Lecture Note: Theories of the Provision and Payment of General Skills Training

David H. Autor
MIT and NBER

October 20, 2003

1 Goal of this lecture

- Review basic Becker story
- Consider stylized model of general skills training in a competitive labor market.
- Add various labor market perfections to the model to consider how they alter firms' training decisions:
  - Search costs and other wedges
  - Search costs that differ by worker skill level
  - Exogenous turnover
  - Nash bargaining
  - Minimum wages
  - Asymmetric information→Adverse selection
- Goal is to understand who pays for general skills training, and what are the efficiency consequences. As it turns out, market structure and training provision are tightly linked, an insight originally due to Becker. Deviating from either the perfectly competitive or pure monopsony cases studied by Becker gives rise to interesting and general insights about wage structure, market institutions, private information, and training provision.
- An interesting feature of the modern literature is its implicit focus on the limitations of contracting. Becker considered one specific contracting problem – holdup of firm by workers. It turns out, this problem is richer than Becker may have seen at that time.

1.1 Basic model

Definition 1 *Generals skills – Skills useful to many employers ("many buyers").*

Definition 2 *Specific skills – Skills with exactly one buyer.*

- Timeline

1. $t = 1$. Training and production (human capital investments sunk). Workers produce output y and receive training τ at cost $c(\tau)$.
2. $t = 1.5$. Workers face option to leave firm.
3. $t = 2$. Production takes place again. Worker produces output $y + f(\tau)$ at any firm (general training).

- Assume: $f'(\tau) > 0, c(0) = c'(0) = 0$ and $c'(\tau) > 0, c''(\tau) > 0, \underset{\tau \to \infty}{c'(\tau)} \to \infty$.

1.2 Basic Becker insight: Holdup

- FOC social optimum for training? $c'(\tau^*) = f'(\tau^*)$.
- How do we know that $\tau^* > 0$?
- Can firm pay worker $w_1 = y$, $w_2 = y + f(\tau) - c(\tau)$?
- Basic insight of Becker model: In a competitive labor market, employers will not pay for general human capital. The problem is "holdup." Since the worker's opportunity (outside) wage rises by $f(\tau)$ with training, firm must pay $w_2 = y + f(\tau)$ or lose employee to a competitor.
- Q: Since employers unwilling to fund socially productive training, is this a market failure? A: Worker is full residual claimant to returns from training investments $\Rightarrow$ Efficient incentive for worker to make the investments.
- Two canonical solutions to this problem:
  1. Direct investments by workers – School, training programs, etc.
  2. Indirect payment to employer: Training wage. Hence:
     $w_1 = y - c(\tau^*)$, $w_2 = y + f(\tau^*)$.
- Just to check intuition:

- How do high rates of labor turnover affect training investment incentives? Not at all.
- How do worker credit constraints impact this model? Quite a bit.
  * Cannot typically borrow against future stock of human capital (except from government, e.g., student loans).
  * Wages may not go low enough (e.g., negative) to cover efficient training expenditures.

- How relevant is this model?
  - Clearly, covers considerable territory: College education is paid for by families, society; Professional schools do not generally subsidize tuition: Law, medicine, business.
  - What about the case of military education? Military flight education costs multiple millions to provide. The skills are completely portable (to commercial aviation).

1.3 Is that all there is to it?

- A persistent suspicion that there is more going on – employers are paying for general skills training.
  - German apprenticeships: Acemoglu and Pischke QJE (1998) Table 1.
  - BLS 1994 study of employer training. Found that employers spent $3.4 billion in 1994 on direct tuition reimbursements and outside training funds (almost certainly general). Was all of this paid for by workers?
  - Example of Temporary Help Supply (THS) firms. Demonstrably clear that THS firms do pay up-front costs of training (about $150 per trainee) prior to job assignment, without any contractual obligation to perform work.

- These examples have spurred continued interest in the provision and payment of general skills training.

- There are 2 seminal papers that give rise to discussion of modern theories of training. One is Grout (1984) *Econometrica,* the second is Greenwald (1986), *Review of Economic Studies*. I'll touch on results from both of these papers later on.
- Continue with the setup from above with.
  - $f(\tau)$ is productivity with $f'(\tau) > 0$ and $f''(\tau) < 0$, and $c(\tau)$ is as above.
  - Denote $\nu(\tau)$ as the outside wage paid to a person with training τ.
  - Assume for now that for whatever reasons, workers do not obtain outside training. Hence $\tau = 0 < \tau^*$ initially.

1.3.1 Case 1: Constant wedge

- See Figure 1.
- Take a case initially where $f'(\tau) = \nu'(\tau)$ and $f''(\tau) = \nu''(\tau)$.
- Consider firm's maximization problem:

$$\max_{\tau} \pi(\tau) = f(\tau) - \nu(\tau) - c(\tau) \tag{1}$$

  Q: What does perfectly competitive market about relationship between $f(\tau)$ and $\nu(\tau)$?
  Q: Why $\nu(\tau)$ not $f(\tau)$ in the firm's maximization?

