Bilateral Worker-Firm Training Decisions and an Application to Discrimination

Stéphane Carcillo, Etienne Wasmer

ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 71-72 – 2003

Bilateral Worker-Firm Training Decisions and an Application to Discrimination*

Stephane CARCILLO, Etienne WASMER*****

ABSTRACT. – A large part of group differences in wages comes from unobserved or unverifiable characteristics such as the intensity of human capital investments on-the-job. This is notably the classical argument to account for gender differentials. We build a framework in which training decisions are bilateral, in the presence of frictions in the labor market generated by a flow-matching model. Workers make learning efforts while firm invest by paying direct training costs. Under complementarity of the learning function of these two inputs (effort and direct costs), the outcome of training decisions requires coordination between the firm and the worker. We exhibit cases in which a high investment in training/high effort Nash equilibrium is a dominant strategy, which makes discrimination between groups difficult, and cases in which there is a coordination failure. We define *coordination discrimination* as a case in which observed characteristics of workers (gender, race, diploma) help to coordinate on an equilibrium. We explore the case of unobservable types of workers, and study under which conditions the (common knowledge) priors of firms do not affect the equilibrium strategies, and under which conditions they play a crucial role instead.

Décisions bilatérales employeur-employé de formation et une application à la discrimination

RÉSUMÉ. – Une partie importante des différences de salaires entre groupes provient de caractéristiques inobservables telles que l'intensité d'investissement en capital humain dans l'emploi. Nous développons un modèle d'appariement avec frictions sur le marché du travail dans lequel les décisions de formation sont bilatérales. Les travailleurs font des efforts d'apprentissage alors que les firmes paient les coûts directs de formation. Nous montrons qu'il existe des cas où investissement important en formation / effort élevé est une stratégie dominante, ce qui implique que la discrimination entre groupes est difficile et d'autres cas où il y a des défauts de coordination.

---

* We benefitiated from useful discussions with N. SAHUGUET, M. CASTANEIRA, A. JACQUES and P. WALBROECK. We are grateful to some participants of the Eurequa (Université Paris 1) labor market seminar for their suggestions. We also thank two anonymous referees for very helpful comments.
**Affiliations: ECARES, ULB.
***Affiliations: Univ. Metz, ECARES, ULB, CEPR and IZA.

1 Introduction

A large part of individual differences in wages or employability comes from unobserved characteristics. The conventional estimates of Mincer's wage equations and more sophisticated extensions, in absence of worker's fixed effects, capture at best 30 to 40 % of the variance in log wages. Inclusion of proxy variables for ability (such as IQ tests, grades, etc...) does not add much to the explanatory power of the extended Mincer equations. The fact that there is a huge unexplained component makes it difficult to understand gender or race differences in earnings. One obvious determinant of wages and wage differences across groups is the intensity of past human capital investments of the individuals observed in the samples of surveys. The first component of human capital, *i.e.* schooling, is usually observed, but other on-the-job investments in human capital are typically not observed and have to be proxied by experience in the labor market, itself proxied by "potential experience", *i.e.* basically age minus education.[1]

A part of the unexplained wage earnings gap between men and women can thus be attributed to the poor quality of the proxy for experience. However, in surveys where "true experience" is available, not much is added in terms of explanatory power of wages.[2] This suggests that what matters is not so much the exact time spent working, but rather how much a worker and a firm have invested in training to improve the productivity of the worker-firm association. Here, we investigate the determinants of these bilateral training decisions, and notably we focus on how racial or gender differences may lead to large differences in the intensity of training, and under which circumstances discrimination between groups can be observed ex-post in wage differences.

Our analysis is based on a simple two-person game. Two partners (interpreted here as a firm and a worker) match and produce some output, which depends on *bilateral training decisions*, namely training *efforts* by workers and training *investments* by firms. We base the analysis on the absence of contractibility on these choice variables, due to unverifiability, as usual in information economics. The way the inputs of both partners are combined into the payoff structure of each agent is reflected by two dimensions, complementarity and returns to scale.

Our results are as follows. High investments from both partners are associated with high returns to scale, low investments from both partners are associated with low returns to scale, asymmetry of training decisions are associated with substitutability of training and efforts, symmetry being instead reinforced by complementarity. Finally, extreme complementarity leads to a coordination incentive game for all values of returns to scale. We then argue that our analysis of strategies of firms and workers can be seen as an extension of the concept of statistical discrimination introduced by PHELPS [1972]

1. See for instance ANGRIST and KRUEGER [1999]
2. See ALTONJI and BLANK [1999] for an exhaustive discussion and references of these ideas and those of the previous paragraph.

and ARROW [1985a and b].[3] We notably introduce a concept of "coordination discrimination", which arises when there are multiple Nash-equilibria if coordination on one equilibrium is biased towards a group, and we discuss why coordination by "cheap talk" is difficult in light of the analysis of LANG [1986)] on language discrimination.[4]

We develop the homogenous worker case, and undertake the comparative statics of several economically relevant parameters. To do so, we need a more fully integrated intertemporal framework, and use a matching framework à la PISSARIDES [1990]. This is the only way to introduce in a parsimonious way the outside option of firms vis-à-vis the workers, the expected discounted value of unemployment of workers and the bargaining power of workers in the Nash-bargaining game over wages. We also derive the average earnings of a group of worker, as a function of the parameters of the game and of the equilibrium transition probability. The framework is extended to the homogenous (two types) workers case, and we study the role of priors in a Bayesian context.

Then we extend the discussion of discrimination, by linking our paper to taste discrimination and statistical discrimination. We also suggest an alternative form of discrimination: playing non-cooperative when the other plays cooperative may be associated with a penalty, that might be lower if the opponent belongs to another social group. Such "default penalties" may help move away from asymmetric equilibria or prisoner's dilemma games for majority workers, but would not affect minorities.

The paper is organized as follows. Next section gives the main intuitions in a simple set-up and discusses our notion of discrimination. Then we detail the notations of the model in a more specific setup with discrete investment and develop the equilibrium concepts (the continuous case is treated in Appendix 1.2). Section 4 derives some comparative statics with homogenous workers, Section 5 introduces unobservable differences between workers and Section 6 further discusses discrimination in our setup. Section 7 briefly explores potential extensions and concludes.

2 Intuitions

Two agents (a firm and a worker) share a surplus depending on their initial training decisions, hereafter denoted by e (effort of the worker) and i (investment of the firm). Let's denote by λ the function that associates those inputs. We chose a CES investment function

$$\lambda(e,i) = (a_0 + a_1 e^{\rho} + a_2 i^{\rho})^{\alpha/\rho} \tag{1}$$

3. Arrow [1985a and b], and many others, for instance COATE and LOURY [1993], SCHWAB [1986], LUNDBERG and STARTZ [1983], discrimination arises because of a coordination failure. The typical mechanism is that workers, expecting a lower return to their human capital investments (*e.g.* because of noise in a productivity test, PHELPS [1972]), under-invest in human capital. This investment cannot be directly observed, and firms have a subjective probability that workers have indeed invested. If workers in a given group tend to under-invest, firms attribute them a lower probability of being skilled, and thus rationally pay them lower wages.

4. COATE and LOURY [1993] have a similar notion of coordination failure (based on group coordination, as opposed to bilateral coordination in our case), which may justify affirmative action.

where the elasticity of substitution is $\sigma_{ei} = 1/(1-\rho)$ and α captures scale effects. This function is consistent with our assumptions, namely that a productive job is the outcome of luck and bilateral training decisions, whereas the levels of effort e and investement i is not verifiable and thus can not be contracted. At this stage, simply consider that the payoff of each agents is an increasing revenue function $\mathcal{R}_k$, $k = f, w$ of λ minus the cost of training, assumed linear for each agent, *i.e.* $-e$ or $-i$. An interpretation of high returns to scale α could be the complexity of tasks in the job, while complementarity of inputs can be seen as the fact that worker's learning without specific help of the firms (example: use of machinery during training) is difficult. In other words, complementarity may be a proxy for capital intensity of the business, while substituability may reflect the case of service industries.

