ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Training and Search On the Job

Authors	Rasmus Lentz, and Nicolas Roys
Working Paper Number	2016-025A
Creation Date	October 2016
Citable Link	https://doi.org/10.20955/wp.2016.025
Suggested Citation	Lentz, R., Roys, N., 2016; Training and Search On the Job, Federal Reserve Bank of St. Louis Working Paper 2016-025. URL https://doi.org/10.20955/wp.2016.025

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Training and Search On the Job*

Rasmus Lentz[†] Nicolas Roys [‡]

October 30, 2015

Abstract

The paper studies human capital accumulation over workers' careers in an on the job search setting with heterogenous firms. In renegotiation proof employment contracts, more productive firms provide more training. Both general and specific training induce higher wages within jobs, and with future employers, even conditional on the future employer type. Because matches do not internalize the specific capital loss from employer changes, specific human capital can be over-accumulated, more so in low type firms. While validating the Acemoglu and Pischke (1999) mechanisms, the analysis nevertheless arrives at the opposite conclusion: That increased labor market friction reduces training in equilibrium.

Keywords: Wage contracts, human capital, training, wage dispersion, frictional labor markets, optimal contract design, firm heterogeneity, sorting.

JEL codes: D21, D43, D83, E24, J24, J31, J33, J41, J62, J63, J64

*We are grateful for helpful comments from Melvyn Coles, John Kennan, Chris Taber, as well as seminar and conference participants at UW-Milwaukee, Penn State, McGill University, the Barcelona 2015 Summer Workshop on Matching and Sorting, The 2015 Econometric Society World Congress, The NYU 2015 Search and Matching Workshop, and The 2015 Århus University Conference on Labour Market Models and Their Applications. Rasmus Lentz acknowledges financial support from the National Science Foundation, grant SES-1459897.

[†]University of Wisconsin-Madison; NBER; LMDG; CAP; E-mail: rlentz@ssc.wisc.edu

[‡]University of Wisconsin-Madison; E-mail: nroys@ssc.wisc.edu

1 Introduction

In the United States, worker reallocation between firms most commonly happens without an intervening unemployment spell.[1] Job-to-job transitions involve a direct competition between the two employers, whereas reallocation through unemployment resets the worker's bargaining position to her valuation of unemployment.[2] Christensen et al. (2005) show that job-to-job reallocations are motivated by the worker's search for more productive positions and higher wages. In this paper, we study the investment in workers' human capital, be it general or specific, during their careers in a frictional setting with on-the-job search and heterogeneous firms. Employment contracts specify employment history conditional training and wage paths, as well as responses to outside competition for the worker. The inclusion of on-the-job search and heterogenous firms in the frictional environment offers new insights into the interaction between training and frictions as well as the implications for wage dynamics.

The impact of human capital accumulation on within and between firm wage dynamics does not in qualitative sense vary by its specificity. Both perfectly general and perfectly specific training result in increasing wages within the job as well as between jobs.[3] The classic distinction between labor market experience and firm tenure effects in wage regressions as studied in among others Altonji and Shakotko (1987), Topel (1991), Altonji and Williams (2005), Dustman and Meghir (2005), and Buchinsky et al. (2010) is therefore a murky reflection of the specificity of accumulated human capital during careers.

The employment contracts in the paper are required to be renegotiation proof and as a result they respond to outside competition in a manner identical to the offer matching process in Postel-Vinay and Robin (2002). General training results in an increased valuation of the worker across all potential employers. The occasional arrival of outside offers delivers the valuation increase to the worker through higher wages either within the job if the outside

[1] From Rogerson and Shimer (2011), the Current Population Survey implies an annual employment to employment rate of about 0.31 for the period 2000-05. The corresponding employment to unemployment hazard is 0.24. Finally, the unemployed job finding rate is 4.3 at annual frequency. These numbers are based on a competing hazards interpretation of the monthly transition probabilities. As documented in Fujita and Moscarini (2013), the numbers over-estimate reallocation to new employers through the unemployment channel because a large part of the outflow from unemployment are recalls.

[2] Recent notable exceptions are Fujita and Moscarini (2013) and Carrillo-Tudela and Smith (2014).

[3] The one exception to this result is that wages with future employers after an unemployment spell are unaffected by specific training. But even that exception is eliminated in settings with for example recall, such as Fujita and Moscarini (2013) and Carrillo-Tudela and Smith (2014).

employer is not capable of matching the current employer's willingness to pay, or in a new match if it can.

Specific training increases the current employer's valuation of the worker, only. But due to productive heterogeneity across firms, the match is not insulated from increased competitive pressure. More productive outside firms can challenge the current firm's increased valuation of the worker. Thus, similar to the case of general training, the occasional meeting with more productive outside employers gradually delivers the valuation increase to the worker either within the job if the outside employer's valuation of the worker does not match that of the current employer, or through a better bargaining position in a job with the new employer. Thus, specific training delivers higher wages with future employers through a better bargaining position, even though the training has no productive impact on the future match.[4] Furthermore, the competitive pressure from more productive employers forces an increasing wage profile within the job as well.[5]

Frictional competition implies that the rents from productive gains caused by either general or specific training are partially delivered to the worker ex post training. In line with Becker (1964), firms therefore make workers "pay" for training up front through lower wages during training, but the result now extends to specific training, as well. The analysis imposes a hold-up problem on training through the introduction of worker risk aversion which makes backloaded wage profiles costly. The hold-up problem reduces training intensity, and again, the effect applies to both general and specific training.

Firm heterogeneity allows the analysis to consider variation in training across firms and the possible impact of mismatch on training. Training can vary across firms purely due to production function complementarities between firm productivity and worker human capital. In the risk neutral case where the hold up problem is absent from the analysis, training is increasing (decreasing) in the firm type if and only if the production function is supermodular (sub-modular). This applies both to general and specific training. The hold-up

[4]The model also implies that specific training has a positive selection effect on future employers; conditional on moving, specific training increases the expected productivity type of the future employer. This is a well known concern that confounded the analyses in Altonji and Shakotko (1987) and Topel (1991). The effect is distinct from the increased bargaining position whereby wages increase conditional on employer type.

[5]Thus, in contrast to Felli and Harris (1996) we obtain the result that purely specific training implies increasing wages within the job.

problem by itself implies that training intensity is increasing in the generosity of a given firm's employment contract, and in our calibration to the U.S. economy, training is also on average increasing in firm type.

That training varies across firms and in generosity cause us to reconsider the conclusions in Acemoglu and Pischke (1999) regarding the relationship between training and frictions. They argue that labor market friction relieves the hold-up problem associated with general training through a reduction in the upward wage pressure ex post training. Therefore, in a second best setting subject to a hold-up problem, increased labor market friction results in more general training. Our analysis includes and validates this mechanism as a partial equilibrium result, which now pertains both to specific and general training. In particular, for a given employer type and generosity, an employment contract will provide more training as labor market friction increases. However, increased labor market friction implies that in equilibrium, workers have worse bargaining positions and they are matched with worse firms. Given the imposed hold-up problem, training is decreasing in both of these dimensions. In the calibrated model, the equilibrium effects turn out to dominate. And we reach the opposite conclusion; that training is decreasing in the extent of labor market friction.

The worker's human capital accumulation is employment history dependent. This is not an unusual feature in models with learning by doing or with human capital depreciation during unemployment. But in these cases, it is typically time spent in unemployment that matters. In our setting, variation in human capital accumulation across workers is tied to the degree of mismatch the worker has experienced throughout her career. Thus, there is a positive association between favorable employment history realizations and human capital accumulation.[6] A wage variance decomposition that measures the contributions from human capital variation as well as labor market frictions must in this case account for the positive covariance between the two contributions.[7] It also implies a positive sorting effect where workers matched with better firms end up with more human capital, ex post. In our cali-

[6] Recent results in Guvenen et al. (2015) find that wages are decreasing in the accumulation of mismatch in the worker's employment history, which is consistent with the mechanisms in our model with training subject to hold-up and/or a super-modular production function.

[7] An example of such a decomposition that accounts for a positive correlation between the two can be found in Bagger and Lentz (2013). They emphasize that a positive correlation between favorable match outcomes and worker human capital can arise through the reverse causality: that higher human capital workers lose relatively more by mismatch and therefore search more intensely for favorable matches.

brated model, the positive correlation between worker human capital and firm productivity turns out to be quite strong despite the absence of assortative matching.

General and specific training have sharply different implications for mobility, and they also differ in terms of social efficiency. General training does not impact future worker mobility, at all: If the worker meets a more productive firm, the worker will move to it, regardless. However, specific training implies that the current match moves up the "willingness to pay for the worker" hierarchy. Therefore a worker may forego the opportunity to move to a more productive firms and only improve his bargaining position with his current employer. Specifically-trained workers have lower job-to-job hazard, everything else equal.

The two kinds of training differ in terms of social efficiency as well. In the risk neutral case (that is, in the absence of the hold-up problem), general training is socially efficient. However, specific training is too high and the problem is greater for lower productivity firms. This is a feature of the wage determination process whereby a future firm always perfectly compensates the current match for its destruction in case of a job-to-job move. The social planner understands that a low productivity match is more likely to be destroyed in favor of a better match and consequently invests less in specific training. But privately, the match does not internalize the loss of specific capital associated with the worker moving since the future firm is fully compensated for the loss. The analysis includes consideration of side payment schemes that can resolve the inefficiency, although these schemes might be difficult to implement in practice.

Section 2 contains the model description and the analysis. The general optimal contract is analyzed in Section 2.3, and the risk neutral case is studied in Section 3 with a particular focus on efficiency. Section 4 calibrates the model to the U.S. economy. Section 5 presents the calibrated steady state equilibrium. In Section 6 we discuss the counter factual of changing the level of frictions in the economy and relate the result to Acemoglu and Pischke (1999). Section 7 concludes.

2 Model

There is a unit measure of workers who can be either employed or unemployed. Matches between workers and firms are formed through a frictional search process. They produce an output that generates a revenue stream. And, the firm can invest in two kinds of human capital: General and specific.

Competition over workers takes place as in Lentz (2014). If a firm meets an unemployed worker, it can make the worker a take-it-or-leave-it lifetime utility offer. The contract specifies a response to the event that the worker meets an outside firm. The worker will accept employment with the outside firm if it can match the worker's current contract. There is limited commitment on the part of both firms and workers: Each can without cost leave the relationship for their respective outside options. And, the contract is required to be renegotiation proof: In any state of nature, the contract must not be Pareto dominated by another feasible contract. Consequently, contracts respond to outside meetings in a way identical to outside offer matching.[8] If the worker meets an outside firm that is willing to improve on the worker's current contract terms, the worker's current contract responds by a continuation utility promise equal to the minimum of the two firms' willingness to pay for the worker.

A firm designs an employment contract to maximize its profits subject to the utility promise it has made to the worker. An employment relationship delivers utility to a worker in three ways: (1) The contract promises a wage payment stream as a function of the employment history, which includes arrivals of outside meetings. (2) The contract promises employment history contingent general and specific human capital training rates. (3) The contract may facilitate rent extraction with future employers.

Denote by h the worker's general human capital level, which we will also refer to as the worker's skill. Let m be the match specific capital level. Match specific capital affects only the productivity of the current match, whereas general human capital applies to all matches. While we refer to the accumulation of match specific capital as training, it accommodates a broader interpretation of investments in the surplus of the current match, including intangibles like goodwill.[9]

Skills have two support points: Skilled ($h = 1$) and unskilled ($h = 0$). And, for match specific capital, $m \in \{0, 1\}$. All workers are born into the labor market unskilled and unemployed. Human capital does not depreciate. It is straightforward to relax these assumptions, but we maintain them throughout the paper for the sake of exposition and to ease the nu-

[8] For details, see Lentz (2014).

[9] Many of the paper's conclusions regarding the incentives to accumulate match specific capital apply to a setup where the parties are learning about the productivity of the match.

merical solution of the model. Let $f_{hm}(p)$ be the output of a match between a productivity p firm and a skill h worker with match specific capital m. It is strictly increasing in h, m, and p.