- FOC for (1) is

$$c'(\tau) = f'(\tau) - \nu'(\tau) \tag{2}$$

  which is satisfied at $\tau^* = 0$. This echoes the Becker model.

- Now instead assume that $f(\tau) - \nu(\tau) = \Delta(\tau) = \Delta > 0$. Can think of Δ as a search friction. How does this search affect τ^* in (2)? Why?

1.3.2 Case 2: Increasing wedge

- See Figure 2.

- Now consider a new search friction whereby

$$f(\tau) > \nu(\tau) \text{ and } f'(\tau) > \nu'(\tau).$$

  Continue to assume concavity of $f(\cdot)$ and $\nu(\cdot)$.

- Assume for now:

  - No exogenous turnover. All workers remain with the firm in period 2 if receive $v(\tau)$.
  - All bargaining power rests with the firm – that is, firm is the full residual claimant of any surplus $(f(\tau) - \nu(\tau) - c(\tau))$.

- First order condition is the same as (2), $f'(\tau) - \nu'(\tau) = c(\tau)$, but now there will be an interior solution at $\tau^* > 0$. This follows because $\Delta(\tau) > 0$ and $\Delta'(\tau) > 0$ (and recall that $c'(\cdot), c''(\cdot) > 0$).

- So notice that this search friction raises training, whereas the other does not. How does the assumption that $f'(\tau) > \nu'(\tau)$ change the holdup problem?

- Q: Is this solution socially efficient? A: No, inefficiently low.

- Social maximization is $f'(\tau) = c(\tau)$, whereas firm only maximizes $(f'(\tau) - \nu'(\tau)) = c(\tau)$. In other words, firm maximizes 'the wedge,' not output net of costs.

- So, firm would only choose the social optimum if $\nu'(\tau) = 0$, i.e., wage invariant to training level.

- This is something of a paradox: More 'inefficient' market structure – greater wedge between productivity and wages – encourages greater firm training. (Echoes Becker result on labor market monopsony, or firm specific capital also has the feature that $\nu'(\tau^{specific}) = 0$.)

- Let's look at two other wrinkles on this result: Nash bargaining and exogenous post-training turnover.

Nash bargaining and turnover

- Nash bargaining: A fraction of the surplus $\beta \in [0,1]$ goes to the worker.
- Nash bargaining corresponds to the absence of contracts. The assumption is that ex ante wage agreements are not feasible, so firms and workers bargain over the surplus ex post. Each demands at least its threat point: for workers, that's $\nu(\tau)$, for firms, that's also $\nu(\tau)$, since they could hire identical workers to produce $f(\tau)$ at wage $\nu(\tau)$ if these workers quit.
- So the maximization is now

$$c'(\tau^\beta) = (1-\beta)\left[f'(\tau^\beta) - \nu'(\tau^\beta)\right], \tag{3}$$

  with the worker's wage set by

$$w(\tau^\beta) = \nu(\tau^\beta) + \beta\left[f(\tau^\beta) - \nu(\tau^\beta)\right],$$

  and profits

$$\pi(\tau^\beta) = (1-\beta)\left[f(\tau^\beta) - \nu(\tau^\beta)\right] - c\left(\tau^\beta\right)$$

- Now, some important questions raised by the Grout article:

1. Why doesn't $c(\tau)$ enter (3) in parallel to $f(\tau)$ or $\nu(\tau)$?
   - Answer: Because $c(\tau)$ is sunk. And this bargain is set ex post.
   - That is, wages are determined by worker's outside opportunity $\nu(\tau)$ and her claim β on firm's quasi-rents, which are $f'(\tau^\beta) - \nu'(\tau^\beta)$.
   - Q: Why is $f'(\tau^\beta) - \nu'(\tau^\beta)$ a quasi rent?
2. Are workers necessarily better off with β large?
   - Notice that with τ constant, it is clearly the case that

$$\nu(\tau) + \beta\left[f(\tau) - \nu(\tau)\right] > \nu(\tau).$$

- But τ is not constant. As Grout's article demonstrates (not for this case), since since τ^β that satisfies $c'\left(\tau^\beta\right) = (1-\beta)\left[f'(\tau^\beta) - \nu'(\tau^\beta)\right]$is strictly less than τ^0 that satisfies $c'\left(\tau^0\right) = \left[f'(\tau^0) - \nu'(\tau^0)\right]$.This follows because in the $\beta > 0$ case, workers are expropriating part of the firm's quasi-rents from the sunk training investment. Anticipating this fact, the firm invests less in training.
- Note that this result also depends upon the convexity of the training cost function. By reducing scale of training, firm increases marginal returns.
- As β rises, workers get a larger share of a smaller pie. In the extreme case where $\beta = 1$, we have $\tau^\beta = 0$, no training occurs.
- This brings up a key point from Grout's article (Theorem 3.2):

$$-\frac{\partial \pi(\cdot)}{\partial \beta} > \frac{\partial w}{\partial \beta},$$

that is, the fall in profits for an increase in worker bargaining power is strictly less than the gain in worker wages – so the bargaining solution reduces social efficiency.

