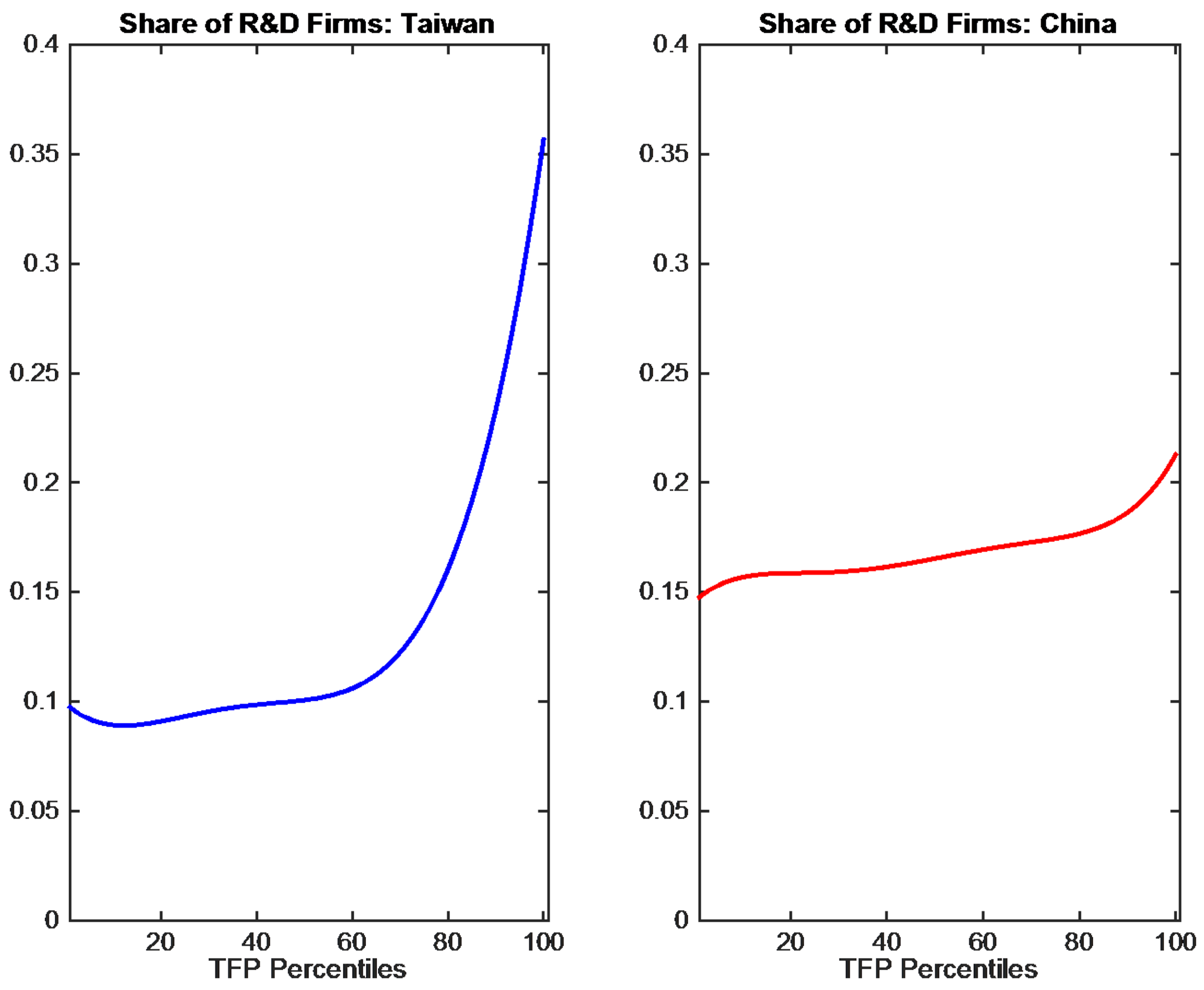
1. Growth rates for non-R&D firms is falling in TFP
   - Roughly the same rate of decline in China and Taiwan
2. R&D firms grow faster than non-R&D firms
   - The gap is growing in the TFP level
3. Taiwan has larger growth difference R&D vs. non-R&D than China
   - Especially so for high TFP firms
4. R&D probability is increasing in TFP
   - More steeply so in Taiwan
5. Firm revenue $(P_i Y_i)$ is positively correlated with R&D
   - Similar patterns in China and Taiwan



TFP Growth: Taiwan
Non-R&D Firms
R&D Firms
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
20
40
60
80
100
Initial TFP Percentiles
TFP Growth: China
Non-R&D Firms
R&D Firms
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
20
40
60
80
100
Initial TFP Percentiles

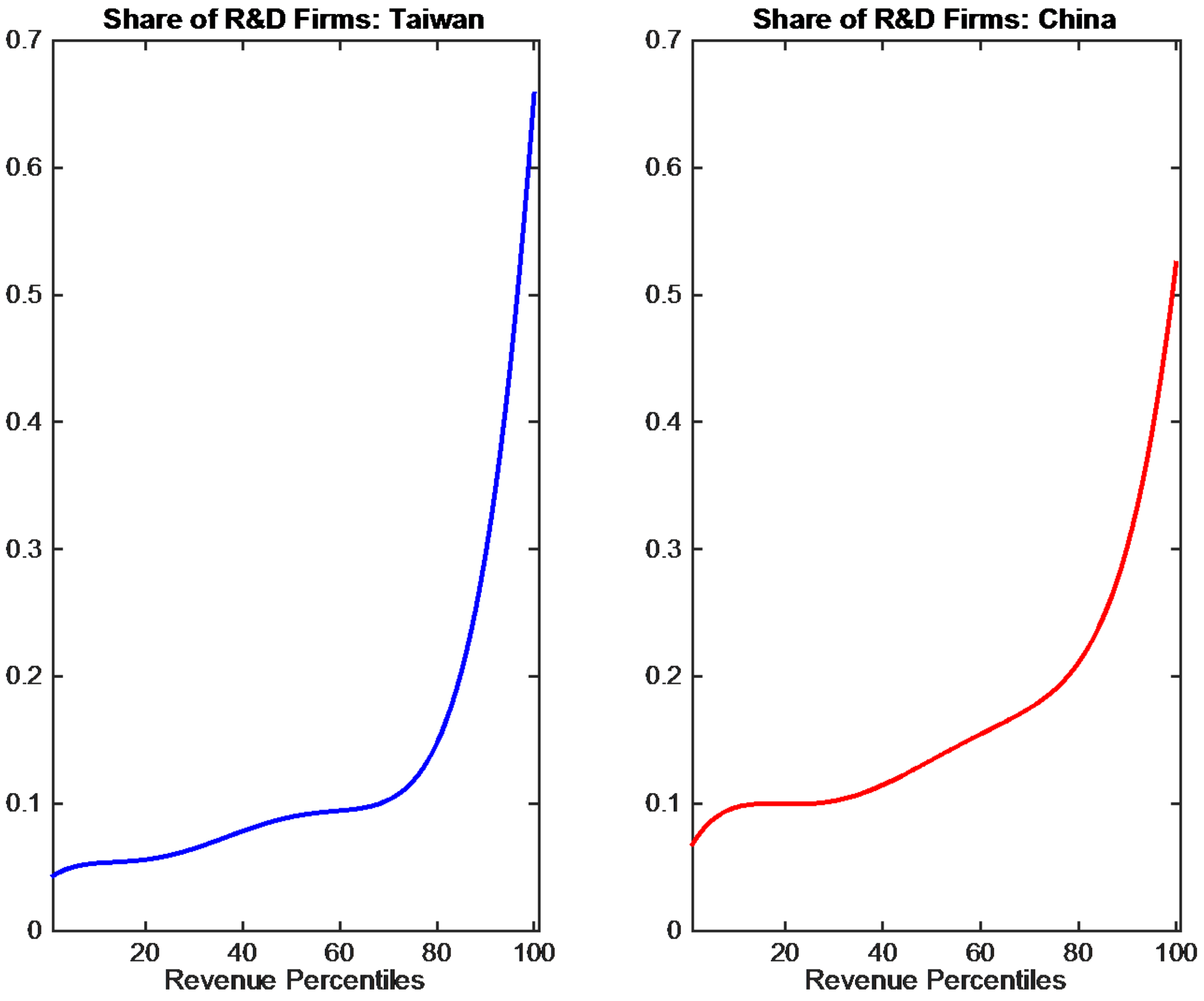
Stylized facts

1. Growth rates for non-R&D firms is falling in TFP
   - Roughly the same rate of decline in China and Taiwan
2. R&D firms grow faster than non-R&D firms
   - The gap is growing in the TFP level
3. Taiwan has larger growth difference R&D vs. non-R&D than China
   - Especially so for high TFP firms
4. R&D probability is increasing in TFP
   - More steeply so in Taiwan
5. Firm revenue ($P_i Y_i$) is positively correlated with R&D
   - Similar patterns in China and Taiwan


Share of R&D Firms: Taiwan
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
20
40
60
80
100
TFP Percentiles
Share of R&D Firms: China
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
20
40
60
80
100
TFP Percentiles

Stylized facts

1. Growth rates for non-R&D firms is falling in TFP
   - Roughly the same rate of decline in China and Taiwan
2. R&D firms grow faster than non-R&D firms
   - The gap is growing in the TFP level
3. Taiwan has larger growth difference R&D vs. non-R&D than China
   - Especially so for high TFP firms
4. R&D probability is increasing in TFP
   - More steeply so in Taiwan
5. Firm revenue ($P_i Y_i$) is positively correlated with R&D
   - Similar patterns in China and Taiwan



Share of R&D Firms: Taiwan
Share of R&D Firms: China
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
20
40
60
80
100
Revenue Percentiles

Building Blocks

- A model with both innovation and imitation (cf. AAZ 2006, KLZ 2016)
- R&D expenditure proxies for investment in innovation
  - Simplification: R&D is an extensive margin (binary) choice
- Distance to the ***local*** frontier determines imitation success rate
  - Implication: high-TFP firms invest in R&D because of low return on imitation
- Adding firm heterogeneity
  - (i) output wedges; (ii) innovation capacities; (iii) R&D costs …
- Obtain predictions about which firms do R&D and how fast they grow

Model

- Continuum of goods. Each good i is produced by a monopolist

- Cobb-Douglas production function. Inputs: capital and labor

- Capital and labor markets are competitive (but possibly distorted)

- A firm-specific OUTPUT WEDGE $(1 - \tau_i)$,
  - A combination of tax/subsidies on capital and labor
  - Later, we add further heterogeneity (for quantitative reasons):
    - Heterogeneity in productivity of R&D
    - R&D costs, tax/subsidies

Measuring Output Wedge and TFP

- Using firms' optimality conditions

Output wedge: $$(1-\tau_i) \propto \frac{(rK_i)^{\alpha}(wL_i)^{1-\alpha}}{P_iY_i}$$

TFP: $$A_i \propto \frac{(P_iY_i)^{\frac{1}{1-\vartheta}}}{(rK_i)^{\alpha}(wL_i)^{1-\alpha}}$$

from which we can estimate $1-\tau_i$ and A_i (cf. Hsieh and Klenow 2009)

- Note: Profit is increasing in $1-\tau_i$ and A_i

Firm’s Life Cycle

- Firms are run by two-period lived OLG of (non-altruistic) entrepreneurs
- Firms are transmitted from parents to children
  - cf. Song, Storesletten & Zilibotti 2011
- Young entrepreneur inherits TFP of parent’s firm
- Young entrepreneur decides on innovation/imitation
- Old entrepreneurs rent capital and labor
  - Produce. Pay back debt. Consume. Die …
- R&D decisions depend only on CURRENT productivity distribution
  - A simplification that eases analysis and estimation…

Imitation vs. Innovation

- Firms' productivity increases over time via innovation and imitation
- Improvement step in log-TFP are fixed and denoted by μ
- Binary choice: either imitate or innovate (cum passive imitation)

- **ACTIVE IMITATION**
  - **No cost**
  - **Success with prob. $q \times (1 - F(A_i))$ [meet a better firm]**
  - **If successful, TFP increases by one step**
- **VALUE OF ACTIVE IMITATION (to the entrepreneur)**

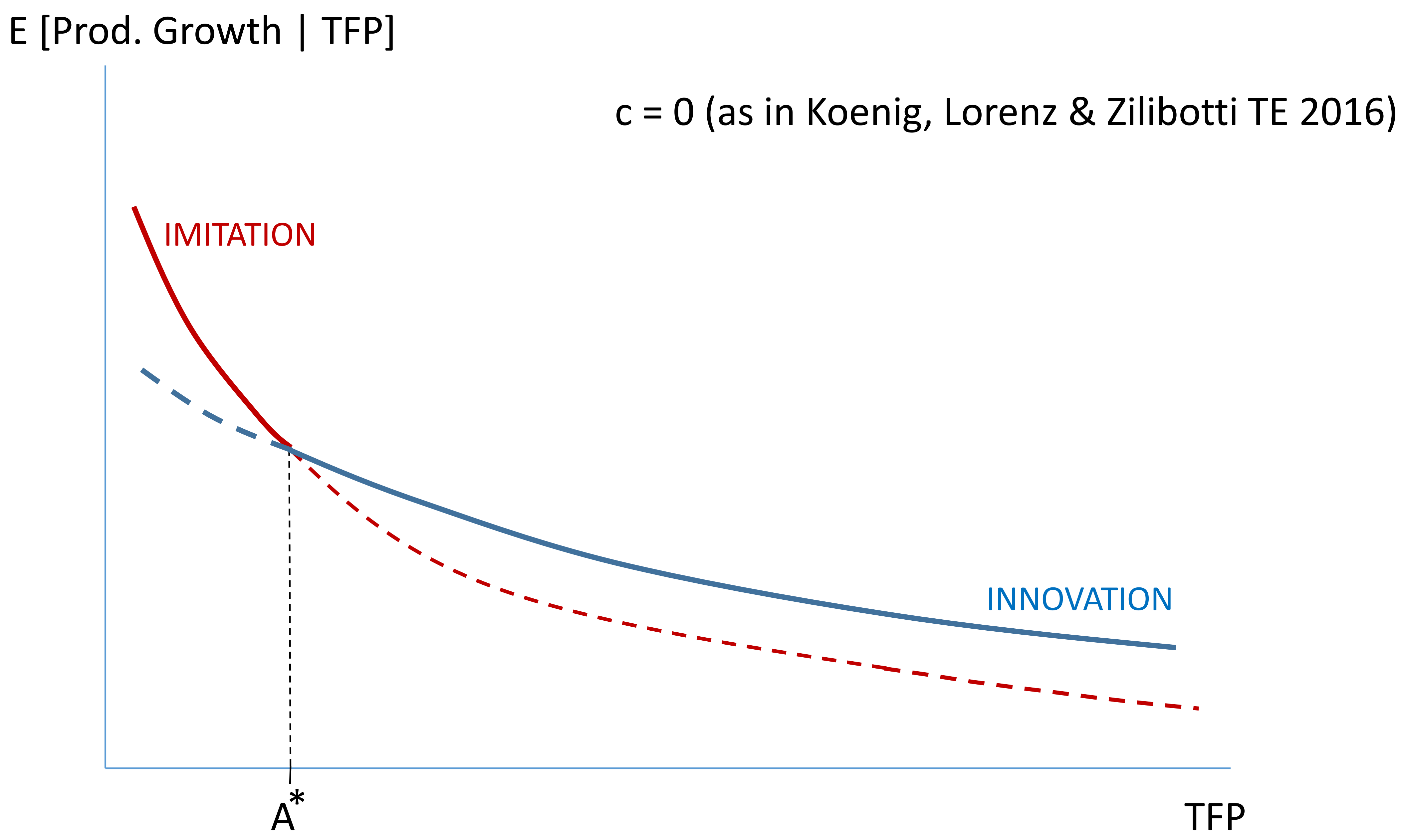
$$\beta \times \begin{bmatrix} q\big(1 - F(A)\big)\pi(\tau, (1+\mu)A) \\ \\ +\Big(1 - q\big(1 - F(A)\big)\Big)\pi(\tau, A) \end{bmatrix}$$