The firm's training decision is modeled as a choice that controls the stochastic process of the worker's human capital evolution. The firm picks the Poisson arrival rate η for which the unskilled worker becomes skilled, and the Poisson rate μ by which a $m = 0$ match transitions to $m = 1$. All firms have access to the same training technology, which is reflected in the monetary general training cost $c_h(\eta)$ and specific training cost $c_m(\mu)$. Both cost functions are increasing and convex.

As emphasized by Becker (1964), privately efficient training decisions can be made if the contract can make the worker "pay" for training up front through low initial wages. Reduced training may result if it is costly to implement such a contract, and Acemoglu and Pischke (1999) argue that in such a case labor market frictions help reduce the hold up problem in general training. In our model, making the worker pay for training up front is made costly through the assumption of risk aversion in worker preferences, and the degree of risk aversion affects the magnitude of the cost.

2.1 Worker lifetime utility

Time is continuous and both firms and workers discount time at rate ρ. Workers die at rate d. Workers are risk averse and consume whatever income they have at a given instant which delivers utility, $u(w)$, where w is the wage level. An unemployed worker with human capital h receives benefits b_h. All employed workers are laid off at exogenous layoff rate δ. Unemployed and employed workers meet employment opportunities at rates λ_u and λ_e, respectively.

A firm's willingness to pay $\bar{V}_{hm}(p)$ as a function of its productivity, worker skill h, and specific capital m is defined as the utility promise such that the firm's discounted stream of future profits from the match is exactly zero. It is a result in the analysis that this object is monotonically increasing in p. Define the inverse function $p_{hm}(V)$ by $V = \bar{V}_{hm}(p_{hm}(V))$. Let $F_h(V)$ be the distribution of willingness to pay levels in the vacancy offer distribution for skill h workers and $m = 0$. Furthermore, use the shorthand $\hat{F}_h(V) = 1 - F_h(V)$.

Denote by $O\left(\bar{V}', V|\bar{V}\right)$ an employment contract's offer response to the worker meeting an outside firm that has willingness to pay $\bar{V}'$, conditional on the firm's own willingness to pay $\bar{V}$ and the contract currently delivering V. The optimal renegotiation proof strategy for the firm is,

$$O\left(\bar{V}', V|\bar{V}\right) = \max\left[V, \min\left[\bar{V}', \bar{V}\right]\right]. \tag{1}$$

Thus, the contract matches the outside firm's willingness to pay up to the firm's own willingness to pay. In the case where the outside firm has willingness to pay less than V, the contract continues unaffected.[10] Alternatively, if outside meetings were contractable, the results in Section 2.3 show that risk aversion implies that it is indeed optimal to keep the utility promise constant in this case. Whenever $V \leq \bar{V}' \leq \bar{V}$, the firm sets the continuation utility promise exactly sufficient to make the worker stay, and no more. For $\bar{V}' > \bar{V}$, it is efficient for the worker to move to the outside employer. The firm would prefer to be able to commit to $O\left(\bar{V}', V \mid \bar{V}\right) = \bar{V}'$ so as to allow for maximal rent extraction from the future employer. However, the firm's participation constraint, imposing that the discounted profit value of the contract cannot be negative, and renegotiation proofness constrain the firm's continuation utility promise to $O\left(\bar{V}', V \mid \bar{V}\right) = \bar{V}$.

Consider a skill h worker currently employed at a productivity p firm and match specific capital m. Define the worker's effective discount rate $r = \rho + d$. The lifetime utility of the contract can be expressed recursively,

$$\begin{aligned}
(r + \delta + \eta + \mu) V &= u(w) + \eta H + \mu M + \delta U_h + \dot{V} + \\
&\quad + \lambda_e \left[\int_V^{\bar{V}_{hm}(p)} (V' - V)\, dF_h(V') + \hat{F}_h\left(\bar{V}_{hm}(p)\right)\left(\bar{V}_{hm}(p) - V\right)\right] \\
&= u(w) + \eta H + \mu M + \delta U_h + \dot{V} + \lambda_e \int_V^{\bar{V}_{hm}(p)} \hat{F}_h(V')\, dV',
\end{aligned} \tag{2}$$

where the second line follows by integration by parts and the contract dictates the collection $\left(w, \dot{V}, H, M, \eta, \mu\right)$. The worker receives flow utility $u(w)$ from the wage rate w. At Poisson rate η, the worker's human capital increases in which case the contract continues with utility promise H. Similarly, at rate μ match specific capital increases and the contract continues with utility promise M. At rate λ_e the worker meets an outside firm and the continuation

[10] This is an immediate implication of a setting where the worker can choose to keep outside meetings hidden. See Lentz (2014).

utility promise of the contract is specified by $O\left(\bar{V}', V|\bar{V}\right)$ in equation (1). H and M are of course only relevant if $h = 0$ and $m = 0$, respectively. Since only in these cases will training rates η and μ be positive. Finally, the value of unemployment U_h is given by,

$$rU_h = u\left(b_h\right).$$

The worker meets firms at rate λ_u. However, regardless of the type of the meeting, the associated employment contract delivers utility promise U_h. Indeed, when a firm meets an unemployed worker, it make her a take-it-or-leave-it lifetime utility offer, which she accept as long as it exceeds the value of unemployment.

2.2 Firm profits

Let $\Pi_{hm}(V, p)$ be the optimal profit value to a firm with productivity p in an (h, m) state match subject to a utility promise V. Following Spear and Srivastava (1987) and Thomas and Worrall (1988), the firm's valuation of the discounted stream of profits associated with a contract can be written recursively by,

$$\begin{aligned}(r+\delta)\,\Pi_{hm}(V,p) = \max_{\left(w,\dot{V},\eta,\mu,H,M\right)\in\Gamma_{hm}(V,p)} \Bigg\{ & f_{hm}(p) - w - c_h(\eta) - c_m(\mu) \\ & + \lambda_e \int_V^{\bar{V}_{hm}(p)} \Pi'_{hm}(V',p)\,\hat{F}_h(V')\,dV' + \eta\left[\Pi_{1m}(H,p) - \Pi_{hm}(V,p)\right] \\ & + \mu\left[\Pi_{h1}(M,p) - \Pi_{hm}(V,p)\right] + \Pi'_{hm}(V,p)\,\dot{V} \Bigg\}, \end{aligned} \quad (3)$$

where $\Pi'_{hm}(V,p) = d\Pi_{hm}(V,p)/dV$. The set of feasible contract design choices is given by,

$$\begin{aligned}\Gamma_{hm}(V,p) = \Bigg\{ & \left(w, \dot{V}, \eta, \mu, H, M\right) \Big| \\ & u(w) + \eta H + \mu M + \delta U_h + \dot{V} + \lambda_e \int_V^{\bar{V}_{hm}(p)} \hat{F}_h(V')\,dV' = (r+\delta+\eta+\mu)\,V \\ & U_h \le M \le \bar{V}_{h1}(p) \\ & U_1 \le H \le \bar{V}_{1m}(p) \Bigg\}. \end{aligned}$$

The last two set of constraints reflect the participation constraints of the worker and firm both. The worker's valuation of the contract cannot be less than the value of unemployment and the discounted profit value of the contract cannot be negative.

It is assumed that production technology is such that the willingness to pay of a firm is never strictly below the value of unemployment for a given human capital level. The first constraint is the promise keeping constraint. The willingness to pay for a given human capital level is given by,

$$\Pi_{hm}\left(\bar{V}_{hm}(p), p\right) = 0. \tag{4}$$

2.3 Optimal contract design

Let the Lagrange multiplier on the promise keeping constraint be $\gamma_{hm}(V,p)$, where $\gamma_{hm}(V,p) > 0$ is a sufficient condition for the recursive formulation of the contracting problem to be valid. Denote by $\varphi_m(V,p)$ the Lagrange multiplier on the worker's participation constraint, $U_{1m} \leq H$. It is verified that the other constraints are not binding for the optimal contract. Since unemployment benefits depends on general human capital, the worker participation constraint might be binding when the worker becomes generally skilled. Since an increase in match specific capital involves an increase in joint match value and neither the worker's or firm's outside options are affected, the participation constraints will not bind in the case when the worker becomes specifically skilled. In the following, we discuss the properties of the optimal employment contract with a given productivity firm.

2.3.1 Profit function and willingness to pay

The profit function is decreasing and strictly concave. Lentz (2014) provides proof of concavity in the more general setting with hidden search (absent training decisions). It is also straightforward to show that the firm's willingness to pay $\bar{V}_{hm}(p)$ is monotonically increasing in all three arguments, which is a direct implication of the monotonicity of the production function.

2.3.2 Wages

In the absence of minimum wages or other constraints on the wage design, the slope of the profit function satisfies,

$$\Pi'_{hm}(V,p) = -(r+\delta)\,\gamma_{hm}(V,p) = \frac{-1}{u'\left(w_{hm}(V,p)\right)}, \tag{5}$$

which follows from the first order conditions on the choices of w and $\dot{V}$. By implication, wages $w_{hm}(V,p)$ are strictly increasing in the utility promise given concavity of the profit function.

2.3.3 Tenure conditional utility promise and wage paths.

By the derivative of the Lagrangian $\partial\mathcal{L}/\partial V = \Pi'(V)$, and the envelope theorem, one obtains,

$$\Pi'_{hm}(V,p) + (r+\delta)\gamma_{hm}(V,p) = \frac{\Pi''_{hm}(V,p)}{r+\delta+\lambda\hat{F}_h(V)+\eta_m(V,p)+\mu_h(V,p)}\dot{V}_{hm}(V,p). \tag{6}$$

Together with equation (5), it must be that in the absence of outside offers and skill increases, the optimal employment contract is flat,

$$\dot{V}_{hm}(V,p) = 0. \tag{7}$$

Thus, in the absence of outside offer arrivals or changes in human capital, the contract's wage profile is flat in tenure.[11] In the limit case where the worker is risk neutral, the flat contract remains optimal, but there is now a multitude of optimal paths. The analysis uses the flat contract in the limit case where the worker is risk neutral.

Unconditionally, the employment contract involves expected changes in the utility promise over job duration through two channels. The contract matches the willingness to pay of firms that the worker meets, which in isolation implies an expected increasing utility promise path in duration. For a given utility promise V, the expected growth rate in the utility promise within the job due to on-the-job search is $\lambda_e \int_V^{\bar{V}_{hm}(p)} (V'-V)\,dF_h(V') \geq 0$. By wages increasing in the utility promise, on-the-job search in isolation implies an increasing wage path in tenure. This is a simple replication of the offer matching process in Postel-Vinay and Robin (2002). We include the argument here for completeness.

Furthermore, the continuation utility promise may change when the worker becomes generally skilled or specifically skilled. The following lemma characterizes the optimal contract's response to human capital changes.

[11] As documented in Lentz (2014), the optimal contract will itself be backloaded ($\dot{V} > 0$) in the case where the worker can engage in hidden search in order to increase the chance of an outside meeting.

Lemma 1. *The optimal contract is for any $p \in [0,1]$ and $V \in \left[U_h, \bar{V}_{hm}(p)\right]$ characterized by*

1. $V < H_m(V,p) < \bar{V}_{1m}(p)$ *and* $V \leq M_h(V,p) < \bar{V}_{h1}(p)$ *with strict inequality if* $p < 1$.

2. *Wages are smooth across human capital capital increases unless the participation constraint is binding:* $w_{h0}(V,p) = w_{h1}(M_h(V,p),p)$ *and* $w_{0m}(V,p) = w_{1m}(H_m(V,p),p)$ *for* $\varphi_m(V,p) = 0$. *If* $\varphi_m(V,p) > 0$ *then* $H_m(V,p) = U_1$ *and* $w_{0m}(V,p) < w_{1m}(H_m(V,p),p)$.