- How do we know this to be true? The change in profits for an increment to β is:

$$\frac{\partial \pi(\tau^\beta)}{\partial \beta} = -f(\tau^\beta) + \nu(\tau^\beta) + \left[(1-\beta)\left(f'(\tau^\beta) - \nu'(\tau^\beta)\right) - c'\left(\tau^\beta\right)\right]\partial\tau^\beta/\partial\beta,$$

whereas the increase in wages is

$$\frac{\partial w(\tau^\beta)}{\partial \beta} = f(\tau^\beta) - \nu(\tau^\beta) + \beta\left[f'(\tau^\beta) - \nu'(\tau^\beta)\right]\partial\tau^\beta/\partial\beta.$$

Putting these two together, the total social gain/loss is

$$\frac{\partial \pi(\tau^\beta)}{\partial \beta} + \frac{\partial w(\tau^\beta)}{\partial \beta} = \left[f'(\tau^\beta) - \nu'(\tau^\beta) - c'\left(\tau^\beta\right)\right]\partial\tau^\beta/\partial\beta < 0,$$

since $\partial\tau^\beta/\partial\beta < 0$ and $f'(\tau^\beta) - \nu'(\tau^\beta) - c'\left(\tau^\beta\right) > 0$ as long as t^β is below the social optimum, which will always be true here. Notice that the social loss here is exactly the contraction in the pie caused by the reduction in training $\partial\tau^\beta/\partial\beta$ (the rest is a transfer from firm to worker).

- So, as soon as we deviate into a case where contracts are not enforceable and costs are sunk, the 2nd welfare theorem no longer holds – there is a direct conflict between efficiency and distribution. This is a primary insight from Grout's article.

- Also briefly consider the case of exogenous worker turnover at rate q. From the firm's perspective, this works exactly like an increase in bargaining power. The FOC is

$$c'(\tau^q) = (1-q)\left[f'(\tau^q) - \nu'(\tau^q)\right].$$

  Training again drops with a rise in q. The reason is that there are fewer trainees around in the 2^{nd} period from whom the firm can capture the returns from training.

1.3.3 Summary so far

The discussion of training in the "increasing wedge" case gives rise to three perverse comparative statics. Training is increasing in:

1. The degree of distortion in the wage structure. The greater is $\Delta'(\tau)$, the more firms train.
2. The more bargaining power that rests with the firm. The smaller is β (worker's bargaining power), the more firm trains.
3. The lower is worker mobility. Less turnover $q \downarrow$, more training.

Hence, anything that increases firm's marginal share of returns to training also increases training in this setup. Yet, you should not think this is a complete accounting; a higher training economy is not necessarily globally more efficient:

- Training is closer to optimal
- But labor supply is sub-optimal due to inefficient wage incentives (workers paid below marginal product).
- The limited monopsony present when turnover is low may dampen efficient reallocation of workers to jobs (speaking loosely).

1.4 Structure of wages and investment in general training

- It is a commonplace belief that worker turnover is inimical to employer training. And indeed, industries with low turnover also have high training.

- Is it the simple duration of the attachment between workers and firms that gives rise to this pattern?

- Consider the constant wedge case, where $\Delta(\tau) > 0$ and $\Delta'(\tau) = 0$. The firm's FOC, as above, is

$$c'(\tau^q) = (1-q)\left[f'(\tau^q) - \nu'(\tau^q)\right],$$

  which is satisfied at $t^q = 0$.

- Hence, the commonplace assertion that "labor market frictions" are sufficient for firm investment in training is incorrect. A "frictional" labor market is not necessarily a high training labor market.

- Frictions must be of a specialized kind: The compression of the wage structure – that is, the gap between wages and marginal products – must be greater for more skilled/trained workers.

- If $f(\tau) - \nu(\tau) = \Delta$, then there is no firm-sponsored training, even with zero exogenous turnover and no worker bargaining power ($q = 0, \beta = 0$).

- This is a key insight because it directs your gaze towards features of the labor market that might give rise to a situation where $f(\tau) - \nu(\tau)$ is increasing in τ, that is $\Delta'(\tau) > 0$. This point is made in full generality by the Acemoglu-Pischke (1999) *JPE* paper.

- Now, we want to consider a two possible reasons why the wage structure might be compressed as needed to generate incentives for firm investment in general skills: 1) Minimum wages; 2) Adverse selection.