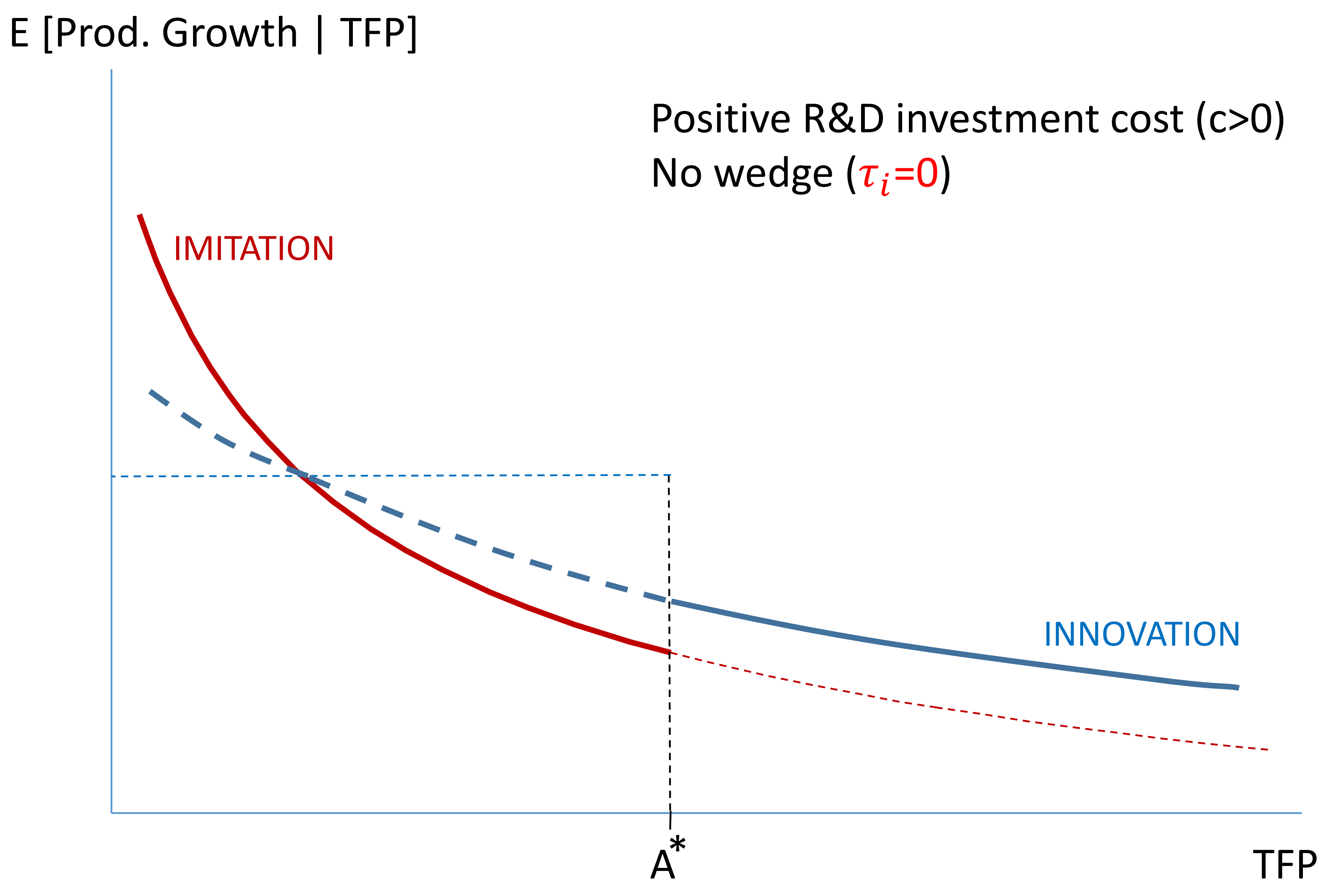
Imitation vs. Innovation

- Firms' productivity increases over time via innovation and imitation
- Improvement step in log-TFP are fixed and denoted by μ
- Binary choice: either imitate or innovate (cum passive imitation)

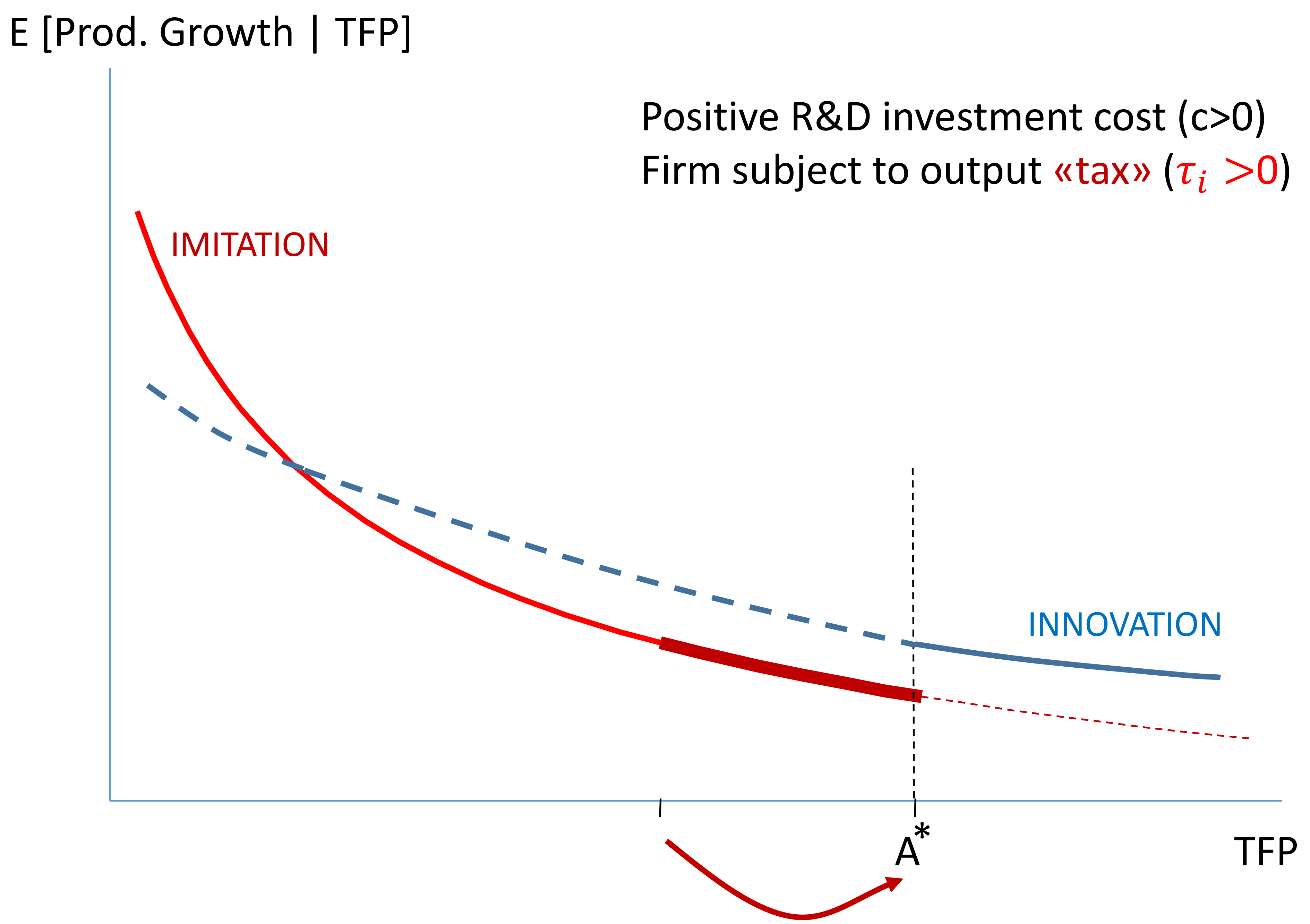
- **INNOVATION**
  - **Pay the R&D cost $\mathrm{C} = c \times A$**
  - **R&D Success with probability p, R&D Failure with probability 1- p**
  - **Passive imitation: Success with prob. $(1-p)\delta q\big(1-F(A_i)\big)$**
- **VALUE OF INNOVATION (to the entrepreneur)**

$$-c + \beta \times \left[\begin{array}{l} \Big(p + (1-p)\delta q\big(1-F(A)\big)\Big) \times \pi(\tau, (1+\mu)A) \\ \\ +\Big((1-p)\big(1-\delta q\big(1-F(A)\big)\big)\Big) \times \pi(\tau, A) \end{array} \right]$$


E [Prod. Growth | TFP]
c = 0 (as in Koenig, Lorenz & Zilibotti TE 2016)
IMITATION
INNOVATION
A*
TFP



E [Prod. Growth | TFP]
Positive R&D investment cost (c>0)
No wedge (τ_i=0)
IMITATION
INNOVATION
A*
TFP


E [Prod. Growth | TFP]
Positive R&D investment cost (c>0)
Firm subject to output «tax» (τ_i >0)
IMITATION
INNOVATION
A*
TFP

Equilibrium Productivity Dynamics

- The productivity distribution evolves endogenously
- State space (in log): $1, 2, \cdots, \infty$
- One step corresponds to a log-productivity increase by μ
- Probability distribution: $f_1, f_2, \cdots$
- Cumulative distribution: $F_n = \sum_{i=1}^{n} f_i$
- Under some conditions (sufficiently high q), there exists a stationary distribution of log-productivity (normalized by the growth rate of the economy) with left and right Pareto tails.

Dynamics and Stationarity

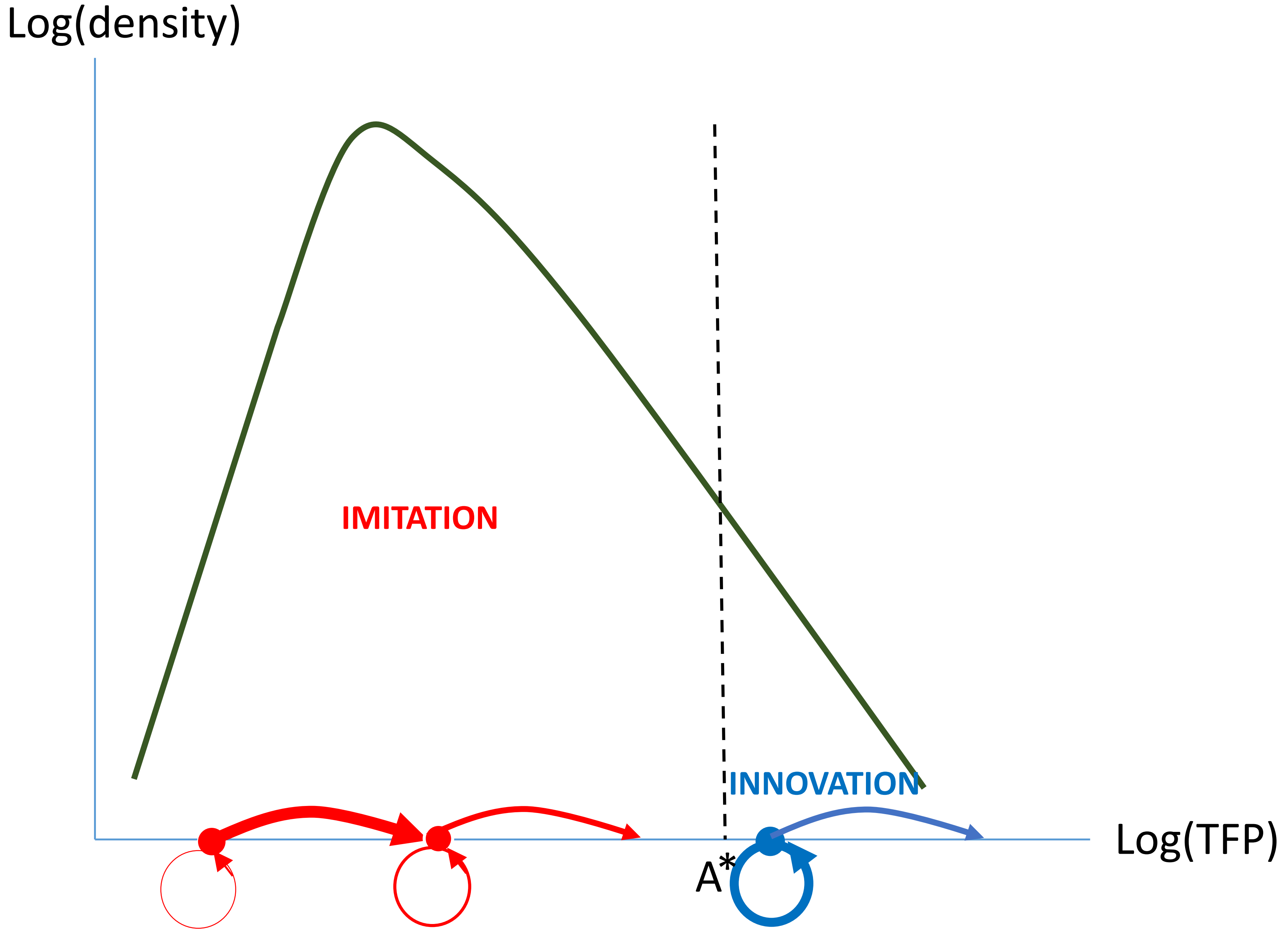
- Define $\chi_n(t)$ as the share of *n*-firms that imitate. Then:

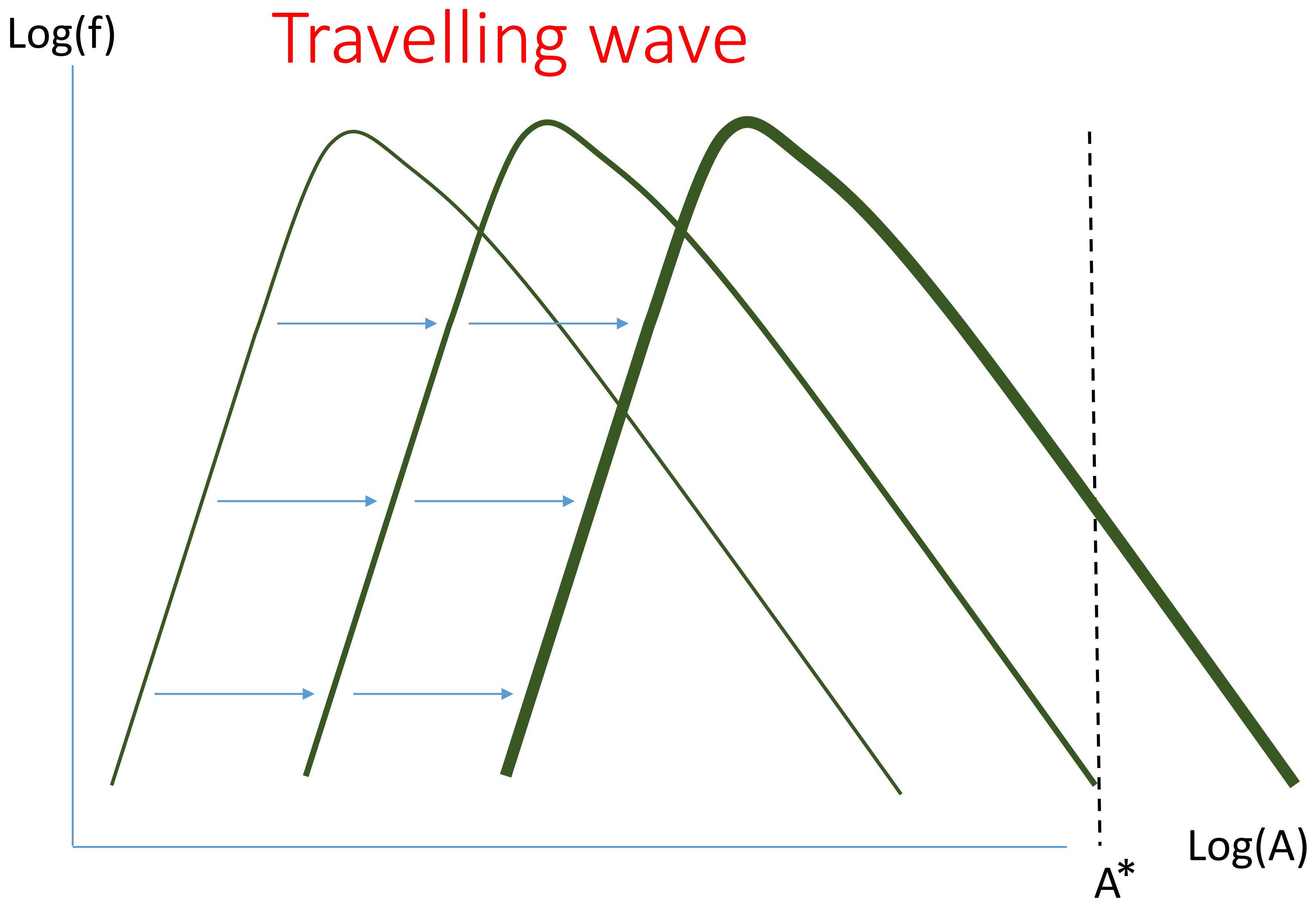
$$f_n(t+1) = q\big(1 - F_{n-1}(t)\big)\chi_{n-1}(t) f_{n-1}(t) + \quad \text{Imitation Success}$$
$$\Big(1 - q\big(1 - F_n(t)\big)\Big)\chi_n(t) f_n(t) + \quad \text{Imitation Failure}$$

$$+\Big(p + (1-p)\delta q\big(1 - F_{n-1}(t)\big)\Big)\big(1 - \chi_{n-1}(t)\big) f_{n-1}(t) \quad \text{Innovation Success}$$
$$+ (1-p)\Big(1 - \delta q\big(1 - F_n(t)\big)\Big)\big(1 - \chi_n(t)\big) f_n(t) \quad \text{Innovation Failure}$$

.

- Under some conditions (sufficiently high q), there exists a stationary distribution of log-productivity (normalized by the growth rate of the economy) with left and right Pareto tails.

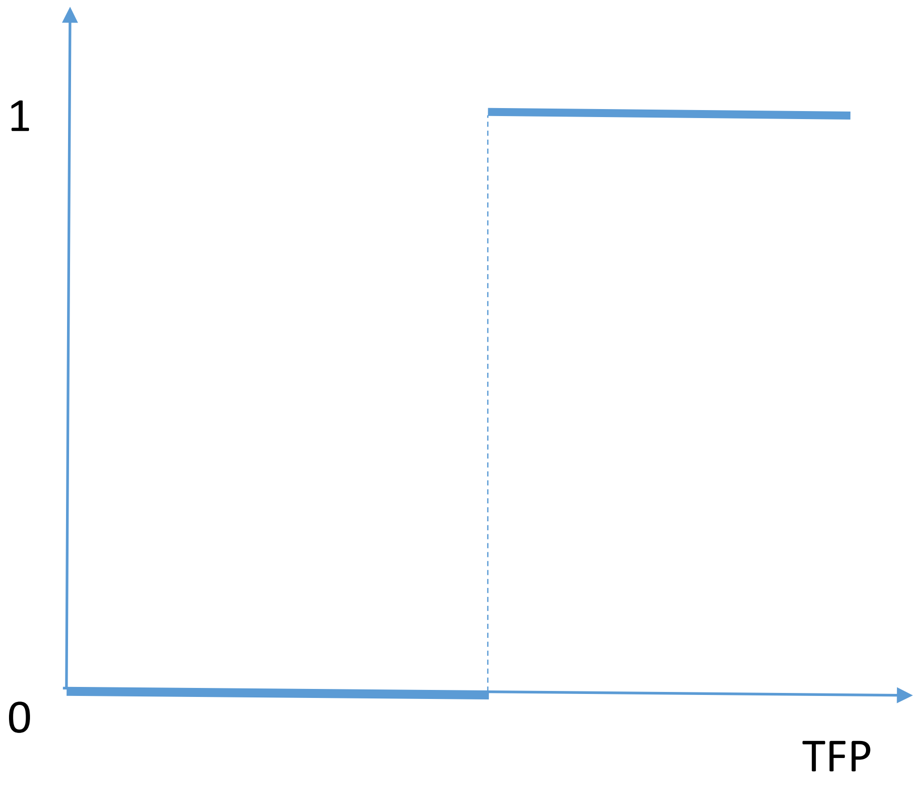

Log(density)
IMITATION
INNOVATION
Log(TFP)
A*


Log(f)
Travelling wave
Log(A)
A*

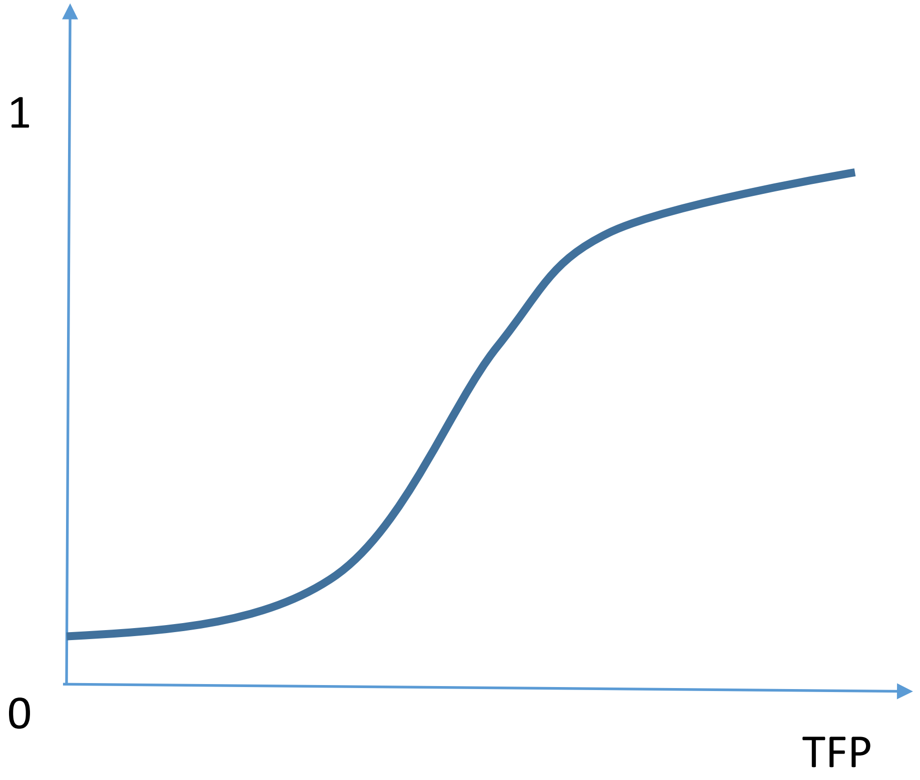
Effect of Heterogeneous Wedges & Shocks

TFP-R&D Profile

The Fraction of R&D Firms w/o heterogeneity (KLZ 2016)



The Fraction of R&D Firms with heterogeneity



Data

- Industrial Firm Survey Data for China and Taiwan (census)

- Taiwan: 1988-1993 balanced panel with 11,000 firms.
  - Taiwan is used for the benchmark estimation

- Later, China: 2001-2007 balanced panel with 78,000 firms.

- Analysis based on data after removing industry fixed effects

Towards Estimating the Model

STEP 1: infer wedges and TFP

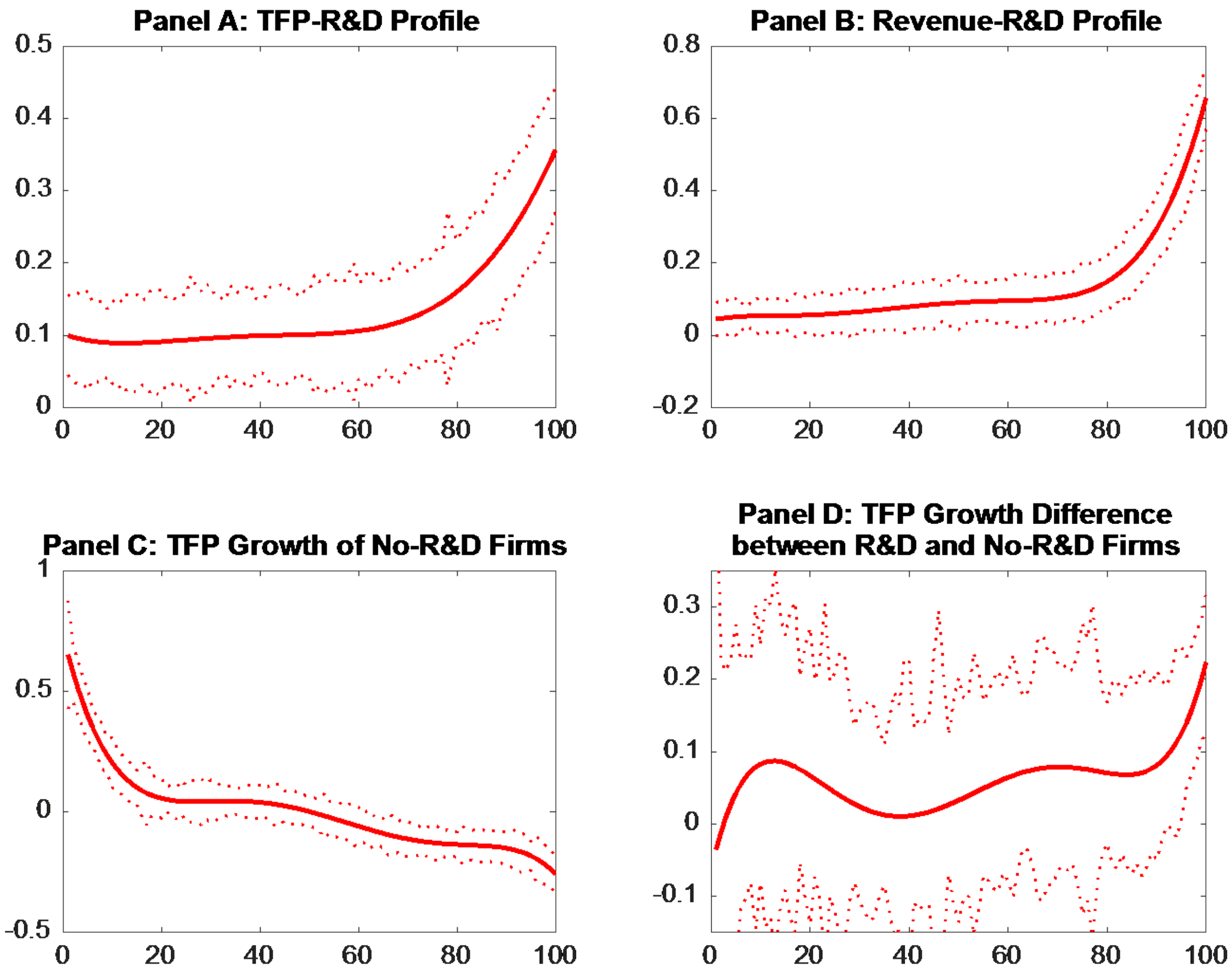
- Retrieve empirical joint distribution of τ and A

STEP 2: derive moments

- Sort firms on estimated TFP (A). For each TFP percentile, calculate
    1) R&D probability (extensive margin)
    2) TFP growth rate conditional on zero R&D
    3) TFP growth rate conditional on R&D > 0
- Sort firms on estimated wedges ($1- \tau$). For each percentile, calculate
    4) R&D probability (extensive margin)