We also assume that e and i can be chosen in a finite, discrete subset, and notably that

$$
\begin{aligned}
& e \in \{e_0, e_1\} \text{ with } 0 \leqslant e_0 < e_1 \\
& i \in \{i_0, i_1\} \text{ with } 0 \leqslant i_0 < i_1
\end{aligned} \tag{2}
$$

If we simplify the notation $\lambda(e_h, i_k)$ to λ_{hk} for $h,k = 0,1$, then the payoff matrix is

$$
\left(\begin{array}{|c|c|}
\hline
\mathcal{R}_w(\lambda_{00}) - e_0; \mathcal{R}_f(\lambda_{00}) - i_0 & \mathcal{R}_w(\lambda_{01}) - e_0; \mathcal{R}_f(\lambda_{01}) - i_1 \\
\mathcal{R}_w(\lambda_{10}) - e_1; \mathcal{R}_f(\lambda_{10}) - i_0 & \mathcal{R}_w(\lambda_{11}) - e_1; \mathcal{R}_f(\lambda_{11}) - i_1 \\
\hline
\end{array}\right)
$$

To give the main intuition, we simply focus of a few numerical examples, derived from the more complex model of next section, but without entering into the details. Those details will be provided in equation (7) and table 7, representing the payoff structure and the parameter values respectively. At this stage, it is simply useful to note that we use a symmetric parametrization ($a_0 = 0,\ \ a_1 = a_2 = 0.5$) for the CES investment function, letting the payoff functions bear the asymmetries between workers and firms. We can then naturally classify those examples into two categories, the symmetric cases and the asymmetric cases, according to FUDENBERG and TIROLE's [1995] typology.

2.1 Symmetric equilibria

2.1.1 Single Pareto-optimal equilibrium (HH : high, LL : low)

Here, two cases can arise, referred to as the HH and the LL situations. In the high investments equilibrium HH displayed in table 1, both parties have individually interest to invest in training. This case arises when training efforts are complements and when average productivity and/or returns to scale are high enough. Indeed, in this case, both players have interest to coordinate on symmetric strategies (efforts are complements), and the most rewarding strategy is the higher level of investment (high returns on investment). Note that this equilibrium is also Pareto-optimal.

TABLE 1
HH : Single Nash, Pareto-dominant (compared to (0,0)) equilibrium ($\rho = -1$, $\alpha = 1.2$)

	$\mathbf{i}_0$	$\mathbf{i}_1$
$\mathbf{e}_0$	18.5; 13.8	19.7; 14.8
$\mathbf{e}_1$	18.3; 16.2	$\boxed{21.3;\ 20.7}$

In the low investment equilibrium LL displayed in table 2, on the contrary, both parties have interest not to invest in training, and this is the only Nash equilibrium. This arises when the technology shows very low returns to scale, when investment and efforts are complements, but also when they turn out to be substitutes instead, if returns to scale are low enough.

TABLE 2
LL : Single Nash, Pareto-dominant (compared to (1,1)) equilibrium ($\rho = -1$,$\alpha = 0.3$)

	$\mathbf{i}_0$	$\mathbf{i}_1$
$\mathbf{e}_0$	$\boxed{21.7;\ 20.3}$	21.9; 19.2
$\mathbf{e}_1$	20.5; 20.6	21.1; 20.3

2.1.2 Prisoners' dilemma (DP)

The DP case is simply a special case of the LL equilibrium, in which the only equilibrium is sub-optimal. The worker and the firm would jointly have interest to coordinate on high investment levels, but they know the other party would individually have interest to deviate in this case. This is a classic prisoners'dilemma problem. It arises whenever efforts are rather complements (or weakly substitutes) and returns to scale are not too strong (otherwise HH can arise as another equilibrium), but not too low either, (otherwise LL can arise as an optimal equilibrium, or at least one party could be better off in this situation compared to HH). See table 3 for an illustration.

TABLE 3
DP : Single Nash, Pareto-dominated equilibrium ($\rho = -1.5$,$\alpha = 0.5$)

	$\mathbf{i}_0$	$\mathbf{i}_1$
$\mathbf{e}_0$	$\boxed{21;18.9}$	21.4;18.2
$\mathbf{e}_1$	20;19.6	21.1;20.4

2.1.3 Coordination incentive (CI)

Another case arises when both the "high" and "low" equilibria exist. See table 4 for an example. In such a case, one equilibrium is Pareto-dominant (in our simulations, this is always the high equilibrium), which should give incentives to both parties to coordinate in order to reach the latter. This case arises whenever efforts are strongly complements and returns to scales not too small. The intuition is that, if efforts were not sufficiently complements, LL could not be a possible equilibrium. On the other hand, if returns to scale happened to be too small, HH could not be a possible equilibrium. This case is one of the example in which discrimination may arise, as argued later in this section.[5]

TABLE 4
CI : Two Nash, Pareto-ranked equilibria ($\rho = -2.5, \alpha = 0.8$)

	$\mathbf{i}_0$	$\mathbf{i}_1$
$\mathbf{e}_0$	$\boxed{20;16.8}$	20.4;16.1
$\mathbf{e}_1$	19;17.5	$\boxed{21.2;20.5}$

2.2 Asymmetric cases

There are also a few asymmetric situations that arise in equilibrium. These situations, as we argue later on, are also useful in the analysis of discrimination.

2.2.1 Behavior convention (DH : Dove / Hawk)

In this case, two off-diagonal equilibria exist with an asymmetry in training investments/efforts: one of the two parties bears all the training investment cost. Which of both equilibria will apply depends on the behavior convention assumed for both players. This case arises whenever efforts are strongly substitutes and returns to scale high enough (otherwise there would be no incentive to invest for one or the other party). Table 5 illustrates.

5. Finally, one could wonder if it is possible to observe another type of coordination problem, the classical Battle-of-the-Sexes situation (BoS). This is a typical case requiring horizontal differences in tastes, in which both parties prefer symmetric choices to anything else, but cannot a priori coordinate on one or the other because they have different tastes. For instance, in case of two Nash equilibria, LL and HH, the worker could be better off in the LL equilibrium but the firm would prefer the HH equilibrium. We have not explored this situation here.

TABLE 5
DH : Two Nash, unranked equilibria ($\rho = 0.8, \alpha = 1.2$)

	$\mathbf{i}_0$	$\mathbf{i}_1$
$\mathbf{e}_0$	18.5;13.8	$\boxed{22.0;19.4}$
$\mathbf{e}_1$	$\boxed{20.6;20.8}$	21.3;20.7

2.2.2 Lonely rider (LH : low / high)

Still in the asymmetric cases, there are situations where only one equilibrium prevails: the firm invests in training but the worker does not. This can happen when efforts are substitutes or slightly complements, for different levels of returns to scale, as in table 6. One could also exhibit the symmetric HL (high/low) case in which the worker does invest but the firm does not.[6]

TABLE 6
LH : Single Nash equilibrium ($\rho = 0.5, \alpha = 0.5$)

	$\mathbf{i}_0$	$\mathbf{i}_1$
$\mathbf{e}_0$	21.1;18.9	$\boxed{22.1;19.7}$
$\mathbf{e}_1$	20.7;21.1	21.1;20.3

2.3 Partial conclusion

It is thus now clear that bilateral uncontractible decisions lead to several different situations with drastically different economic outcomes. Our setup is very general, and in partial equilibrium, we could interpret our model as two parties undertaking an initial effort that can not be verified by a third party. For instance, IRA and the British government making "peace efforts", or, Corsican nationalist and the French government, or, Palestinian leaders and the Israeli government, are situations that could be modelled within our game. The returns to scale and the substituability of these efforts will predict the likely outcome of this game in each situation. We can now explain how discrimination will emerge from our set-up.

6. This cannot be obtained with the benchmark values of the parameters, but it is a possible outcome whenever the marginal impact of efforts is larger than the marginal impact of investments (as with $a_1 = .8$ and $a_2 = .2$ for instance), or when the payoff function of the worker is sufficiently above the payoff function of the employer. This arises later on when the bargaining power of workers is significantly above .5 (the intuition is that this raises the marginal impact of worker's efforts) and when their outside option is large, for reasons discussed in Section 4.1.