1.4.1 Minimum wages and investment in general training

- It is a standard view that minimum wages will reduce training. Why? Function like a credit constraint – reduce workers' ability to take a pay cut to pay for general training.
- This result has been "verified" in papers by Blackburn and Neumark among others – but this work does not exploit within-state variation in minimum wages, so it's hard to have great confidence in these papers.
- Consider a case as depicted in Figure 3. Here $\Delta'(\tau) = 0$ by assumption, so normally not an incentive for firm investment in worker training (however, we still assume that $\Delta(\tau) > 0$, which is not standard).
- Imposition of a binding minimum wage compresses wage distribution below the minimum, but leaves it unaffected above. If the following things are true:
  - Firms don't shut down
  - Positive profits are still made on these workers
  - $c(\tau_{\min}) < [f(\tau_{\min}) - \nu(\tau_{\min})] - (\tau_{\min}) < [f(0) - w_{\min})]$
    ...firms will want to increase training.
- A table from the Acemoglu-Pischke (1999) paper on minimum wages and general training provides some evidence that this might be going on.
- At a minimum, their paper does not support the contention that the impact of binding minimum wage on training provision is negative.

1.4.2 Adverse selection, monopsony, and training

- Greenwald's seminal 1986 *ReStud* paper offered the general observation that if workers are heterogeneous and potential employers cannot observe this heterogeneity perfectly, incumbent employers are likely to have an informational advantage over their employees relative to outside employers. This informational advantage can give rise to adverse selection.

- This insight is directly analogous to the Akerlof 1970 "Market for Lemons" paper applied to the labor market – where it could easily be more relevant. (Q: Unlike used cars, workers can tell you about their 'quality.' Why doesn't this solve the problem?)

- Let's embellish the current model to incorporate the case of worker heterogeneity and see how it gives rise to adverse selection. Assumptions:

  1. Two types of workers: $\eta \in \{H, L\}$. $\Pr(\eta = H) = \rho$.
  2. Type H produces $f(\tau)$, type L produces 0 (a normalization).
  3. Ability η is unknown to employers at time of hire. ρ and $f(\tau)$ are common knowledge.
  4. Ability for each employee is observed by incumbent employer in 1^{st} period
  5. There is *exogenous* turnover after end of 1^{st} turnover. A fraction q of all workers turns over immediately after training.
  6. In addition, there is *endogenous* turnover if $w < \nu(\tau)$.

- Time line is as above:

  1. $t = 0$: Training and production
  2. $t = 1.5$: Turnover
  3. $t = 2$: Production

- What are equilibrium wages and training in this model? Solve by backward induction.

  - Consider wages in 2^{nd} period *assuming* that $\tau^* > 0$.
  - The expected productivity of trained type H workers is $f(\tau^*)$
  - The expected productivity of trained type L workers is 0.

- How are wages set in the outside labor market? Our outside wage concept is $\nu(\tau^*)$. This is the opportunity wage for a trained worker (remember, we are assuming all workers are trained).

- Since outside firms cannot distinguish H from L workers, it must be the case that $\nu(\tau^*)$ is independent of η. So $\nu(\tau^*)$ must equal the *expected* productivity of workers available for hire at wage $\nu(\tau^*)$. (This is the Perfect Bayesian equilibrium concept that is normally applied to models with information uncertainty).

- It's easiest to think about the expected productivity of workers who have turned over, since they are in the market (but we can generalize to the case of 'raids' on firms). Expected productivity of departed workers is

$$E(f(\cdot)\,|\text{separate}) = \frac{(\#H \text{ separators})}{\#H \text{ separators} + \#L \text{ separators}} \cdot f(\tau^*)$$

- What is the composition of the outside pool:

  - Exogenous turnover q of H workers $\rightarrow q\rho$
  - Exogenous turnover q of L workers $\rightarrow q(1-\rho)$
  - All type L workers who don't turn over by themselves will be offered a wage of 0 by their incumbent employer after the 1^{st} period (this is the informational advantage). This fraction is $(1-q)(1-\rho)$. These group of endogenous separators will 'pool' with the exogenous separators. Hence, productivity of outside pool is

$$\nu(\tau^*) = \frac{q\rho}{q\rho + (1-\rho)} \cdot f(\tau^*) \tag{4}$$

- Consequently, the incumbent employer need only pay $\nu(\tau^*)$ to retain its type H workers. Incumbent employer benefits from "ex post monopsony" stemming from private information.

  Why? Because this is the wage that they would command if they quit the firm (in equilibrium, they do not, of course). Moreover, if an outside firm tried to bid them away, the incumbent employer could simply match wage offers up to $f(\tau^*)$ for its H workers while not matching this wage for its L workers. Hence, the outside firm would always overpay for workers it poached ('winner's curse').

- Now, return to firm's maximization for training in 1^{st} period,

$$\max_{\tau} \pi = (1-q)\,\rho\,[f(\tau) - \nu(\tau)] - c(\tau),$$

with FOC

$$c'(\tau^*) = (1-q)\,\rho\,[f'(\tau^*) - \nu'(\tau^*)]\,.$$

Since $\nu'(\tau) < f'(\tau)$, this maximization will have a solution with $\tau^* > 0$.

