Who Pays for Training?

Theory and Evidence on Firm-Level Differences in Training Investments

Xiao Ma
Peking Univ.

Alejandro Nakab
Univ. Torcuato di Tella

Daniela Vidart
Univ. of Connecticut

Banco de España and Banco de Portugal Conference on Labour Markets

Motivation

- Lifetime human capital has two main components: Schooling & **On-the-Job**
  - Key for lifecycle earning dynamics & dispersion (Becker (1962), Rubinstein and Weiss (2006))
  - Allocation of costs between firms and workers hard to observe
  - Also unclear what firm/worker characteristics matter for on-the-job training provision
- **This paper**: Examine how training changes with firm characteristics, and use this to inform allocation of learning costs
  - Empirical evidence: Training lower in smaller firms, driven by TFP and labor share differences
  - Theory: Use empirics to shed light on cost-sharing + examine inefficiencies in training investments across firms

Overview

- **Empirics**: Show how training changes with firm characteristics
  - Harmonized on-the-job training definition capturing key margins of worker learning
  - Data from 100+ countries: Training opportunities are consistently lower in smaller firms
  - Admin data from Mexico and China: TFP and labor share key in driving this
    - Firms with higher TFP invest more in training
    - Firms with lower labor share invest less

- **Analytical model**: Consider 4 cost-sharing schemes for explicit (out-of-pocket) costs
  1. Firms bear all explicit training costs & determine training
  2. Workers bear all explicit training costs & determine training
  3. Max joint surplus; cost-sharing determined by Nash bargaining
  4. Workers & firms pay fixed share; training determined by party with lower affordability

$\Rightarrow$ Only cases where firms choose training match TFP & labor share results

Overview

- **Empirics**: Show how training changes with firm characteristics
  - Harmonized on-the-job training definition capturing key margins of worker learning
  - Data from 100+ countries: Training opportunities are consistently lower in smaller firms
  - Admin data from Mexico and China: TFP and labor share key in driving this
    - Firms with higher TFP invest more in training
    - Firms with lower labor share invest less

- **Analytical model**: Consider 4 cost-sharing schemes for explicit (out-of-pocket) costs
  1. Firms bear all explicit training costs & determine training
  2. Workers bear all explicit training costs & determine training
  3. Max joint surplus; cost-sharing determined by Nash bargaining
  4. Workers & firms pay fixed share; training determined by party with lower affordability

$\Rightarrow$ Only cases where firms choose training match TFP & labor share results

Overview

- **Quantitative Model**: Calibrate to US & quantify size of inefficiencies + scope of policy
  - Firms paying a calibrated fixed share of explicit costs fits evidence on training returns best
  - Inefficiency: Firms fail to internalize gains to workers, & other firms after separation
  - Optimal training subsidy larger in smaller firms but still substantial in large firms
  - However, even providing same subsidy to all firms generates 7% increase in net output in US

Literature review

1. On-the-Job Training & Funding

- Will firms fund general training?
  (Becker (1964), Acemoglu (1997), Acemoglu and Pischke (1999), Moen and Rosén (2004))
- Our contribution: Use data and theory to inform cost-sharing schemes

2. Firm-level differences in training investments

- Training & human capital acquisition varies across firms
  (Black et al. (1999) and Braga (2018), Gregory (2019), Friedrich et al. (2021), Jarosch (2022))
- Importance of firm size
  (Barron et al. (1987), Frazis et al. (1995), Harris (1999))
- Importance of productivity
  (Acemoglu and Pischke (1998) and Bagger et al. (2014))
- Our contribution: Extend analysis to many countries + separate role of TFP & labor share

Data and Empirical Findings: Differences in Training Investments by Firm Size

Definition of training

- Organized and structured learning activity outside education system; two categories:
- *Formal training*
  - High degree of organization, planned in advance, typically separated from workplace
  - Examples: Classes, seminars, presentations
- *Informal training*:
  - Less structured, typically connected to the workplace, and adapted to specific-worker needs
  - Examples: Learning through colleagues, learning circles, job rotation
- Our definition of training:
  - Excludes schooling, informal learning (e.g learning thru media), and learning-by-doing
  - Relates to job-related training: Improve work skills, or opportunities for promotion.

Other sources | Formal Training Def | Informal Training Def | Job-relation Def

Data

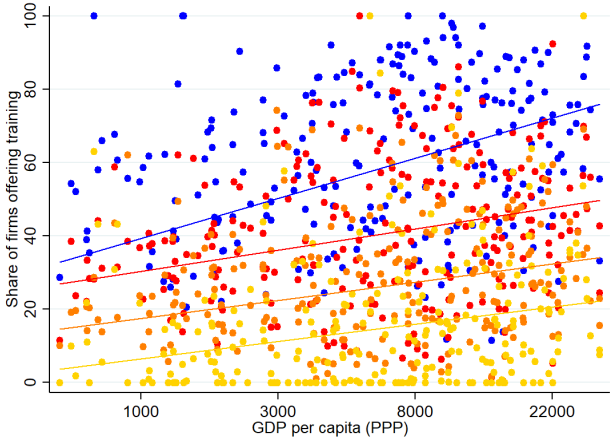
Firm-Level Data

- World Bank Enterprise Survey (76 countries; 2005–2017)
- EU Continuing Vocational Training (26 countries; 2005, 2010, 2015)
- Construct harmonized cross-country representative measures
- 400,000 firms in broad range of development: GDP per capita $1,000 - $60,000; split by firm-size categories

Worker-Level Data

- OECD PIAAC Survey (34 countries; 2011–2017)
- EU Continuing Vocational Training (26 countries; 2005, 2010, 2015)
- Construct harmonized cross-country representative measures
- 50,000 workers in broad(ish) range of development; split by firm-size categories

Firm-level data: Share of firms offering formal training by firm size

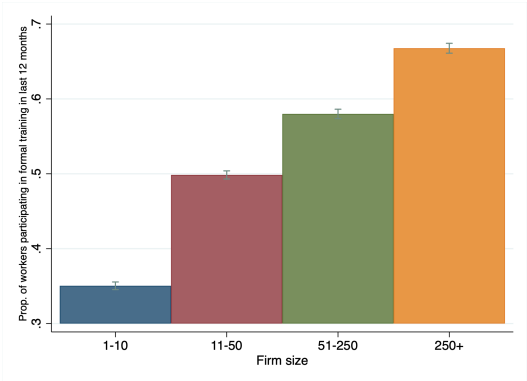


Each dot represents the share of firms in a specific firm size category offering formal training in each country. The firm sizes considered are: 1-5 (Gold), 6-20 (Orange), 21-100 (Red), and 100+ (Blue). Data on training comes from the WB-ES. Data on GDP per capita come from the Penn World Tables.