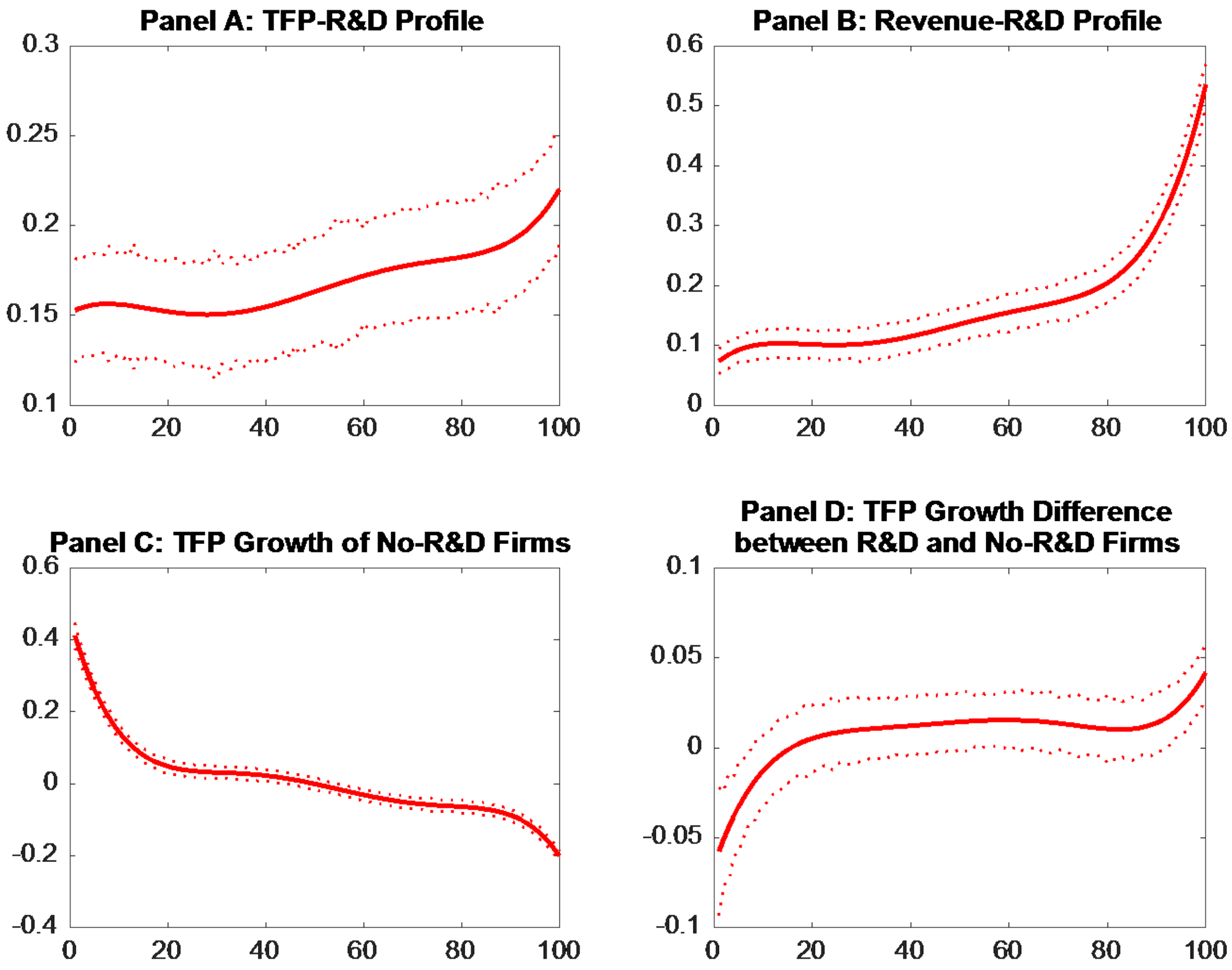
Taiwan data

(dotted lines are +/- 2*std)



China data

(dotted lines are +/- 2*std)



Estimating the model (SMM)

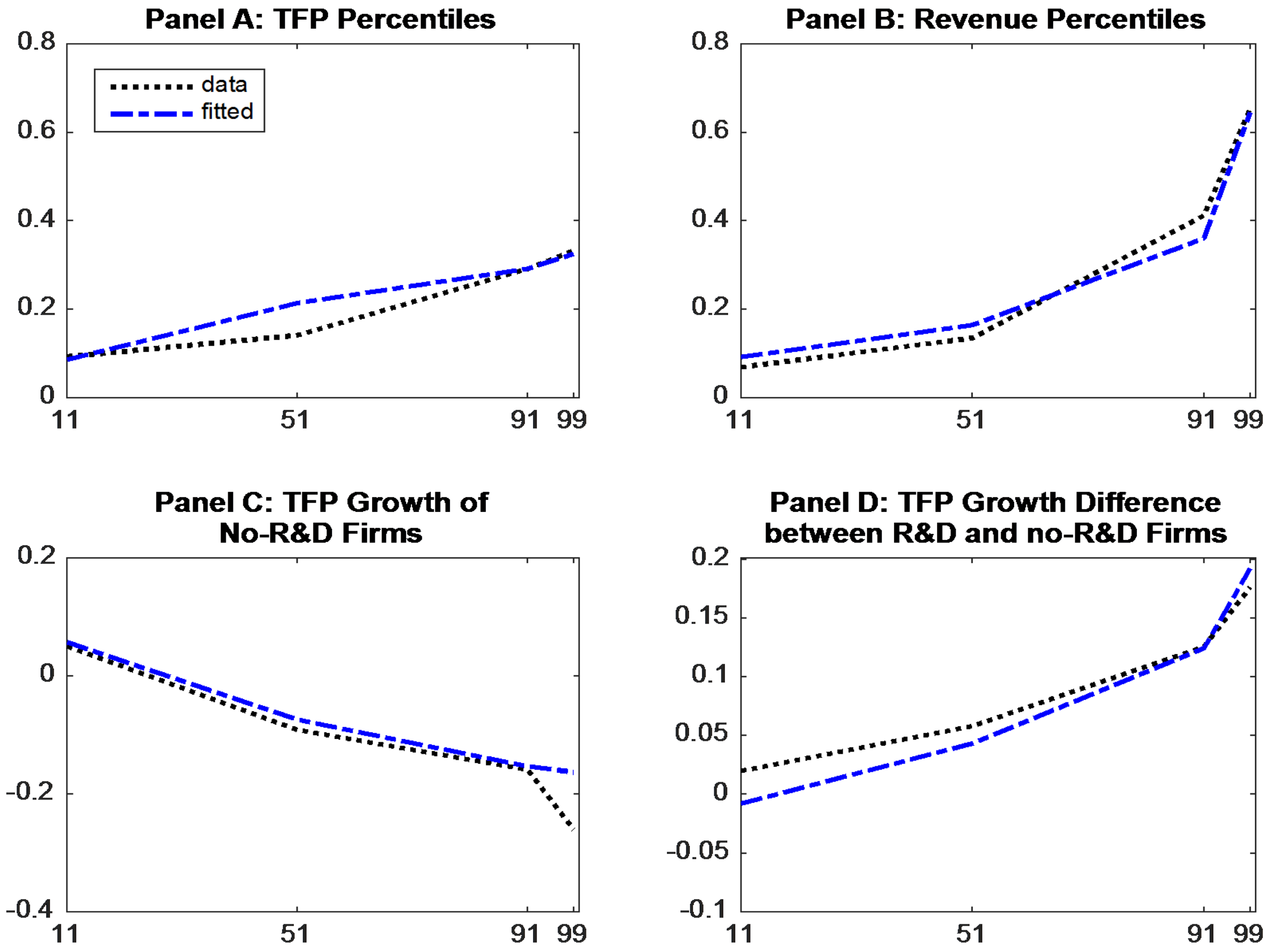
Estimate model by Simulated Method of Moments (for Taiwan)

- Estimate four parameters:
  - p distribution (probability of success of innovation). Uniform $[0, \overline{p}]$
  - q imitation parameter
  - δ passive imitation parameter
  - c R&D cost: estimate mean and variance
    - Assume $c_i = c + \eta_i$ where η_i is i.i.d. normal: $\mathrm{N}(0, \mathrm{var}(\eta))$
  - Target 16 moments, efficient weighting

  (based on percentiles of distributions in 4 panels above, drop bottom 10%)

Taiwan Estimates

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
mean of c	0.52
std of c	0.59



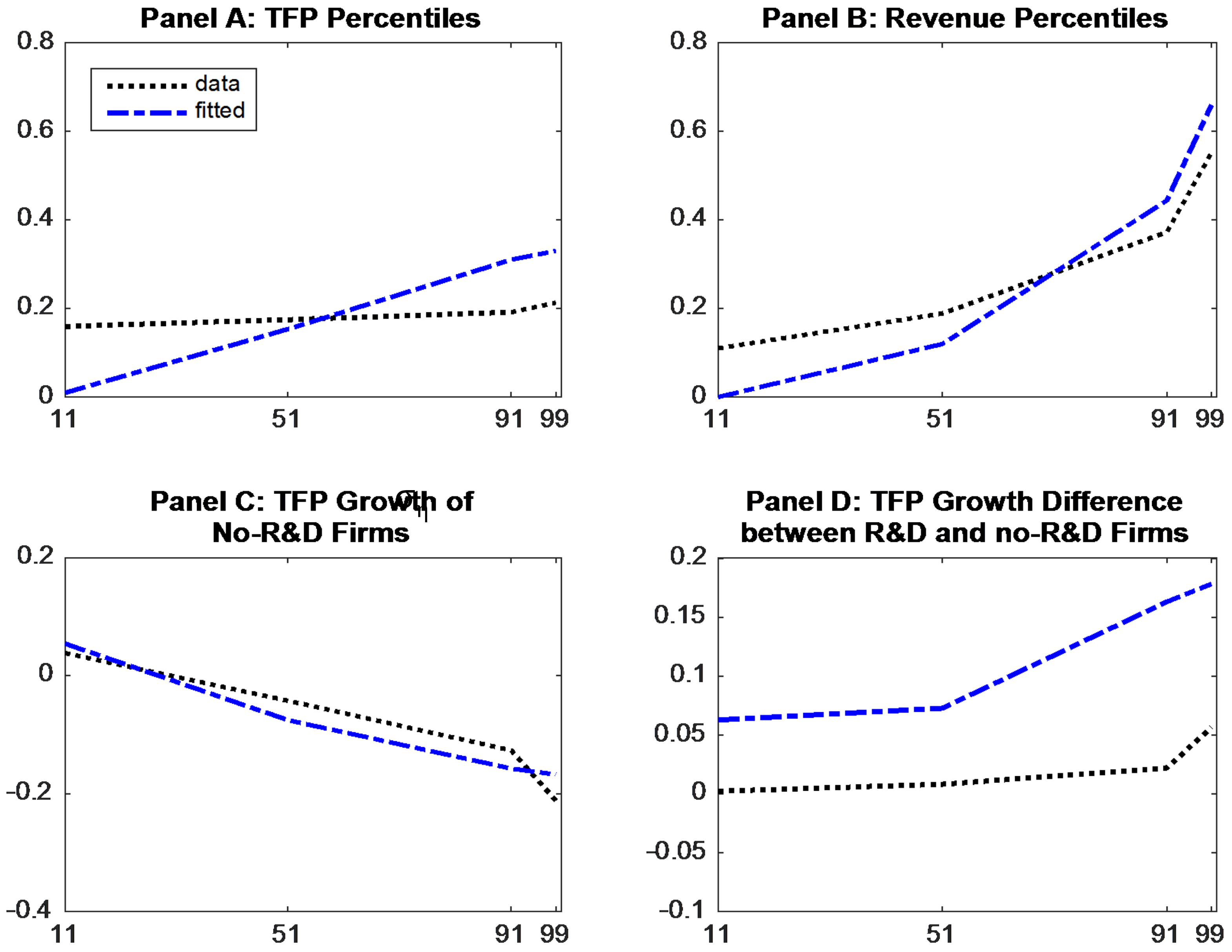
China benchmark (Taiwan based)

- Impose technological parameters estimated for Taiwan:
  - q, δ, $\overline{p}$ (max R&D success prob.), and σ_η (variance of R&D cost c)
- Impose estimated Chinese tax wedges
- Reestimate mean R&D cost c (to match average R&D probability)

China: benchmark

(Taiwan parameters, Chinese wedges, reestimate mean c)

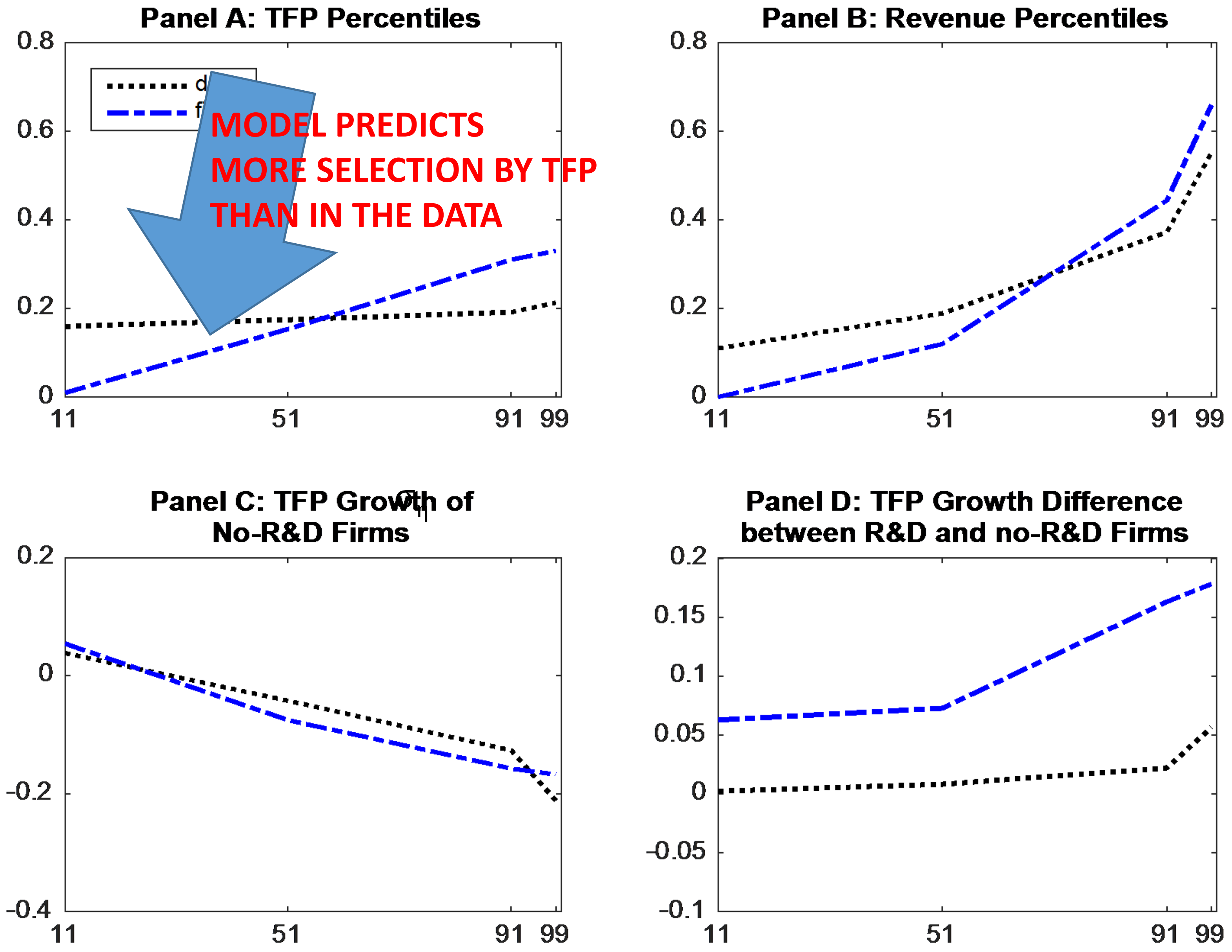
	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
std of c	0.59
	Re-estimated for China
mean of c	1.50