2.4 **Coordination failures**

Let us propose an interpretation of discrimination, quite natural in our context because linked to coordination failure. In particular, there are two types of situations in which coordination is required: the coordination incentive case and the dove-hawk case. In both examples there is a multiplicity of equilibria, and discrimination may arise if agents use gender or ethnic origin as a coordination mechanism.

DEFINITION 1. *Coordination discrimination arises when there are two equilibria, and the coordination on one equilibrium is biased towards one group of workers.*

2.4.1 **Symmetric equilibria, single type**

According to this definition, with a single type of worker, only equilibria CI and DH may be associated with Coordination Discrimination. In principle, in the case of CI, the coordination on the high equilibrium is easy, since one of the two equilibria is Pareto-superior to the other. It is sometimes considered that a simple "cheap talk" can lead to the elimination of the Pareto-inferior equilibrium. Assuming that the Pareto-inferior equilibrium is the LL equilibrium,[7] then the resolution of the coordination problem leads to higher investments from both parts, faster productivity gains and thus higher wages in the cross-section. We tend to disagree on how easily cheap talk can resolve the coordination problem, and three arguments can be opposed to this view.

First, the literature on discrimination has introduced the idea that communication might not be easy nor cheap in all cases, and notably that "language discrimination" may arise. In his seminal contribution, LANG [1986] discusses the effect of transaction costs introduced by the existence of language differences, that may reflect either true differences in language, but also, more pervasively but not less importantly, of dialects, of differences in intonations, attitudes, gestures, that increase the noise of communication, leading to more costly communication, or to communication misinterpretations. Combining Lang's insight with our concept of coordination discrimination, we have an alternative explanation of equilibrium differences in wages if, for instance, white employers and employees coordinate on a Pareto-superior equilibrium, while black employees and white employers have more difficulties to coordinate and end-up randomly in one or the other equilibrium.

Second, FUDENBERG and TIROLE [1995] argue, following Schelling's concept of focal points, that one can imagine several situations in which there exists focalness of some strategies depending on the player's culture and past experience. If it is difficult to explain why two lost persons in Paris would decide to wait for each other in front of the Eiffel Tower and not in front of the Sacré Coeur, it is similarly difficult to explain why minority groups may face focalness on the low equilibrium. This does not necessarily mean that this is not what happens in several occasions.

7. which our simulation confirm in all cases we have studied.

Finally, in order to help select an equilibrium when there are several, one can also introduce the notion of risk-dominance, following HARSANYI and SELTEN [1988]: some people may prefer an outcome to another if they believe that the opponent randomizes $(a; 1-a)$ on its own choice. We can easily exhibits situations in which for a group of workers, there is obvious coordination on HH, while for the other, risk-dominance leads to the choice of the LL equilibrium. Again, the perception of a may differ across groups, in a way adversely affecting minorities.

To sum up, the last three paragraphs have developed arguments against the view that cheap talk can easily lead to the coordination on the right Pareto-dominant strategy. Instead, there are several, subjective aspects and dimensions along which specific groups can be adversely affected by the bilateral decisions made in the game.

2.4.2 Asymmetric equilibria and reputation

Further, this concept of coordination discrimination can also be applied to asymmetric cases. Indeed, in the DH equilibrium, it may be that minorities spontaneously anticipate a low investment from the firm and wish to undertake the high investment, while majority workers feel secure that the firm makes the high investment while they do the low investment. Who will support the entire costs of the training investments/efforts is entirely a matter of convention.[8]

Of course, one can argue that in such an asymmetric situation, it would be difficult to observe wage differences between groups, if the effort by black workers in the LH case leads to as fast transition to high productivity as the investment made by firms for white workers. Indeed, it is only if there is an asymmetry in the function $\lambda(e,i)$ and more precisely, if the returns to effort are lower than the returns to investment by firms, that one would observe differences in wages by groups. This is true, but there would still be utility differences between groups, since $\mathcal{R}_w(\lambda_{01}) - e_0$ would be higher than $\mathcal{R}_w(\lambda_{10}) - e_1$.[9]

A last comparative statics exercise may be to introduce a psychological or a pecuniary penalty for misconduct of agents, such as playing L in CI, DH or even DP. Assume now that the penalty for such a behavior is lower when the two agents belong to different ethnic, gender or sociological groups. This penalty (denote it by z_w and z_f) will shifts the curves 1 and 2 to the right in figure 1. It thus reduces the likelihood of LH and DH, and raises CI. And, again, if CI leads less frequently to HH for black workers than for white workers, the existence of larger penalties in case of misconduct towards majority members leads to more discrimination.

8. Interestingly, the latter case, an asymmetric situation where workers are favored by employers at their own expenses, can be seen as another view of the concept of nepotism (introduced by GOLDBERG in [1982]) in which employers have a positive rate of substitution between profits and the number of workers of the favored group (for instance, they wish to overemploy beautiful women without any explicit profitability).

9. Later on introducing wages, and interpreting e as hours worked, one could also interpret the entry wage $\widetilde{w}/e$ as the hourly entry wage, and in this case it would be higher for white than for black workers.

2.4.3 Other links with the literature on discrimination

The idea of discrimination as coordination failure is present in ARROW [1985a] and COATE and LOURY [1993]. In Coate and Loury notably, workers have to choose their skill level taking into account the prior probability of firms, which affect the expected payoff of the skills investments. COATE and LOURY show cases in which there are two equilibria, one with low investment/low prior and another one with high investment/high prior, leading to a coordination problem. In our case, the investment choice by firm play a similar role as the prior of the firm in their analysis. However, an important difference with COATE and LOURY is that the game they model is sequential, a party doing a pre-market investment first, while our analysis is based on simultaneous decisions being made. As such, in our model, *coordination failures arise not only because of firms' beliefs, but also because of workers' beliefs about how firms will behave toward them*. An additional difference between us and COATE and LOURY is the nature of investments: specific and on-the-job in our case, and general and pre-market in COATE and LOURY.

Another relevant remark is that, as argued above, our concept of discrimination is close to Lang's discrimination by language, since sometimes language cannot be used to better coordinate. However, in Section 6, we discuss how our setup can be made consistent with other explanations of discrimination: discrimination based on taste (BECKER [1957]) and statistical discrimination (PHELPS [1972], and ARROW [1985a and b]).

3 A more specific setup

The remainder of the paper is devoted to understand under which exact conditions these different situations occur. Time is continuous and agents are risk-neutral. Let y be the optimal production of a job, and let $\widetilde{y} < y$ be the output produced by an untrained worker or a new entrant. It is assumed that both financial resources and efforts lead to switch from the low productive to the high productive state. Consistently with the previous section, we denote by e the *learning* effort of the worker, and by i, the investment in training by the firm at the entry level. Hereafter, "investment" will mean investment by the firm in the training of the worker, while effort will mean the investment in human capital by the worker.[10]

In absence of training, a worker may remain lowly productive for a long time. As explained earlier, a CES investment function.

$$\lambda(e,i) = (a_0 + a_1 e^{\rho} + a_2 i^{\rho})^{\alpha/\rho}$$

captures the main channels through which e and i interact: the trick here is to assume that λ is the Poisson intensity at which the match becomes produc-

10. Note that training leads workers to acquire *specific* human capital. *i.e.* skills which cannot be used in another firm. See BECKER [1964] for the fruitful distinction between general and specific skills.

tive.[11] Note that this is consistent with the assumption of unverifiability of training decisions. The stochastic assumption of the efficiency of learning effort is in the spirit of the Bayesian inference of JOVANOVIC and NYARKO [1996]. The function $e \mapsto \lambda(e,.)$ can be thought as a learning curve, the function $i \mapsto \lambda(.,i)$ as a kind of production function. Assuming a_1 and a_2 to be positive parameters, the higher the two inputs (effort and investment), the lower the expected duration of the low productivity state.

The effort and investment decisions are made at the entry level, and cannot be observed at the time of decision, or verified in any way afterwards. The firms' motivation for providing training is profit maximization. The workers'-motivation for making the effort of learning is to get a higher wage when the surplus of the match is higher: the higher the anticipated wage increase, the higher the learning effort.

