Labour and Innovation Economics

Fabian Waldinger (LMU Munich)

Lecture Content

1. The signaling model
2. *Methodology: Differences-in-Differences*
3. Evidence on the signaling model:
    1. Tyler, Murnane, and Willett (2000)
    2. Student presentation: Clark and Martorell (2014)
4. General training (Becker, 1964)
5. Why do firms pay for general training? (Acemoglu and Pischke, 1999)
6. Firm-specific training
7. Evidence for firm-specific training:
    1. Jacobson, LaLonde, and Sullivan (1993)

Education as a Signal of Productivity

- While the human capital model suggests that schooling increases productivity and therefore earnings, the signaling model postulates that even if education does not increase productivity one may observe a positive correlation of schooling and earnings if employers see schooling as a signal for productivity
- *Human Capital Model:*

$$\text{Schooling} \Rightarrow \text{Productivity} \Rightarrow \text{Earnings}$$

- *Signaling Model:*

$$\text{Productivity} \Rightarrow \text{Earnings} \Uparrow \text{Schooling}$$

The Signaling Model

- Key aspects of the model:
  1. Employers do not observe true productivity
     $\Rightarrow$ they have to rely on education as a signal of true productivity
  2. Workers know their true productivity
  3. The cost of obtaining education must be lower for high productivity individuals to ensure a separating equilibrium

- There are two possible equilibria:
  1. Pooling equilibrium:
     every worker gets the same schooling and earns the same wage.
  2. Separating equilibrium:
     higher productivity workers get more education and earn higher wages than low ability workers.

- Workers have the following utility function:

$$U(w, e, \eta) = w - c(e, \eta)$$

  - w = wage
  - e = education
  - η = ability which affects productivity
  - c = schooling costs

- with: $c_e > 0$; $c_{ee} > 0$
- and: $c_{e\eta} < 0$ (This is the crucial assumption to ensure a separating equilibrium: the cost of schooling increases less for the more able)

- Firms's productivity depends on e and η :

$$y(e, \eta)$$

- with: $y_\eta > 0$
- and $y_e \geq 0$; $y_{ee} \leq 0$
  if human capital would not affect productivity at all: $y_e = 0$

- Workers hypothesize a wage schedule that depends on e (the wage cannot depend on η because the firm does not observe it)
- The worker then chooses the optimal level of education:

$$\max w(e) - c(e, \eta)$$

- FOC:

$$w'(e) - c_e(e, \eta) = 0$$

- SOC:

$$w''(e) - c_{ee}(e, \eta) < 0$$

- Assuming that the first order condition of the workers results in a distinct e for each η it needs to be the case that wages w(e) equal productivity at every e. (if competition among firms lead to zero profits).

$$w(e^*) = y(e, \eta(e))$$

- where $w(e^*)$ is the solution to the worker's maximization problem.

The Signaling Model - Overeducation in Equilibrium

- The signaling models predicts that workers get more schooling than optimal (we show this in the following bit)
- Differentiating $w(e^*) = y(e, \eta(e))$ w.r.t. e:

$$w'(e) = y_e + y_\eta \frac{d\eta}{de}$$

- Using this and the FOC of the workers $(w'(e) - c_e(e, \eta) = 0)$ we get:

$$y_e + y_\eta \frac{d\eta}{de} - c_e(e, \eta) = 0$$

$$y_e - c_e(e, \eta) = -y_\eta \frac{d\eta}{de}$$

- To figure out whether $y_e - c_e(e, \eta)$ is positive or negative we have to figure out the sign of $\frac{d\eta}{de}$

The Signaling Model - Overeducation in Equilibrium

- Take the worker's FOC ($w'(e) - c_e(e,\eta) = 0$) and differentiate w.r.t e

$$w''(e) - c_{ee} - c_{e\eta}\frac{d\eta}{de} = 0$$

$$\Rightarrow \frac{d\eta}{de} = \frac{w''(e) - c_{ee}{<}0 \text{ (because of SOC)}}{c_{e\eta} < 0 \text{ (by assumption)}} > 0$$

- There is thus over-investment in education in equilibrium:

$$y_e - c_e(e,\eta) = -y_\eta \frac{d\eta}{de} < 0 \;\; \text{(because } \frac{d\eta}{de} > 0 \text{ and } y_\eta > 0\text{)}$$

- With perfect information the level of investment in education would be optimal: $y_e - c_e(e,\eta) = 0$

The Signaling Model - A Concrete Example

- Assume the cost function looks as follows:

$$c(e, \eta) = k\frac{e}{\eta}$$

with $k > 0$

- Also assume that the lowest ability types $\underline{\eta}$ get no education $e(\underline{\eta}) = 0$; and that $\underline{\eta} = 1$.
- Properties of this cost function (see above):
  - $c_e = \frac{k}{\eta} > 0$
  - $c_{e\eta} = -\frac{k}{\eta^2} < 0$ (crucial assumption for separating equilibrium)
- Let's start with the worker's problem. Her FOC is:

$$w'(e) - \frac{k}{\eta} = 0$$

The Signaling Model - A Concrete Example

- Let's now look at the firm assuming that the solution to the worker's problem results in a distinct e for each η it needs to be the case that $w(e) = \eta$ at every e

$$w(e^*) = \eta$$

- Taking the first derivative w.r.t. η we get:

$$\frac{\partial w(e^*)}{\partial e}\frac{\partial e^*}{\partial \eta} = 1$$

- Substituting in the workers' FOC ($w'(e) = \frac{k}{\eta}$):

$$\frac{\partial e^*}{\partial \eta} = \frac{\eta}{k}$$

The Signaling Model - A Concrete Example

- We now have the derivative of the optimal e but not e^* itself. To find $e^*(\eta)$ we need to integrate the derivative from $\underline{\eta}$ to η

$$e^*(\eta) - e^*(\underline{\eta}) = \int_{\underline{\eta}}^{\eta} \frac{\eta}{k} d\eta = \frac{1}{2k}\eta^2 - \frac{1}{2k}\underline{\eta}^2$$

- using the fact that $\underline{\eta} = 1$ and $e(\underline{\eta}) = 0$ by assumption, we obtain an expression for the equilibrium level of education $e^*(\eta)$:

$$e^*(\eta) = \frac{\eta^2}{2k} - \frac{1}{2k}$$

- To obtain the equilibrium level of wages we first solve this for η:

$$\eta = \sqrt{e^*(\eta)2k + 1}$$

- In equilibrium we have $w(e^*) = \eta$ (see above) and therefore $w(e^*) = \sqrt{e^*(\eta)2k + 1}$

The Signaling Model - A Concrete Example

- In equilibrium we therefore have:

$$e^*(\eta) = \frac{\eta^2}{2k} - \frac{1}{2k}$$

and

$$w(e^*) = \sqrt{e^*(\eta)2k + 1}$$

- Suppose $k = 1/2$; and we have two individuals one has $\eta = 1$ and the other has $\eta = 2$
- What is the equilibrium level of education for these two individuals?

$$e^*(1) = 0; e^*(2) = 3$$

- What are their equilibrium wages?