China: benchmark

(Taiwan parameters, Chinese wedges, reestimate mean c)

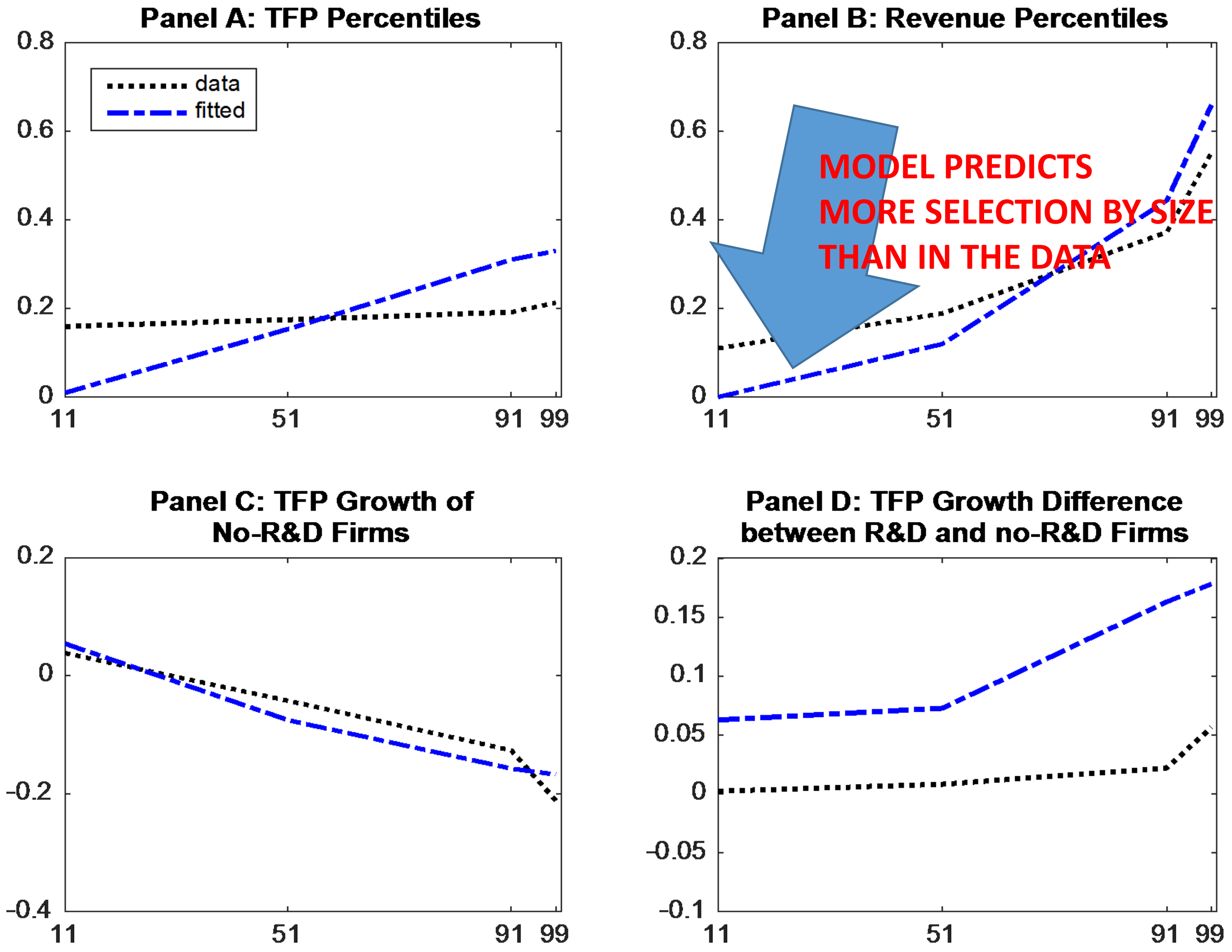
	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
std of c	0.59
	Re-estimated for China
mean of c	1.50



China: benchmark

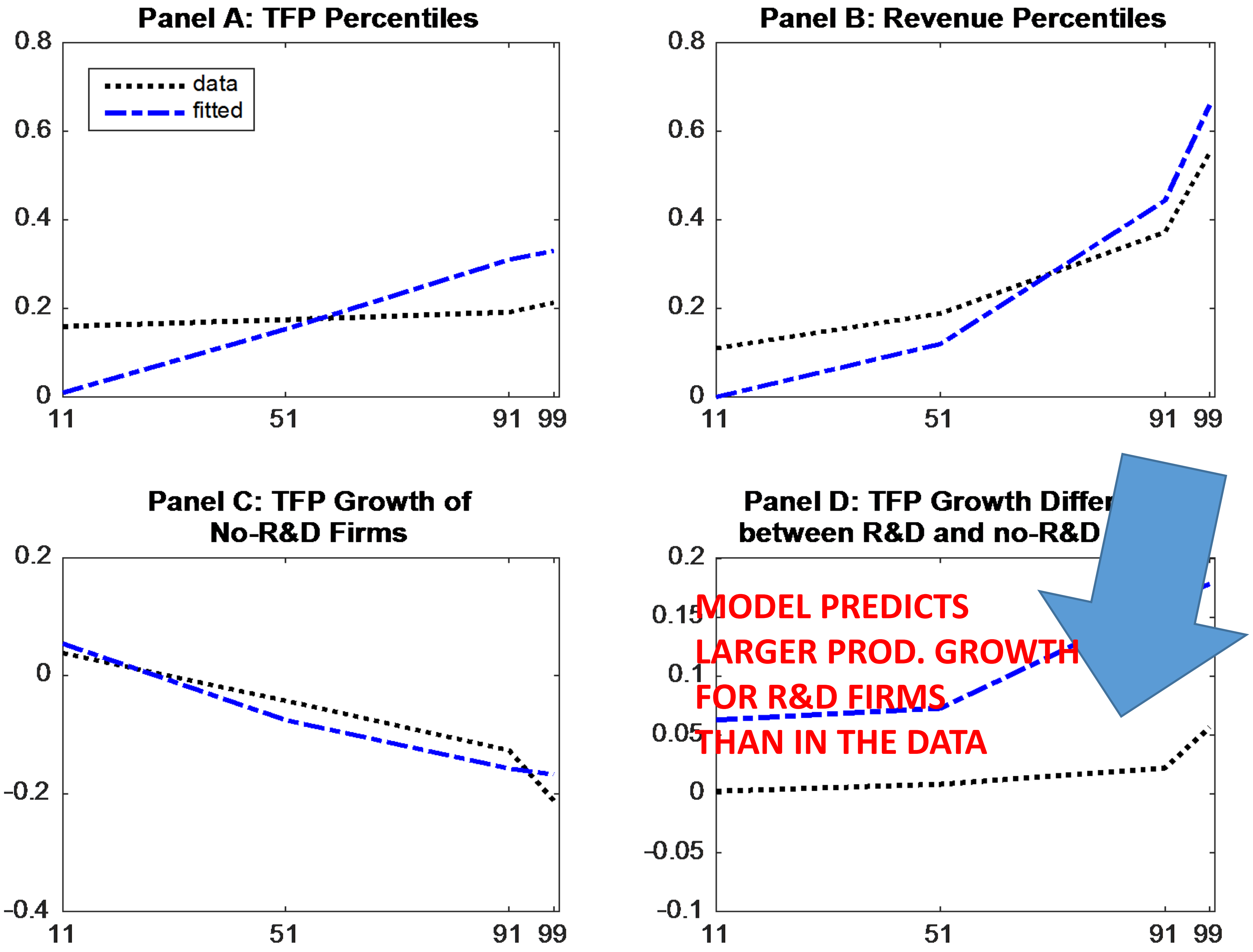
(Taiwan parameters, Chinese wedges, reestimate mean c)

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
std of c	0.59
	Re-estimated for China
mean of c	1.50



China: benchmark
(Taiwan parameters, Chinese wedges, reestimate mean c)

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
std of c	0.59
	Re-estimated for China
mean of c	1.50



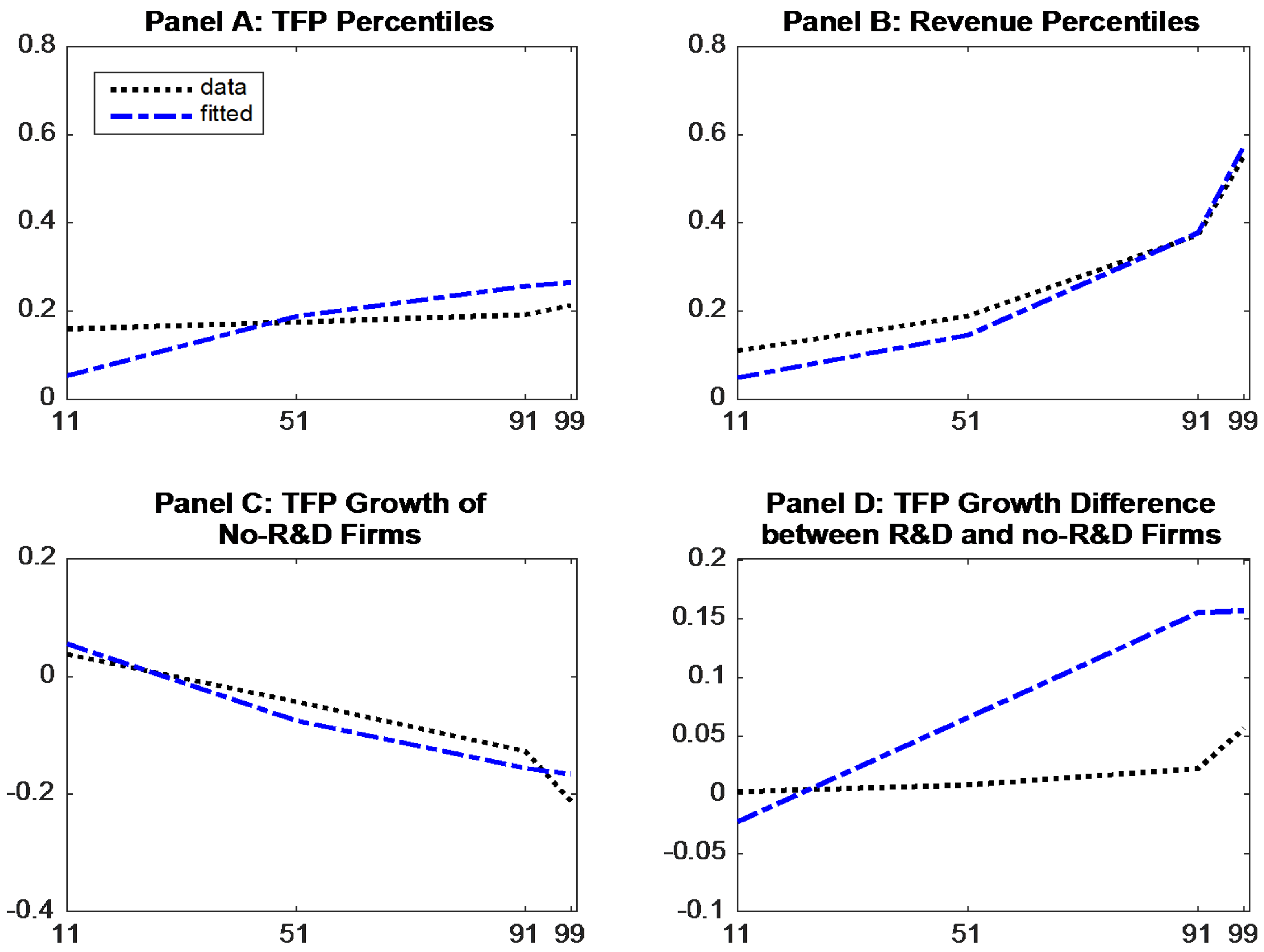
Counterfactuals

A. Quantitative failure of Taiwan model for China:
   i. Model predicts that R&D firms grow faster than in data
   ii. Model predicts steeper selection into R&D by TFP & revenue than in data

B. Candidate additional mechanisms
   1. Policy distortions scramble decisions (increased dispersion in C)
   2. Scarcity of innovative talent in China (lower p relative to Taiwan)
   3. Moral hazard in R&D

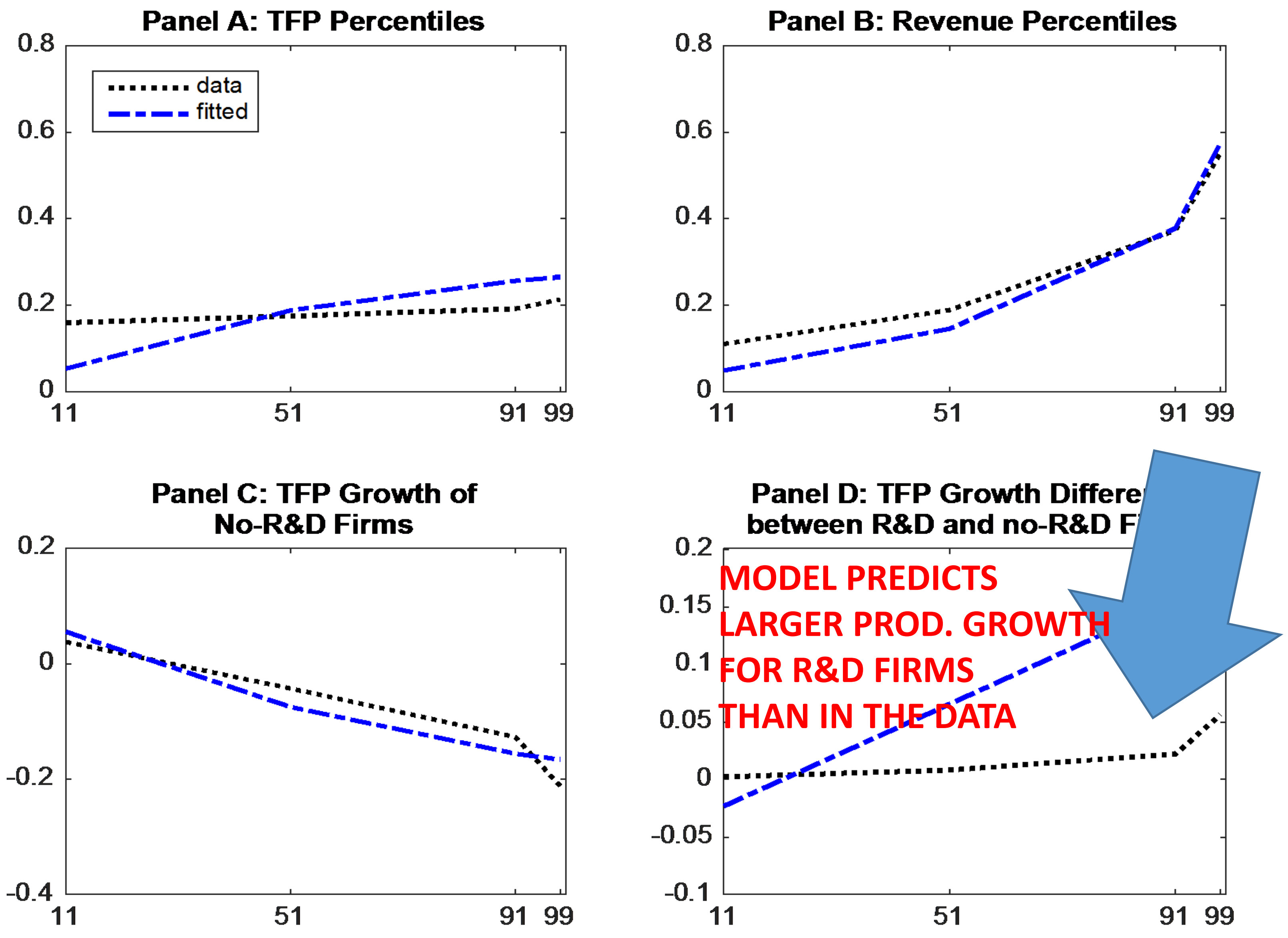
China: scrambling experiment
(increasing <u>**variance**</u> c)