Proof. The human capital change conditional utility promises satisfy the first order equations,

$$\Pi'_{1m}(H_m(V,p),p) - \Pi'_{0m}(V,p) = \frac{-(r+\delta)\varphi_m(V,p)}{\eta_m(V,p)} \tag{8}$$

$$\Pi'_{h1}(M_h(V,p),p) - \Pi'_{h0}(V,p) = 0. \tag{9}$$

If the worker participation constraint is not binding, $\varphi_{hm}(V,p) = 0$, the wage profile is flat over human capital jumps,

$$\Pi'_{0m}(V,p) = \Pi'_{1m}(H_m(V,p),p)$$
$$\Pi'_{h0}(V,p) = \Pi'_{h1}(M_h(V,p),p),$$

which implies,

$$w_{0m}(V,p) = w_{1m}(H_m(V,p),p)$$
$$w_{h0}(V,p) = w_{h1}(M_h(V,p),p).$$

If the worker's participation constraint in case of a skill increase is binding, wages jump up in case of a skill increase. This follows directly from the concavity of the profit function and that wages are increasing in the utility promise. The participation constraint forces the firm to offer a greater utility promise than the one that makes wages smooth across the skill jump. By concavity the binding participation implies, $\Pi'_{0m}(V,p) > \Pi'_{1m}(H_m(V,p),p)$ and therefore $w_{0m}(V,p) < w_{1m}(H_m(V,p),p)$.

Now, consider the claim that $V < M_h(V,p) < \bar{V}_{h1}(p)$ for $p < 1$. Proof is by contradiction. Suppose first that $M_h(V,p) = \bar{V}_{h1}(p)$. For the sake of simplicity, take the case where $h = 1$. Trivially, it must be that $w_{11}\left(\bar{V}_{11}(p),p\right) = f_{11}(p)$ since there is no possibility of future wage gains within the contract. It must then be that $w_{10}(V,p) \leq w_{10}\left(\bar{V}_{10}(p),p\right) \leq$

$f_{10}(p) - c_m\left(\mu\left(\bar{V}_{10}(p), p\right)\right) < f_{11}(p)$. This is because at $M_1(V,p) = \bar{V}_{11}(p)$ the firm hands over all gains to specific training to the worker. Hence, $\Pi_{10}\left(\bar{V}_{10}(p), p\right) = 0$ implies that wages w_{10} cannot exceed production less training cost. Thus, $M_h(V,p) = \bar{V}_{h1}(p)$ implies that $w_{10}(V,p) < w_{11}(M_h(V,p), p)$, violating (9). Suppose instead, also by contradiction that $M_1(V,p) \leq V$. By the utility promise constraint we have that,

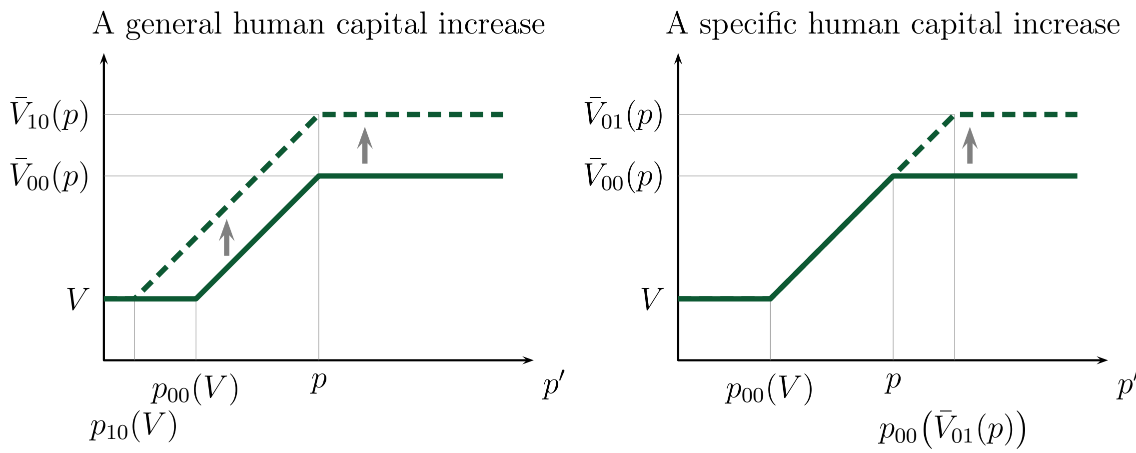
$$\begin{aligned}
(r+\delta) M_1(V,p) &= u\left(w_{11}\left(M_1(V,p), p\right)\right) + \delta U_1 + \lambda_e \int_{M_1(V,p)}^{\bar{V}_{11}(p)} \hat{F}_1(V')\, dV' \\
&= u\left(w_{10}(V,p)\right) + \delta U_1 + \lambda_e \int_{M_1(V,p)}^{\bar{V}_{11}(p)} \hat{F}_1(V')\, dV' \\
&> u\left(w_{10}(V,p)\right) + \mu(V,p)\left[M_1(V,p) - V\right] + \delta U_1 + \lambda_e \int_{V}^{\bar{V}_{10}(p)} \hat{F}_1(V')\, dV' \\
&= (r+\delta) V.
\end{aligned}$$

The second equality follows from (9). The inequality follows directly from the presumption that $M_1(V,p) \leq V$ and that $\bar{V}_{11}(p) > \bar{V}_{10}(p)$. Therefore $M_1(V,p) \leq V$ is contradicted. The basic intuition is that since wages are smooth across the human capital change, a utility promise $M_1(V,p) \leq V$ implies greater future utility promise growth than prior to the human capital increase. At an unchanged current wage level, a greater future utility promise growth is inconsistent with a reduction in the utility promise. Hence it must be that $V < M_h(V,p) < \bar{V}_{h1}(p)$. The $p = 1$ case is the exception. In this case $M_1(V,1) = V$. The reason being that $\bar{V}_{10}(1)$ is the upper bound on the support of $F_1(V)$. The fact that the firm's willingness to pay increases from $\bar{V}_{10}(1)$ to $\bar{V}_{11}(1)$ does not result in an increase in the worker's expected utility promise growth rate for any given utility promise, because there are no outside firms to challenge the increase.

Arguments for $h = 0$ as well as the skill increase conditional utility promise $V < H_m(V,p) < \bar{V}_{1m}(p)$ go along the same lines. □

Lemma 1 states that gains from human capital growth are strictly shared between firm and worker. In the case of general human capital the worker will not manage to extract all of the gains, confirming the arguments in Acemoglu and Pischke (1999) that frictions allows the firm to extract some of the rents due to training. The more novel part of Lemma 1 is that it turns out that the same is true for gains from specific training: The frictional competition environment with heterogenous firms forces the firm to hand over part of the

Figure 1: Expected Gains from On-the-Job Search



gains from specific training to the worker. Our setup provides a particular formalization of Becker's (1964) argument that competition can induce the firm to share the ex post returns to even perfectly specific training.

An increase in either general or specific capital is for a given utility promise associated with an increase in the worker's expected net utility gains from outside meetings. However, the exact source and magnitude of the gains depend on whether the capital gain is general or specific. Consider an $(h, m) = (0, 0)$ worker with a current utility promise of V who is employed with a firm that has willingness to pay $\bar{V}_{00}(p)$. Using the productivity distribution in the vacancy pool, $\Phi(p)$, define the worker's expected utility growth rate from on-the-job search,

$$\Omega_{00}(V, p) = \lambda_e (1 - \Phi(p)) \left[\bar{V}_{00}(p) - V\right] + \lambda_e \int_{p_{00}(V)}^{p} \left(\bar{V}_{00}(p') - V\right) d\Phi(p') . \tag{10}$$

The first term on the right hand side of equation (10) reflects job-to-job transitions from the current firm to an outside firm where the receiving firm will deliver an employment contract with utility promise $\bar{V}_{00}(p)$. The second term reflects how competition forces transfers from the firm to the worker within the current match. If the worker meets a firm with productivity $p' \in [p_{00}(V), p]$, the meeting results in the worker staying in the current match with an increased utility promise of $\bar{V}_{00}(p') > V$.

Figure 1 illustrates the increased expected gains from on-the-job search. The figures are stylized in that one would not generally expect $\bar{V}_{hm}(p)$ to be linear. Suppose the worker becomes generally-skilled. This event is illustrated in the left hand panel. Holding the current utility promise fixed at V, the increased competitive pressure on the match is reflected in the now greater expected utility growth rate from on-the-job search,

$$\Omega_{10}(V,p) = \lambda_e (1-\Phi(p)) \left[\bar{V}_{10}(p) - V\right] + \lambda_e \int_{p_{10}(V)}^{p} \left(\bar{V}_{10}(p') - V\right) d\Phi(p').$$

All outside firms are now willing to pay more for the worker, $\bar{V}_{10}(p') > \bar{V}_{00}(p')$. In addition, the support of firm types that can impose competitive pressure on the match expands downward from $p_{00}(V)$ to the lower firm type $p_{10}(V)$. Should the worker move, she will move with a higher utility promise, $\bar{V}_{10}(p)$.

Consider alternatively an increase in specific capital from $m = 0$ to $m = 1$. This is illustrated in the right hand panel. Holding the utility promise fixed at V, the competitive pressure on the match increases to

$$\Omega_{01}(V,p) = \lambda_e \left(1 - \Phi\left(p_{00}\left(\bar{V}_{01}(p)\right)\right)\right) \left[\bar{V}_{01}(p) - V\right] + \lambda_e \int_{p_{00}(V)}^{p_{00}\left(\bar{V}_{01}(p)\right)} \left(\bar{V}_{00}(p') - V\right) d\Phi(p').$$

In this case, the incumbent firm's willingness to pay for the worker increases to $\bar{V}_{01}(p) > V_{00}(p)$. Outside firms do not change their willingness to pay for the worker, but the upper bound on the set of firms that will force up the utility promise within the job increases to $p_{00}\left(\bar{V}_{01}(p)\right) > p$. Furthermore, when the worker moves to a better firm, she moves with a greater utility promise of $\bar{V}_{01}(p)$.

Thus, both in the case of a general and a specific capital increase, if the firm were to hold the worker's utility promise constant at V, the faster expected growth rate in future utility promises would have to imply a drop in the current wage. By worker risk aversion, this is suboptimal. Therefore, the optimal contract has an increase in the utility promise as either general or specific capital increases.

2.3.4 Job-to-job mobility, tenure and wages

Both Altonji and Shakotko (1987) and Topel (1991) emphasize that tenure effects in wages may be associated with a selection effect on the type of future firms which complicates the distinction between experience and tenure effects in their analyses. Our analysis exhibits exactly this effect in the case of specific human capital accumulation. As specific capital increases, the firm type threshold such that the worker is indifferent between moving to it and staying with the current firm goes up. Thus, conditional on moving, the expected firm type of the new firm increases as specific capital goes up, and consequently increased specific capital will have a positive wage impact beyond the current match through this selection effect.

In addition, our analysis contains another important channel through which specific training will result in higher wages with future firms: Even though specific capital is not portable between firms, bargaining position carries over. Specific training increases the willingness to pay of the worker's current employer, which means that conditional on moving, the worker will do so with a greater utility promise with the new firm. Hence, even conditional on the type of the future employer, specific training in the current firm raises wages with future employers - this despite the fact that the willingness to pay of the future employer is unchanged.

Thus, specific training raises wages with future employers and within the current match. General training does as well.

2.3.5 Training rates

The first order conditions for the two training rates are given by,

$$[\Pi_{1m}(H_m(V,p),p) - \Pi_{0m}(V,p)] - \Pi'_{0m}(V,p)[H_m(V,p) - V] = c'_h(\eta_m(V,p)) \tag{11}$$

$$[\Pi_{h1}(M_h(V,p),p) - \Pi_{h0}(V,p)] - \Pi'_{h0}(V,p)[M_h(V,p) - V] = c'_m(\mu_h(V,p)), \tag{12}$$

where the capital change conditional utility promises satisfy equations (8) and (9).

The first order conditions on training state that the marginal cost of training must equal the marginal profit gain from the increase in either general or specific human capital. The first bracketed term on the right hand sides of the first order conditions (11) and (12) is the direct jump in profits due to the capital increase. The second term reflects the profit value of the change in worker's utility promise, where by equation (5), $\Pi'_{hm}(V,p) = -1/u'(w_{hm}(V,p))$ is the profit impact of a one unit increase in the utility promise.