$$w^*(1) = 1; w^*(2) = 2$$

- Do we have a separating equilibrium? Does anybody want to change her education?
- The workers' utility is:

$$w(e) - k\frac{e}{\eta} = w(e) - \frac{e}{2\eta} \text{ (if } k = 1/2)$$

- Utility if *low* ability worker changes her education:

$$w(3) - c(3,1) = 2 - 3/2 = 0.5$$

- Utility if *low* ability worker if she keeps her education level:

$$w(0) - c(0,1) = 1 - 0/2 = 1$$

- $\Rightarrow$ she does not want to change

The Signaling Model - A Concrete Example

- Do we have a separating equilibrium? Does anybody want to change her education?
- The workers' utility is:

$$w(e) - k\frac{e}{\eta} = w(e) - \frac{e}{2\eta}(if k = 1/2)$$

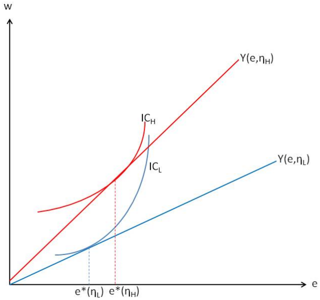
- Utility if *high* ability worker changes her education:

$$w(0) - c(0, 2) = 1 - 0/2 = 1$$

- Utility if *high* ability worker if she keeps her education level:

$$w(3) - c(3, 2) = 2 - 3/4 = 1.25$$

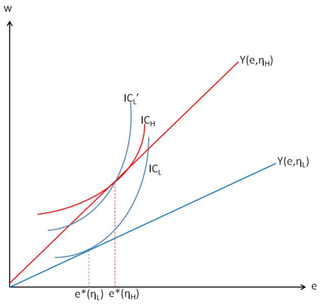
- $\Rightarrow$ she does not want to change



- Here we assume only two ability levels high (H) and low (L)

Graphical Analysis

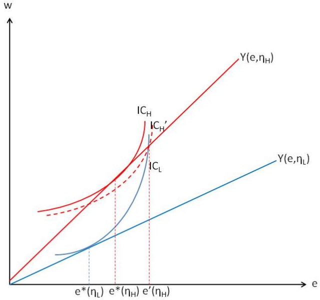
What would Low Ability Individuals Do in this Case?



- If high ability people did not react, low ability people could increase their schooling and reach the higher indifference curve IC_L'

Graphical Analysis

Separating Equilibrium: High Ability People get More than the Optimal Level of Schooling



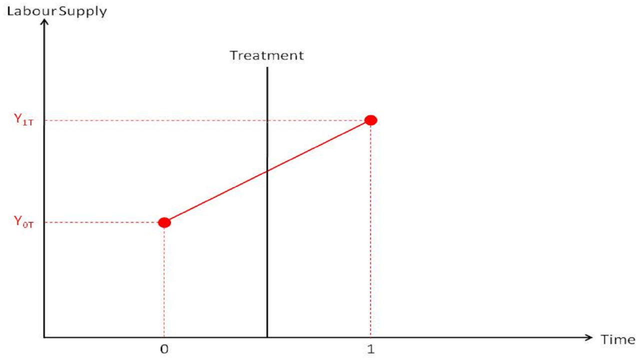
- High ability individuals obtain schooling until it does no longer pay of for low ability people to increase schooling
- Note: high ability ICs are flatter than low ability ICs where they cross, because education is less costly for the high ability people

Empirical Predictions of the Human Capital and Signaling Models

1. People with more schooling are more productive
   Both HC and S models predict this
2. People with more schooling receive higher wages
   Both HC and S models predict this
3. People will attend school while they are young (i.e. before they enter the workforce)
   Both HC and S models predict this
4. The rate of return to schooling should roughly be equal to the rate of interest
   Only HC predicts this

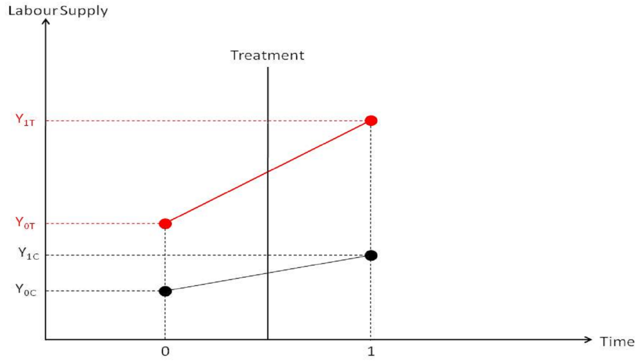
Methodology: Differences-in-Differences

- We often want to evaluate the effect of a certain program using pre and post-treatment data
- Common problem: other factors (which affect treatment outcomes) also change from the pre to the post period (e.g. changes in the business cycle)



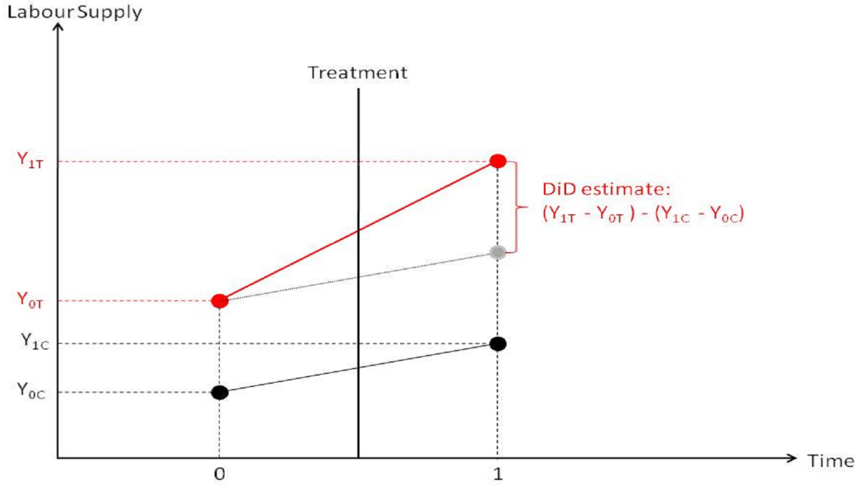
Methodology: Differences-in-Differences

- Solution: find a control group that is unaffected by the treatment but otherwise behaves exactly the same
- In that case we control for other changes between the pre- and the post period using the changes in the control group



Methodology: Differences-in-Differences

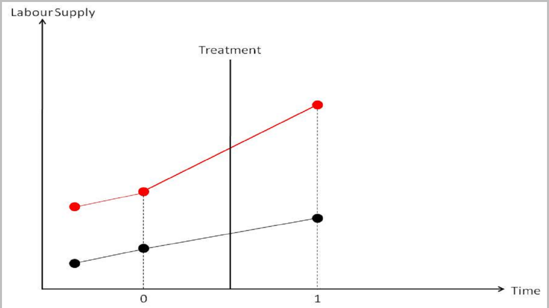
The Differences-in-Differences Estimator



Methodology: Differences-in-Differences

Differences-in-Differences Estimator: Crucial Assumptions

- The key assumption is that treatment and control group would have the same time trend in the absence of the treatment
- This does not mean that they have to have the same mean of the outcome!
- Difficult to verify but one usually uses pre-treatment data to show that the trends are the same



Methodology: Differences-in-Differences

Differences-in-Differences Regression

- We can estimate the differences-in-differences estimator in a regression framework
- Advantages:
  - It is easy to calculate standard errors
  - We can control for other variables which may reduce the residual variance (reduces standard errors)
  - It is easy to include multiple periods
  - We can study treatments with different treatment intensity. (e.g. varying increases in the marginal tax rate for different people)
- Simplest DiD regression model:

$$Outcome_{it} = \beta_1 + \beta_2 Treatment_i + \beta_3 Post_t + \beta_4 (Treatment \times Post)_{it} + \varepsilon_{it}$$

  - *Treatment*: dummy variable = 1 if individual in treatment group.
  - *Post*: dummy variable = 1 after treatment.