	Estimates for Taiwan
q	0.45
δ	0.50
$\bar{p}$	0.26
	Re-estimated for China
mean of c	**7.50**
std of c	**7.60**



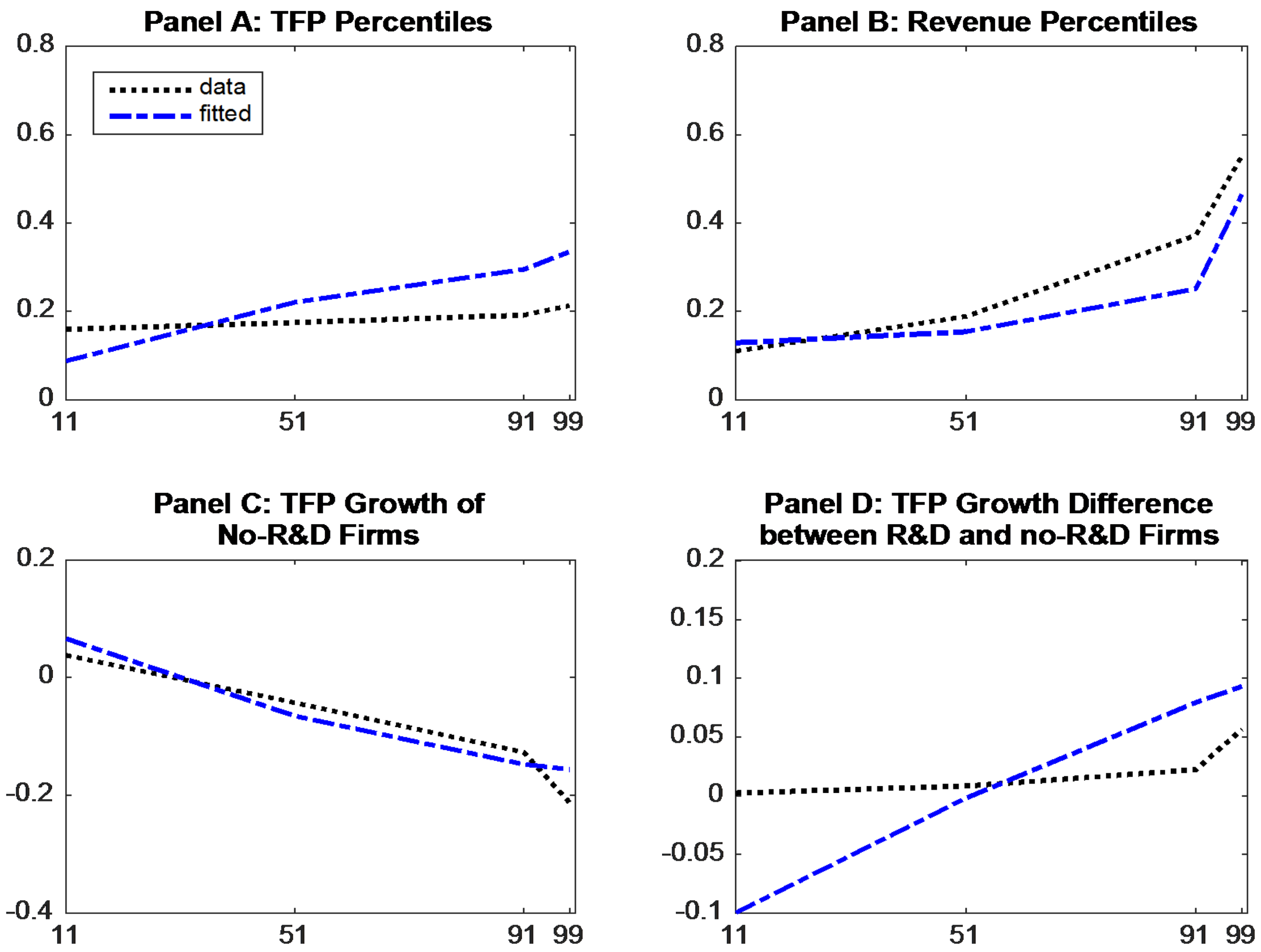
China: scrambling experiment
(increasing <u>**variance**</u> c)

	Estimates for Taiwan
q	0.45
δ	0.50
$\bar{p}$	0.26
	Re-estimated for China
mean of c	**7.50**
std of c	**7.60**



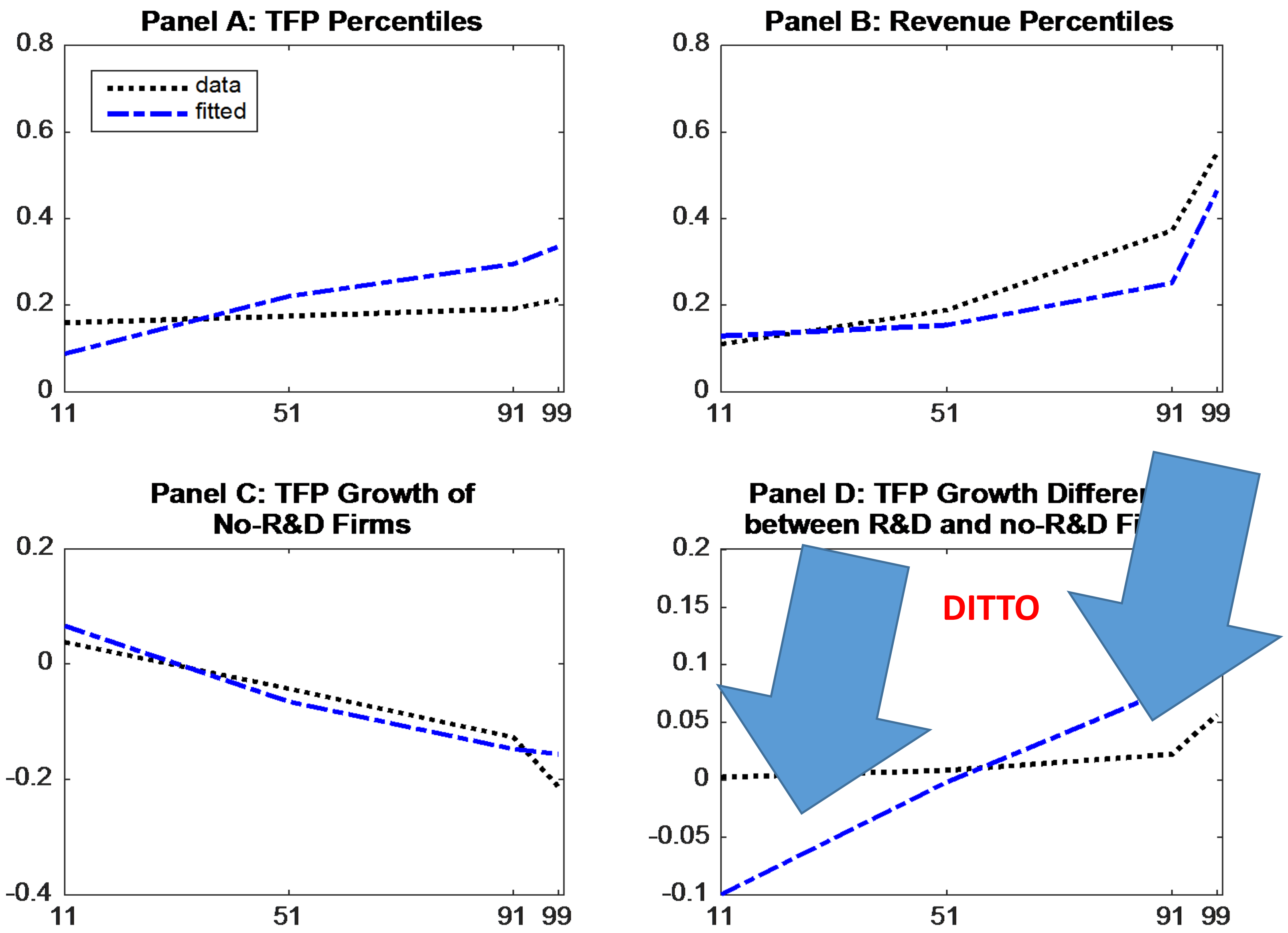
China: talent scarcity experiment
(lower $\overline{p}$)

	Estimates for Taiwan
q	0.45
δ	0.50
	Re-estimated for China
$\bar{p}$	**0.15**
mean of c	2.00
std of c	2.40



China: talent scarcity experiment
(lower $\overline{p}$)

	Estimates for Taiwan
q	0.45
δ	0.50
	Re-estimated for China
$\bar{p}$	**0.15**
mean of c	2.00
std of c	2.40



Moral hazard in R&D

- Assume $C_i = c_i A + \varepsilon_i = (c + \eta_i)A + \varepsilon_i$ where
  - η_i captures dispersion in technology (same $\mathrm{var}(\eta_i)$ in China and Taiwan)
  - ε_i is a tax/subsidy to R&D (only in China)

Moral hazard in R&D

- Assume $C_i = c_i A + \varepsilon_i = (c + \eta_i)A + \varepsilon_i$ where
  - η_i captures dispersion in technology (same $\mathrm{var}(\eta_i)$ in China and Taiwan)
  - ε_i is a tax/subsidy to R&D (only in China)
- Moral hazard: Firms can fake R&D
  - cash a subsidy and do imitation instead (avoiding cost and benefits of R&D)
  - Note: firms with low p and negative ε are likely to fake R&D

Moral hazard in R&D

- Assume $C_i = c_i A + \varepsilon_i = (c + \eta_i)A + \varepsilon_i$ where
  - η_i captures dispersion in technology (same $\mathrm{var}(\eta_i)$ in China and Taiwan)
  - ε_i is a tax/subsidy to R&D (only in China)
- Moral hazard: Firms can fake R&D
  - cash a subsidy and do imitation instead (avoiding cost and benefits of R&D)
  - Note: firms with low p and negative ε are likely to fake R&D
- Allow ε_i to be correlated with τ_i and A_i. Motivation:
  - Government supports more productive firms (subsidizes R&D in high-A firms)
  - Government supports its darlings (subsidizes R&D in low-τ firms, e.g. SOE)

$$\varepsilon_i = \varepsilon_{av} + c_1 A_i + c_2(1 - \tau_i)$$

China

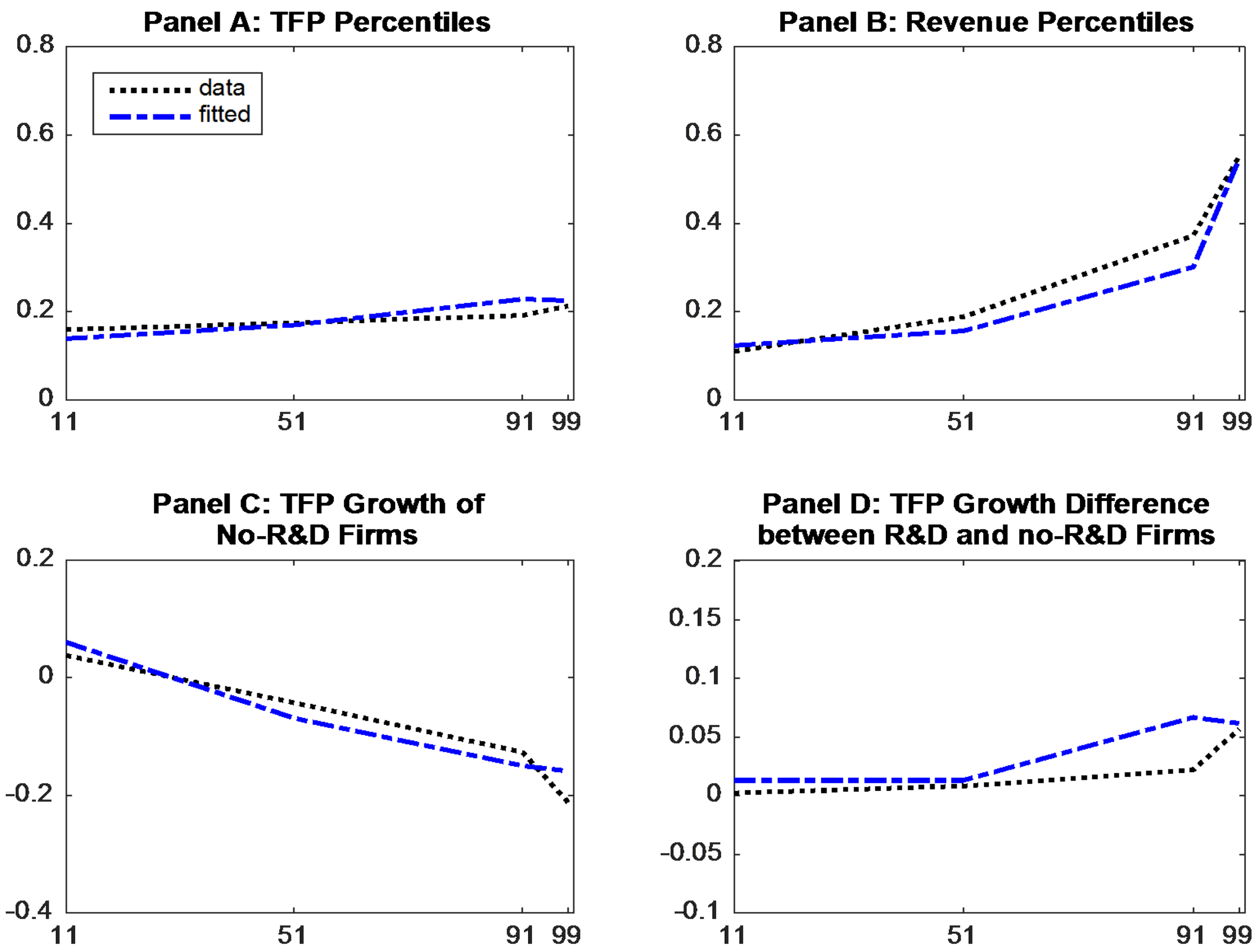
SOEs have higher propensity to R&D

Dep. Variable: R&D Dummy (extensive margin)
Balanced Panel 2001-2007

Variables	(1)	(2)	(3)
Log_tfp	**0.0265*** **	**0.362*** **	**0.0655*** **
	(0.00309)	**(0.0376)**	**(0.00385)**
Investm. wedge		-0.161***	-0.0271***
		(0.0191)	(0.00185)
Labor wedge		-0.230***	-0.0410***
		(0.0251)	(0.00299)
SOE	**0.190*** **	**0.0833*** **	**0.0109*** **
	(0.0272)	**(0.0170)**	**(0.00323)**
Log_tfp SOE	0.0371***	0.196***	0.00433
	(0.00548)	(0.0311)	(0.00575)
Investm. wedge SOE		-0.113***	-0.00005
		(0.0139)	(0.00341)
Labor wedge SOE		-0.0936***	-0.00330
		(0.0245)	(0.00322)
Industry Dummies	+	+	+
Year Dummies	+	+	+
Firm Dummies	-	-	+
Number of obs.	424,784	424,784	424,784