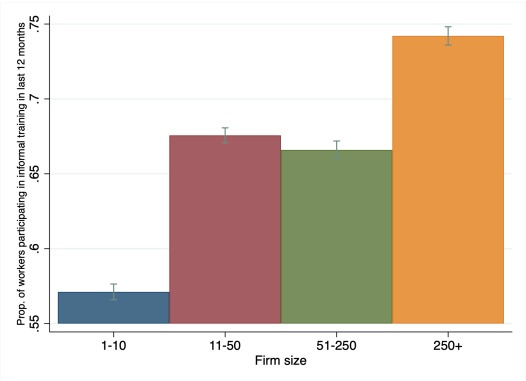
Table Training Purpose

Worker-level data: Share of workers participating in training in last 12 months by firm size

(a) Formal Training



(b) Informal Training



Hours of Formal Training Table

Data and Empirical Findings: Role of TFP and Labor Shares in Training Investments

Data and definition of training, TFP, labor share

- Chinese Annual Survey of Manufacturing (revenue $>$ 5 million RMB; 2001 & 2005-2007)
  - Training: Expenditures in formal training > 0
- Mexican Economic Census (firms not in agriculture operating in fixed location; 2019)
  - Training: Share of employees formally/informally trained > 0
- TFP: Residual of Cobb-Douglas production function of capital and labor, using firm-level data on revenue, payroll, and fixed capital stock (Hsieh and Klenow (2009)) Details
  1. Set elasticity w.r to labor $1 - \alpha = \frac{2}{3}$
  2. Let $1 - \alpha$ vary across industries (preferred)
- Labor share: Ratio between firm's payroll and sales
- Sample: Firms with 1+ paid workers, positive value of sales, VA, payroll & fixed capital

Correlation between Training, TFP and Labor Share

Dep. Variable:	China Firm Offers Training				Mexico Firm Offers Training	
	(1)	(2)	(3)	(4)	(1)	(2)
Labor Share	-0.150***	-0.024*	-0.188***	-0.031**	-0.038***	-0.10***
	(0.008)	(0.013)	(0.008)	(0.013)	(0.002)	(0.002)
log TFP (labor share = 2/3)	0.007***	0.010***			0.014***	
	(0.001)	(0.001)			(0.0003)	
log TFP (HK)			0.002***	0.009***		0.004***
			(0.0004)	(0.001)		(0.0002)
log(Firm Size)	0.079***	0.038***	0.080***	0.037***	0.096***	0.102***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.0003)	(0.0003)
Age FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y		
Industry FE	Y	Y	Y	Y	Y	Y
Firm FE		Y		Y		
Observations	772,599	658,495	768,475	654,775	1,561,690	1,375,834
R-squared	0.064	0.704	0.064	0.704	0.212	0.218

Intensive margin | Labor share results in WB-ES | Correlation bw. firm size and labor share in WB-ES

Analytical Model: Which Cost-Sharing Schemes Fit our Findings?

Model setup

- **Workers**: Ex ante homogeneous, accumulate human capital throughout life
  - Live for 2 periods; population size of each age normalized to 1
  - Supply one unit of labor inelastically each period
  - Max discounted lifetime consumption:

$$\max_{\{c^Y, c^O\}} \; c^Y + \frac{c^O}{1+\rho},$$
$$s.t. \quad c^Y + \frac{c^O}{1+r} = w^Y + \frac{w^O}{1+r},$$

- **Firms**: Post vacancies and wages to attract workers
  - Heterogenous in productivity $z \sim G(z)$
  - Firm j-worker i specific production is $y_{ji} = z_j h_i$

- Wage posting allows for:
  - Meaningful variation in allocation of explicit (or out-of-pocket) training costs
  - Variation in labor-shares across firms

Search and matching

Job searchers:

- Job destruction: probability δ of job-to-unemployment
- On-the-job searchers: η portion of workers look for new jobs
- Young workers are born unemployed

$$\tilde{U} = (1 + \eta(1 - \delta) + \delta)$$

Vacancies:

- Firm posts vacancies $v(z)$ every period with costs $\psi(v) = \frac{c_v v^{1+\gamma_v}}{1+\gamma_v}$, $\gamma_v > 0$
- Contract stipulates wage rate per efficiency unit $w(z)$ and working period (2 periods for young, 1 period for old)

$$V = \int v(z) dG(z)$$

Matching process:

- Matching function: $M(U, V) = \min\{U, V\}$

Human capital accumulation

- Human capital accumulation depends on training s:

$$h^O = h^Y + \zeta s^{\gamma_s}$$

- Training costs:
  - Direct costs $c_s \bar{w}$ per unit of training time; proportional to average wage $\bar{w}$
  - Opportunity cost; 1 unit of training time causes a 1 unit decrease in production time
- Firms can offload some of expected training costs thru lower wages posted (*implicit* costs)
- But, after match and observing worker's h, *explicit* (out-of-pocket) costs & training decision remain
- We consider 4 options for how training is decided & paid:
  1. Firms bear all explicit training costs & determine training
  2. Workers bear all explicit training costs & determine training
  3. Max joint surplus; cost-sharing determined by Nash bargaining
  4. Workers & firms pay fixed share; training determined by party with lower affordability