Methodology: Differences-in-Differences

Differences-in-Differences Regression

$$Outcome_{it} = \beta_1 + \beta_2 Treatment_i + \beta_3 Post_t + \beta_4 (Treatment \times Post)_{it} + \varepsilon_{it}$$

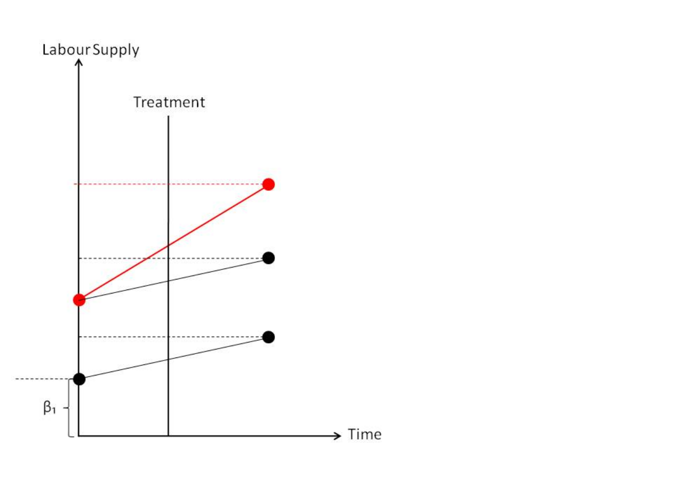
- β_4 is the differences-in-differences estimate.
- In control group:
  - Pre-treatment: $Outcome_{it} = \beta_1$
  - Post-treatment: $Outcome_{it} = \beta_1 + \beta_3$
- In treatment group:
  - Pre-treatment: $Outcome_{it} = \beta_1 + \beta_2$
  - Post-treatment: $Outcome_{it} = \beta_1 + \beta_2 + \beta_3 + \beta_4$
- Differences-in-Differences:

$$\begin{aligned} & [y_{1T} - y_{0T}] - [y_{1C} - y_{0C}] \\ =& [(\beta_1 + \beta_2 + \beta_3 + \beta_4) - (\beta_1 + \beta_2)] - [(\beta_1 + \beta_3) - \beta_1] \\ =& \beta_4 \end{aligned}$$

Methodology: Differences-in-Differences

Differences-in-Differences Regression

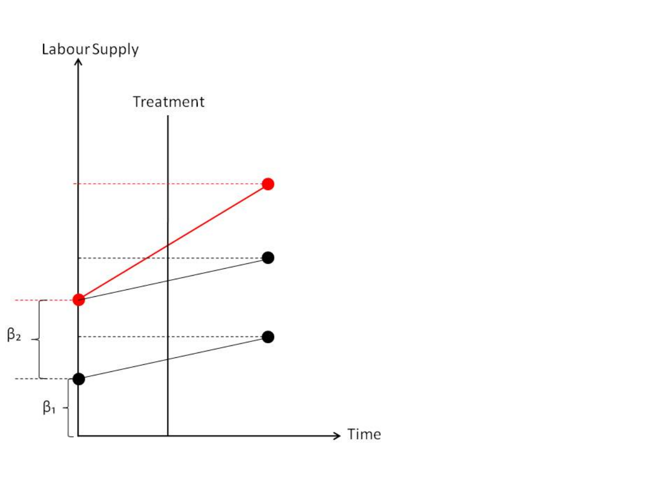
$$Outcome_{it} = \beta_1 + \beta_2 Treatment_i + \beta_3 Post_t + \beta_4 (Treatment \times Post)_{it} + \varepsilon_{it}$$



Methodology: Differences-in-Differences

Differences-in-Differences Regression

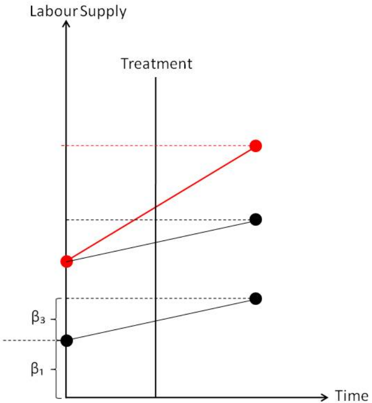
$$Outcome_{it} = \beta_1 + \beta_2 Treatment_i + \beta_3 Post_t + \beta_4 (Treatment \times Post)_{it} + \varepsilon_{it}$$



Methodology: Differences-in-Differences

Differences-in-Differences Regression

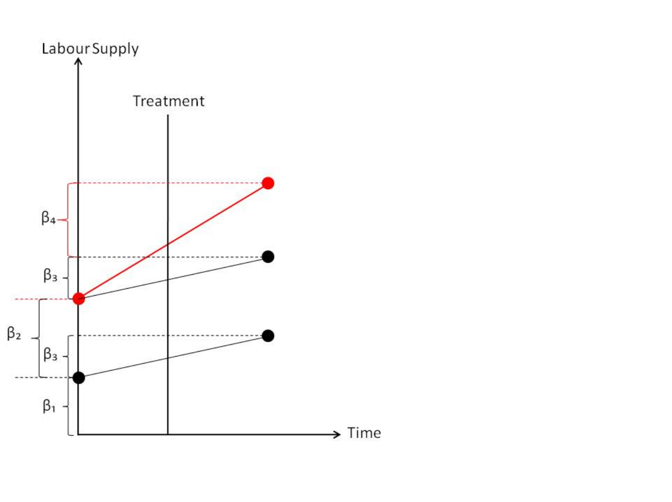
$$Outcome_{it} = \beta_1 + \beta_2 Treatment_i + \beta_3 Post_t + \beta_4 (Treatment \times Post)_{it} + \varepsilon_{it}$$



Methodology: Differences-in-Differences

Differences-in-Differences Regression

$$Outcome_{it} = \beta_1 + \beta_2 Treatment_i + \beta_3 Post_t + \beta_4 (Treatment \times Post)_{it} + \varepsilon_{it}$$



The Signaling Value of the GED

- Tyler, Murnane, and Willett (2000) estimate the signaling value of the GED
- The GED (General Education Development) is a diploma that high-school dropouts can obtain (about 1/3 of dropouts obtain it)
- Tyler, Murnane, and Willett use interstate variation in GED passing standards for the GED to identify the signaling value of the GED
- They compare individuals who receive the same GED score (have the same human capital) but some of them get the certificate while others do not get it

Is the GED a Signal Similar to Education in the Theory

- The crucial assumption for education to work as a signal according to the theory is that high productivity people find it less costly to obtain the signal
- While the average study time for the GED is about 20 hours, some individuals spend much more time preparing the exam. So it seems that higher ability people have lower costs preparing for the exam

OLS Returns to the GED

- Obtaining a GED may not only be a signal to employers but could also measure human capital differences
- In a simple OLS where you regress earnings on whether you receive a GED, or a high school degree (compared to dropping out of high-school without degree) you estimate a combination of signaling value and differences in human capital reflected by the GED

	Whites[a]	Minorities[b]
Intercept	9.518**	9.320**
	(0.057)	(0.079)
Female	−0.397**	−0.306**
	(0.019)	(0.030)
GED	0.162*	0.164
	(0.072)	(0.109)
High school graduate	0.536**	0.581**
	(0.057)	(0.073)

Different States Have Different Passing Standards for the GED

- Their identification strategy exploits the fact that different states have different passing standards for the GED. They can thus compare people with the same score but some of them get the GED while others do not get it
- Different states specify different passing cutoffs according to the scores in each of the five subtest of the GED (Mathematics, English,...) and according to the overall mean score.
- They exploit differences in passing standards that existed in 1990. They consider 3 groups of states with the following passing standards:
  - a minimum score of at least 40 *or* a mean score of at least 45.
  - a minimum score of at least 35 *and* a mean score of at least 45.
  - a minimum score of at least 40 *and* a mean score of at least 45.