China: R&D Moral hazard experiment
(add ε_i)

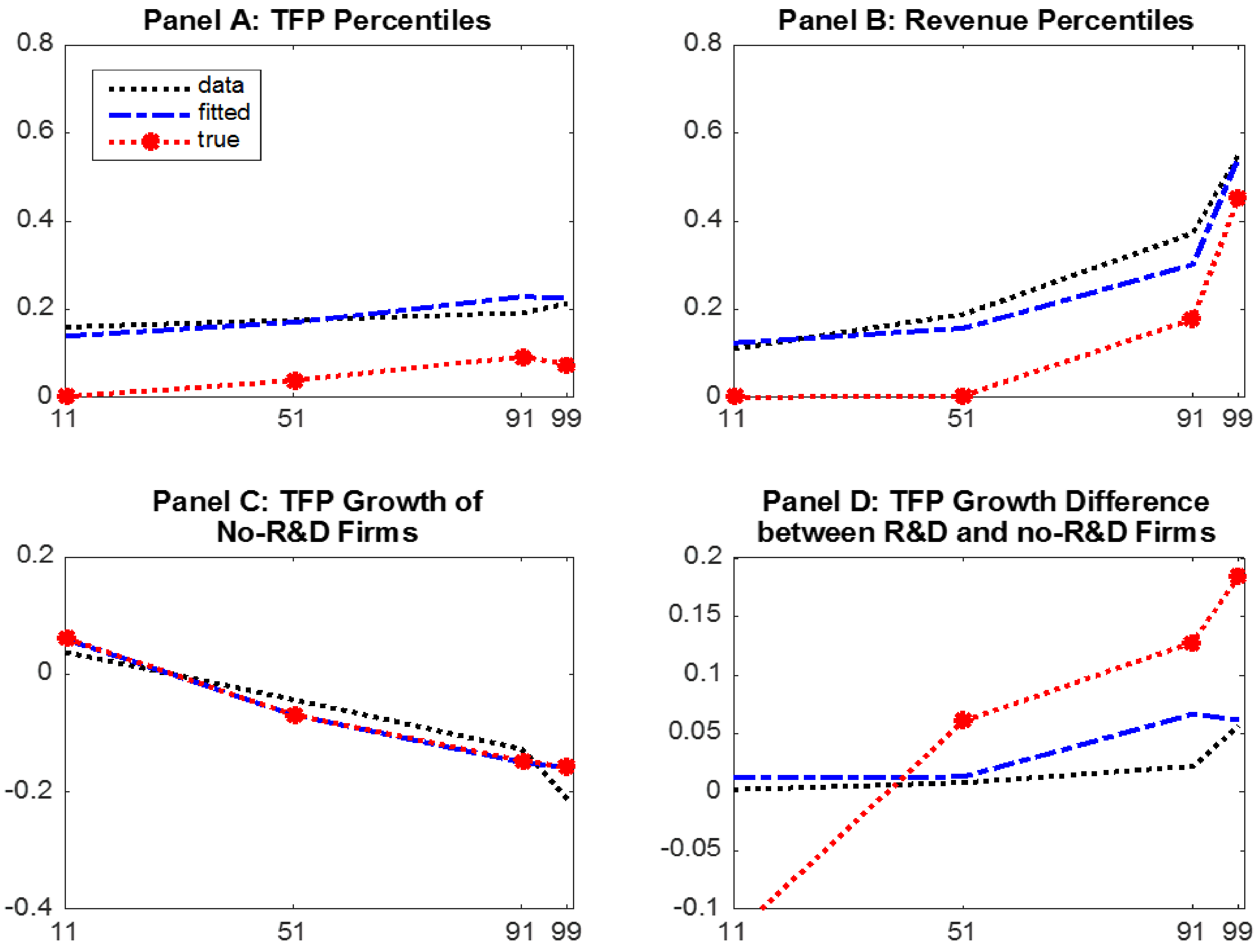
	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
	Re-estimated for China
mean of c	**3.50**
	Fake R&D
mean of ε	0.50
std of ε	**0.95**
c_1	**-0.21**
c_2	**-0.23**



China: R&D Moral hazard experiment

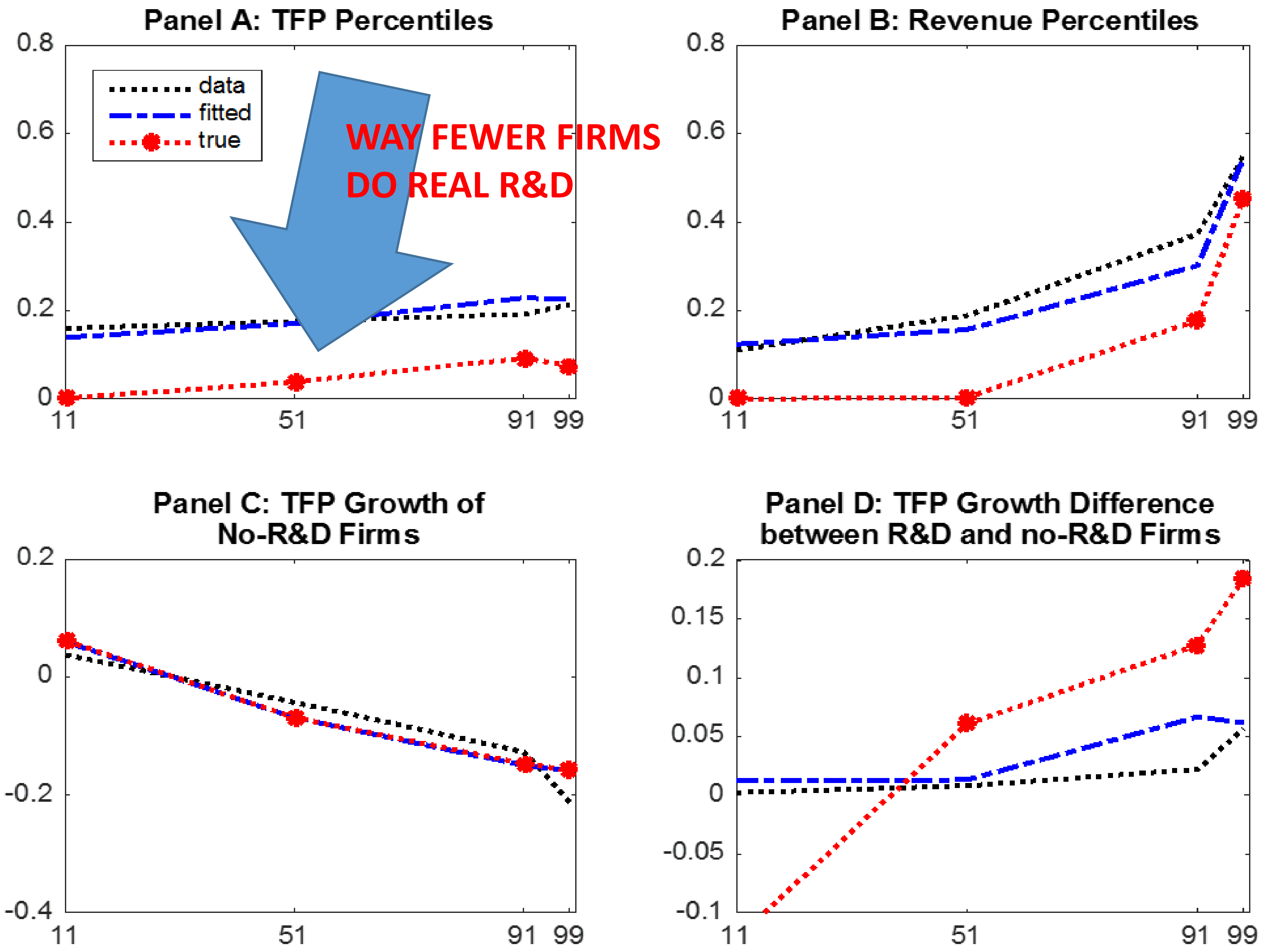
(add ε_i)

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
	Re-estimated for China
mean of c	**3.50**
	Fake R&D
mean of ε	0.50
std of ε	**0.95**
c_1	**-0.21**
c_2	**-0.23**



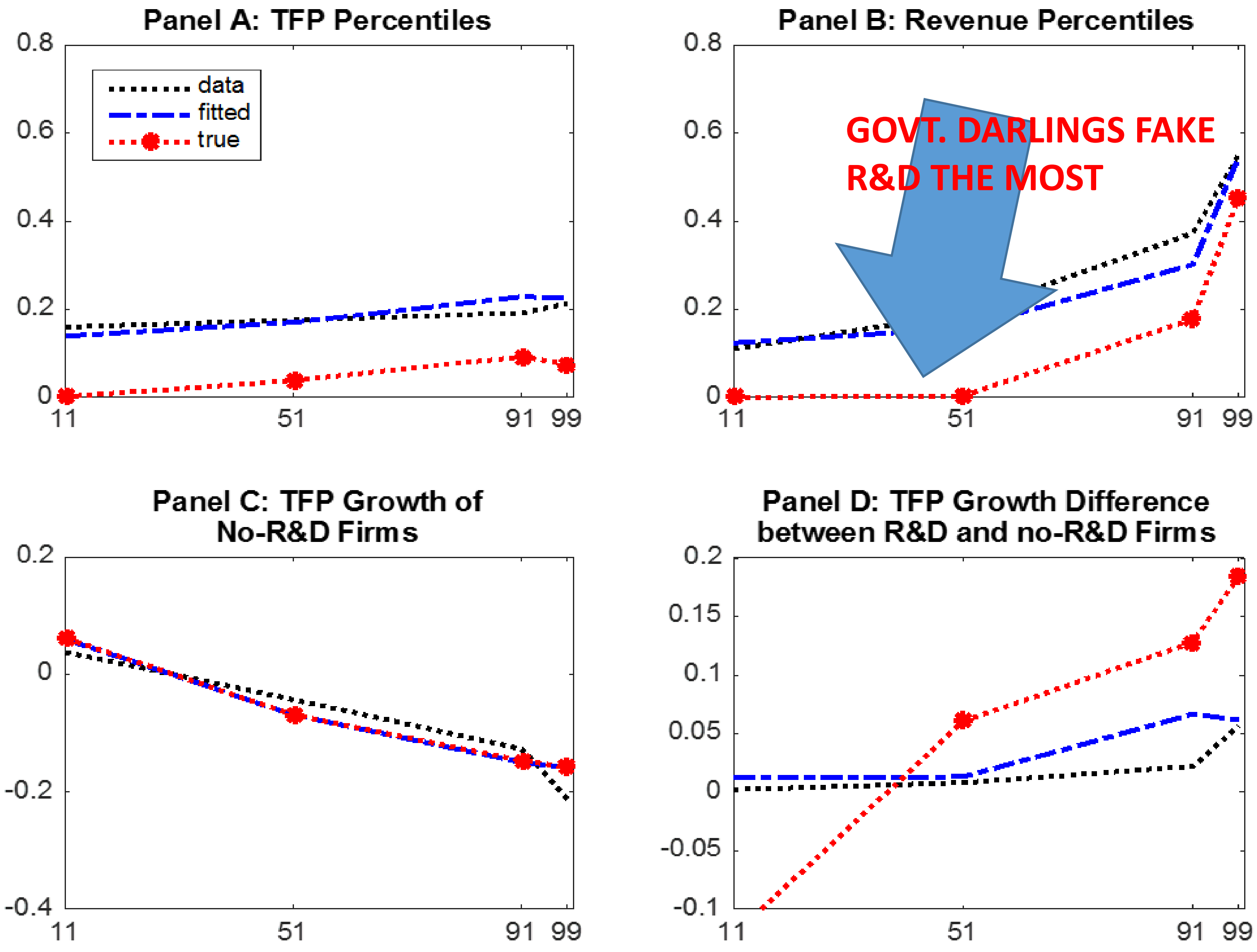
China: R&D Moral hazard experiment
(add ε_i)

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
	Re-estimated for China
mean of c	**3.50**
	Fake R&D
mean of ε	0.50
std of ε	**0.95**
c_1	**-0.21**
c_2	**-0.23**