- In this example, adverse selection compresses the wage structure. Why does this work? Due to private information, the expected productivity of the pool of separators is below the expected productivity of the randomly drawn hire during the 1^{st} period. That's because good workers predominately stay with their own firm, while the bad ones turnover after their low productivity is discovered by incumbent employer. Although good workers would like to earn $f(\tau^*)$, their actual productivity, rather than $\nu(\tau)$, the opportunity wage, they would be cursed by adverse selection if they turned over. Hence, incumbent employer has limited monopsony power.

- Q: The model has a hidden assumption – unrelated to the information structure of the model – that made this result work. What is it? There is a complementarity between training and ability; specifically, $\partial^2 f(\cdot)/\partial\tau\partial\eta > 0$: H workers get $\partial f(\cdot)/\partial\tau > 0$ from training, whereas L workers get $\partial f(\cdot)/\partial\tau = 0$.

- Consider instead if we had assumed that $f(\tau) = \eta + \gamma(\tau)$. Then equation (4) becomes

$$\nu(\tau) = \frac{q\rho}{q\rho + (1-\rho)} + f(\tau), \tag{5}$$

and so $f'(\tau) - \nu(\tau) = 0$, and no general training is provided.

- So, this assumed complementarity between training and ability (positive cross-partial derivative) was necessary to generate the "increasing wedge" that makes it profitable for firms to fund general training. Whether this assumption is reasonable – in addition to being necessary – is a question that the model cannot answer.

- In fact, there is considerable evidence that better educated and better paid workers – in addition to more "able" workers within education cells – receive more training. Is this evidence convincing for existence of complementarity? If you believed in complementarity, how would you explain the fact that Instrumental Variables (IV) returns to education are

higher than OLS returns? IV returns, after all, are identified from behavior of less-educated workers.

1.4.3 Testing the adverse selection model

- For many economists, the adverse selection model has something of the ring of truth. That is, incumbent employers do have better information about their own employees than do other potential employers – this is hard to dispute. And this informational advantage might make it rational to invest in their general skills if it gave rose to some type of wedge between productivity and outside wages.

- Problem: A model that is inherently about unobservable information is difficult to test using standard data sources.

- One idea: Look for exogenous shocks to employment that break adverse selection equilibrium.

    1. Gibbons and Katz (1991) *Journal of Labor Economics* (not about training). Layoffs versus plant closings. It follows very generally from the adverse selection argument that workers who separate from a firm endogenously (layoffs) should be worse than workers who separate due to exogenous shocks (plant closings – not necessarily exogenous, but affect bad and good workers in the plant simultaneously).
       Hence, G&K compare these two groups, show that in fact the laid off workers do considerably worse (esp. for white collar workers). This is particularly striking because you might assume that workers losing jobs in plant closing would fare worse due to concentrated job loss in one geographic area.
    2. Acemoglu and Pischke (1999, *QJE*). They view "military quitters" as a group like those who experience plant closings. Exogenously separate, not cursed by adverse selection. So, even though military quitters are lower in ability than average workers, their post-military wages are higher. This is interesting, though perhaps not a strong est.

- Another idea: Study a labor market where...

- Lots of general training given.
- Essentially no firm-specific capital (since workers sent to multiple client sites), so training most likely "general."
- High turnover. (An outlier in the industry plot of training versus turnover.)
- Training provided up-front during unpaid, non-production hours and is not contracted. Hence, there is no opportunity for firms to pay a training wage that is below marginal product.
- Heterogeneity of policies – not all firms train (but the 'high quality' ones do).

- Autor, 2001 *QJE* attempts to understand why this training is profitable for THS firms. Motivations for training (interviews):

  1. 'Recruitment' – Attract skilled/motivated workers
  2. Testing – Training is a skill screen
  3. Skills development – Human capital acquisition, particularly in office software

- Insight of this model is that if training and ability are indeed complements, offering training should induce positive self-selection on unobserved ability – that is, high ability workers value training more than low ability workers since it is complementary with ability. (See Salop and Salop 1976 *QJE* for a similar structure applied to a different benefit.)

- Setup of model. Building on previous information model, but now giving workers some *ex ante* knowledge of own ability (necessary for self-selection to be relevant):

  1. Workers have some private information about own ability.
     - $\eta = \{H, L\} = \{1, 0\}$. $\Pr(\eta = H) = \rho$.
     - Workers have beliefs about their own ability ex ante: $b = \{H, L\}$
     - $\Pr(\eta = H|b = h) = \delta_h > \rho$. $\Pr(\eta = H|b = l) = \delta_l < (1 - \rho)$. So $1 > \delta_h > \rho > \delta_l > 0$. Beliefs are informative but not infallible.

2. Firms learn about ability by training. That is, training incorporates testing, but it is not only testing.
3. Ability and training are publicly observed in 3^{rd} period. So, period of monopsony is brief.
4. Production $f(\tau) = \eta(1+\tau)$. Training and ability are complements.