Overview of GED Score Groups

TABLE III

GED SCORE GROUPS FORMED BY COMBINING MINIMUM AND MEAN SCORES (OUTLINED CELLS = VARIATION IN GED-STATUS BY STATE, DARK SHADING = ALL POSSESS GED, NO SHADING = NONE HAVE GED.)

	Mean score	
Minimum score	<45	> = 45
20–34	Score group 1	
35–39	Score group 2	Score group 4
40–44	Score group 3	Score group 5
45–46		Score group 6
47–48		Score group 7
49–50		Score group 8
51–52		Score group 9
53+		Score group 10

Identification Strategy

- To estimate the signaling value of the GED they use differences-in-differences:
- Signaling value of GED $=(Y_{T-Low} - Y_{C-Low}) - (Y_{T-High} - Y_{C-High})$
    - Y_{T-Low} = earnings in treatment state (low passing standards) who have a low score and get GED
    - Y_{C-Low} = earnings in control state who have a low score and do not get the GED
    - Y_{T-High} = earnings in treatment state who have a high score and get GED
    - Y_{C-High} = earnings in control state who have a high score and get GED

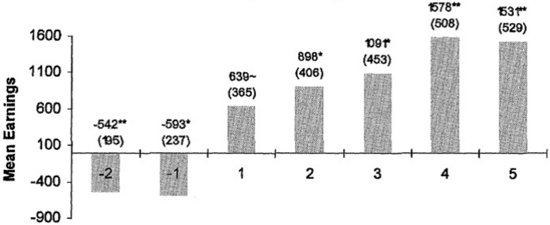
Differences-in-Differences Results

	Experiment 4			Experiment 3			Experiment 3*		
	State passing standard is		Low-High standard contrast	State passing standard is		Low-High standard contrast	State passing standard is		Low-High standard contrast
	Low	High		Low	High		Low	High	
Panel A: Whites									
Test score is									
Low	9628	7849	1779	9362	7843	1509	9362	8616	746
	(361)	(565)	(670)	(400)	(312)	(507)	(400)	(219)	(456)
High	9981	9676	305	9143	9165	−23	9143	9304	−162
	(80)	(65)	(103)	(135)	(63)	(149)	(135)	(135)	(150)
Difference-in-differences for whites			1473*			1531**			907~
			(678)			(529)			(481)
Panel B: Minorities									
Test score is									
Low	6436	8687	−2252	7005	7367	−363	7005	6858	147
	(549)	(690)	(882)	(347)	(347)	(495)	(347)	(290)	(452)
High	7560	8454	−894	7782	8375	−593	7782	7568	214
	(184)	(96)	(207)	(214)	(93)	(233)	(214)	(133)	(252)
Difference-in-differences for minorities			−1357			231			−67
			(906)			(548)			(518)

How does the Effect Vary over Time

Plot of the differences-in-differences estimates

Figure II: Experiment 3



- The effect of the GED takes some years to manifest itself
- Reassuringly, there are no positive effects BEFORE taking the test (common trends assumption seems satisfied)

Potential Problems in Interpreting the Effects as Evidence in Favor of Signaling

1. After obtaining the GED individuals get more education that increases their human capital
   They suggest that these effects are not important in their case
2. States with lower GED standard may have higher minimum wages
   This is not the case
3. Different passing standards may be correlated with ability because they may affect:
   1. The decision to attempt the test.
      (they argue this should lead to a downward bias)
   2. The decision to migrate to another state
      (not an issue - they show migration patterns)
   3. The decision about how much effort to exert on the test.
      (could be a problem if in low standards states some actually better individuals put very little effort and just pass as opposed to obtaining a high pass)

Potential Problems in Interpreting the Effects as Evidence in Favor of Signaling

- They adjust their estimates by assuming that people who score low in the low standard states were actually of higher quality (and thus enter Y_{T-High} instead of Y_{T-Low}).

	Adjusted numbers of observations					Adjusted estimates under different skill composition assumptions	
	p_c [a]	N_{treat} [b]	N_T [c]	$p_c(N_{\text{treat}})$ [d]	$[N_T - p_c(N_{\text{treat}})]$ [e]	Treatment group contamination assumption of . . .	Adjusted estimate
Experiment 4	0.011	18332	653	202	451		
						no contamination[f]	1473* (678)
						the contamination is 50% true 5s and 50% true 6s.	1166 (1416)
						all contamination is from true 6s.	612 (1526)

- Becker (1964) was the first to think carefully about on-the-job training and distinguished two types of skills that may be valuable to employers:
    1. General skills - skills useful to many employers
    2. Specific skills - skills useful to only one employer
- This distinction is important because it affects who will be willing to pay for training

Who Will Pay for General Training?

- Formalizing the Becker idea. Consider two periods:
  1. $t = 1$: *Training and production:*
     - Workers produce output y
     - Workers receive training τ at cost $c(\tau)$
  2. $t = 1.5$: *Workers may leave the firm.*
  3. $t = 2$: *Production:*
     - Workers produce output $y + f(\tau)$ at any firm (because they received *general* training)
- We assume the following about the $f(\tau)$ and $c(\tau)$ functions:
  - $f'(\tau) > 0$
  - $c(0) = c'(0) = 0$; $c''(\tau) > 0$; $\underset{\tau\to\infty}{c'(\tau)} \to \infty$ (convex training cost function)

Becker Insight: Holdup

Firms Will Not Pay for General Training in a Competitive Labor Market

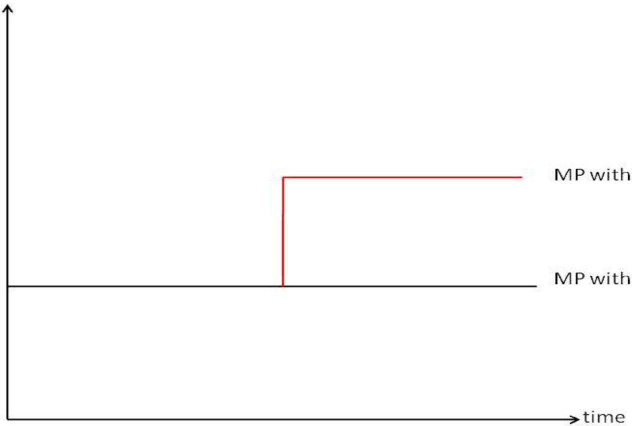
- The social optimum for training is given when: $c'(\tau^*) = f'(\tau^*)$.
- Suppose the firm pays for training. Can it offer the following wage schedule?
  1. $w_1 = y$
  2. $w_2 = y + f(\tau) - c(\tau)$
- In a competitive labor market other firms would offer the worker $w_2^{Other} = y + f(\tau)$. Therefore the incumbent firm would have to pay the same or lose the worker
- The incumbent firm would therefore not be willing to pay for general training in a competitive labor market

Can We Achieve the Socially Optimal Level of General Training?