The sequence of events is the following: workers and firms meet. The worker receives an entry wage $\widetilde{w}$ that is exogenous (it may be interpreted as a minimum wage). Workers and firms anticipate that, after the training period, the wage will be bargained through a surplus sharing rule. Anticipating this wage, workers and firms choose their reciprocal investment taking the reaction of the other player as given: both e and i are chosen such as to maximize the present discounted value of the job for the worker and the firm respectively.

Given the unverifiability of efforts and investments, and the fact that they cannot be deduced for sure from the time of transition to the higher productive state (this is a typical situation of contract theory in which the observed outcome is a combination of luck and effort), we don't allow in our model for the possibility of a contract that would insure the optimality of training decisions. In fact, and in line with the incomplete contract literature (*e.g.* see MALCOMSON [2000]), it can be demonstrated that in this theoretical setup, decentralized investments are always lower than efficient investments. The only case in which one of the partner invests the efficient amount of training is when he/she receives the integrality of the productivity increase, a classical hold-up case.[12]

At a steady-state, the asset values of the state "unemployment", "employed low productive", "employed high productive" and of a job with low productivity and a job with high productivity are respectively:[13]

$$\begin{aligned} r\widetilde{E} &= \widetilde{w} + \lambda(e,i)(E - \widetilde{E}) + s(U - \widetilde{E}) \\ rE &= w + s(U - E) \end{aligned} \tag{3}$$

11. See WASMER [1999] for this setup and some macroeconomic consequences of training decisions.

12. This result was proved within this setup by WASMER [1999].

13. Continuous time is not needed here, since there are actually three periods: the entry level where investment is made, the low productive state and the high productive state occuring randomly. However, we decided to keep this continuous time structure to introduce the usual matching framework that would allow us to obtain unemployment as the outcome of job creation/job destruction, and also to conveniently introduce the impact of training decisions. A discrete time setup is more involving, since we need to check that the number of match per period is below the initial stocks of job seekers and job vacancies.

and

$$
\begin{aligned}
r\widetilde{J} &= \widetilde{y} - \widetilde{w} + \lambda(e,i)(J - \widetilde{J}) + s(V - \widetilde{J}) \\
rJ &= y - w + s(V - J)
\end{aligned} \tag{4}
$$

where $\widetilde{w}$ and w are the wages in each state. These arbitrage equations read as the equality between the equity value of being in a given state and the flow income given instantaneous transition probabilities of moving to other states.[14]

The wage is bargained in the productive stage according to a generalized Nash-bargaining game. The bargaining parameter of workers is denoted by β, and the wage follows the conventional rule

$$w = (1 - \beta)rU + \beta(y - rV)$$

TABLE 7
Baseline parameter values

rU	rV	$\widetilde{y}$	y	$\widetilde{w}$	β	r	s	a_0	a_1	a_2	$e_0 = i_0$	$e_1 = i_1$
1	1	0	3	0.5	0.5	0.05	0.1	0	0.5	0.5	0.1	1.5

Replacing, we have that

$$E = U + \beta(y - rU - rV)/(r + s) \tag{5}$$

$$J = V + (1 - \beta)(y - rU - rV)/(r + s) \tag{6}$$

Given the discrete structure of bilateral training decisions according to assumption (2), we have a typical game theory setup in which workers and employers play strategically. Agents face a finite number of possible strategies.[15] Then, comparing payoffs in each case, we examine all the possible solutions of the training investment game.

The payoff of this game in each four cases is easy to compute. It is given by $\widetilde{E}(e,i) - e$ (resp. $\widetilde{J}(e,i) - i$). Note that, given that the wage is bargained after the realization of the transition to the high productivity state, the wage and accordingly, both E and I are independent of (e,i).

Denoting by $\omega = \widetilde{w} + sU$ and $\pi = \widetilde{y} - \widetilde{w} + sV$, and condensing the notation $\lambda(e_h,i_k)$ by λ_{hk} for $h,k = 0,1,$ the payoff matrix is given by the following table :

14. Our model focusses on the partial equilibrium properties of this game and we take U and V as given. However, in a general equilibrium setup, one would have that U and V depend on transition probabilities (the exit rate of unemployment and the recruitment rate of firms).

15. An analysis of the case in which the decisions of firms and workers are taken in a continuous set is derived in Appendix.

$$(7)\quad \left(\begin{array}{|c|c|}\hline \dfrac{\omega+\lambda_{00}E}{r+s+\lambda_{00}}-e_0;\ \dfrac{\pi+\lambda_{00}J}{r+s+\lambda_{00}}-i_0 & \dfrac{\omega+\lambda_{01}E}{r+s+\lambda_{01}}-e_0;\ \dfrac{\pi+\lambda_{01}J}{r+s+\lambda_{01}}-i_1 \\ \hline \dfrac{\omega+\lambda_{10}E}{r+s+\lambda_{10}}-e_1;\ \dfrac{\pi+\lambda_{10}J}{r+s+\lambda_{10}}-i_0 & \dfrac{\omega+\lambda_{11}E}{r+s+\lambda_{11}}-e_1;\ \dfrac{\pi+\lambda_{11}J}{r+s+\lambda_{11}}-i_1 \\ \hline \end{array}\right)$$

We can now detail how the payoff matrix of the previous section was calculated, by showing in table 7 the values of the parameters. Note also that, if $\widetilde{y}$ was too small, one may have that $\widetilde{E}-e<U$ or that $\widetilde{J}-i<V$; in other words, we need to check that the participation constraints hold for both the worker and the firm. In fact, these constraints are satisfied provided that $\lambda(e_0,i_0)$ is above zero and thus that the gain from the transition to the productive state is large enough. This is why, in table 7, $e_0=i_0>0$. We are now able to formally derive the frontiers and identify the areas of the different equilibrium strategies in the space parameter.

3.1 Frontiers

In adequately choosing the parameter values, we can draw, in the space parameter (ρ,α), the areas of each equilibrium and various frontiers. See figure 1. These equations defining the frontiers are:

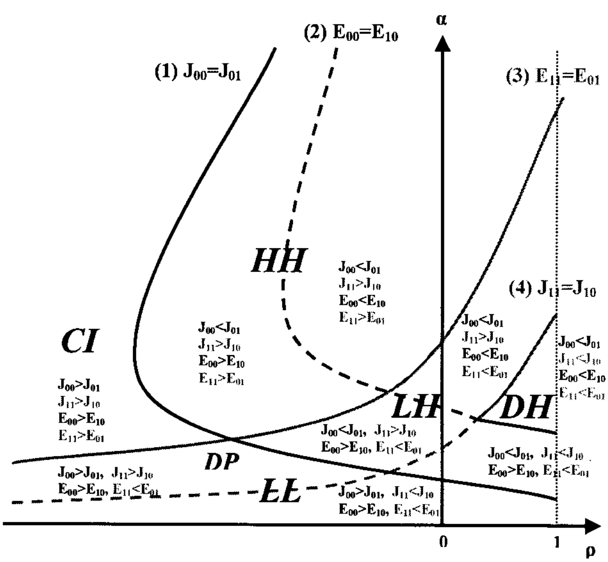
(1) $\widetilde{J}_{00}=\widetilde{J}_{01}\Leftrightarrow i_1(r+s+\lambda_{01})-i_0(r+s+\lambda_{00})=(\lambda_{01}-\lambda_{00})[J(r+s)-\pi]$

(2) $\widetilde{E}_{00}=\widetilde{E}_{10}\Leftrightarrow e_1(r+s+\lambda_{10})-e_0(r+s+\lambda_{00})=(\lambda_{10}-\lambda_{00})[E(r+s)-\omega]$

(3) $E_{11}=E_{01}\Leftrightarrow e_1(r+s+\lambda_{11})-e_0(r+s+\lambda_{01})=(\lambda_{11}-\lambda_{01})[E(r+s)-\omega]$

(4) $\widetilde{J}_{11}=\widetilde{J}_{10}\Leftrightarrow i_1(r+s+\lambda_{11})-i_0(r+s+\lambda_{10})=(\lambda_{11}-\lambda_{10})[J(r+s)-\pi]$

FIGURE 1
Frontiers and equilibria with benchmark values of the parameters



Their interpretation is quite straightforward: the left hand-side is a measure of the marginal costs of investment while the right-hand side is a measure of the marginal gain of this investment (simply write $i_1 = i_0 + di$ and $e_1 = e_0 + de$ and take a first-order approximation for small di and de to see this). The sketch of the calculation of the slope of these curves is in Appendix 1.1.