Increases in the human capital change conditional utility promises, $H_m(V,p)$ and $M_h(V,p)$, reduce the direct profit gains from training, but the loss is compensated by the worker's greater expected utility gains, which are translated into current profits through reduced wages today. In a risk neutral setting these two effects exactly offset each other and the training decisions are unaffected by the particular choice of $H_m(V,p)$ and $M_h(V,p)$. Thus, Becker's (1964) insight that even though a perfectly competitive environment dictates that the firm has to deliver all match surplus to the skilled worker, $H_m(V,p) = \bar{V}_1(p)$, the training choice remains privately efficient since the firm is perfectly compensated via lower wages during the training period.

When the worker is risk averse, future utility promises can no longer be translated into profits one-to-one through a lowering of current wages. Therefore, training comes to depend on both the current utility promise V and the contract's optimal choice of $H_m(V,p)$ and $M_h(V,p)$. It follows by differentiation of the first order conditions (11) and (12) that,

$$\frac{\partial \mu_h(V,p)}{\partial V} = \frac{-\Pi''_{h0}(V,p)\left[M_h(V,p) - V\right]}{c''_\mu(\mu_h(V,p))}$$
$$\frac{\partial \eta_m(V,p)}{\partial V} = \frac{-\Pi''_{0m}(V,p)\left[H_m(V,p) - V\right]}{c''_\eta(\eta_m(V,p))}.$$

By Lemma 1, both specific and general training are increasing in the utility promise iff the profit function is strictly concave in V, which is the case given $u(\cdot)$ strictly concave. We discuss variation in training across firm types in detail in Section 4.

Risk aversion imposes the classic hold-up problem on the training choice: Increased human capital (both general and specific) implies increased competitive pressure on the match and consequently greater rents to the worker. The firm will want to reduce current wages to capture the ex post rents flowing to the worker. However, risk aversion imposes a cost on this mechanism. The concavity of the profit function is a reflection that the hold-up problem is more severe for low utility promises because the contract is operating on a higher curvature part of the worker's utility function.

2.4 Steady state

Denote by e_{hm} the mass of employment of general skill h workers in jobs with match specific capital, m. Let u_h be the mass of unemployed general skill h workers. Normalize the population at unity, $1 = \sum_h (u_h + \sum_m e_{hm})$. Furthermore, denote by $G_{hm}(V,p)$ the cumulative distribution of match states for type (h,m) matches, where by definition $G_{hm}\left(\bar{V}_{hm}(1), 1\right) = 1$. The steady state conditions on the employment and unemployment stocks follow the simple logic that the flow into the stock must equal the flow out.

The steady state condition on $e_{00}G_{00}(V,p)$ is given by,

$$\lambda_u u_0 \Phi(p) + \lambda e_{01} \int_0^{\bar{p}_{01}(V)} \int_{U_0}^{\bar{V}_{01}(p')} \left[F_0\left(\bar{V}_{00}(p)\right) - F_0\left(\bar{V}_{01}(p')\right)\right] g_{01}(V',p')\, dV'dp' =$$

$$e_{00}\Bigg\{ \int_0^{\bar{p}_{00}(V)} \int_U^{\bar{V}_{00}(p')} \left[d + \delta + \eta_0(V',p') + \mu_0(V',p') + \lambda \hat{F}_0\left(\bar{V}_{00}(p)\right)\right] g_{00}(V',p')\, dV'dp' +$$

$$\int_{\bar{p}_{00}(V)}^{p} \int_U^{V} \left[d + \delta + \eta_0(V',p') + \mu_0(V',p') + \lambda \hat{F}_0(V)\right] g_{00}(V',p')\, dV'dp' \Bigg\}.$$

The first term on the left hand side is the flow into the $e_{00}G_{00}(V,p)$ pool from unemployment. The second term is the flow in from the pool of matches with high match specific capital where the worker nevertheless receives a better offer and consequently moves into low match specific capital. The integral is over types of matches with high match specific capital. The outer integral is over firms that have willingness to pay less than V. Any firm with a willingness to pay more than V may be beat, but the worker would move into the e_{00} pool with a utility promise greater than V. The inner integral is then all the possible utility promises that workers may have in these firms. The term $\left[F_0\left(\bar{V}_{00}(p)\right) - F_0\left(\bar{V}_{01}(p')\right)\right]$ is the probability that a worker in a type p' firm will receive an offer that is better than her current firm's willingness to pay, but is from a type firm less than p. If that happens, the worker moves into the $e_{00}G_{00}(V,p)$ pool. The terms on the right hand side are standard: The worker leaves the pool upon death, unemployment, general and specific skill acquisition, and if the worker receives an outside offer that takes her out of the pool. The latter can happen in two ways: If a worker is currently employed with a firm that has willingness to pay less than V then an outside offer must be from a firm better than p to make her leave the pool. If she is with a firm with willingness to pay greater than V, then it is sufficient that the outside offer be better than V.

The steady state conditions on $e_{01}G_{01}(V,p)$, $e_{10}G_{10}(V,p)$, and $e_{11}G_{11}(V,p)$ follow the same type of argument and are given in Appendix A.

3 Risk neutral case

It is illustrative to consider the special case of risk neutrality. We obtain two new analytical results. First, we find that in the risk neutral case, variation in training across firms are only driven by technological properties, and in particular complementarities in the production functions. Second, we find that specific training tends to be inefficiently high in low type firms.

3.1 Training decisions

Assume $u'' = 0$ and without loss of generality transform the utility function so that $u'(w) = 1$. The profit function takes the form $\Pi_{hm}(V, p) = \bar{V}_{hm}(p) - V$. The first order conditions for the optimal contract's training rates are then by equations (11) and (12),

$$\begin{aligned} c_h'(\eta_m(V, p)) &= \bar{V}_{1m}(p) - \bar{V}_{0m}(p) \\ c_m'(\mu_h(V, p)) &= \bar{V}_{h1}(p) - \bar{V}_{h0}(p). \end{aligned}$$

It is immediately seen that the training rates do not depend on the particular utility promise in the contract. The risk neutral case eliminates the hold-up problem from the analysis, and in particular the variation of the severity of the problem as a function of the utility promise.

The firm's willingness to pay solves,

$$\begin{aligned} (r+\delta)\bar{V}_{hm}(p) = f_{hm}(p) + \delta U_h + (1-h)\left[\eta_m(p)\left[\bar{V}_{1m}(p) - \bar{V}_{hm}(p)\right] - c_h(\eta_m(p))\right] \\ + (1-m)\left[\mu_h(p)\left[\bar{V}_{h1}(p) - \bar{V}_{hm}(p)\right] - c_m(\mu_h(p))\right], \end{aligned}$$

where the dependency of the training rates on V has been eliminated. By differentiation it follows that,

$$\begin{aligned} \eta_1'(p) &= \frac{f_{11}'(p) - f_{01}'(p)}{[r+\delta+\eta_1(p)]\, c_h''(\eta_1(p))} \\ \mu_1'(p) &= \frac{f_{11}'(p) - f_{10}'(p)}{[r+\delta+\mu_1(p)]\, c_m''(\mu_1(p))}. \end{aligned}$$

The expressions for $\eta_0'(p)$ and $\mu_0'(p)$ account for possible complementarities between general and specific training both direct and through firm productivity,

$$\begin{aligned} c_h''(\eta_0(p))\,\eta_0'(p) &= \frac{f_{10}'(p) - f_{00}'(p) + \mu_1(p)\, c_m''(\mu_1(p))\,\mu_1'(p) - \mu_0(p)\, c_m''(\mu_0(p))\,\mu_0'(p)}{r+\delta+\eta_0(p)} \\ c_m''(\mu_0(p))\,\mu_0'(p) &= \frac{f_{01}'(p) - f_{00}'(p) + \eta_1(p)\, c_h''(\eta_1(p))\,\eta_1'(p) - \eta_0(p)\, c_h''(\eta_0(p))\,\eta_0'(p)}{r+\delta+\mu_0(p)}. \end{aligned}$$

In the case of a modular production function one immediately obtains that, $\eta_m'(p) = \mu_h'(p) = 0$. Thus, in the risk neutral case, if the production function does not have complementarities between firm productivity and training, then training is constant across firm types. Competitive pressure varies across firms, but whatever the share of ex post gains to training it delivers to the worker, the firm can translate it into profits through lower wages at the time of training without any efficiency loss. Specifically, notice that the meeting rates λ_u and λ_e do not affect $\bar{V}_{hm}(p)$ and therefore do not impact the training levels.

In the risk neutral case, variation in training across firms is driven purely by complementarities in production between firm productivity and training. Positive complementarities imply that training is increasing in firm type. Negative complementarities induce a negative relationship between training and firm type.

Finally, the offer matching characteristic of the wage determination process implies that competitive pressure (λ_e) does not affect training in the risk neutral case: Whenever the worker moves to another firm, that firm fully compensates the destruction of the previous match. In particular, this includes the value of human capital investments that were made in the match. Hence, the probability that a worker will move to another firm does not impact the discount rate on human capital investments.

This result highlights the importance of the wage determination process as to whether the firm's position in the firm hierarchy directly affects training through the implied job-to-job transition rate. In our setting it does not. However, if the worker's gains associated with a move to another firm fall short of the old firm's losses, one would expect that general training be decreasing in the degree of competitive pressure on the match, since it now raises the effective discount rate on the returns to human capital investments. Furthermore, specific investment is stimulated since it is a way to reduce the match surplus destruction associated with job-to-job transitions. In the Supplemental Appendix, Sanders and Taber (2012) discuss such a case in an environment where wages are statically bargained based on an outside worker option of unemployment and the current firm cannot provide side payments to avoid the destruction of match surplus. Fu (2011) presents an analysis with a super-modular production function and a piecewise wage posting environment where matches are not necessarily fully compensated for their destruction when workers reallocate.

3.2 Social planner

General training is in the risk neutral case socially efficient.[12] The proof is relegated to Appendix B. For specific training, the environment has an intriguing inefficiency: In the modular production function case, the decentralized economy provides the same level of specific training everywhere on the ladder. However, the social planner solution for specific training is increasing in firm type: The planner discounts match specific capital in low productivity firms at a greater rate because workers are more likely to reallocate to better firms. Therefore, the social planner invests more in specific training further up the ladder. The inefficiency in the decentralized economy is a result of future employers perfectly compensating the old match for its destruction, which includes the value of the match specific capital. Thus, there is a private return to match specific capital investment that is not present in the social returns. It implies that there tends to be too much specific training in low type firms. The following formalizes the argument.