- Since employers are not willing to pay for general training will nobody get general training?
- Two solutions of how to achieve the socially optimal level of training:
  1. Workers pay for general training directly (School, university, etc.)
  2. Workers pay indirectly to employer by accepting a lower training wage. The wage schedule will be:
     1. $w_1 = y - c(\tau^*)$
     2. $w_2 = y + f(\tau^*)$
- Does higher turnover affect training incentives? No, because the wage in the second period is $y + f(\tau^*)$ anyway
- How do credit constraints for workers impact the model?
  - workers cannot borrow against the future stock of human capital (except from government: e.g. student loans)
  - wages in the first period may not go low enough (e.g. negative) to cover efficient training expenditures

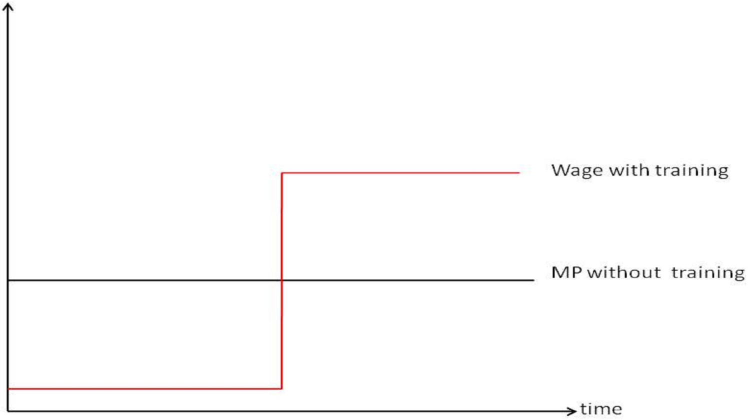
Graphical Analysis

Marginal Products



Graphical Analysis

Firms Paying a Training Wage in the First Period

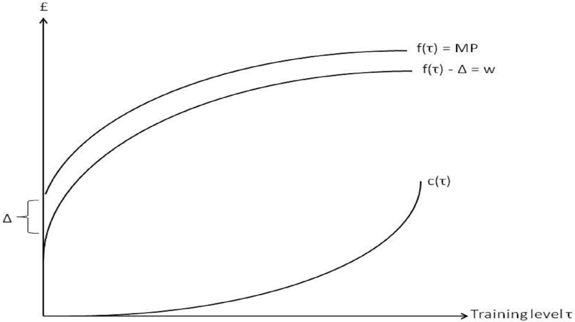


Relevance of the Becker Insight

Is it True that Only Workers Pay for General Human Capital?

- We observe that a lot of general training is paid for by workers. They pay for attending school, university and so on. This is in line with Becker's insight
- In practice, however, we also see that firms pay considerable amounts for general training of their workers:
  - Autor (2001) for example shows that temporary help firms offer a lot of general training (e.g. IT skills) to their workers. By definition this must be general training as the workers are used by other firms
  - In Germany, firms train a large number of apprentices (about 1/3 of a cohort) during 3 year apprenticeships. A large fraction of that training is in general skills
- How can we explain that firms pay for general training?

- Firms would only pay for general training if wages were lower than the MP
- Suppose wages are always lower than MP by a constant wedge for every level of training:



When Would Firms Pay for General Training?

Constant Wedge

- Firms always get Δ as a rent. They would not invest in general training because Δ is constant and the costs of training are increasing
- Showing this mathematically:
- With a constant wedge we have $f'(\tau) = v'(\tau)$ and $f''(\tau) = v''(\tau)$ where $v(\tau)$ is the outside wage paid to a person with training τ.
- The firm maximizes profits:

$$\max \pi(\tau) = f(\tau) - v(\tau) - c(\tau)$$

- FOC:

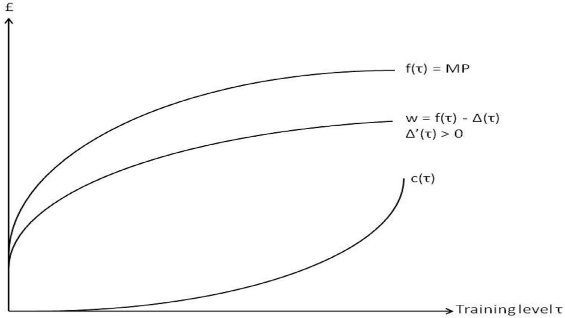
$$f'(\tau) - v'(\tau) = c'(\tau)$$

- which is satisfied at $\tau = 0$.
  (Like in the Becker model firms would not pay for general training).

When Would Firms Pay for General Training?

Increasing Wedge

- Firms would only invest in training if Δ increases with training:



- Showing mathematically that with an increasing wedge firms would pay for general training
- The FOC is the same as above: $f'(\tau) - v'(\tau) = c'(\tau)$
- But now there will be an interior solution at $\tau^* > 0$. Because $\Delta'(\tau) > 0$ and therefore $f'(\tau) - v'(\tau) > 0$ and thus $c'(\tau) > 0$ in the optimum. This occurs at positive training levels
- Is this solution socially efficient?
- No, it will be too low. The socially optimal level of training would be where $f'(\tau) = c'(\tau)$ but the firm only maximizes $f'(\tau) - v'(\tau) = c'(\tau)$
- Thus the firm would only choose the social optimum if $v'(\tau) = 0$; i.e. the wage is invariant to the training level

Why Could the Outside Wage be Lower Than Productivity?

- Firms would only pay for general training if the wedge increases with the level of training. Why could that be the case?
  1. Adverse selection
  2. Outside firms don't know τ. (In that case $v'(\tau) = 0$ and the firm can pay everybody the untrained wage; see above)
  3. Search:
     If there is a possibility that workers are unemployed for a period the discounted present value of the outside wage is lower than productivity. Therefore firms can pay below thee marginal product
  4. Complementaries between specific and general training

Adverse Selection Can Make Employers Pay For General Training

- The Acemoglu and Pischke (1999) model shows how adverse selection can lead firms to pay for general training
- The basic insight is that incumbent firms have an informational advantage in knowing the ability of their workers compared to outside firms. They will therefore get rid of bad workers which reduces the quality of workers for outside firms. Outside firms will thus pay lower wages
- This insight is analogous to the Akerlof (1970) "Market for Lemons" paper
- Here we are going to look at a slightly simplified version of their model: instead of using a distribution of abilities workers are either high or low productivity. All the insights of the model stay the same

Acemoglu & Pischke Model - Timing

- t = 1:
  - Firm hires worker and decides whether to train the worker
  - The firm observes the worker's ability η (this does not depend on whether it trains the worker or not)
  - In the first period there is no production (a normalization)
- t = 1.5:
  - Market offers outside wage $v(\tau)$. They do not observe ability
  - Firm makes counter offer $w(\tau, \eta)$
  - Worker can decide to leave or might have to move for exogenous reasons (this will occur with probability q)
- t = 2:
  - Production takes place

Acemoglu & Pischke Model - Model Details:

- Production:
- The production function is $f(\tau, \eta) = \tau\eta$
  where τ is the level of training and η is the worker's ability
- The worker's ability is: $\eta = \begin{cases} 0 & \text{with probability } p \\ 1 & \text{with probability } (1-p) \end{cases}$
- Probability that a worker quits $= \begin{cases} 1 & \text{if } w(\tau, \eta) < v(\tau) \\ q & \text{if } w(\tau, \eta) \geq v(\tau) \end{cases}$
  where q is the exogenous turnover probability.