- Time line of this model:
  - $t = 1$: Workers select a firm (training versus non-training) and receive any training offered.
  - $t = 1.5$: Exogenous + endogenous turnover
  - $t = 2$: Deployed by THS firm to client site
  - $t = 3$: Hired into non-THS sector (monopsony ends)

- To find equilibrium, backward induction. Start at 3^{rd} period. Assume we are at separating equilibrium. H workers go to training firms, L workers do not.

- Third period:

$$w_3 = \eta(1+\tau).$$

  Since ability and training are publicly observed in 3^{rd} period, workers are paid their marginal product, which depends on η and where they worked in 1^{st} period and hence how much training they received.

- Second period:

$$w_2 = \left\{ \begin{array}{ll} \nu(0) = \delta_l & \text{if } b = l \\ \nu(\tau) = \frac{q\delta_h(1+\tau^*)}{q\delta_h+(1-\delta_h)} & \text{if } b = h \end{array} \right\}.$$

  Notice that there is no adverse selection for low ability workers (who go to non-training firms); since non-training firms don't train, don't get private information. So, wage of workers at non-training firms is simply the share who are in expectation high ability, which is δ_l.

  By contrast, wage of trainees $\nu(\tau)$ is set by adverse selection due to private information.

- Comparison of $\nu(0) \lesseqgtr \nu(\tau)$: Could easily be the case that $\nu(0) > \nu(\tau)$, even though ability, productivity higher at training firms (since high ability share is $\delta_h > \delta_l$). Why? Consider comparison with "military quitters."

- Now, we can get separating condition for low ability to choose non-training firms, high ability to choose training firms. Expected earnings are:

	ω_1	ω_2	ω_3
No train	0	$\nu(0)$	δ_l
Train	0	$\nu(\tau^*)$	$\delta_l(1+\tau^*)$

- For high ability, expected earnings are:

	ω_1	ω_2	ω_3
No train	0	$\nu(0)$	δ_h
Train	0	$\nu(\tau^*)$	$\delta_h(1+\tau^*)$

- So, the separating condition under which high ability workers choose training firms and low ability workers choose non-training firms is simply,

$$\delta_l \tau^* < \nu(0) - \nu(\tau^*) < \delta_h \tau^*.$$

  The expected gain to training for high ability workers exceeds short term wage costs in period 2, and vice versa for low ability workers.

- Testable empirical implication: Wages lower (post-training) at training establishments, despite up-front training. This is particularly surprising since wages of trainees should be higher ex post in Becker setup – and this is almost universally found in literature. This would be true in general even if workers were paying for training if ability and training are complements. (See tables from QJE 2001)

- So, in this model, training is not solely a human capital formation mechanism. It also elicits otherwise unobservable information about worker (beliefs about) ability, which firms can use profitably (for a short period) while they hold information privately.

- (A contribution of this paper: May help to understand the role that THS industry plays as an information broker in labor market, rather than as simply a provider of spot market labor services.)

1.5 Conclusions

- View of general training has come some ways since Becker, though its fair to say that these models are standing on Becker's shoulders rather than standing Becker on his head.
- View of training has to some extent caught up with what economists probably once believed but did not have tools to express: private information, poor contracting, distorts costs and benefits of training. It is neither wholly paid for by workers nor necessarily efficiently provided by firms.
- Does this literature have a future? It's possible to argue that training has been over-studied by labor economists. It's no longer interesting to ask what are the 'returns to training.' However, it is interesting to ask about how market structure affects training provision (the question implicitly raised by the A&P models on wage compression and training provision).
- Interesting example: training provision and AIDS in Africa. How would you expect AIDS to affect training investment? Overall? By age group? By geography?
- Next topic: Specific capital. The Becker view and some contemporary versions.