Firms' and workers' marginal benefits from human capital

$$MR_W(z) = \underbrace{(1-\delta)\left(1-\eta\bar{F}(w)\right)}_{\text{prob worker stays in firm}} \underbrace{\beta(z)z}_{\text{worker's share of surplus}} + \underbrace{(1-\delta)\eta\int_w^\infty \beta(z')z'dF(w(z'))}_{\text{expected share of surplus in new firm}} + \underbrace{\delta\int \beta(z')z'dF(w(z'))}_{\text{expected share of surplus after unemp}}$$

$$MR_F(z) = \underbrace{(1-\delta)\left(1-\eta\bar{F}(w)\right)}_{\text{prob worker stays in firm}} \underbrace{(1-\beta(z))z}_{\text{firm's share of surplus}}$$

- $F(w)$ is wage distribution of offers
- $\beta(z)$ is labor share where $\beta(z) = w(z)/z$
- $\bar{F}(w) = \int_w^\infty dF(w(z))$ is the probability of obtaining an offer with higher wage than w

Firms' and workers' gains from human capital

$$MR_W(z) = \underbrace{(1-\delta)\left(1-\eta\bar{F}(w)\right)}_{\text{prob worker stays in firm}} \underbrace{\beta(z)z}_{\text{worker's share of surplus}} + \underbrace{(1-\delta)\eta\int_w^\infty \beta(z')z'dF(w(z'))}_{\text{expected share of surplus in new firm}} + \underbrace{\delta\int \beta(z')z'dF(w(z'))}_{\text{expected share of surplus after unemp}}$$

$$MR_F(z) = \underbrace{(1-\delta)\left(1-\eta\bar{F}(w)\right)}_{\text{prob worker stays in firm}} \underbrace{(1-\beta(z))z}_{\text{firm's share of surplus}}$$

- $F(w)$ is wage distribution of offers
- $\beta(z)$ is labor share where $\beta(z) = w(z)/z$
- $\bar{F}(w) = \int_w^\infty dF(w(z))$ is the probability of obtaining an offer with higher wage than w
- Connection to empirical facts
  - Both firms and workers have higher benefits from training if z higher (though current firm's z more important for firms)
  - But, only firms benefit more from training if $\beta(z)$ lower
  - Firms must determine training for consistency with empirical facts!

Connection to empirical facts

- Consider productivity z (affects labor prod and poaching prob) & labor share $\beta(z)$ separately

- **When firms choose**, training increases with z ✓ and decreases with $\beta(z)$ ✓
  - Link to z: Supermodularity of production; alleviation of hold-up problem
  - Link to $\beta(z)$: Less revenue perceived by firm

- **When worker chooses**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Current z less important as care about future; opp cost of training may dominate
  - Link to $\beta(z)$: More revenue perceived by worker

- **When maximize joint surplus**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Same as when worker chooses
  - Link to $\beta(z)$: Higher $\beta(z)$ reduces J-J transitions making training more desirable

⇒ **Firms must determine training for consistency with empirical facts!**

- This will happen when firms pay all costs; or pay fixed share of costs and $\mu_F > \mu_W \frac{MR_F(z)}{MR_W(z)}$
  - Will be often true, as $MR_F(z) << MR_W(z)$

Connection to empirical facts

- Consider productivity z (affects labor prod and poaching prob) & labor share $\beta(z)$ separately

- **When firms choose**, training increases with z ✓ and decreases with $\beta(z)$ ✓
  - Link to z: Supermodularity of production; alleviation of hold-up problem
  - Link to $\beta(z)$: Less revenue perceived by firm

- **When worker chooses**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Current z less important as care about future; opp cost of training may dominate
  - Link to $\beta(z)$: More revenue perceived by worker

- **When maximize joint surplus**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Same as when worker chooses
  - Link to $\beta(z)$: Higher $\beta(z)$ reduces J-J transitions making training more desirable

⇒ **Firms must determine training for consistency with empirical facts!**

- This will happen when firms pay all costs; or pay fixed share of costs and $\mu_F > \mu_W \frac{MR_F(z)}{MR_W(z)}$
  - Will be often true, as $MR_F(z) << MR_W(z)$

Connection to empirical facts

- Consider productivity z (affects labor prod and poaching prob) & labor share $\beta(z)$ separately

- **When firms choose**, training increases with z ✓ and decreases with $\beta(z)$ ✓
  - Link to z: Supermodularity of production; alleviation of hold-up problem
  - Link to $\beta(z)$: Less revenue perceived by firm

- **When worker chooses**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Current z less important as care about future; opp cost of training may dominate
  - Link to $\beta(z)$: More revenue perceived by worker

- **When maximize joint surplus**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Same as when worker chooses
  - Link to $\beta(z)$: Higher $\beta(z)$ reduces J-J transitions making training more desirable

⇒ **Firms must determine training for consistency with empirical facts!**

- This will happen when firms pay all costs; or pay fixed share of costs and $\mu_F > \mu_W \frac{MR_F(z)}{MR_W(z)}$
  - Will be often true, as $MR_F(z) << MR_W(z)$

Connection to empirical facts

- Consider productivity z (affects labor prod and poaching prob) & labor share $\beta(z)$ separately

- **When firms choose**, training increases with z ✓ and decreases with $\beta(z)$ ✓
  - Link to z: Supermodularity of production; alleviation of hold-up problem
  - Link to $\beta(z)$: Less revenue perceived by firm

- **When worker chooses**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Current z less important as care about future; opp cost of training may dominate
  - Link to $\beta(z)$: More revenue perceived by worker

- **When maximize joint surplus**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Same as when worker chooses
  - Link to $\beta(z)$: Higher $\beta(z)$ reduces J-J transitions making training more desirable

⇒ **Firms must determine training for consistency with empirical facts!**

- This will happen when firms pay all costs; or pay fixed share of costs and $\mu_F > \mu_W \frac{MR_F(z)}{MR_W(z)}$
  - Will be often true, as $MR_F(z) << MR_W(z)$

Connection to empirical facts

- Consider productivity z (affects labor prod and poaching prob) & labor share $\beta(z)$ separately

- **When firms choose**, training increases with z ✓ and decreases with $\beta(z)$ ✓
  - Link to z: Supermodularity of production; alleviation of hold-up problem
  - Link to $\beta(z)$: Less revenue perceived by firm