- To obtain equilibrium wages and training we solve the model by backward induction:
- In period 2:
- Market offers $v(\tau)$
- Firms best response:
  - offer $w(\tau, \eta) = v(\tau)$ if $v(\tau) \leq \tau\eta$ *(good worker)* (they don't have to pay more because that would not make him more likely to stay)
  - offer $w(\tau, \eta) = 0$ if $v(\tau) > \tau\eta$ *(bad worker)*
- The firm lets low η workers go

- What is the overall probability that a worker moves?

$$\text{prob. of move} = \underbrace{P(\tau\eta < v)}_{\text{"lemons"}} + \underbrace{qP(\tau\eta \geq v)}_{\text{exogenous movers}}$$

- The probability that $(\tau\eta < v) = p$ (see above)

$$\Rightarrow \text{prob. of move} = p + q(1 - p)$$

- Expected productivity of movers:

$$E(\tau\eta|move) = \underbrace{\frac{p}{p+q(1-p)}}_{lemons} 0 \times \tau + \underbrace{\frac{q(1-p)}{p+q(1-p)}}_{exogenous\ movers} 1 \times \tau$$

$$= \frac{q(1-p)}{p+q(1-p)}\tau$$

- Outside firms pay workers according to their expected productivity:

$$v(\tau) = \frac{q(1-p)}{p+q(1-p)}\tau$$

Acemoglu & Pischke Model - Equilibrium

First Period Subgame - Optimal Amount of Training

- Training firm will only have good workers in period 2. ($\eta = 1$). Its profits are therefore:

$$\begin{aligned} \pi &= (1-q)(1-p)[f(\tau,\eta) - w(\tau,\eta)] - c(\tau) \\ &= (1-q)(1-p)[\tau - v(\tau)] - c(\tau) \end{aligned}$$

  The training firm pays the training cost for everybody because it does not yet know who the good workers are

- The firm chooses the optimal level of training to maximize profits

- FOC:

$$\pi' = (1-q)(1-p)[1-v'(\tau)] - c'(\tau) = 0$$

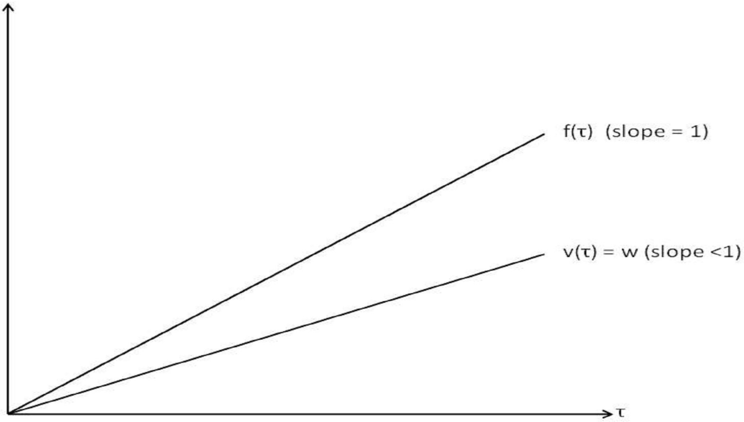
$$(1-q)(1-p)[1-v'(\tau)] = c'(\tau)$$

- This has a solution with $\tau^* > 0$ because $v'(\tau) = \frac{q(1-p)}{p+q(1-p)} < 1$ (see two slides above)
- Thus $v(\tau)$ has a slope smaller than 1.
- What is the slope of $f(\tau, \eta) = \tau\eta$?
- Only the high productivity types are left with the firm

$$\Rightarrow f(\tau, \eta) = \tau \Rightarrow f'(\tau, \eta) = 1$$

Acemoglu & Pischke - Graphical Analysis

- We are in a situation where there is an increasing wedge between the MP and the wage due to adverse selection and therefore the firm pays for general training:



Acemoglu & Pischke - Empirical Evidence

- Acemoglu & Pischke investigate whether they can find evidence for adverse selection in the German apprenticeship market
- German apprentices leave the training firm for 3 possible reasons:
  1. Employer does not offer a permanent contract ("lemons")
  2. Worker quits voluntarily (exogenous movers)
  3. To do compulsory military service
- What does the model predict for the earnings of these 3 groups of leavers?

1. *Stayers* earn at least as much as the movers: $w(\tau) \geq v(\tau)$
   (In the model $w(\tau) = v(\tau)$ because firms get all the rents from the adverse selection problem)
2. *Military Quitters* earn more than other movers: $v^m(\tau) > v(\tau)$
   Because military quitters have only a random fraction of lemons
3. *Military Quitters* may earn more than stayers: $v^m(\tau) \gtreqless w(\tau)$
   Military quitters have lower average ability than stayers but once they have separated from the training firm they can obtain their marginal product while the incumbent firm extracts rents from the stayers

Acemoglu & Pischke - Regression Results

Independent variable	Qualification and career survey (1)	(2)	1984 SOEP (3)
Attended 10th grade	0.160	0.162	0.153
	(0.011)	(0.007)	(0.028)
Experience	0.123	0.108	0.028
	(0.024)	(0.017)	(0.098)
Experience^2/100	−0.694	−0.532	−0.086
	(0.182)	(0.126)	(0.848)
Experience^3/10,000	1.840	1.229	0.161
	(0.558)	(0.379)	(3.007)
Experience^4/1,000,000	−1.806	−1.063	−0.123
	(0.597)	(0.399)	(3.740)
Stayer	0.012	0.027	—
	(0.015)	(0.008)	
Military quitter	0.045	0.011	—
	(0.025)	(0.014)	
Ever did military service	—	—	−0.022
			(0.024)
R^2	0.384	0.337	0.126

- Column (1) uses firms with $\geq$ 50 employees, column (2) uses firms of all sizes, but drops size controls to be comparable to the SOEP

Specific Skills

- As we have seen before Becker distinguished between 2 types of skills:

  1. General skills
  2. Specific skills

- Who is going to pay for training in firm-specific skills?
- Rewriting the productivity and outside wage functions to include specific human capital we get:

$$y = f(g, s) \text{ with } f_g(g, s) > 0 \text{ and } f_s(g, s) > 0$$
$$w = v(g, s) \text{ with } v_g(g, s) > 0 \text{ and } v_s(g, s) = 0$$

(the outside wage cannot increase with specific human capital).

Specific Skills – Would Firms Pay for Specific Training?

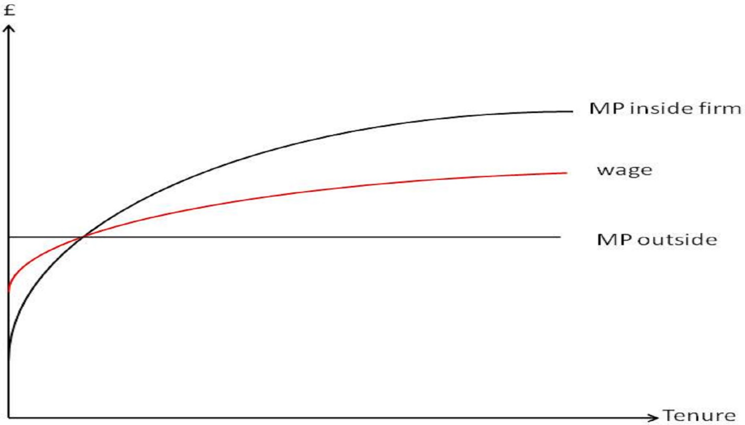
- Suppose the firm paid for specific training
- The worker would get his MP which only depends on general training in the outside market and therefore at the training firm as well
- Because the inside and the outside wage are exactly the same, the worker has no incentive to stay at the firm
- He can threaten to leave but then the firm would lose out and so the firm would be willing to pay a higher wage than the outside wage to keep the worker

Specific Skills – Would Workers Pay for Specific Training?