4 The single worker's type case

4.1 Comparative statics

We have so far treated the outside option of agents (U for workers and V for firms) as parameters, in a partial equilibrium context.[16] Consistently, we analyze the impact of these parameters on the equilibrium strategies.

4.1.1 Higher equity value of unemployment

First, when U increases the value of high productivity employment for workers increases directly and also indirectly through the rise in the negotiated wage. Thus, for workers to be indifferent between making high or low effort, knowing the firm will provide low training i_0 ($\widetilde{E}_{00} = \widetilde{E}_{10}$), efforts and investments should be more complements: frontier (2) shifts to the left. Equivalently, if one is initially on the frontier (2) and that U rises, workers won't be indifferent any longer and will invest ($\widetilde{E}_{00} < \widetilde{E}_{10}$) : they now stand to the right of frontier (2). Similarly, for workers to be indifferent between making low or high effort, knowing this time that firms will provide high training ($\widetilde{E}_{11} = \widetilde{E}_{01}$), efforts and investments should be more substituable (which also means returns to scale should be smaller since the curve is monotonously increasing): frontier (3) shifts to the right.

As for firms, an increase in U reduces profits of high productivity jobs. Thus, for firms to be indifferent between investing or not, knowing the worker will supply low effort ($\widetilde{J}_{00} = \widetilde{J}_{01}$), efforts and investments should be more substitutable: frontier (1) shifts to the right. Finally, for firms to be indifferent between investing or not, knowing workers will provide high effort ($\widetilde{J}_{11} = \widetilde{J}_{10}$), efforts and investments should be more complements (which also means returns to scale should be higher since the curve is monotonously increasing) : frontier (4) shifts to the left.

To sum up, frontiers (1) and (3) in figure 1 shifts to the right while frontiers (2) and (4) shifts to the left. Consequently, the areas where respectively CI and DH prevail enlarge, and the area where LH prevails shrinks. So, an increase in the outside option of workers induces a relative increase in the value of high productivity jobs, inducing equilibria CI and DH to happen

16. Those values can be calculated in general equilibrium using a free-entry condition for firms and introducing an endogenous job finding rate workers. This is left to future work.

more often for various values of α and ρ than before the increase. The effect on the likelihood of HH and LL is ambiguous. Note that for high values of the outside option, the case were only workers invest HL is a possible outcome of the game: the reason is simply that, the higher U, the higher the return to workers'efforts because of higher wage w.

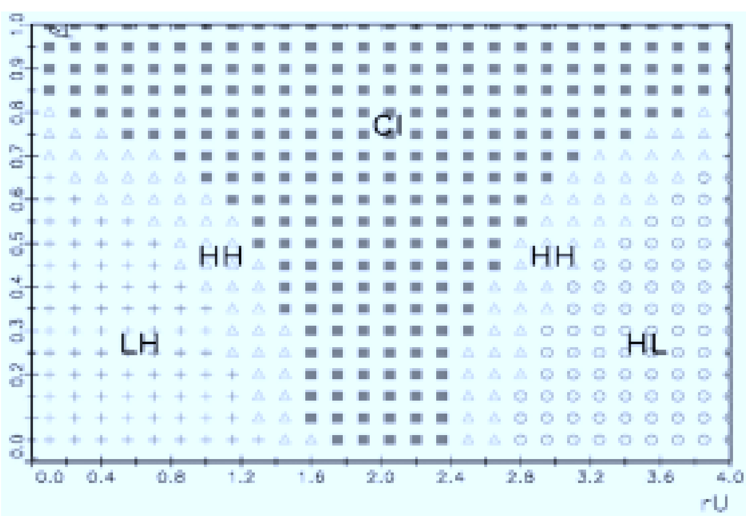
4.1.2 Higher threat point of firms

When V increases, the value of high productivity jobs for firms increases directly and also indirectly through the fall in the negotiated wage. So, frontiers (1) and (3) in figure 1 shift to the left while frontiers (2) and (4) shift to the right for reasons exactly similar to those explained in the analysis of the effect of higher U. Consequently, the areas where respectively CI and DH prevail tend to shrink, and the area where LH prevails tends to enlarge. An increase in the outside option of firms induces a relative increase in the value of high productivity jobs, inciting firms to invest more in training even though workers do not. But by reducing the associated wage, it decreases the value of high productivity employment for workers which, in turn, induces them not to invest in training. The impact on LL and HH is still ambiguous.

4.1.3 Asymmetries in bargaining strength

We can analyze the effect of bargaining power of workers β: an increase in β will have the same impact on the frontiers as an increase in U or a decrease in V, the intuition being that the returns to firms' investments are reduced while payoffs of efforts of workers are raised. To combine the insight of the previous paragraphs, we can derive an alternative representation of the equilibrium: the equilibrium strategies can be represented in the space parameter $(rU; \beta)$ in figure 2. Note that equilibrium HL is now observed as an outcome for large U, as discussed in footnote 6.

FIGURE 2
Different equilibria in the space parameter (rU, β) ***for*** $\rho = -1$ ***and*** $\alpha = 0.75$.



4.1.4 **Productivity of the match**

Finally, when job production y increases, the value of jobs increases for both workers (increased wage) and firms (increased profits). In turn, for both workers and firms to be indifferent between not investing or investing knowing the other party would not invest anyway would require investments and efforts to be more complements (frontiers (1) and (2) shifts leftward). Symmetrically, for both workers or firms to be indifferent between investing or not knowing the other party would invest anyway would take efforts to be more substitutes or returns to scale to decrease (frontiers (3) and (4) shift rightward). So, overall, the area where the HH equilibrium prevails enlarges when economic conditions improve for jobs associated to training. The comparative statics of discount factors $r+s$ is postponed to Section 5.2 to avoid repetitions.

4.2 Average wage earnings

To introduce the discussion of group earnings differentials of the next sections, let us now calculate the average wage of a given group of worker. We denote by $\widetilde{N}$ the number of workers at the entry level of productivity and by $N-\widetilde{N}$ the number of workers with productivity y; N is thus the total number of employed workers. Let's denote by ϕ_{kl}, $k,l=0,1$ the fraction of entry workers in a match where investments k,l are chosen. Denoting by $\Lambda=\sum_{k,l}\phi_{kl}\lambda_{kl}$ the expected probability of transition towards the high productivity state of a given population of entry workers, we have:

$$\partial(N-\widetilde{N})/\partial t=\Lambda\widetilde{N}-s(N-\widetilde{N})$$

In a steady state, flows are in equilibrium and thus $\widetilde{N}/(N-\widetilde{N})=s/\Lambda$ or equivalently,

$$\widetilde{N}/N=\frac{1}{1+\Lambda/s}.$$

The average wage of workers is given by

$$\begin{aligned}\overline{w}&=(\widetilde{N}/N)\widetilde{w}+(1-\widetilde{N}/N)w\\&=w+\frac{1}{1+\Lambda/s}(\widetilde{w}-w)\end{aligned}$$

Given $w>\widetilde{w}$, it is clear that the lower Λ, the lower the average wage, or, naturally, the higher the investments on-the-job, the higher the average wage. Note that, when there is a single Nash equilibrium, $\phi_{kl}=0$ except for k,l of the equilibrium strategies. When there are no pure equilibrium or conversely, multiple pure equilibria, ϕ_{kl} may reflect randomness of choices as in mixed strategies for instance. When there are multiple types of workers, as in the next lines, ϕ_{kl} may reflect the fraction of the groups choosing between the different strategies.

In what follows, we will introduce two types of agents in a Bayesian equilibrium context and analyze wage differences. Then we will discuss various concepts of discrimination in light of our bilateral investment setup.