- **When worker chooses**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Current z less important as care about future; opp cost of training may dominate
  - Link to $\beta(z)$: More revenue perceived by worker

- **When maximize joint surplus**, training may not increase with z × and increases with $\beta(z)$ ×
  - Link to z: Same as when worker chooses
  - Link to $\beta(z)$: Higher $\beta(z)$ reduces J-J transitions making training more desirable

⇒ **Firms must determine training for consistency with empirical facts!**

- This will happen when firms pay all costs; or pay fixed share of costs and $\mu_F > \mu_W \frac{MR_F(z)}{MR_W(z)}$
  - Will be often true, as $MR_F(z) << MR_W(z)$

Quantitative Model: Quantifying Training Inefficiencies Across Firms & Policy

Quantitative setup extensions

- Workers live for $J > 2$ periods
- Human capital depreciates at rate d every period, but remains above lower bound
- Firms' productivity is Pareto-distributed: $G(z) = 1 - z^{-\kappa}$ (Axtell, 2001)
- Matching function: $M(\tilde{U}, V) = c_M \tilde{U}^{\psi} V^{1-\psi}$

Calibration

- Model period = quarter
- Working life $J = 160$ periods (40 years)
- Calibrated to US. Two sets of parameters:

1. Externally calibrated
2. Internally calibrated

Calibration

- Model period = quarter
- Working life $J = 160$ periods (40 years)
- Calibrated to US. Two sets of parameters:

1. Externally calibrated

Parameter	Model	Source
ρ - Discount rate	0.01	Annualized interest rate of 0.04
b - Ratio of lowest wage to average wage	0.6	Hornstein et al. (2011)
ψ - Elasticity of matches to searchers	0.7	Shimer (2005)
γ_v - Convexity of vacancy costs	1	Dix-Carneiro et al. (2019)

2. Internally calibrated

Internal calibration

– For each cost-sharing scenario minimize distance between moments in data and model

Moments	Data	Model			
		Firms Pay	Workers Pay	Maximize Joint Surplus	Constant Cost Shares
Moments: labor market					
Unemployment rate (%)	6.5	6.6	6.5	6.4	6.3
Ratio of #Vacancies to #Unemployed	0.55	0.55	0.54	0.58	0.52
Pareto parameter of firm size distribution	1.06	1.05	1.12	1.03	1.02
Share of employed people remaining in the same firm after one quarter	0.88	0.89	0.88	0.88	0.89
Share of employed people remaining employed after one quarter	0.94	0.95	0.94	0.94	0.94
Moments: training intensity					
Average training intensity (% time)	2.20	2.17	2.21	2.13	2.19
Ratio of training costs to wage costs of training	0.24	0.24	0.24	0.23	0.25
Moments: training across firms					
Ratio of training intensity in firms with 100-499 employees to that with 50–99 employees	1.19	1.28	0.95	0.92	1.20
Percent wage increase of 20 years' experience (%)	89	-	-	-	89

– Constant cost share case matches training gradient best

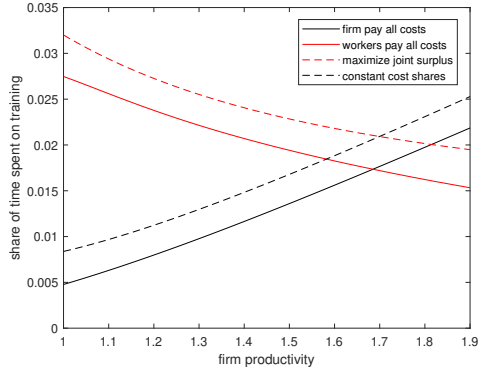
Internal calibration

– For each cost-sharing scenario minimize distance between moments in data and model

Moments	Data	Model			
		Firms Pay	Workers Pay	Maximize Joint Surplus	Constant Cost Shares
Moments: labor market					
Unemployment rate (%)	6.5	6.6	6.5	6.4	6.3
Ratio of #Vacancies to #Unemployed	0.55	0.55	0.54	0.58	0.52
Pareto parameter of firm size distribution	1.06	1.05	1.12	1.03	1.02
Share of employed people remaining in the same firm after one quarter	0.88	0.89	0.88	0.88	0.89
Share of employed people remaining employed after one quarter	0.94	0.95	0.94	0.94	0.94
Moments: training intensity					
Average training intensity (% time)	2.20	2.17	2.21	2.13	2.19
Ratio of training costs to wage costs of training	0.24	0.24	0.24	0.23	0.25
Moments: training across firms					
Ratio of training intensity in firms with 100-499 employees to that with 50–99 employees	1.19	1.28	0.95	0.92	1.20
Percent wage increase of 20 years' experience (%)	89	-	-	-	89

– Constant cost share case matches training gradient best

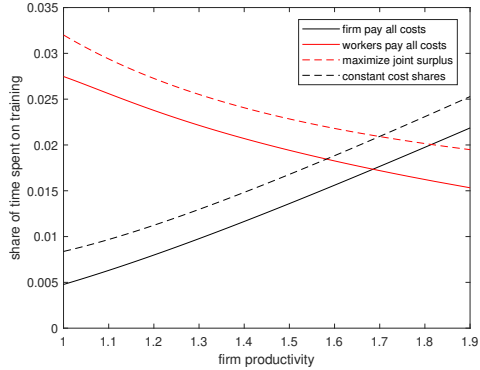
Training gradient in each scenario



- Training is higher in more productive firms when firms determine training due to joint effects of z and $\beta(z)$
- Matches evidence that more productive firms foster faster rates of skill acquisition (Engbom (2021), Arellano-Bover (2020), Arellano-Bover and Saltiel (2023))

Further intuition

Training gradient in each scenario



– Training is higher in more productive firms when firms determine training due to joint effects of z and $\beta(z)$

– Matches evidence that more productive firms foster faster rates of skill acquisition (Engbom (2021), Arellano-Bover (2020), Arellano-Bover and Saltiel (2023))