- Suppose the worker paid for specific training and got all the rents
- In that case the firm would not gain from employing the trained worker
- So it may just employ another worker
- If the firm fires the worker before returns = costs the worker would lose out
- The worker therefore has an incentive to give the firm some of his rents to be able to stay in the firm

Specific Skills – Would Workers Pay for Specific Training?

- This is a typical case where some form of Nash bargaining seems plausible.
- The wage schedule would therefore look as follows:



Estimating the Returns to Specific Human Capital

- One way to estimate whether specific human capital matters is to estimate the following regression:

$$\ln(w) = \beta_1 + \beta_2 S + \beta_3 EXP + \beta_4 TENURE + \varepsilon$$

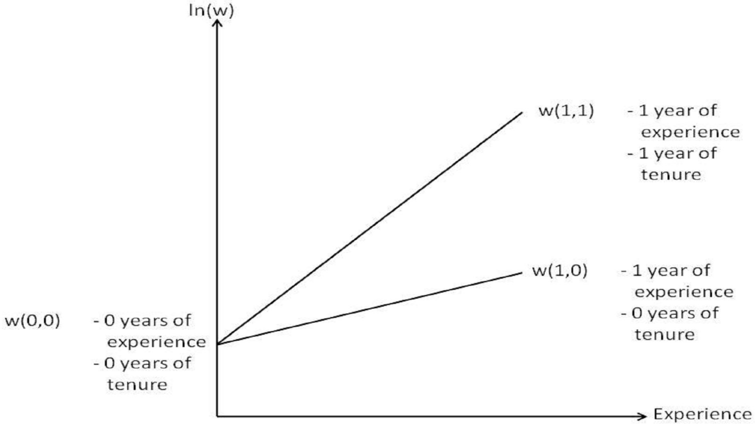
Estimating the Returns to Specific Human Capital

- One way to estimate whether specific human capital matters is to estimate the following regression:

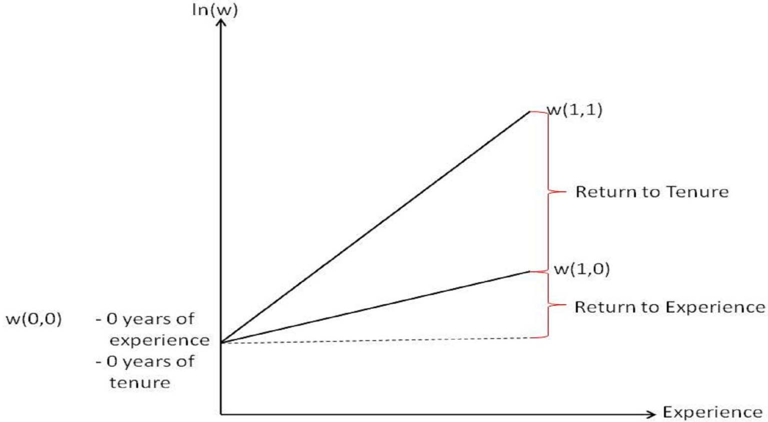
$$\ln(w) = \beta_1 + \beta_2 S + \beta_3 EXP + \beta_4 TENURE + \varepsilon$$

- This is, however, problematic because workers with more tenure may be different from other workers due to the following reasons:
  1. Because they are very good the firm will always try to keep the worker
  2. Because their search has helped them to find a particularly good firm they do not want to leave
- Because of these reasons the coefficient on tenure is likely to be upward biased if one estimates this model in a cross-sectional sample of workers

Estimating the Returns to Specific Human Capital



Estimating the Returns to Specific Human Capital

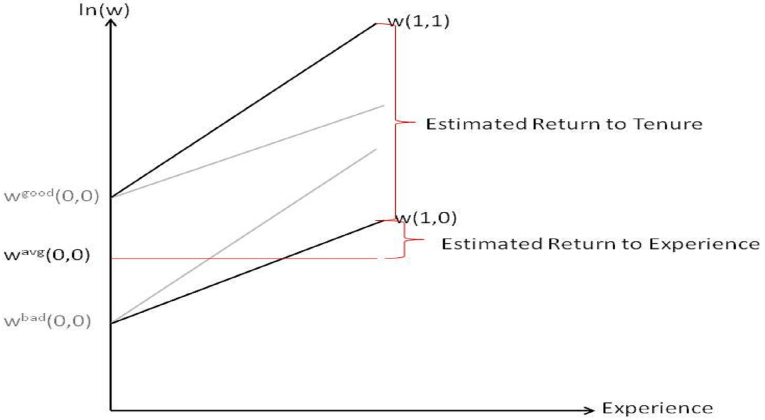


- The problem is that we never observe all 3 wages for the same person.

Estimating the Returns to Specific Human Capital

- The cross-sectional estimates of returns to job tenure and experience would be:
  - experience:
$$w(1,0) - w(0,0)$$
  - tenure:
$$w(1,1) - w(1,0)$$
- This would be biased. Illustrating this with a stark example:
- Suppose everybody with a bad job quits and everybody in a good job stays

Estimating the Returns to Specific Human Capital



Estimating the Returns to Specific Human Capital

- Cross sectional estimates of tenure would therefore be upward biased
- During the 1980s, researchers have started to estimate returns to tenure using panel data. This allows to include individual specific effects which addresses part of the problem
- A particularly careful and revolutionary study is the study on earnings losses of displaced workers by Jacobson, LaLonde, and Sullivan (1993)

Estimating the Earnings Losses of Displaced Workers

- Jacobson, LaLonde, and Sullivan (1993) investigate returns to specific human capital by looking at displaced workers.
- There are a two important reasons why displaced workers may have lower earnings because they lose:
  1. *Firm specific human capital*. (They would initially lose this job-specific HC but their earnings may recover if they work in the new jobs for long enough)
  2. *Search* capital: workers may have searched for higher paying jobs for some time before ending up in their old job. Now they have to start again
- They focus on high-tenure workers because they are more likely to have accumulated substantial amounts of firm specific human capital or "match" capital prior to their job loss
- They use data from administrative records covering 5 percent of the workforce in Pennsylvania for the years 1974 to 1986

What happens to Earnings After Job Separation?