For simplicity, we disregard general human capital an assume a modular production function. Without loss of generality assume $f_m(p) = f(p) + m$. Consider a utilitarian social planner problem of maximizing the contribution of a worker in a low match specific capital match. Denote the contribution by,

$$
\begin{aligned}
(r+\delta)\mathcal{V}_0(p) &= \max_{\mu}\left[f(p) - c_m(\mu) + \delta\mathcal{U} + \mu\left(\mathcal{V}_1(p) - \mathcal{V}_0(p)\right) + \lambda_e \int_p^1 \left[\mathcal{V}_0(p') - \mathcal{V}_0(p)\right] d\Phi(p)\right] \\
&= f(p) + \delta\mathcal{U} + \mathcal{M}(p) + \lambda_e \int_p^1 \frac{\left[f'(p') + \mathcal{M}'(p')\right]\hat{\Phi}(p')}{r+\delta+\lambda_1\hat{\Phi}(p')} dp',
\end{aligned}
$$

where the value of a high match specific capital match is,

$$
\begin{aligned}
(r+\delta)\mathcal{V}_1(p) &= f(p) + m + \delta\mathcal{U} + \lambda_e \int_{\tilde{p}(p)}^1 \left[\mathcal{V}_0(p') - \mathcal{V}_0(\tilde{p}(p))\right] d\Phi(p) \\
&= f(p) + m + \delta\mathcal{U} + \lambda_e \int_{\tilde{p}(p)}^1 \frac{\left[f'(p') + \mathcal{M}'(p')\right]\hat{\Phi}(p')}{r+\delta+\lambda_1\hat{\Phi}(p')} dp'.
\end{aligned}
$$

The threshold $\tilde{p}(p)$ is defined by $\mathcal{V}_1(p) = \mathcal{V}_0(\tilde{p}(p))$. Since $\mathcal{V}_0(p) < \mathcal{V}_1(p)$ and the value is increasing in p, it must be that $\tilde{p}(p) > p$. The loss of firm specific capital that is associated with switching firms must be compensated by a sufficiently large gain in firm type. The

[12] With the exception of the case with production function complementarities in production: Specific training is generally not provided efficiently which will carry over to general training in this case.

value of the investment option is,

$$\mathcal{M}(p) = \max_{\mu} \left[-c_m(\mu) + \mu\left(\mathcal{V}_1(p) - \mathcal{V}_0(p)\right)\right].$$

And the socially optimal specific investment choice solves,

$$c_m'(\mu(p)) = \mathcal{V}_1(p) - \mathcal{V}_0(p).$$

Some algebra yields,

$$\mathcal{V}_1(p) - \mathcal{V}_0(p) = \max_{\mu} \frac{m + c_m(\mu) - \int_p^{\tilde{p}(p)} \frac{\left[f'(p') + \mu\left[\mathcal{V}_1'(p') - \mathcal{V}_0'(p')\right]\right]\hat{\Phi}(p')}{r+\delta+\lambda_1\hat{\Phi}(p')} dp'}{r+\delta+\mu}. \tag{13}$$

Differentiation and the envelope theorem leads to,

$$\mathcal{V}_1'(p) - \mathcal{V}_0'(p) = \frac{f'(p)}{r+\delta}\left[\frac{\hat{\Phi}(p)}{r+\delta+\lambda_e\hat{\Phi}(p)} - \frac{\hat{\Phi}(\tilde{p}(p))}{r+\delta+\lambda_e\hat{\Phi}(\tilde{p}(p))}\right].$$

By $\tilde{p}(p) > p$ it follows that $\mathcal{V}_1'(p) - \mathcal{V}_0'(p) > 0$. Therefore, the social planner's choice of specific investment is increasing in p, $\mu'(p) = \left[\mathcal{V}_1'(p) - \mathcal{V}_0'(p)\right]/c_m''(\mu(p)) > 0$.

3.3 Commitment

The inefficiency in specific human capital training in low-type firms arises because a future employer fully compensates the destruction of the match if the worker moves. Therefore, if the current match can extract all rents from future employer meetings, it will internalize the value of the destruction of match specific capital in case the worker moves. We propose an instrument that achieves this outcome. The current firm can issue the following obligation: If the worker moves, the firm will pay the holder of the obligation the difference between the outside firm's willingness to pay and its own willingness to pay, that is $B = \bar{V}(p') - \bar{V}(p)$ where $p' > p$ is the type of the type of the outside firm and p is the type of the firm itself. In a competitive market the firm can sell this obligation at flow rate $\lambda_e \int_{\bar{V}(p)}^{\bar{V}(1)} \left[V - \bar{V}(p)\right] dF(V)$. With the obligation, the firm's willingness to pay for the worker comes to equal that of the outside firm.[13] Thus, the obligation allows for efficient separation and the current match extracts all the rents from future employers.

[13]This implies that with the obligation a firm type p will be setting a continuation value conditional on a higher type outside firm meeting of $V^o\left(\bar{V}'\right) = \bar{V}' > \bar{V}$, which would involve a violation of the firm's participation constraint should the worker decide to stay with the firm. Thus, the obligation needs to state that in case the worker ends up staying with the current firm and it subsequently lays off the worker due to a violation of the participation constraint, then the firm must honor the payment, B, to the holder of the obligation in this case as well.

Returning to the example above where general human capital has been eliminated from the analysis. Subject to the obligation, the value of the current contract to the worker is,

$$(r+\delta+\mu)V = w + \mu M + \delta U + \lambda_e \int_V^{\bar{V}(1)} \hat{F}(V')\,dV'.$$

The firm is maximizing the profit expression,

$$(r+\delta)\Pi_0(V,p) = f_0(p) - w - c_m(\mu) + \mu\left[\Pi_1(M,p) - \Pi_0(V,p)\right] - \lambda_e \int_V^{\bar{V}(p)} \hat{F}(V')\,dV',$$

where the expected liability payment from the obligation due to the worker quitting is perfectly offset by the revenue flow from the sale of the obligation. Furthermore, linearity of the profit function simplifies the profit loss integral from outside offers.

Insert the utility promise expression into the firm's profits to obtain,

$$(r+\delta)\bar{V}_0(p) = f_0(p) - c_m(\mu) + \delta U + \mu\left[\bar{V}_1(p) - \bar{V}_0(p)\right] + \lambda_e \int_{\bar{V}_0(p)}^{\bar{V}_0(1)} \hat{F}(V')\,dV',$$

where the optimal specific training rate solves,

$$c_m'(\mu) = \bar{V}_1(p) - \bar{V}_0(p).$$

The firm's willingness to pay in the $m=1$ case satisfies,

$$(r+\delta)\bar{V}_1(p) = f_1(p) + \delta U + \lambda_e \int_{\bar{V}_1(p)}^{\bar{V}_0(1)} \hat{F}(V')\,dV'.$$

After a change of variable one obtains,

$$\begin{aligned}(r+\delta+\mu)\left[\bar{V}_1(p) - \bar{V}_0(p)\right] &= f_1(p) - f_0(p) + c_0(\mu) - \\ &\quad \lambda_e \int_p^{\tilde{p}(p)} \frac{f_0'(p) + \mu\left[\bar{V}_1'(p) - \bar{V}_0'(p)\right]}{r+\delta+\lambda_e\hat{\Phi}(p)} \hat{\Phi}(p')\,dp'. \qquad (14)\end{aligned}$$

Equation (14) is identical to the planner solution (13). Hence, the privately optimal specific training intensity coincides with that of the social planner.

Variations on the style of obligation as that above can in our analysis undo the limitations to commitment that are implied by the renegotiation proofness restriction. We rule out that markets exist for such instruments. Nevertheless, the mechanisms of the obligation instrument above are instructive: Efficiency is obtained by adoption of side payments not within the match, but rather with a third party so as to ensure a credible bargaining position with a possible future employer of the worker. Side payments within the match will be undone by other side payments in a renegotiation proof contract. Consider for example an instrument such as an unvested pension scheme where the firm's pension liability to the worker is eliminated should the worker leave for another firm. At first glance, one might expect that such a scheme would allow additional rent extraction from a future employer, however the current firm has incentives to provide side payments so as to ensure efficient separation and an elimination of its pension liability, and for practical purposes, the pension scheme ends up being vested.

4 Model parametrization and calibration

The model is calibrated to the U.S. economy. We use the following functional forms:

$$
\begin{aligned}
u(w) &= \frac{c^{1-\theta}}{1-\theta} \\
c_h(\eta) &= \frac{\left(c_0^h \eta\right)^{1+c_1}}{1+c_1^h} \\
c_m(\mu) &= \frac{\left(c_0^m \mu\right)^{1+c_1}}{1+c_1^m} \\
f_{ij}(p) &= h_i + m_j + p, \qquad (i,j) \in \{0,1\}^2
\end{aligned}
$$

Firm productivity is Pareto truncated below and above. The vector of parameters (ρ, θ, m) is set a priori using estimates from the literature. We choose a coefficient of risk aversion $\theta = 2$. The death rate reflects an average working life of 40 years, $m = 0.025$. The discount rate is set to a 5% annual rate, $\rho = 0.05$.

The remaining parameters are chosen to fit salient features of the U.S. labor market. To discipline the model's accumulation processes, we reproduce the age-earning profile. The model's distribution of firm productivity is calibrated to fit to the firm-average wage distribution. In addition, the job destruction rate $\delta = 0.24$ is set to match the U.S. monthly layoff rate of 2%, and the calibration of $\lambda_0 = 4.3$ is set the fit the U.S. job finding rate out of unemployment. We also match the U.S. annual job-to-job transition rate of 0.31.[14] Finally, we impose that both types of human capital are provided at the same intensity on average.

[14]This corresponds to the employment to employment hazard rate reported by Rogerson and Shimer (2011).

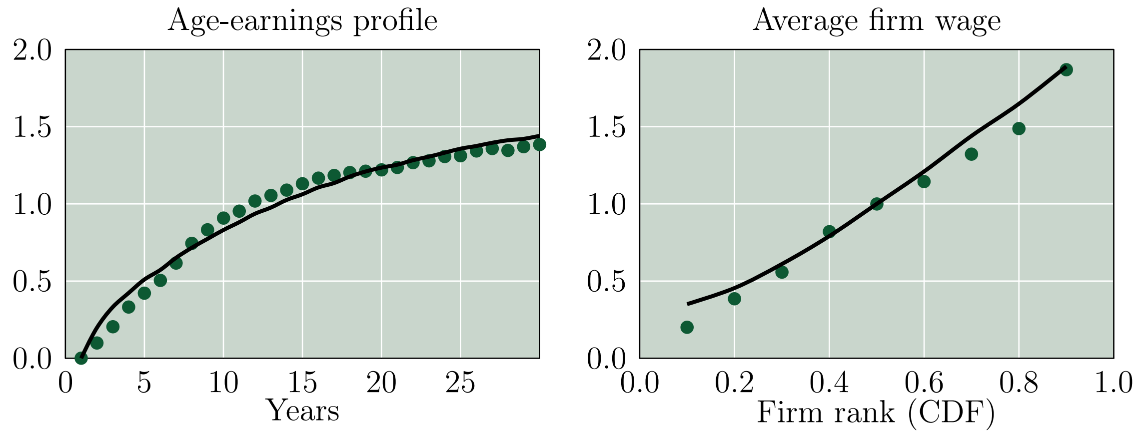
The age-earnings profile is obtained through the 1979–2010 survey years of the the National Longitudinal Survey of Youth, 1979. It is a representative sample of US households that was administered yearly from 1979-1994 by the Bureau of Labor Statistics, and once every two years since. We measure wages as the hourly pay rate at the time of the interview and deflate it using a CPI index. The sample is restricted to individuals' wage observations after they left school and never went back. Potential experience is defined as age minus age of entry in the labor market. We focus on males from both the core sample and the supplemental sample using race specific weights to have a sample representative of the population. Years that individuals spend in the military are excluded. Furthermore, individuals that spent more than two years in the military are excluded, entirely. We trim the top and bottom 1% of wage observations. Our final sample contains 4,866 individuals and 65,484 and individual-year observations. The average number of observations per individuals is 13.5. Wages are regressed on a full set of dummies for potential experience and individual fixed effects. We ask our model to match the dummies for potential experience. The first year of wages are excluded from the calibration due to the model's sharp assumption that all workers start their careers with a lifetime utility equal to the value of unemployment.

The firm wage distribution is obtained from Compustat, which provides annual accounting data on publicly listed US firms. We focus on the year 2014 as it is the year with the largest number of observations and we drop firms with less than 10 employees. Our final sample contains 1,997 firms. We calculate average wage per firms using the number of employees in the firms and the total wage bill.[15] The model parameters are chosen to reproduce the dispersion in average-wage per worker weighting each observation by the number of employees.

Figure 2 reports the fit of the model. The solid line is the model and the dots represents data-points. The left-panel reports the age-earnings profile coefficients from the first year of entry into the labor market experience until potential experience reaches 30. The right panel report the deciles of average weighted by the number of employees. Our parsimonious model does a good at reproduce the data patterns. We fit better the dispersion in average wage at the top than at the bottom of the distribution.

[15]The wage bill measure ""represents salaries, wages, pension costs, profit sharing and incentive compensation, payroll taxes and other employee benefits."

Figure 2: Model Fit



Note: Data points in green dots. Model estimate in solid black line.