Further intuition

Constant cost share case also matches training returns evidence best

Parameter	Firm Pay All Costs	Worker Pay All Costs	Maximize Joint Surplus	Constant Cost Shares
c_M - Constant in matching function	0.78	0.87	0.82	0.85
η - On-the-job search intensity	0.22	0.32	0.23	0.27
c_s - Ratio of training costs per time to wage	0.28	0.24	0.22	0.28
c_v - Constant in vacancy function	0.68	0.27	0.31	0.52
ζ - Constant in training function	0.20	0.02	0.02	0.06
γ_s - Convexity of training function	0.44	0.28	0.31	0.21
κ - Parameter of Pareto productivity dist	5.14	9.69	4.92	5.96
δ - Exogenous separation rate	0.07	0.07	0.07	0.07
$\mu_F(z)$ - Share of training costs paid by firm	1	0	surplus share	0.30

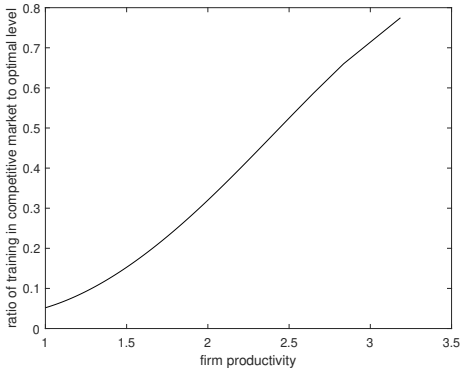
- Evidence indicates $3\% - 5\%$ wage returns from 60 hours training
- Calibrated cost share returns (assuming 480 work hours in quarter): $\zeta \times \left(\frac{60}{480}\right)^{\gamma_s} = 0.06 \times \left(\frac{60}{480}\right)^{0.21} = 3.6\%$
- These returns are too low or too high in other three scenarios

Inefficiencies and Subsidies

Inefficiencies in provision of training

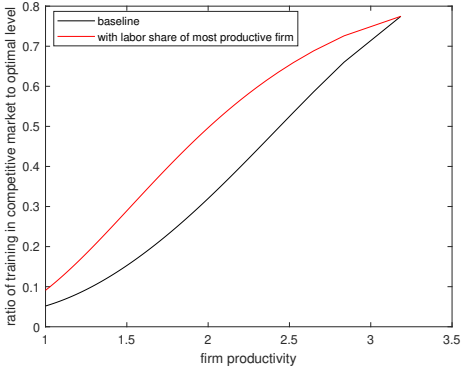
- **Inefficiency**: Firms don't internalize benefits to workers, or other firms after separation
- Assess extent & gradient of inefficiencies by characterizing social planner's choices
- Constrain social planner to take competitive vacancy & wage distributions as given
  - $\Rightarrow$ Focus on training inefficiencies rather than inefficiencies from frictional labor markets

Training in CE relative to SP



- All firms provide less training than SP, but worse in unproductive firms
- Why? Higher labor shares & probability of losing worker aggravate inefficiency

Examine relative importance of hold-up and labor shares to inefficiencies



- **Red line**: Fix labor share to be that of most productive firm for all firms
- Inefficiencies decrease but remain large $\Rightarrow$ hold-up problem is primary driver

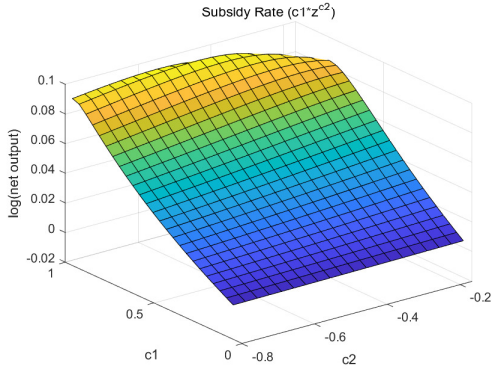
Subsidies and optimal policy

- Consider scope of subsidies financed thru lump-sum taxes to correct inefficiencies
  1. Policy that can set different subsidy rates to firms of every productivity level separately
     - Set subsidy as $s(z) = c_1 z^{c_2}$
  2. Policy that can set different subsidy rates for different firm size brackets
     - Set different subsidies for firms above and below median size
  3. Policy that sets same subsidy rates for all firms

Subsidies by productivity level

- Set subsidy as $s(z) = c_1 z^{c_2}$

Figure: Net Output Gain as Function of Subsidy Parameters

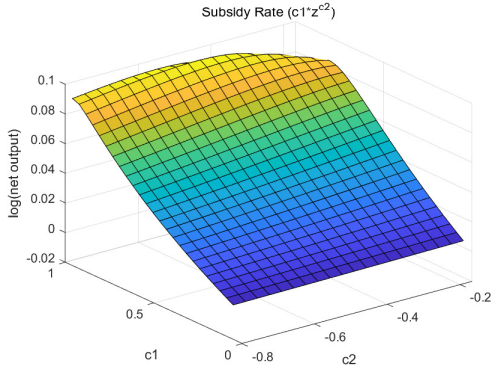


- Optimal policy: $c_1 = 0.92$, $c_2 = -0.5$, net output gain $= 10\%$
  - Subsidize unproductive firms more due to large inefficiencies
  - Also subsidize productive firms due to some inefficiencies + stop reallocation

Subsidies by productivity level

- Set subsidy as $s(z) = c_1 z^{c_2}$

Figure: Net Output Gain as Function of Subsidy Parameters

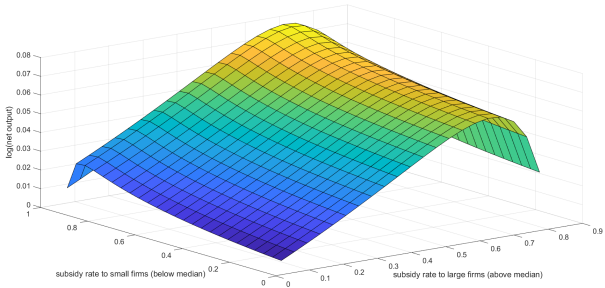


- Optimal policy: $c_1 = 0.92$, $c_2 = -0.5$, net output gain $= 10\%$
  - Subsidize unproductive firms more due to large inefficiencies
  - Also subsidize productive firms due to some inefficiencies + stop reallocation