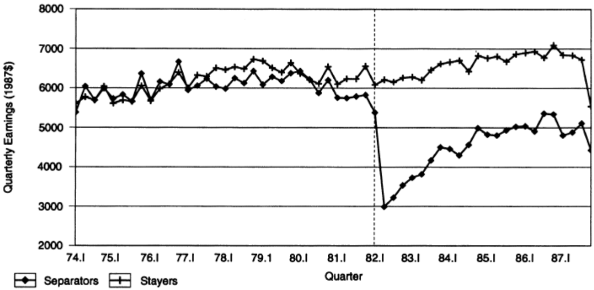


FIGURE 1. QUARTERLY EARNINGS (1987 DOLLARS) OF HIGH-ATTACHMENT WORKERS SEPARATING IN QUARTER 1982:1 AND WORKERS STAYING THROUGH QUARTER 1986:4

- Earnings of separators fall sharply after separation and then rise quickly for the first quarters after separation. They stay, however, about 20 percent below their pre-separation level
- Already before the separation wages of separators are slightly lower than wages of stayers. It will therefore be important to control for the heterogeneity of workers

Defining Earnings Losses

- Previous studies had often measured the losses of displaced workers by subtracting their post-displacement earnings from their pre-displacement earnings. This is problematic due to a number of reasons:
  1. This does not control for macroeconomic factors that cause changes in workers' earnings regardless of the displacement
  2. This does not account for the earnings growth that would have occurred in the absence of the job loss
  3. Firm's declining fortunes may adversely affect workers' earnings several years prior the the displacement
- They define displaced workers' earnings losses to be the difference between their actual and expected earnings had the events that led to their job losses not occurred.

Econometric Model

- They estimate the following regression:

$$y_{it} = \alpha_i + \gamma_t + \mathbf{x}_{it}\beta + \sum D_{it}^k \delta_k + \varepsilon_{it}$$

  - y_i is income of individual i in quarter t.
  - α_i is an individual FE. (controlling for heterogeneity across workers)
  - γ_t is a quarter FE. (controls for macroeconomic effects affecting all workers)
  - $\mathbf{x}_{it}$ consists of observed time varying characteristics of the worker (such as age)
  - D_{it}^k is a full set of dummies covering 20 quarters prior and many quarters after separation interacted with whether a worker is displaced
- This is essentially a DiD estimator but they estimate the separation effect for several pre and post-separation quarters
- They estimate their results for a mass layoff sample (more than 30 percent of the workforce dismissed) and other layoffs

Econometric Model – Crucial Assumption

- As with any DiD type strategy we have to be careful that treatment and control group would evolve similarly in the absence of treatment
- They address this concern by including linear worker-specific time trends (a separate time trend for each worker):

$$y_{it} = \alpha_i + \omega_i t + \gamma_t + \mathbf{x}_{it}\beta + \sum D_{it}^k \delta_k + \varepsilon_{it} (\omega_i t \text{ is the trend})$$

- As the key parameters of interest is the large set of quarterly dummies interacted with whether you are dismissed they plot the estimated regression coefficients:

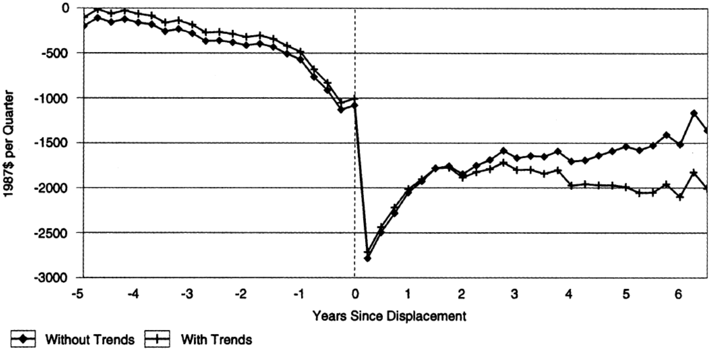


FIGURE 2. EARNINGS LOSSES FOR SEPARATORS IN MASS-LAYOFF SAMPLE

Mass Layoff Results

- As we can see displaced workers have falling earnings already about 3 years before the displacement
- In the displacement year earnings drop sharply
- Earnings recover somewhat in the first 3 years after displacement but are flat and about 25 percent lower than pre-displacement earnings for a long time after displacement
- If one controls for worker-specific time trends the estimated losses are even larger in the long run. (this suggests that firms are not displacing workers with more slowly growing earnings)

Mass Layoff Results – Within Firm Results

- An alternative way to measure the earnings loss due to displacement is to compare workers within a same firm (j) where some have been displaced and others have not

$$y_{ijt} = \alpha_{ij} + \gamma_{jt} + \sum D_{it}^{k}\delta_k + \varepsilon_{ijt}$$

  Note: the individual FE now include the j subscript because the sample contains only workers who have stayed in the same firm

  - Advantage: we avoid potential bias if particular firms make workers with a certain skill set especially productive
  - Disadvantage: we may underestimate (or overestimate) the effect of dismissal as the struggling firm may reduce wage of remaining workers because it is struggling (or overestimate if the firm keeps only those workers with a particularly good firm-worker skill match)
- The sample excludes workers from firms that shut down

Mass Layoff Results – Within Firm Results

- Now they compare displaced workers' earnings to earnings of workers in the same firm:

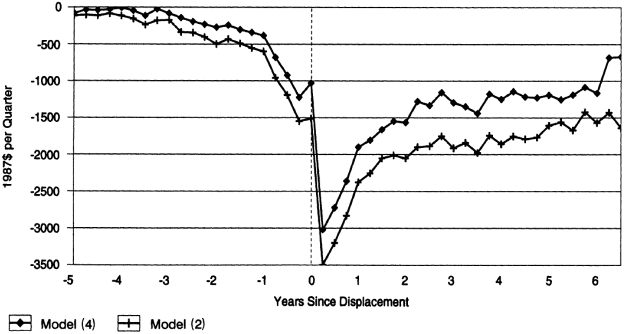


FIGURE 4. SENSITIVITY OF EARNINGS-LOSS ESTIMATES FOR MASS-LAYOFF SAMPLE TO DIFFERENT COMPARISON GROUPS

Non-Mass Layoff Results

- They also investigate the evolution of earnings for separators in a non-mass layoff sample

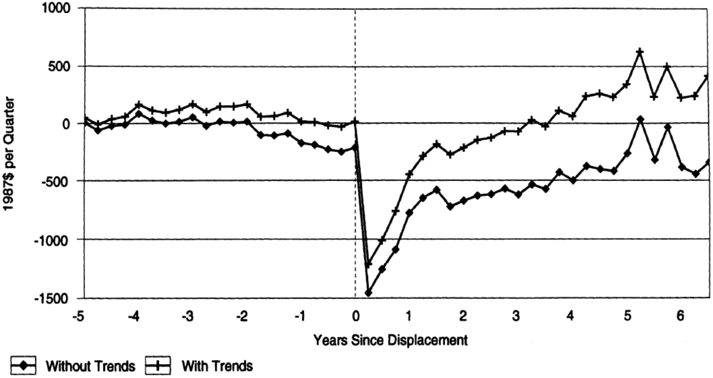


FIGURE 3. EARNINGS LOSSES FOR SEPARATORS IN NON-MASS-LAYOFF SAMPLE

Non-Mass Layoff Results

- The evolution of earnings for separators in the non-mass layoff sample look very different:
  - They do not decline before separation
  - In the quarter of separation they decline sharply but by less than those in the mass layoff sample and after that earnings recover
  - This is not surprising as the sample includes a larger fraction of workers who quit their jobs or who had fewer firm-specific skills

Is this Evidence that Specific Human Capital Matters?

- The study is extremely careful and thoughtful. It does, however, not fully inform us whether these earnings losses are due to the loss of firm-specific human capital or the loss of "search" capital
- Job loss may also affect worker's health, marriage, and so on. All these factors will affect the wage profile after job losses
- It is very difficult to obtain empirical evidence on the importance of specific human capital!