Table 1: Model parameters

	I. Fixed	
$h_0 = m_0$	0.00	unskilled productivity
m	0.03	death rate
ρ	0.05	discount factor
θ	2.00	risk aversion
$\underline{p}$	1.00	Pareto distribution: lower-bound
	II. Calibrated	
δ	0.24	job destruction rate
λ_0	4.30	job offer rate: unemployed
λ_1	1.79	job offer rate: employed
c_0^h	37.41	training cost: constant
c_0^m	12.20	training cost: variable
$c_1^h = c_1^m$	0.81	training cost: variable
$h_1 = m_1$	3.83	skilled productivity
α	0.29	Pareto distribution: slope
$\overline{p}$	24.60	Pareto distribution: upper-bound

The calibrated parameter values are reported in Table 1. We find that the unemployed meeting rate is close to 2.5 times that of the employed meeting rate. Compared to for example Jolivet et al. (2006), this implies a relatively small difference in the efficiency of unemployed relative to employed search. In part, it is a reflection of the model's accumulation of match specific capital which in isolation reduces the job-to-job transition rate. Specific training is provided at a third of the cost of general training. As we emphasize below, the model implies significant positive sorting between firm productivity and both worker skill and match specific capital. Hence, the model's fit to the average firm wage distribution reflects both productivity differences across firms, but also a sorting of more productive workers into more productive firms.

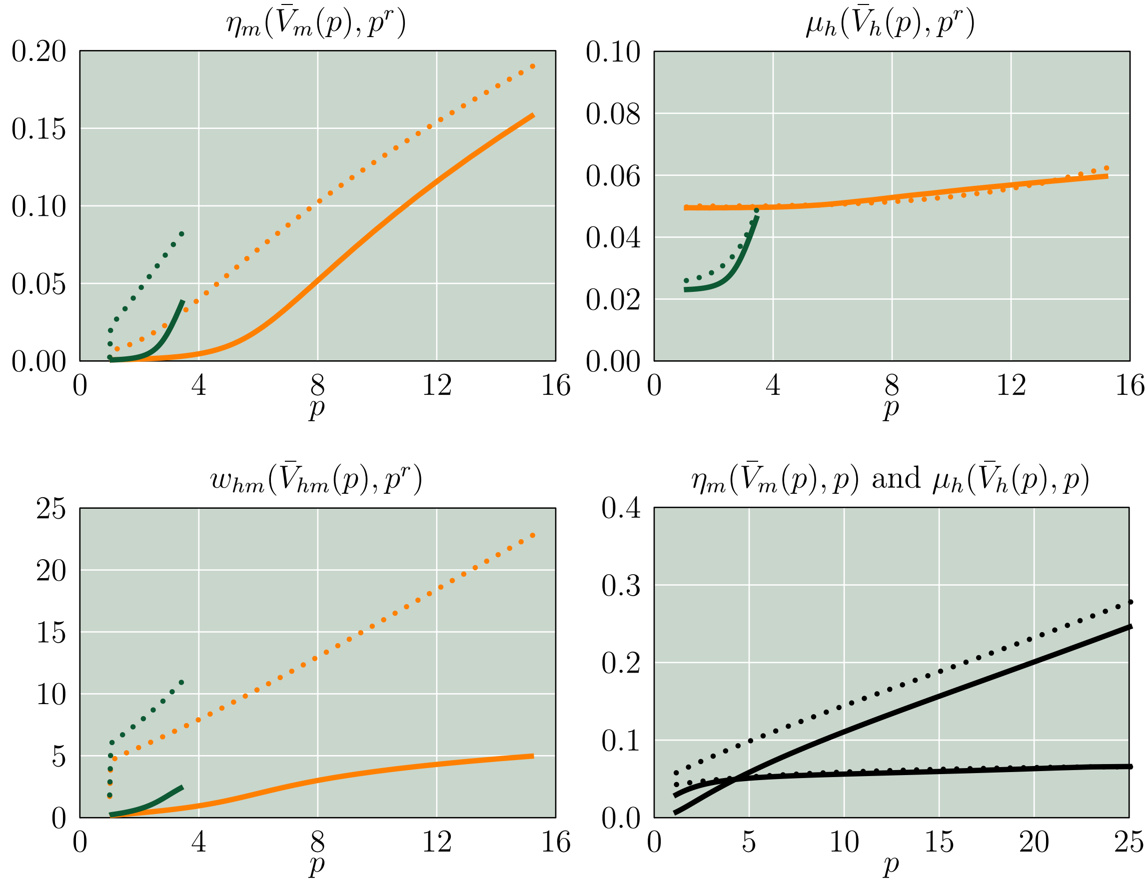
5 Calibrated contracts and steady state

Figure 3 shows the employment contracts for the 50^{th} and 90^{th} percentile firm productivity types as a function of the utility promise in the contract. The figure expresses the utility promise in terms of the willingness to pay of a firm with a given productivity p. This is done to facilitate comparison across contracts.

Within a contract, the wage is increasing in the utility promise. Holding the utility promise constant wages are decreasing in the firm type. The latter is a well-known feature of the outside offer matching feature of the wage mechanism, also seen in Postel-Vinay and Robin (2002). For a given utility promise, an increase in firm type implies greater expected gains from the on-the-job search process, which the firm can translate into higher profits through lower current wages. The lower right hand panel in figure 3 shows the wage at full surplus extraction across firm types, which is seen to be monotonically increasing in firm type. Production function complementarities between human capital and firm productivity can introduce a compensating differential between wages and training, but such considerations are not relevant given the modular production function specification in the current calibration. However, even in this case, whether higher type firms on average give higher wages depends on the composition of utility promises across its workers. We explore this in the next section.

As expected, both general and specific training are increasing in the utility promise within

Figure 3: Employment contracts by firm type.



Note: Firm type conditional contracts drawn for $r = 0.5$ and $r = 0.9$. Upper left panel: Solid lines for $m = 0$ and dotted lines for $m = 1$. Upper right panel: Solid lines for $h = 0$ and dotted lines for $h = 1$. Lower left panel: Dashed lines for $(h, m) = (0, 0)$ and dotted lines for $(h, m) = (1, 1)$. Lower right panel: Training at full rent extraction. Solid lines for $h = 0$ and $m = 0$, respectively. Dotted lines for $h = 1$ and $m = 1$, respectively.

a contract. For a given firm type, a lower utility promise implies a steeper expected future wage path, which increases the severity of the hold up problem. Therefore, training is lower for lower utility promises. The differences in competitive pressure across the two types of training show up in the figures as well. Specific training within the 90^{th} percentile firm is almost constant in the utility promise whereas general training is considerably more sensitive to the utility promise. The competitive pressure on future utility promises associated with specific training is determined primarily by the firm's position in the firm hierarchy: As the match becomes more productive due to an increase in m, competitive pressure on the worker's future utility promises is only affected in the event that the worker meets a more productive firm than the current firm. The wage is lowered up front to reflect the expected utility promise gains associated with training. The only reason the current utility promise does play a role in the provision of specific training is because the surplus loss associated with lowering the worker's wage is proportional to the worker's marginal utility, which is decreasing in the utility promise.

The increased competitive pressure associated with increased general human capital is on the other hand primarily determined by the current utility promise, V. A meeting with any productivity firm greater than $p_{1m}(V)$ is associated with an increased utility promise pressure due to the increase in h. Thus, for a lower V there is a larger mass of outside firms that can exert pressure on the match. In combination with the greater marginal utility of wages associated with the lower utility promise, V, the surplus loss of reducing the worker's wages up front in expectation of the future utility promise gains from general training is more sensitive to V.

Finally, risk aversion is a separate source of positive complementarity between general and specific training. If a worker's skill increases, her utility promise increases and her wages come to increase faster. The lower marginal utility of wages reduces the surplus loss associated with the backloading of wages due to one type of training. This effect is related to the strategic complementarity results in Balmaceda (2005) and Kessler and Lülfesmann (2006) where the existence of non-contractable specific training can counteract the hold-up problem in particular wage bargaining settings.

5.1 Training in steady state

The previous section highlights the importance of the utility promise composition within firms as a determinant of realized training levels. Figure 4 shows the average steady state training levels across firm types. Both general and specific training are on average increasing in firm type. For a given utility promise, training is decreasing in firm type because of the implied lower current wage and more severe hold up problem. However, between firm competition is such that in steady state the higher utility promises in more productive firms result in greater training, both general and specific.

The left panel of Figure 5 shows the accumulation of general, specific, and search capital over the careers of a cohort of workers. All three stocks are accumulating over time, which

Figure 4: Steady state training across firm types

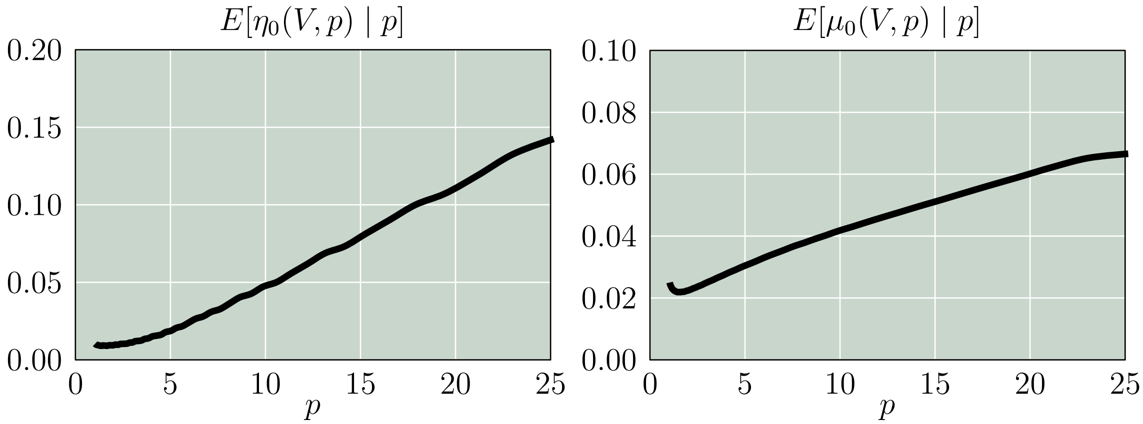
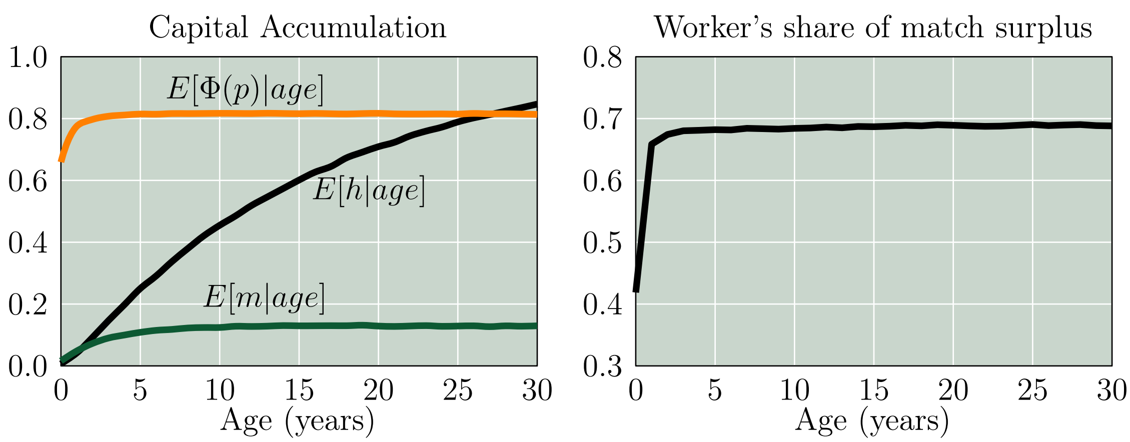
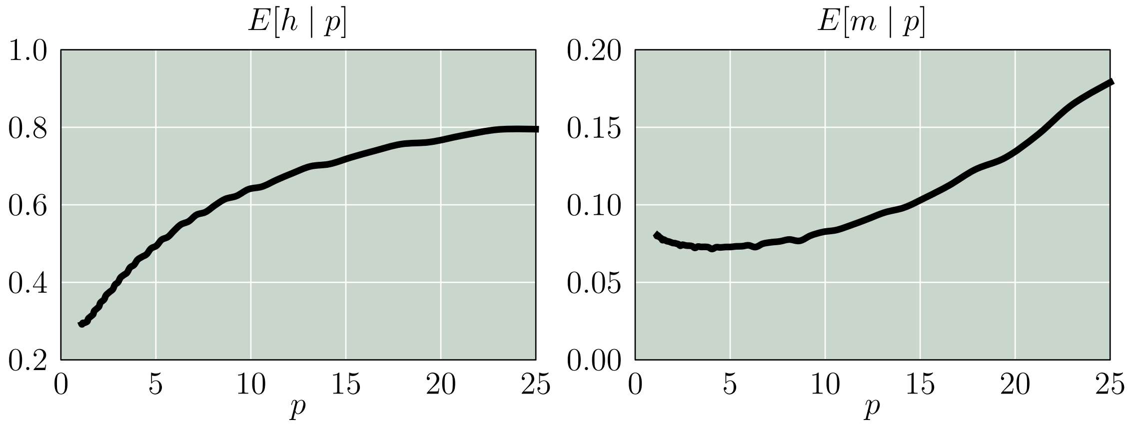


Figure 5: Steady state capital accumulation and surplus share by cohort age



in combination give rise to an increasing wage profile over experience as shown in the right hand panel. Unemployment resets the search and specific capital stocks. Consequently, the fraction of the population that is specifically trained does not go to one. It will also not be the case that all workers will eventually find themselves employed with the best firm type. But conditional on survival, all workers become generally skilled eventually given that there is no depreciation of general human capital in the model. The right panel shows the worker's average share of match surplus by age.