Subsidies by firm-size bracket

– Set different subsidies for firms above and below median size (note: size maps directly into productivity)

Figure: Optimal subsidy for small and big firms

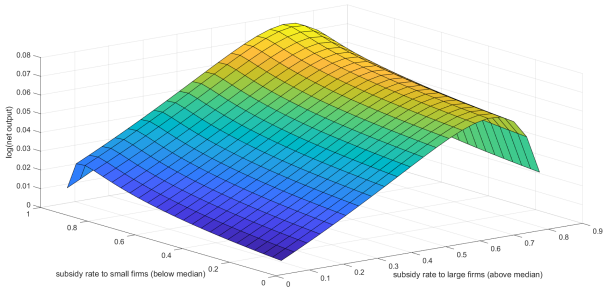


– Optimal policy: 88% subsidy for small firms, 65% subsidy for large firms, net output gain = 8%

Subsidies by firm-size bracket

– Set different subsidies for firms above and below median size (note: size maps directly into productivity)

Figure: Optimal subsidy for small and big firms

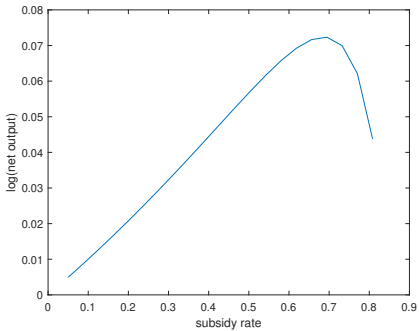


– Optimal policy: 88% subsidy for small firms, 65% subsidy for large firms, net output gain = 8%

Subsidies targeting all firms equally

- Set same subsidy rate for all firms

Figure: Net Output Gain as Function of Subsidy Parameters

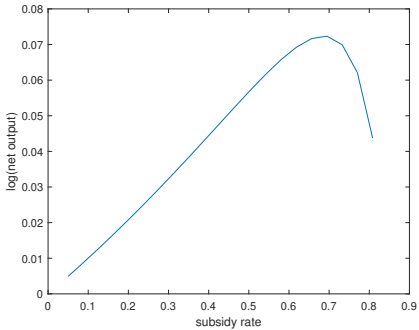


- Optimal policy: 69% subsidy for all firms, net output gain = 7%
  - In US most common training-promoting policy is reimbursement of 50% of training costs
  - Our policy implies larger subsidies desirable

Subsidies targeting all firms equally

- Set same subsidy rate for all firms

Figure: Net Output Gain as Function of Subsidy Parameters



- Optimal policy: 69% subsidy for all firms, net output gain = 7%
  - In US most common training-promoting policy is reimbursement of 50% of training costs >
  - Our policy implies larger subsidies desirable

Conclusions

- Examine how training changes with firm characteristics, and use this to inform allocation of learning costs
- **Empirics**: Training lower in smaller firms, driven by TFP and labor share differences
- **Analytical model**: Consider 4 cost-sharing schemes, find only cases where firms choose training match results
- **Quantitative Model**: Calibrate to US & quantify size of inefficiencies + scope of policy
  - Firms paying a calibrated fixed share of explicit costs fits evidence on training returns best
  - Inefficiencies larger in unproductive firms
  - Optimal training subsidy larger in unproductive firms but still substantial in productive firms
  - 7% increase in net output for US even with same subsidy rate for all firms

Definition and Examples

Table: Human Capital Sources and Examples

How Structured		Firm Sponsored	Non Firm Sponsored
	1.Schooling	MBA paid by firm	MBA self-financed
	2.Formal Training	Firm-organized presentation	Course w/certification
	3.Informal Training	Guided o-t-j Training, Job Rotation	-
	4.Informal Learning	-	Self-learning (Reading Journals)

Back

Formal Training Definition

- Learning activity planned and organized in advance
- Aim to generate "the acquisition of new competences or the development and improvement of existing ones"
- Typically separated from the workplace (e.g. classroom or institution) and show a high degree of organization by a trainer
- Content designed for a group of learners (e.g. curriculum exists)
- Excluded:
  - Employed holding an apprenticeship
  - Initial vocational training (IVT)
  - Employee orientation or training to familiarize workers with equipment or machinery.

Back

Informal Training Definition

- Typically connected to the active work and the active workplace
- Less structured and adapted to specific worker needs
- From the EU-CVT 5 activities are singled out:
  - **Guided on-the-job training.** planned periods of instruction or practical experience in the workplace using the normal tools of work
  - **Job Rotation** within the enterprise and exchanges with other enterprises
  - **Learning circles.** groups of persons employed who come together on a regular basis with the primary aim of learning
  - **Self-directed learning**: planned learning initiative where worker manages the settings of the activity
  - **Participation in conferences**, workshops, trade fairs and lectures

Back

Job Related refers to when participants obtain training to

- Obtain knowledge and/or learn new skills needed for a current or future job,
- Increase earnings,
- Improve job and/or career opportunities in a current or another field, and
- Generally improve opportunities for advancement and promotion

Back

Correlation between Firm Size and Training

Dep. Variable	Formal Training Offered by Firm				
	(1)	(2)	(3)	(4)	(5)
log(Firm Size)	11.6***	11.7***	10.8***	10.9***	11.0***
	(0.24)	(0.24)	(0.23)	(0.24)	(0.25)
Constant	0.91	-5.27***	-5.33***	-2.04	-4.48
Observations	117,103	117,103	117,103	117,103	109,555
R^2	0.083	0.098	0.167	0.172	0.183
Year FE		Y		Y	Y
Country FE			Y	Y	Y
Industry FE					Y