5 Two types of workers

5.1 Setup

Suppose now that there are two types of workers, who only differ when training investments and efforts have been made: the high productive g (good) yield a revenue y^g, the low productive b (bad) yield a revenue $y^b < y^g$. However, we assume the level of discrete efforts is the same for both types without loss of generality. We also assume that firms have a prior about the distribution of the two kinds of workers: a proportion p is of the low type and a proportion $1 - p$ is of the high type. This prior distribution is common knowledge among firms and workers. Given the structure of the investment game, the possible revisions of these priors will take place after investment levels are revealed and thus have no impact on the decisions made (which are sunk).

The present discounted values of employment $\widetilde{E}^x_{hk}$ and E^x for $x = g,b$, are given by substituting y by y^x in equations (3) and (5), and the present discount value of jobs $\widetilde{J}^x_{hk}$ and J^x is given by the same substitution in equations (4) and (6).

Ex-ante, workers know for sure their payoff as a function of h,k but firms ignore the type x and therefore consider two payoff matrices, each one determining the equilibria. They then select their optimal decision based on the expected value of their payoff, conditional on the optimal strategies being played. For each game, the payoff matrix is as before:

$$\left(\begin{array}{|c|c|}\hline \widetilde{E}^x_{00} - e_0;\ \widetilde{J}^x_{00} - i_0 & \widetilde{E}^x_{01} - e_0;\ \widetilde{J}^x_{01} - i_1 \\ \hline \widetilde{E}^x_{10} - e_1;\ \widetilde{J}^x_{10} - i_0 & \widetilde{E}^x_{11} - e_1;\ \widetilde{J}^x_{11} - i_1 \\ \hline \end{array}\right)$$

Note that the cases $p = 0$, and $p = 1$ are equivalent to the previously studied one-type-of-worker case. We will now show how the value of p will drastically affects the nature of the potential equilibrium.

5.2 Strategy of the firm

Let us introduce the benchmark values of $\rho_0 = -1$ and $\alpha_0 = 0.75$ as a reference point, such that the equilibrium lies in HH in one case ($p = 0$) and in LL in the other case ($p = 1$), as displayed in figures 3 and 4. The reference point is labelled A on these graphs, somewhere close to the

FIGURE 3
p equals 0 : all workers produce $y = 3$

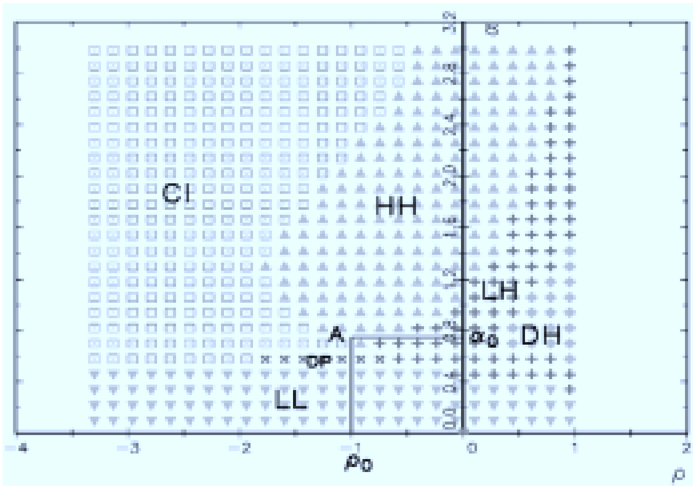
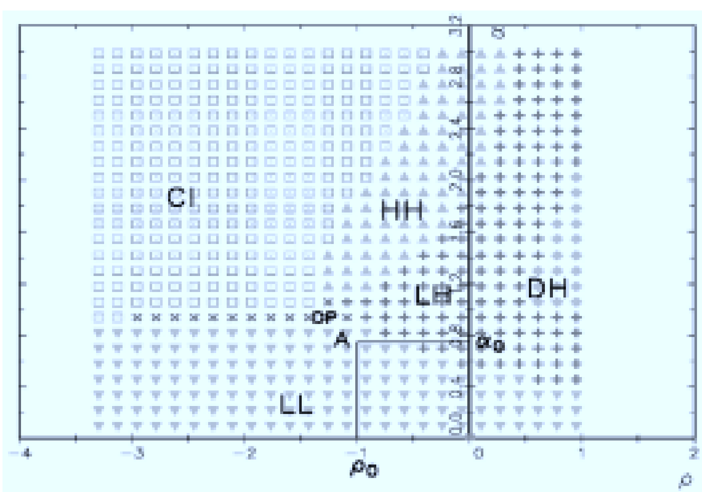


FIGURE 4
p equals 1 : all workers produce $y = 2$



intersection of frontiers (1) and (3) in figure 1. With $p = 0$, we are in HH, and thus have $\widetilde{E}^g_{11} > \widetilde{E}^g_{01}$; $\widetilde{J}^g_{11} > \widetilde{J}^g_{10}$, and additionally $\widetilde{E}^g_{00} > \widetilde{E}^g_{10}$; $\widetilde{J}^g_{00} < \widetilde{J}^g_{01}$. Assume ow that, when $p = 1$, we are in LL with $\widetilde{E}^b_{00} > \widetilde{E}^b_{10}$; $\widetilde{J}^b_{00} > \widetilde{J}^b_{01}$, and additionally $\widetilde{E}^b_{11} < \widetilde{E}^b_{01}$; $\widetilde{J}^b_{11} < \widetilde{J}^b_{10}$.

Accordingly, the payoff faced by a firm if she plays i_1 is given by $p\widetilde{J}_{01}^b + (1-p)\,\widetilde{J}_{11}^g$. Deviating from this strategy and playing i_0 instead, would yield $p\widetilde{J}_{00}^b + (1-p)\,\widetilde{J}_{10}^g$. So the firm will not deviate from strategy i_1 if $p\widetilde{J}_{01}^b + (1-p)\,\widetilde{J}_{11}^g \geqslant p\widetilde{J}_{00}^b + (1-p)\,\widetilde{J}_{10}^g$, which gives, after rearranging, the following condition:[17]

TABLE 8
Conditions on p for low and high equilibria to hold

Equilibrium	$y^b = 1.5$	$y^b = 2$
$\left(\left(e_0^b, e_1^g\right), i_1\right)$	$\forall p$	$\forall p$
$\left(\left(e_0^b, e_0^g\right), i_0\right)$	$p \geqslant 0.36$	$p \geqslant 0.54$

$$(8) \qquad p < \frac{\dfrac{\pi + \lambda_{11} J^g}{r+s+\lambda_{11}} - \dfrac{\pi + \lambda_{10} J^g}{r+s+\lambda_{10}} - (i_1 - i_0)}{\dfrac{\pi + \lambda_{00} J^b}{r+s+\lambda_{00}} - \dfrac{\pi + \lambda_{01} J^b}{r+s+\lambda_{01}} + \dfrac{\pi + \lambda_{11} J^g}{r+s+\lambda_{11}} - \dfrac{\pi + \lambda_{10} J^g}{r+s+\lambda_{10}}}$$

If this is satisfied then there is a unique pure strategy equilibrium $\left(\left(e_0^b, e_1^g\right), i_1\right)$. Note that this condition varies according to the value of parameters, and in particular the difference between y and y^b: the higher the difference between these productivity levels, the lower the threshold level under which the firm will always invest i_0.

Now, playing i_0 would yield $p\widetilde{J}_{00}^b + (1-p)\,\widetilde{J}_{00}^g$ while deviating from this strategy, playing i_1 instead would yield $p\widetilde{J}_{01}^b + (1-p)\,\widetilde{J}_{01}^g$. So, i_0 is an equilibrium strategy for firms if $p\widetilde{J}_{00}^b + (1-p)\,\widetilde{J}_{00}^g \geqslant p\widetilde{J}_{01}^b + (1-p)\,\widetilde{J}_{01}^g$ which gives the following no-deviation condition on strategy i_0 :

$$(9) \qquad p > \frac{\dfrac{\left(\pi + \lambda_{01} J^g\right)}{r+s+\lambda_{01}} - \dfrac{\pi + \lambda_{00} J^g}{r+s+\lambda_{00}} + (i_1 - i_0)}{\dfrac{\lambda_{01}\left(J^b - J^g\right)}{r+s+\lambda_{01}} - \dfrac{\lambda_{00}\left(J^b - J^g\right)}{r+s+\lambda_{00}}}$$

17. This inequality holds as a condition for the firm playing i_1 only if the denominator is positive, as it is actually the case in the simulations below. If it was negative, one would require p to be larger than the ratio of the denominator to the numerator.