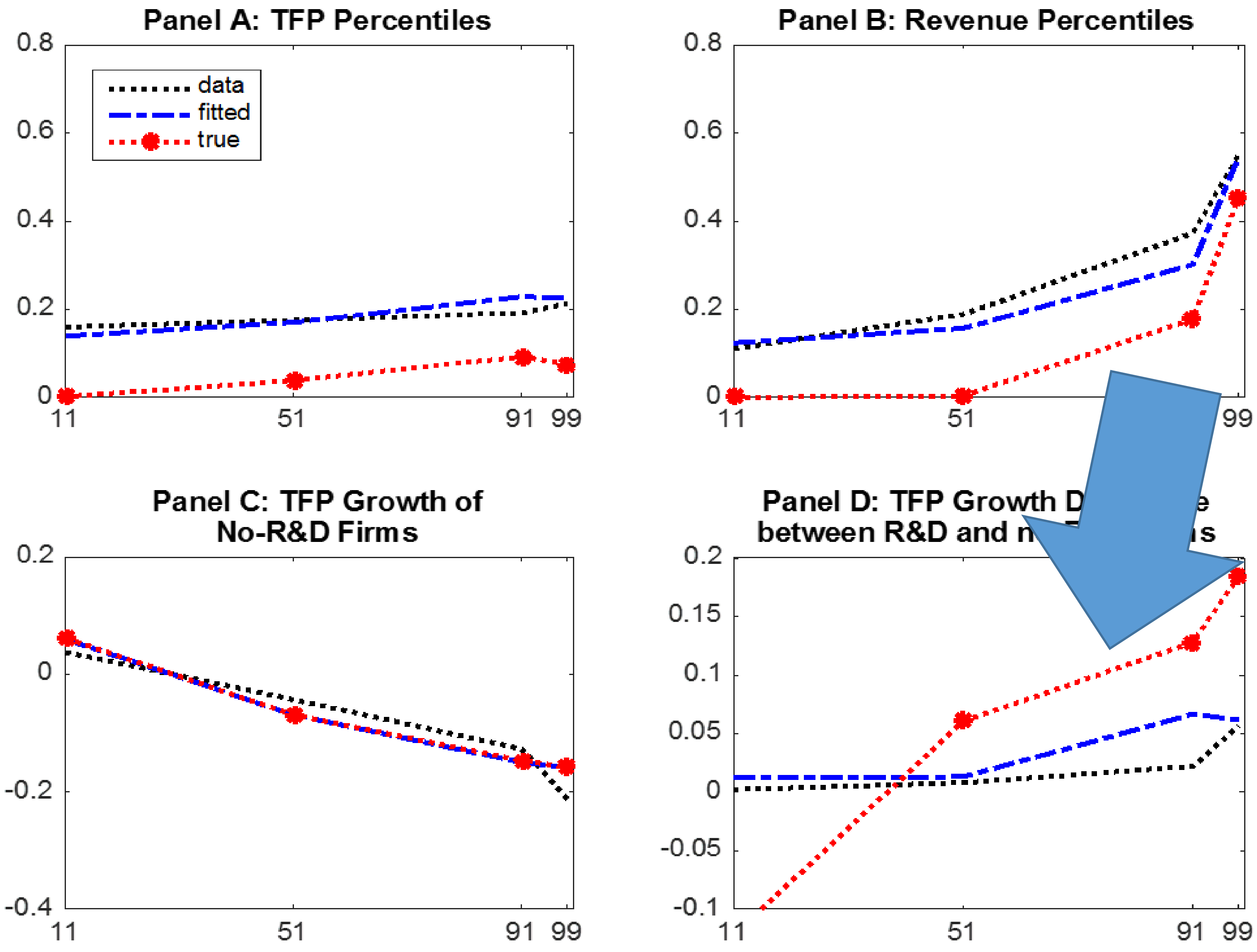
China: R&D Moral hazard experiment
(add ε_i)

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
	Re-estimated for China
mean of c	**3.50**
	Fake R&D
mean of ε	0.50
std of ε	**0.95**
c_1	**-0.21**
c_2	**-0.23**



China: R&D Moral hazard experiment
(add ε_i)

	Estimates for Taiwan
q	0.45
δ	0.40
$\bar{p}$	0.25
	Re-estimated for China
mean of c	**3.50**
	Fake R&D
mean of ε	0.50
std of ε	**0.95**
c_1	**-0.21**
c_2	**-0.23**



FIRMS DOING REA R&D DO MUCH BETTER: HUAWEI, etc.

Macro Effects of Removing R&D Distortions

- Removing R&D distortions estimated for China

1. TFP growth up by 0.8 percentage points (mean of c re-estimated to match the share of R&D firms)

2. TFP growth up by 1.4 percentage points (also adjusting c to Taiwanese level)

Conclusion

- Document evidence on firm-level distribution of R&D and growth in manufacturing industries in China and Taiwan
- Develop a theory of innovation (driven by R&D), imitation, and growth, with a focus on R&D misallocation
- Estimate the model using firm-level data from Taiwan and China
- Evaluate counterfactual: remove R&D distortions in China relative to Taiwan
- Next: extend analysis to Western economies (use data for Norway)

China:
TFP positively Correlated with R&D

Dep. Variable: R&D Dummy (extensive margin)
Balanced Panel 2001-2007

Variables	(1)	(2)	(3)
Log_tfp	**0.0250*** **	**0.426*** **	**0.0827*** **
	(0.00321)	**(0.0465)**	**(0.00408)**
Investm. wedge (log_y/k)		-0.197***	-0.0306***
		(0.0235)	(0.00199)
Labor wedge (log_y/l)		-0.270***	-0.0547***
		(0.0309)	(0.00315)
Industry Dummies	+	+	+
Year Dummies	+	+	+
Firm Dummies	-	-	+
Number of obs.	424,784	424,784	424,784

Notes: Robust standard errors in parentheses

Firms' life cycle (cont.)

- A young entrepreneur inherits the TFP of the parent's firm

  (subject to shocks)

$$\log(A_i(t)) = \log(A_i^{OLD}(t)) + \theta\varepsilon_i \ (t)$$

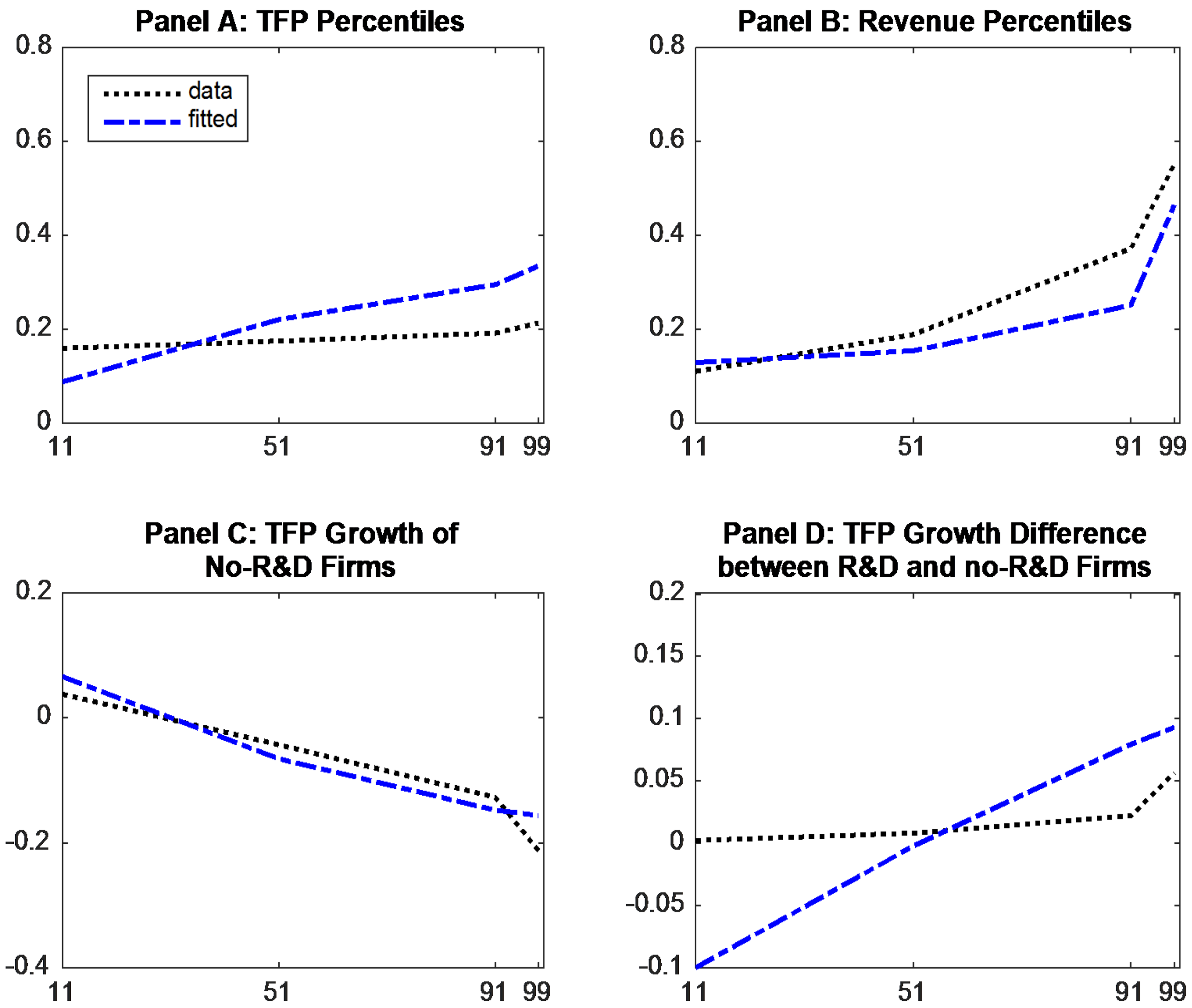
  ε is (discrete) normally distributed

- The young entrepreneur also draws an output wedge τ

  from a distribution $\phi(A)$

China:

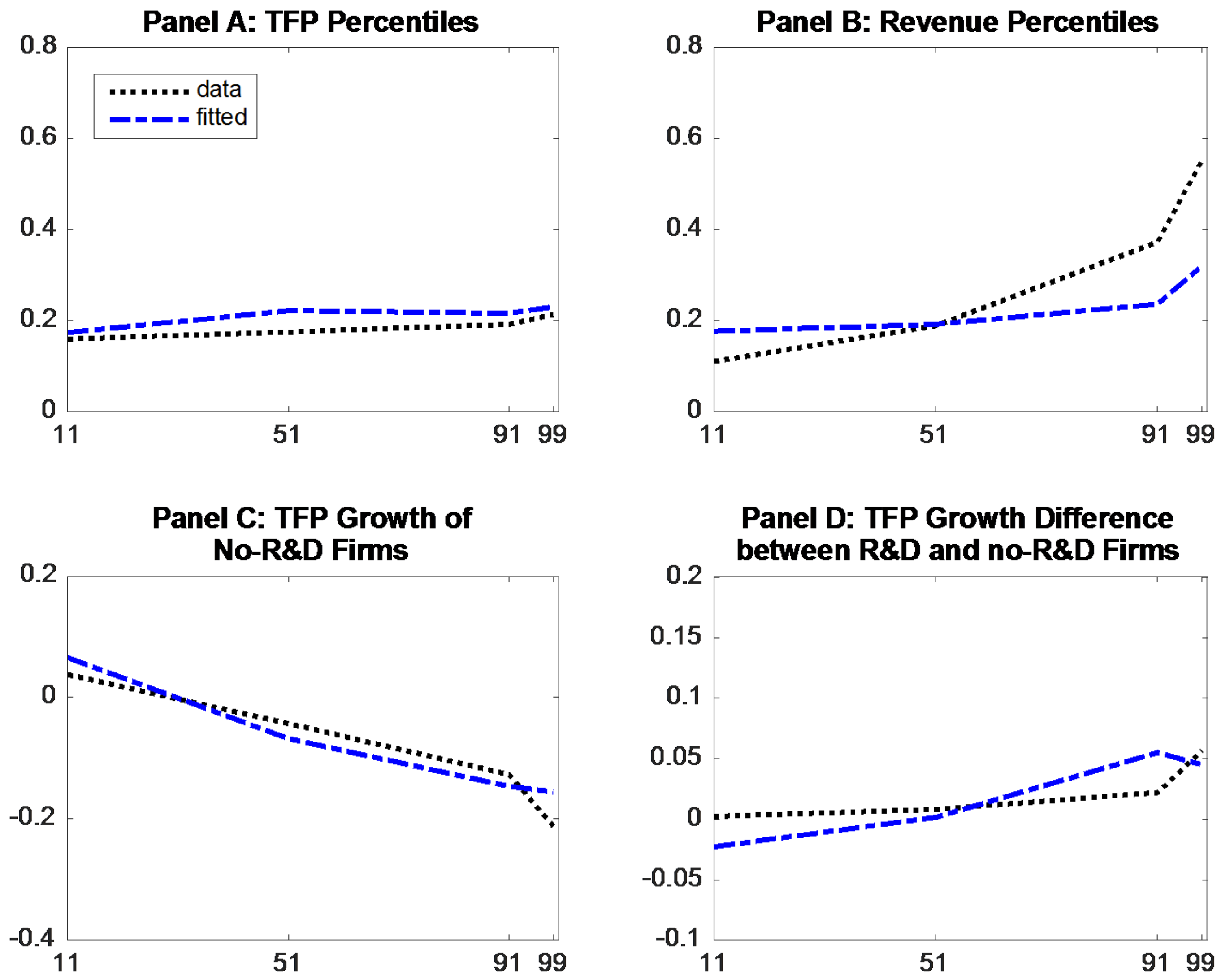
Talent scarcity + scrambling (reestimate $\overline{p}$ and mean, variance of c)



China:

different R&D technology (reestimate all param.)
R&D ≈ imitation

	Estimates for China
q	0.45
δ	**0.90**
$\bar{p}$	**0.05**
mean of c	2.20
std of c	2.40



Related literature

- Technological convergence through innovation/imitation
    - Acemoglu, Aghion & Zilibotti (AAZ 2006), König, Lorenz and Zilibotti (KLZ 2016)

- Endogenous dynamics of productivity distribution
    - Lucas & Moll (2014), Perla & Tonetti (2014), Benhabib, Perla & Tonetti (2017), Luttmer (2007), Ghiglino (2011), König, Lorenz & Zilibotti (2016)

- R&D investments and policy
    - Klette & Kortum (2004), Akcigit & Kerr (2017), Acemoglu, Akcigit, Bloom & Kerr (2013), Hsieh & Klenow (2015), Lentz & Mortensen (2008)

- Misallocation in China
    - Hsieh & Klenow (2009), Song, Storesletten & Zilibotti (2011), Hsieh & Song (2016), Cheremukhin, Golosov, Gurev & Tsyvinski (2016), Tombe & Zhu (2016), Zilibotti (2017)

Model

- Final good production

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

- This yields isoelastic demands for each good

$$P_i(t) = \left(\frac{Y_i(t)}{Y(t)} \right)^{-\vartheta}$$

- Production function of each good is Cobb-Douglas

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

(Static) Equilibrium

$$\pi(\tau_i, A_i) = max_{\{K(i),L(i),Y(i)\}} P_i Y_i - (1+\tau_{Li}) w L_i - (1+\tau_{Ki}) r K_i$$

- Solution:

$$(1-\alpha)(1-\vartheta)\frac{Y_i}{L_i} = (1+\tau_{Li})w$$

$$\alpha(1-\vartheta)\frac{Y_i}{K_i} = (1+\tau_{Ki})r$$

$$Y_i = \propto \left(A_i(1-\tau_i)\right)^{\frac{\vartheta}{1-\vartheta}}$$

where $(1-\tau_i) \stackrel{\text{def}}{=} (1-\tau_{Li})^{\alpha-1}(1-\tau_{Li})^{-\alpha}$