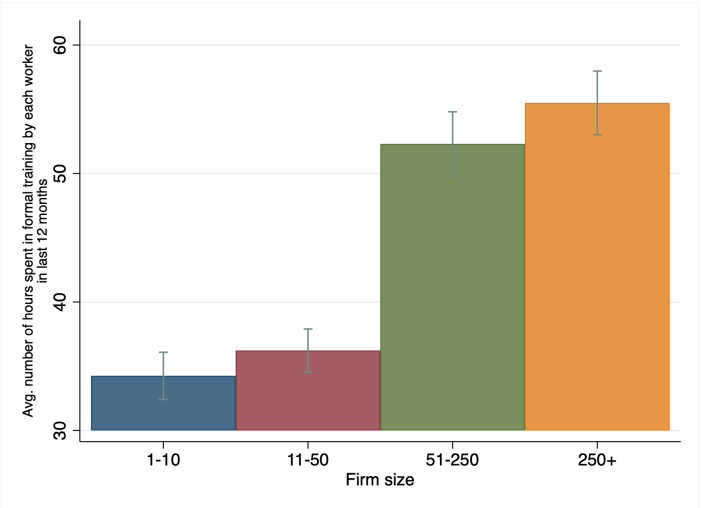
Back

Share of Firms Offering Formal Training by Purpose and Firm Size (EU-CVT)

	Average By firm Size in 2010			
	All	10-49	50-249	250+
General IT	27.3	23.7	34.5	54.7
Professional IT	16.9	14.5	21	37.5
Management	32	26.2	43.7	74.3
Team working	32.5	29	38.3	61.6
Customer handling	38.5	35.4	44.1	62.7
Problem solving	30.1	28.5	31.2	50
Office administration	26.9	24.3	32.3	45.1
Foreign language	15.3	11	24	46.9
Technical or job-specific	69	67.2	73.2	81.2
Oral or written communication	14.7	12.7	16.9	36.5
Numeracy and/or literacy	7	6.7	6.5	14.7
Other skills and competences	11	11.2	10.4	10.3

Back

Hours spent in formal training by firm size



Back

Correlation between training and firm size

Dep. Variable	(1) Formal Training	(2) Informal Training	(3) Hours of Formal Training
Firm of 11–50 workers	0.09***	0.07***	3.70
	(0.01)	(0.01)	(2.60)
Firm of 51–250 workers	0.15***	0.06***	16.79***
	(0.01)	(0.01)	(3.55)
Firm of 250+ workers	0.21***	0.09***	20.95***
	(0.01)	(0.01)	(4.26)
Constant	0.26**	0.72***	10.65
	(0.10)	(0.05)	(13.18)
Age FE	Y	Y	Y
Country FE	Y	Y	Y
Demographic Controls	Y	Y	Y
Worker Type FE	Y	Y	Y
Industry FE	Y	Y	Y
Occupation FE	Y	Y	Y
Wage Controls	Y	Y	Y
Observations	55,502	56,342	56,393
R-squared	0.22	0.12	0.04

Back

TFP measures

- Production function in firm i in industry s: $Y_{i,s} = A_{i,s} K_{i,s}^{\alpha} L_{i,s}^{1-\alpha}$
- Aggregate output in industry s: $Y_s = \left(\sum_{i=1}^{M_s} Y_{i,s}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$
- Measure of TFPQ: $\frac{(P_{i,s} Y_{i,s})^{\frac{\sigma}{\sigma-1}}}{K_{i,s}^{\alpha} (w_{i,s}, L_{i,s})^{1-\alpha}}$
  - $P_{i,s} Y_{i,s}$ is revenue (raised to $\frac{\sigma}{\sigma-1}$ to recover output)
  - $K_{i,s}$ is book value of fixed capital
  - $w_{i,s} L_{i,s}$ is payroll (thus, controls for human capital)
  - σ is set to 3 following Hsieh and Klenow (2009)

1. $1 - \alpha = \frac{2}{3}$ (Gollin (2002))
2. $1 - \alpha =$ average labor share (measured thru ratio between payroll and VA) in industry s

Back

Correlation between Training (intensive margin), TFP and the Labor Share

Dep. Variable:	China Log (Per-Worker Training Expenditures)				Mexico Share Emp. Trained	
	(1)	(2)	(3)	(4)	(1)	(2)
Labor Share	-0.753***	-0.001	-1.050***	-0.041	0.016***	-0.036***
	(0.042)	(0.066)	(0.042)	(0.066)	(0.002)	(0.002)
TFP (labor share = 2/3)	0.111***	0.084***			0.015***	
	(0.002)	(0.004)			(0.0002)	
TFP (HK)			0.081***	0.076***		0.005***
			(0.002)	(0.004)		(0.0002)
log(Firm Size)	0.263***	-0.035***	0.276***	-0.047***	0.053***	0.059***
	(0.003)	(0.011)	(0.003)	(0.011)	(0.0003)	(0.0003)
Age FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y		
Industry FE	Y	Y	Y	Y	Y	Y
Firm FE		Y		Y		
Observations	772,337	658,204	768,216	654,483	1,561,690	1,375,834
R-squared	0.061	0.708	0.060	0.709	0.21	0.22

Correlation between Training and the Labor Share (WB-ES)

Dep. Variable:	Firm Offers Formal Training					
	(1)	(2)	(3)	(4)	(5)	(6)
Labor Share	-0.14***	-0.080***	-0.076***	-0.11***	-0.10***	-0.10***
	(0.019)	(0.018)	(0.018)	(0.020)	(0.020)	(0.020)
Log(Firm Size)		0.11***	0.11***	0.11***	0.11***	0.11***
		(0.0027)	(0.0028)	(0.0029)	(0.0029)	(0.0029)
Constant	0.37***	0.044***	0.18	0.17	0.23	0.23
	(0.0057)	(0.010)	(0.12)	(0.37)	(0.40)	(0.40)
Year FE			Y		Y	Y
Country FE				Y	Y	Y
Industry FE						Y
Observations	92,012	92,012	92,012	92,012	92,012	87,295
R-squared	0.002	0.009	0.031	0.133	0.137	0.157

Back

Correlation between Firm Size and the Labor Share (WB-ES)

Dep. Variable	Labor Share				
	(1)	(2)	(3)	(4)	(5)
Log(Firm Size)	-1.10***	-1.13***	-1.00***	-0.98***	-1.34***
	(0.12)	(0.14)	(0.11)	(0.12)	(0.12)
Constant	25.6***	66.7***	28.2***	66.0***	36.1***
	(0.43)	(11.2)	(1.57)	(11.3)	(7.67)
Year FE		Y		Y	Y
Country FE			Y	Y	Y
Industry FE					Y
Observations	111,375	111,375	111,375	111,375	100,196
R-squared	0.004	0.009	0.049	0.055	0.105

Notes: This table shows different specifications in which we regress the labor share on firm size using data from the WB-ES. The labor share measure is described in app: WB-ES. Industry FE correspond to two-digit industries. Robust standard errors in parentheses. *p ¡0.05, **p ¡0.01, ***p ¡0.001.