In this case there is a pure strategy equilibrium $\left(\left(e_0^b,e_0^g\right),i_0\right)$. Once again, the higher the difference between y and y^b the lower the threshold above which this equilibrium will hold. Using parameter values as in table 7, with $y^g = y$, simulations in table 8 show that above some threshold on p,[18] both types of equilibria can hold.

5.3 Interpretation

One can see that, by superposition of the figures 3 and 4, there are areas of the parameters in which the priors p do not matter, while there are areas in which those priors drastically matter, as is the case for the values α_0 and ρ_0 chosen above. For such values of the parameters, priors of the firms have a huge impact on the behavior of workers. Notably, *even though extreme priors of the firms ($p = 0$ or $p = 1$) lead to a single equilibrium, there are intermediate priors leading to a multiplicity of equilibria*. This indicates the crucial role of the priors of the firms: in the first column of table 8, it is notably shown that, when expectations of the firms are low enough ($p > 0.36$), workers of the good group can anticipate the coordination on a LL equilibrium, while such a case in never possible when expectations are good enough (low p). Of course, the higher y^b, the higher the cutoff point for p above which the second equilibrium occurs. We come back in next Section on the implications of these coordination problems.

We explore more at length in Appendix 1.2 the wage earnings differences between the two groups of workers, and extend this analysis to the case of white workers and black workers.

6 More discussion about discrimination

As argued above, our framework allows for several interpretations in terms of discrimination. We first discuss the most straightforward applications of our setup to the existing literature (taste and statistical discrimination), and then rapidly extend our concept of coordination discrimination in the case of two types of workers.

6.1 Statistical discrimination

There are typically two dimensions of statistical discrimination, one based on adverse selection (truly or wrongly, black workers may be perceived as

18. Inequalities (8) and (9) are based on the fact that, in the subsequent simulations, the term in the denominator is positive for the first inequality, and negative in the second one, which in the latter case induces a change in the direction of the inequality.

less productive) and the other linked to moral hazard (for instance, minority workers or women may exogenously quit more frequently). Let us discuss both dimensions in our model.

That lower productivity leads to lower wage is easy to understand when bargained wages partly reflect productivity. What happens in addition in our analysis is that a lower productivity parameter y is also going to reduce human capital investments. This amplifies wage gaps between black and white workers. Indeed, when y increases, we saw that HH was a less likely outcome: frontiers (1) and (2) shift rightward, frontiers (3) and (4) shift leftward.

Similarly, if firms expect black to be high turnover workers (higher $r + s$), the high investment areas will be reduced. The reason is that an increase in the turnover rates diminishes proportionally more the actualized value of future high productivity jobs than the actualized value of current jobs (the latter being discounted at rate $\lambda + r + s$). So to keep $\widetilde{E}_{00} = \widetilde{E}_{10}$ and $\widetilde{J}_{00} = \widetilde{J}_{01}$, efforts need to be more substitutes: frontiers (1) and (2) shift rightward. Inversely, to keep $\widetilde{J}_{11} = \widetilde{J}_{10}$ and $\widetilde{E}_{11} = \widetilde{E}_{01}$ under these new conditions, efforts should be more complements (or returns to scale should be higher: frontiers (3) and (4)). So, the area where the HH equilibrium prevails shortens to the benefits of the areas where CI, DH and LL respectively prevail.

6.2 Taste discrimination

Becker [1957] introduced taste discrimination as negatively affecting the utility of employers with respect to black employment. In that sense, once a black worker has been matched to an employer, the relative value of the match for the employer with respect to its outside option V is made lower. Let us assume that this is modelled by higher V. We saw in Section 4.1.2 that higher V leads to lower likelihood of the HH equilibria.

In addition, simulations we made in the continuous effort/investment cases, as evidenced in the table displayed in Appendix 8.2, show that the effect of a larger V for a given level of production y leads to lower effort by workers and in many cases, to lower effort by firms, but that the decline in the effort by workers is much larger than the decline in the investment by firms. This is simply due to the fact that a higher V also raises profits of the firms after wage are negotiated, thus raising the payoffs of investments. At the same time, it reduces the payoff of workers, drastically reducing their investments. In other words, when the employer expresses its preference for other workers outside the firm, she is in a better bargaining position, but this discourages her worker and reduces the transition probability to the high productive state. This even reduces its own investment, in virtue of the complementarity between effort and investment.

We have thus an interesting situation in which, because of taste discrimination by employers, workers reduce their learning efforts, while employers do not change their investment. In such a case, punishing discriminatory practices will be very difficult.

6.3 **Coordination under symmetric equilibria but two-type of workers**

In Section 5.2, we saw that, even in situations in which certainty about workers'types leads to single HH or LL equilibria, uncertainty (in that case, $0 < p < 1$) leads to multiple equilibria among which even good types may do the low effort. We can make a parallel between this example and Phelps'[1972] concept of statistical discrimination. In his model, there is some uncertainty about the quality of a group of workers. Let us assume that there are white workers whose post-transition productivity is perceived as being y, and black workers that can be of two types, the low and high types with productivity y^b and y^h described in the previous section, p reflecting the proportion of low types in the group of black workers. In this case, white workers and firms coordinate on the HH equilibrium, but for black workers, some sufficiently bad priors by firms can lead these workers, even of the good type, to anticipate that the firm will chose the low investment, and thus, to select themselves a low level of investment.

7 Concluding comments

We have developed a framework to analyze on-the-job investments by both workers and firms. Competitive forces leading one or the other part to finance all the investment are annihilated by several aspects of the model: first, human capital generated by training is assumed here to be purely specific, while in the case of general human capital investments, BECKER [1964] showed that the workers finance them through lower entry wages. Second, frictions in the labor market reduce the role of outside options of agents in the determination of equilibrium payoffs. Third, the entry wage is exogenous, although this assumption is not crucial here. Finally, effort and investment in training are made at the entry level and, not being verifiable, cannot be contracted upon, leading to incompleteness and inefficiencies. It is firmly believed that these assumptions capture important ingredients of many bilateral relationships, far beyond the labor market. It may for instance be a good benchmark to analyze peace talks in which there are unobservable (or not verifiable by a third-party) efforts really made by negotiators to reduce the activity of their "extremists".

This paper being partly exploratory, we have then discussed the different cases we obtain as a function of the complementarity of the decisions of the two partners as well as of the total returns to scale of these decisions. We have then used this setup to analyze wage earnings differentials and discrimination. We have quite naturally introduced the concept of "coordination discrimination" in which the existence of multiple Nash equilibria in the investment/effort strategies can lead to systematically biased outcomes for some groups. We have discussed how language discrimination can arise in such a case, and also how social penalties for misconduct can reduce/amplify group differences. When further introducing unobserved types of one side of the market (here, workers differing by their productivity), we have showed how perceptions by the other side affect the equilibrium strategies. We finally extend the analysis to statistical and taste discrimination.

Our future work has two main directions. The first one is to derive the general equilibrium properties of the model, in which firms'decisions of creating jobs are endogenous. The second one is to derive the updating of the perceptions from firms of the various fractions of the types of agents in a Bayesian setup. ▼

• References

ALTONJI J. G., BLANK R. (1999). – « Race and Gender in the Labor Market », in *Handbook of Labor Economics* vols 1-3, in Volume 3C, Edited by: Orley C. Ashenfelter and David Card.

ANGRIST J.D., KRUEGER A.B. (1999). – « Empirical strategies in labor economics », – in *Handbook of Labor Economics* Vols 1-3, in Volume 3C, Edited by: Orley C. Ashenfelter and David Card.

ARROW K. (1985a). – « The Theory of Discrimination », in *Collected Papers of Kenneth J. Arrow*, vol. 6. Cambridge, Mass. and London: Harvard University Press, Belknap Press, 1973.

ARROW K. (1985b). – « Models of Job Discrimination », in *Collected Papers of Kenneth J. Arrow*, vol. 6. Cambridge, Mass. and London: Harvard University Press, Belknap Press, 1973.