Figure 6: Average human capital level by firm type.



5.2 Sorting

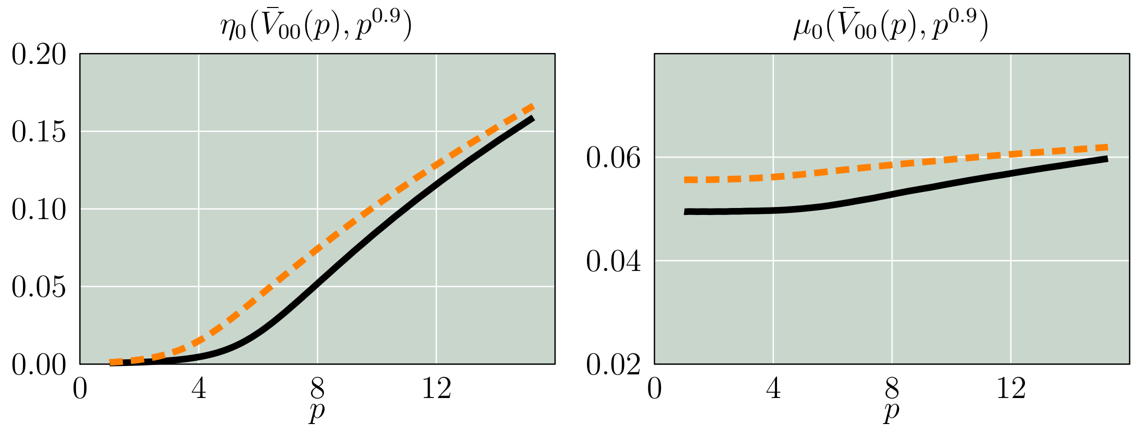
Higher rank firms provide more training and consequently their workers tend to have higher human capital, both general and specific. That is, the steady state match distribution exhibits a positive correlation between firm and worker productivity. Figure 6 shows the average levels of human capital by firm type in the steady state. As can be seen the labor force of higher ranked firms is both more skilled and has higher match specific capital as well. This is not a result of positive assortative matching. In addition, there are no complementarities in production in the calibration. It is a reflection of the state dependence in the model that fortunate employment draws with more productive firms contribute not only to a better contract value but also to a faster development of both general and specific skills.

6 Training and frictions

Acemoglu and Pischke (1999) emphasize that increased labor market friction allows firms to provide more general training when it is costly to resolve the hold up problem by making the worker pay for training up front through lower wages. Wasmer (2006) adds to the argument that increased labor market friction will increase specific training in a setup where matches invest in specific capital to reduce the risk of job destruction.

Figure 7 demonstrates the Acemoglu and Pischke (1999) mechanism within a given firm's contract. It shows the 90^{th} percentile productivity firm's training choices for $\lambda_e = 1.0$ and $\lambda_e = 2.0$. The horizontal axis is the utility promise support of the contract represented by the

Figure 7: Firm type $\Phi(p) = 0.9$ employment contract for $\lambda_e = 1.0$ and $\lambda_e = 2.0$.



Note: Solid line drawn for $\lambda_e = 2.0$ and dashed line drawn for $\lambda_e = 1.0$.

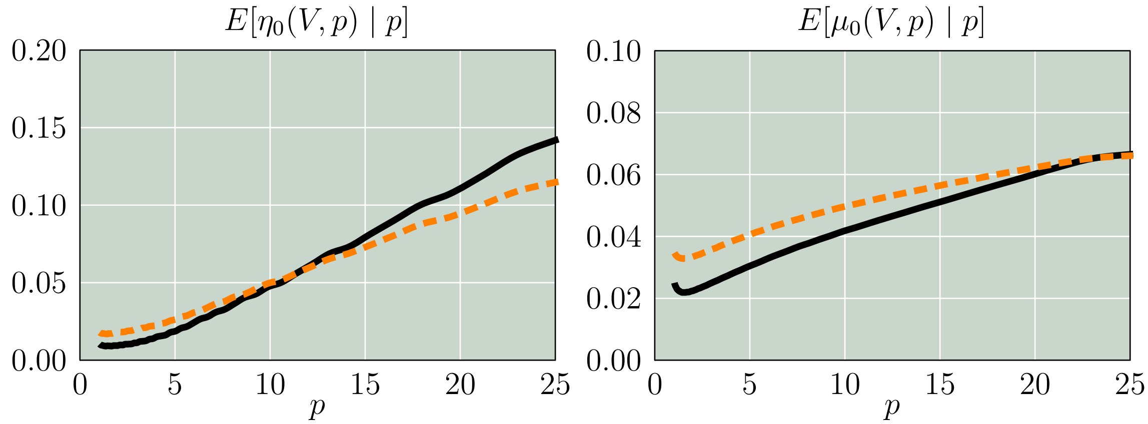
willingness to pay of type p firm. The figure shows that holding the firm hierarchy position associated with a given utility promise fixed, an increase in the contact rate is associated with a decrease in training. For the given competitive position the greater contact rate implies steeper future wages, a lower current wage, and therefore a more severe hold up problem from training. This is exactly the Acemoglu and Pischke (1999) argument and in our environment it applies not only to general training but also to specific training.

But the steady state utility promise composition within a firm's labor force is not constant in changes in the contact rate. Specifically, a higher contact rate implies a right shift of utility promises due to greater competitive pressure between firms. Figure 8 shows the average training levels by firm type in steady state for $\lambda_e = 1.0$ and $\lambda_e = 2.0$. It is seen that this effect by itself substantially modifies the Acemoglu and Pischke (1999) mechanism to such an extent that a substantial range of firm types actually *increase* training in steady state when the contact rate goes up. This effect is more pronounced for general than specific training.

In addition to the composition of utility promises, overall training and accumulation of skills in the economy also depend on the match distribution which is also affected by changes in frictions. As the contact rate increases, worker mismatch will tend to decline in the sense that workers will match with better firms. Since training is generally decreasing in mismatch, one would expect an effect on training from changes in the match distribution as well.

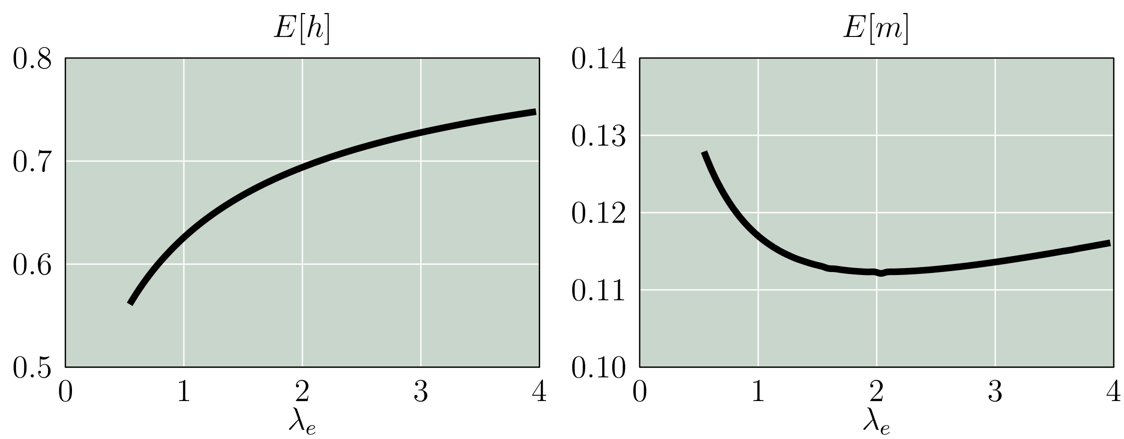
The left hand panel in Figure 9 shows the average human capital levels in the steady

Figure 8: Average steady state firm type conditional training levels for $\lambda_e = 1.0$ and $\lambda_e = 2.0$.



Note: Solid line drawn for $\lambda_e = 2.0$ and dashed line drawn for $\lambda_e = 1.0$.

Figure 9: Average steady state levels of general and specific human capital by λ_e.



state economy for different levels of contact rates. As can be seen, general skill is robustly increasing in the contact rate, which is opposite to the intuition developed in Acemoglu and Pischke (1999). The analysis in this paper embodies the central mechanism in their paper, but it is dominated by composition effects from search on the job and the presence of firm heterogeneity that is a natural consequence of a frictional labor market environment.

For the given calibration, match specific capital is non-monotone in the contact rate. Specific training is increasing in firm type and eventually the improved match distribution will result in more training as mismatch declines. However, for lower contact rates, the lower training levels within firm type for given utility promises dominate and result in less specific training.

7 Concluding remarks

We have in this paper put forth a framework for the study of wage dynamics that allows for search frictions, tim heterogeneity and human capital accumulation. In contrast to passive learning processes, we model the active investment in general human capital and match specific capital in response to the magnitude of the returns. The intensity of labor market competition is a primary factor in the determination of the returns to training, and we carry out the analysis in frictional setting where heterogeneous firms naturally coexist and workers can move directly between firms through a standard on the job search process. Optimally designed employment contracts set wages and training rates conditional on the history of the match.

We find that training varies by firm type. In isolation, the moral hazard problem associated with training implies that more productive firms train more. A super-modular production function in human capital and firm productivity will amplify this relationship. Thus, aggregate human capital accumulation comes to depend on the equilibrium match distribution of worker over firm types, and we show that it is of first order importance in the model calibrated to the US economy. For one, the classic Acemoglu and Pischke (1999) result that increased labor market friction alleviates the hold up problem in training and therefore results in more training is overturned through dominating equilibrium effects: Increased labor market friction results in worse matches and reduced bargaining positions - both of which imply reduced training.

In terms of the classic decomposition of labor market outcomes into luck and skill, the current analysis demonstrates that variation in skill is at least in part a result of variation in luck. The calibrated economy displays substantial sorting despite the absence of complementarities in production and the absence of assortative matching. The positive relationship between worker skill and firm productivity is a result of the faster accumulation of skill and match specific capital in more productive firms. Consequently, more productive firms tend to have more skilled workers as well as higher match specific capital.

The presence of firm heterogeneity also allowed us to point to an important feature of wage dynamics and specific training: The presence of more productive firms than the current firm implies that the match value increase associated with specific training can be contested in the market by these more productive firms. Therefore, specific training is associated with both increasing wages within the job as well as increased wages with future employers. Consequently, the distinction between tenure and experience effects in wage dynamics is not by itself sufficient to evaluate the importance of specific relative to general training. An obvious avenue for future research is to utilize worker reallocation patterns to help with the separate identification of the two processes: Specific training reduces reallocation whereas general training has no impact on mobility in the model.