Back

Optimal training levels in different scenarios

Scenario	Share of Training Costs Paid: Firm ($\mu_F(z)$)	Share of Training Costs Paid: Worker ($\mu_W(z)$)	Optimal Training Level
1. Firms Pay	1	0	$\left(\frac{\zeta\gamma_s MR_F(z)}{(1+r)(c_s\bar{w}+z)}\right)^{\frac{1}{1-\gamma_s}}$
2. Workers Pay	0	1	$\left(\frac{\zeta\gamma_s MR_W(z)}{(1+r)(c_s\bar{w}+z)}\right)^{\frac{1}{1-\gamma_s}}$
3. Maximize Joint Surplus	$\frac{MR_F(z)}{MR_W(z)+MR_F(z)}$	$\frac{MR_W(z)}{MR_W(z)+MR_F(z)}$	$\left(\frac{\zeta\gamma_s(MR_W(z)+MR_F(z))}{(1+r)(c_s\bar{w}+z)}\right)^{\frac{1}{1-\gamma_s}}$
4. Constant Cost Shares	μ_F	μ_W	$\min\left\{\left(\frac{\zeta\gamma_s MR_F(z)}{(1+r)\mu_F(c_s\bar{w}+z)}\right)^{\frac{1}{1-\gamma_s}}, \left(\frac{\zeta\gamma_s MR_W(z)}{(1+r)\mu_W(c_s\bar{w}+z)}\right)^{\frac{1}{1-\gamma_s}}\right\}$

Back

Firms' wage and vacancy problem

Given young workers' training $s(z)$, the firm chooses wage $w(z)$ and vacancies $v(z)$ thru:

$$\max_{\{w(z),v(z)\}} \underbrace{\frac{v(z)}{\theta}\frac{1}{1+\eta(1-\delta)+\delta}\left[z-w(z)-\mu_F(z)(c_s\bar{w}+z)s(z)+\frac{1}{1+r}MR_F(z)(1+\zeta s(z)^{\gamma_s})\right]}_{\text{profits from hiring young workers}}$$

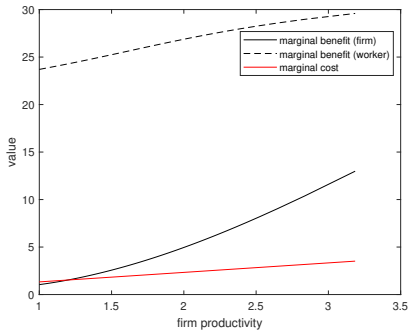
$$+\underbrace{\frac{v(z)}{\theta}\frac{\eta(1-\delta)+\delta}{1+\eta(1-\delta)+\delta}\frac{\eta(1-\delta)F(w(z))\bar{l}(w(z))+\delta\bar{l}}{\eta(1-\delta)+\delta}(z-w)}_{\text{profits from hiring old workers}}-\underbrace{\frac{c_v v(z)^{1+\gamma_v}}{1+\gamma_v}}_{\text{vacancy costs}}$$

$$\text{s.t.} w(z)\geq b\bar{w}$$

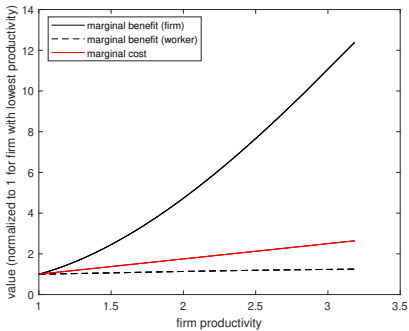
Back

Marginal benefits and costs from training

(a) Absolute Values



(b) Normalized Values



- Marginal returns to training are lower for firms than workers $\forall z$ due to hold-up problem
- Marginal returns increase with z for both parties due to supermodularity in production, but faster for firms
  - As firms become more productive worker less likely to leave as wage rises with z
  - But wage rises slower than marginal profit $(A_M z - w(z))$ with z
- Marginal costs increase with z due to loss of production time from training

Back

Training Subsidies in the United States

Country	Year	Subsidy or Incentive to employer
Alabama	2014 - present	75% of training costs reimbursed
Arizona	2015 - 2020	50-75% of training costs reimbursed
Colorado	2018 - present	60% of training costs reimbursed
Florida	1993 - present	50-75% of training costs reimbursed
Georgia	1994 - present	50% of training costs tax deductible
Hawaii	1991 - present	50% tuition costs reimbursed
Illinois	1992 - present	50% of training costs reimbursed
Kentucky	1984 - present	50% of training costs reimbursed
Maryland	1989 - present	50% of training costs reimbursed
Massachusetts	2008 - present	50% of training costs reimbursed
Mississippi	2013 - present	50% of training costs reimbursed
Montana	2005 - present	Funding of $5,000 for training
Nebraska	2005 - present	Funding of $800-4,000 for training
New Hampshire	2007 - present	50% of training costs reimbursed
New Jersey	1992 - present	50% of training costs reimbursed
New Mexico	1972 - present	50-75% of training costs reimbursed
Pennsylvania	1999 - present	Funding of $600-1,200 per trainee
Rhode Island	2006 - present	50% of training costs reimbursed
Washington	1983 - present	50% of training costs reimbursed
Wisconsin	2012 - present	50% of training costs reimbursed
Wyoming	1997 - present	Funding of $1,000 per trainee

Back