Figure 1

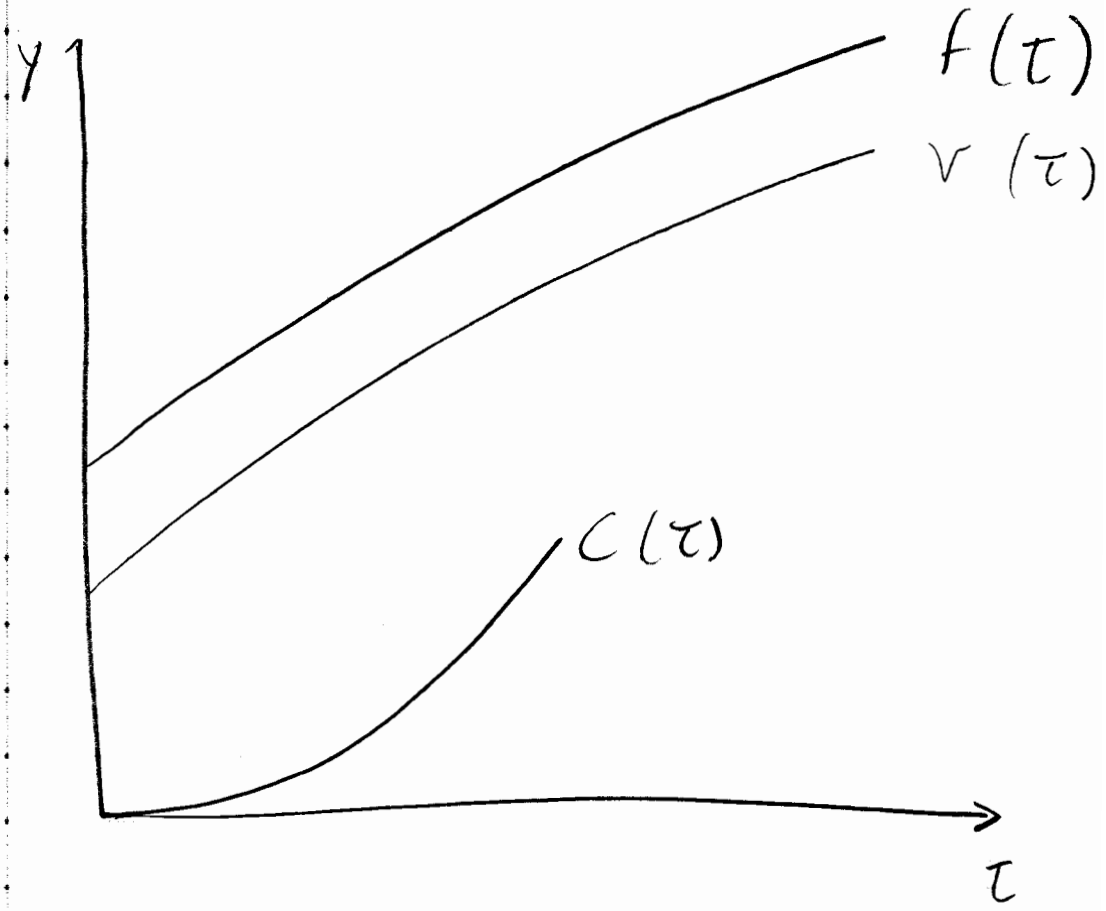


Figure 2

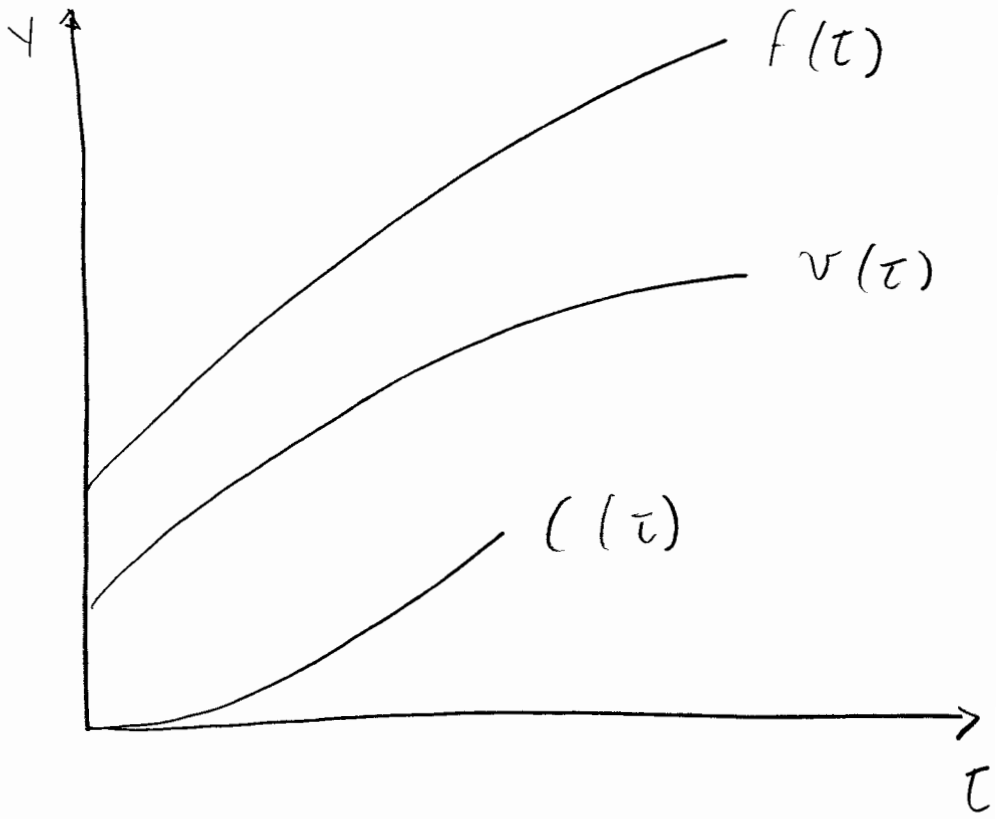
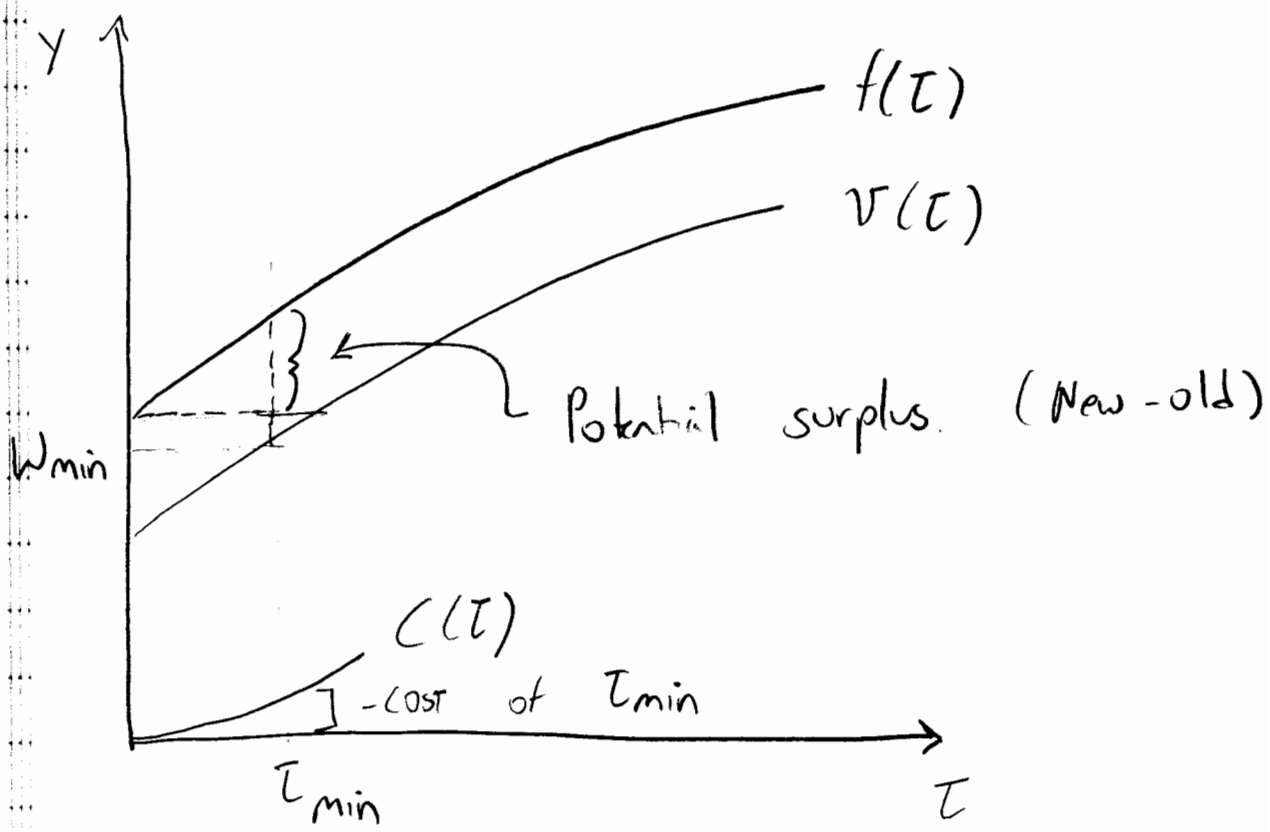


Figure 3


y
f(τ)
v(τ)
Potential surplus (New - old)
w_{min}
c(τ)
- cost of τ_{min}
τ_{min}
τ

Tables removed due to copyright considerations. Please see

Acemoglu, Daron, and Steve Pischke. Table I. In "Why Do Firms Train? Theory and Evidence." Quarterly Journal of Economics 113, no. 1 (February 1998): 99.

Acemoglu, Daron and Jorn-Steffen Pischke. 2002. "Minimum Wages and On-the-Job Training." Center for Economic Performance Working Paper No. 527, April.

Autor, David H. Figure entitled "Industry Frequency of Company Sponsored Training."

Gibbons, Robert, and Lawrence F. Katz. Table 2. In "Layoffs and Lemons." Journal of Labor Economics 9, no. 4 (1991): 364.

Gibbons and Katz. Table 3: 365.

Acemoglu and Pischke (1998). Table III: 109.

Autor, David H. Tables I, II, and III. In "Why Do Temporary Help Firms Provide Free General Skills Training?" Quarterly Journal of Economics 116, no. 3 (2001): 1413, 1431, 1432.

Autor, David, Frank Levy, and Richard J. Murname. Tables 12 and 15. In "Skills Training in the Temporary Help Sector: Employer Motivations and Worker Impacts." MIT mimeo (August 1999).