A Steady state conditions

Assuming that unemployed workers do not turn down any meetings, the steady state conditions on the employment and unemployment stocks are,

$$(d+\lambda_u)\,u_0 = d+\delta\,(e_{00}+e_{01}) \tag{15}$$

$$(d+\lambda_u)\,u_1 = \delta\,(e_{10}+e_{11}) \tag{16}$$

$$(d+\delta+\bar{\eta}_0+\bar{\mu}_0)\,e_{00} = \lambda_u u_0 + e_{01}\int_0^1\int_{U_0}^{\bar{V}_{01}(p')} \lambda\hat{F}_0\left(\bar{V}_{01}(p')\right) g_{01}(V',p')\,dV'dp' \tag{17}$$

$$(d+\delta+\bar{\mu}_1)\,e_{10} = \lambda_u u_1 + \bar{\eta}_0 e_{00} + e_{11}\int_0^1\int_{U_1}^{\bar{V}_{11}(p')} \lambda\hat{F}_1\left(\bar{V}_{11}(p')\right) g_{11}(V',p')\,dV'dp' \tag{18}$$

$$\bar{\mu}_0 e_{00} = \left(d+\delta+\bar{\eta}_1+\int_0^1\int_{U_0}^{\bar{V}_{01}(p')} \lambda\hat{F}_0\left(\bar{V}_{01}(p')\right) g_{01}(V',p')\,dV'dp'\right) e_{01} \tag{19}$$

$$\bar{\eta}_1 e_{01}+\bar{\mu}_1 e_{10} = \left(d+\delta+\int_0^1\int_{U_1}^{\bar{V}_{11}(p')} \lambda\hat{F}_1\left(\bar{V}_{11}(p')\right) g_{11}(V',p')\,dV'dp'\right) e_{11}, \tag{20}$$

where $\bar{\mu}_h = \int_0^1\int_{U_h}^{\bar{V}_{h0}(p')} \mu_h(V',p')\,dG_{h0}(V,p)$ and $\bar{\eta}_m = \int_0^1\int_{U_0}^{\bar{V}_{0m}(p')} \eta_m(V',p')\,dG_{0m}(V,p)$.

The steady state conditions on $e_{01}G_{01}(V,p)$, $e_{10}G_{10}(V,p)$, and $e_{11}G_{11}(V,p)$ are respectively,

$$e_{00}\int_0^p\int_{U_0}^{\bar{V}_{00}(p')} 1\left[M_0(V',p')\le V\right]\mu_0(V',p')\,g_{00}(V',p')\,dV'dp' =$$
$$e_{01}\int_0^p\int_{U_0}^{\bar{V}_{01}(p)} \left[m+\delta+\eta_1(V',p')+\lambda\hat{F}_0\left(\bar{V}_{01}(p')\right)\right] g_{01}(V',p')\,dV'dp'.$$

The steady state condition on $e_{10}G_{10}(V,p)$ is,

$$\lambda_u u_1\Phi(p)+\lambda e_{11}\int_0^{p_{11}(V)}\int_{U_1}^{\bar{V}_{11}(p')}\left[F_1\left(\bar{V}_{10}(p)\right)-F_1\left(\bar{V}_{11}(p')\right)\right] g_{11}(V',p')\,dV'dp'$$
$$+\,e_{00}\int_0^p\int_{U_0}^{\bar{V}_{00}(p')} 1\left[H_0(V',p')\le V\right]\eta_0(V',p')\,g_{00}(V',p')\,dV'dp' =$$
$$e_{10}\left\{\int_0^{p_{10}(V)}\int_U^{\bar{V}_{10}(p')}\left[d+\delta+\mu_1(V',p')+\lambda\hat{F}_1\left(\bar{V}_{10}(p)\right)\right] g_{10}(V',p')\,dV'dp'+\right.$$
$$\left.\int_{\bar{p}_{10}(V)}^p\int_U^V\left[d+\delta+\mu_1(V',p')+\lambda\hat{F}_1(V)\right] g_{10}(V',p')\,dV'dp'\right\}.$$

And finally, the steady state condition on $e_{11}G_{11}(V,p)$ is,

$$e_{10}\int_0^p\int_{U_0}^{\bar{V}_{10}(p')} 1\left[M_1(V',p')\leq V\right]\mu_1(V',p')\,g_{10}(V',p')\,dV'dp' =$$
$$e_{11}\int_0^p\int_{U_0}^{\bar{V}_{11}(p)}\left[d+\delta+\lambda\hat{F}_1\left(\bar{V}_{11}(p')\right)\right]g_{11}(V',p')\,dV'dp'.$$

B Social planner, general human capital

Consider the case where human capital is only general. Assume a modular production function and risk neutral workers. Without loss of generality assume $f_h(p)=f(p)+h$, and $b_h=f_h(0)$. Consider a utilitarian social planner problem of maximizing the contribution of a worker in a low skill match. Denote the contribution by,

$$\begin{aligned}(r+\delta)\mathcal{V}_0(p) &= \max_{\mu}\left[f(p)-c_h(\eta)+\delta\mathcal{U}_0+\eta\left(\mathcal{V}_1(p)-\mathcal{V}_0(p)\right)+\lambda_e\int_p^1\left[\mathcal{V}_0(p')-\mathcal{V}_0(p)\right]d\Phi(p)\right]\\ &= f(p)+\delta\mathcal{U}_0+\mathcal{H}(p)+\lambda_e\int_p^1\frac{\left[f'(p')+\mathcal{H}'(p')\right]\hat{\Phi}(p')}{r+\delta+\lambda_1\hat{\Phi}(p')}dp',\end{aligned}$$

where the value of a high match specific capital match is,

$$\begin{aligned}(r+\delta)\mathcal{V}_1(p) &= f(p)+1+\delta\mathcal{U}_1+\lambda_e\int_p^1\left[\mathcal{V}_1(p')-\mathcal{V}_1(p)\right]d\Phi(p)\\ &= f(p)+1+\delta\mathcal{U}_1+\lambda_e\int_p^1\frac{f'(p')\hat{\Phi}(p')}{r+\delta+\lambda_1\hat{\Phi}(p')}dp'.\end{aligned}$$

The value of the investment option is,

$$\mathcal{H}(p)=\max_{\eta}\left[-c_h(\eta)+\eta\left(\mathcal{V}_1(p)-\mathcal{V}_0(p)\right)\right].$$

And the socially optimal specific investment choice solves,

$$c_h'(\eta(p))=\mathcal{V}_1(p)-\mathcal{V}_0(p).$$

Some algebra yields,

$$\mathcal{V}_1(p)-\mathcal{V}_0(p)=\frac{r+\delta+rc_h(\eta)}{r(r+\delta+\eta)}$$

which implies,

$$\mathcal{V}_1'(p)-\mathcal{V}_0'(p)=0$$

Therefore, the social planner's choice of specific investment is constant in p, $\eta'(p)=0$.

C Numerical Solution

Firm productivity is discretized and each element of the grid $\{p_j\}_{j=1,\ldots,N_p}$ has equal probability $\frac{1}{N_p}$. All integrals are numerically approximated with Gauss-Legendre quadrature. We use linear interpolation to approximate policy functions off the grid. We solve for the optimal contracts using the following iterative algorithm. Use initial guesses for the functions F_1 and F_0.

1. Solve for $\bar{V}_{11}^j$ and the corresponding w using (2), (3) and (4).

2. For any $V_{11} \in \left[U_1, \bar{V}_{11}^j\right]$, solve for the corresponding (w, Π) using (2), (3) and (5).

3. Solve for $\bar{V}_{10}^j$ and the corresponding (w, μ, Π) using (2), (3), (4) and (12).

4. Using the previous step, update F_1. Return to 1. until convergence.

5. For any $V_{10} \in \left[U_1, \bar{V}_{10}^j\right]$, solve for the corresponding (w, η, Π) using (2), (3), (5) and (12).

6. Solve for $\bar{V}_{01}^j$ and the corresponding (w, η) using (2), (3), (4) and (11).

7. For any $V_{01} \in \left[U_0, \bar{V}_{01}^j\right]$, solve for the corresponding (w, μ, Π) using (2), (3), (5) and (11). If the participation constraint is not satisfied, we set $H = U_1$.

8. Solve for $\bar{V}_{00}^j$ and the corresponding (w, η, μ) using (2), (3), (4), (11) and (12). If the participation constraint is not satisfied, we set $H = U_1$.

9. Using the previous step, update F_0. Return to 5. until convergence.

10. For any $V_{00} \in \left[U_0, \bar{V}_{00}^j\right]$, solve for the corresponding (w, η, μ, Π) using (2), (3), (5), (11) and (12).

In step 5 and above, we check whether the participation constraint $H \geq U_1$ is binding.

References

Acemoglu, D. and Pischke, J.-S. (1999). The structure of wages and investment in general training. *Journal of Political Economy*, 107(3):539–572.

Altonji, J. G. and Shakotko, R. A. (1987). Do wages rise with job seniority? *Review of Economic Studies*, 54(3):437–459.

Altonji, J. G. and Williams, N. (2005). Do wages rise with job seniority? a reassessment. *Industrial and Labor Relations Review*, 58:370–397.

Bagger, J. and Lentz, R. (2013). An empirical model of wage dispersion with sorting. *University of Wisconsin–Madison Working Paper*.

Balmaceda, F. (2005). Firm-sponsored general training. *Journal of Labor Economics*, 23(1):115–133.

Becker, G. S. (1964). *Human Capital*. University of Chicago Press, Chicago.

Buchinsky, M., Fougère, D., Kramarz, F., and Tchernis, R. (2010). Interfirm mobility, wages and the returns to seniority and experience in the united states. *Review of Economic Studies*, 77:972–1001.

Carrillo-Tudela, C. and Smith, E. (2014). Search capital. *Working Paper*.

Christensen, B. J., Lentz, R., Mortensen, D. T., Neumann, G., and Werwatz, A. (2005). On the job search and the wage distribution. *Journal of Labor Economics*, 23(1):31–58.

Dustman, C. and Meghir, C. (2005). Wage, experience and seniority. *Review of Economic Studies*, 72:77–108.

Felli, L. and Harris, C. (1996). Learning, wage dynamics, and firm-specific human capital. *Journal of Political Economy*, 104(4):838–868.

Fu, C. (2011). Training, search and wage dispersion. *Review of Economic Dynamics*, 14:650–666.

Fujita, S. and Moscarini, G. (2013). Recall and unemployment. *Working Paper*.

Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2015). Occupational switching and self-discovery in the labor market. *Working Paper*.

Jolivet, G., Postel-Vinay, F., and Robin, J.-M. (2006). The empirical content of the job search model: Labor mobility and wage distributions in europe and the US. *European Economic Review*, 50(4):877–907.

Kessler, A. S. and Lülfesmann, C. (2006). The theory of human capital revisited: On the interaction of general and specific investments. *The Economic Journal*, 116:903–923.

Lentz, R. (2014). Optimal employment contracts with hidden search. *NBER Working Papers*, (19988).

Postel-Vinay, F. and Robin, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6):2295–2350.

Rogerson, R. and Shimer, R. (2011). Search in macroeconomic models of the labor market. In *Handbook of Labor Economics, Volume 4a*. Elsevier Science B.V.

Sanders, C. and Taber, C. (2012). Life-cycle wage growth and heterogenous human capital. *Annual Review of Economics*, 41(1):399–425.

Spear, S. E. and Srivastava, S. (1987). On repeated moral hazard with discounting. *Review of Economic Dynamics*, pages 599–617.

Thomas, J. and Worrall, T. (1988). Self-enforcing wage contracts. *Review of Economic Studies*, pages 541–554.

Topel, R. H. (1991). Specific capital, mobility and wages: Wages rise with job seniority. *Journal of Political Economy*, 99(1):145–176.

Wasmer, E. (2006). Specific skills, search frictions, and firing costs. *American Economic Review*, 96(3):811–31.