BECKER G. (1957). – *The Economics of Discrimination*, Chicago University Press.

BECKER G. (1964). – *Human capital*, NBER and Chicago Univ. Press.

BLAU F., KAHN L. (1992). – « The gender earnings gap: Some International Evidence », NBER wp 4224.

CAIN G. G. (1986). – « The Economic Analysis of Labor Market Discrimination: A Survey » – in *Handbook of Labor Economics* Vols 1-3, Volume 1, Edited by: Orley C. Ashenfelter and Richard Layard.

COATE S., LOURY G. (1993). – « Will Affirmative Action Policies Eliminate Negative Stereotypes? », *American Economic Review*, December 1993.

FUDENBERG D., TIROLE J. (1995).– *Game Theory*, MIT Press.

GOLDBERG M. S. (1982). – « Discrimination, Nepotism, and Long-Run Wage Differentials », *The Quarterly Journal of Economics*, vol. 97, n°. 2. (May), p. 307-319.

HARSANYI J., SELTEN R. (1988). – *A General Theory of Equilibrium Selection in Games*. MIT Press.

JOVANOVIC B., NYARKO Y. (1996). – « Learning by Doing and the Choice of Technology », *Econometrica*, 64(6), November, 1299-1310.

LANG K. (1986). – « A Language Theory of Discrimination », *Quarterly Journal of Economics*, vol. 101, n°. 2. (May, 1986), pp. 363-382.

LUNDBERG S. and STARTZ R. (1983). – « Private Discrimination and Social Intervention in Competitive Labor Markets », *American Economic Review*, June, 73, p. 340-347.

MALCOMSON J. (2000). – « Employment Contracts », *Handbook of Labor Economics*, Vol 1-3, in Volume 3C, Edited by: Orley C. Ashenfelter and David Card.

PHELPS E. S. (1972). – « The Statistical Theory of Racism and Sexism », *American Economic Review* (72), 659-661.

PISSARIDES C. A. (2000). – *Equilibrium Unemployment Theory*, MIT Press, Cambridge Mass.

SCHWAB S. (1986). – « Is Statistical Discrimination Efficient? » *American Economic Review*, vol. 76, n°. 1. (March), p. 228-234.

WASMER E. (1999). – « Screening vs. Training in General Equilibrium Search Models ». Paper presented at the Econometric Society Latin American Society in Cancun, 1999.

APPENDIX

1.1 Slopes of the frontiers in the discrete investment case

To obtain the slopes of these curves, we need first to calculate $d\lambda$ with respect to ρ, α. We have

$$\frac{d\lambda(e,i)}{\lambda(e,i)} = d\alpha\left[\lambda(e,i)^{\rho/\alpha}/\rho\right]$$
$$+d\rho\left[-\frac{\alpha}{\rho^2}\lambda(e,i)^{\rho/\alpha} + \frac{\alpha}{\rho}a_1\ln e.e^{\rho} + \frac{\alpha}{\rho}a_2\ln i.i^{\rho}\right],$$

the sign of which is not obvious. One can however see that, the lower the arguments of λ, the more likely $\partial\lambda/\partial\rho$ has opposite sign as ρ.

It follows that, for negative ρ, the effect of an increase in α will raise the importance of the terms $\frac{\alpha}{\rho}a_1\ln e.e^{\rho}$ (resp. $\frac{\alpha}{\rho}a_2\ln i.i^{\rho}$). If $\lambda(e_0,i_0)$ is very small and close to zero, a negative exponent will lead $-\frac{\alpha}{\rho^2}\lambda(e,i)^{\rho/\alpha}$ to be large and negative.

1.2 Continuous reciprocal investment.

1.2.1 First order conditions

We now assume that e,i can be chosen in a range $[0,+\infty[$ and that $\lambda_{ee} < 0$, $\lambda_{ii} < 0$. Using the notation Δx for the difference between x and $\widetilde{x}$ for $x = \underset{\sim}{w}$, y, J and E, the maximization of respectively $\widetilde{E} - e$ for the workers, and $\widetilde{J} - i$ for the firm actually leads to two first order conditions on effort e and investment i, and given (3) and (4):

$$\frac{\partial\widetilde{E}}{\partial e} = \frac{\lambda_e\Delta w}{(r+s+\lambda)^2} = 1 \tag{10}$$
$$\frac{\partial\widetilde{J}}{\partial i} = \frac{\lambda_i(y-\Delta w)}{(r+s+\lambda)^2} = 1$$

Let us denote by $\Phi^e(e,i)$ and $\Phi^i(e,i)$ the functions determining the best response of workers and firms respectively. We have

$$\Phi^e(e,i) = \lambda_e\Delta w - (r+s+\lambda(e,i))^2 = 0$$
$$\Phi^i(e,i) = \lambda_i(y-\Delta w) - (r+s+\lambda(e,i))^2 = 0$$

Differentiating further, we have

$$\mathbf{A}\begin{pmatrix} de \\ di \end{pmatrix} = \begin{pmatrix} \dfrac{\partial \Phi^e}{\partial e} & \dfrac{\partial \Phi^e}{\partial i} \\ \dfrac{\partial \Phi^i}{\partial e} & \dfrac{\partial \Phi^i}{\partial i} \end{pmatrix} \begin{pmatrix} de \\ di \end{pmatrix} = 0$$

with

$$\frac{\partial \Phi^e}{\partial e} = \lambda_{ee} \Delta w - 2(r+s+\lambda)\lambda_e < 0$$
$$\frac{\partial \Phi^e}{\partial i} = \lambda_{ei} \Delta w - 2(r+s+\lambda)\lambda_i = K\lambda_e(\sigma_{ei} - \frac{2\lambda}{r+s+\lambda})$$
$$\frac{\partial \Phi^i}{\partial e} = \lambda_{ei}(y - \Delta w) - 2(r+s+\lambda)\lambda_e = K\lambda_i(\sigma_{ei} - \frac{2\lambda}{r+s+\lambda})$$
$$\frac{\partial \Phi^i}{\partial i} = \lambda_{ii}(y - \Delta w) - 2(r+s+\lambda)\lambda_i < 0$$

with $K = \dfrac{(r+s+\lambda)^2}{\lambda}$ and $\sigma_{ei} = \dfrac{\lambda\lambda_{ei}}{\lambda_i \lambda_e}$. It is this clear that the workers reaction function is increasing if there is sufficient complementarity $\left(\sigma_{ei} > \dfrac{2\lambda}{r+s+\lambda}\right)$ and decreasing otherwise. Note that $\dfrac{2\lambda}{r+s+\lambda}$ is a function of e,i.

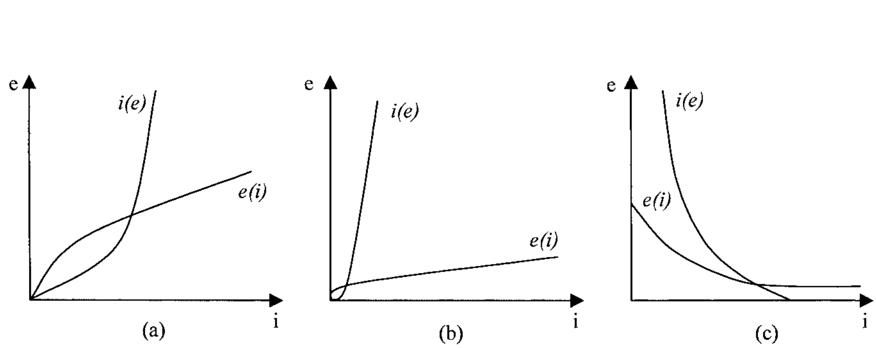
FIGURE 5

Best response curves of the worker and the firm in the continuous case :

(a) training efforts are strongly complement and returns to scale are constant;

(b) training efforts are complement and returns to scale are small;

(c) training efforts are substitute and returns to scale are constant.



One can easily show that an increase in the option value of worker or, equivalently, an increase in the relative bargaining power of workers, would increase the level of wage received when the job become highly productive. Hence, the best response of workers would shift up.

1.2.2 Multiple types, full observability of types

The setup is simply a twofold version of the previous one-population set up. Indeed, upon contracting with workers, we assume that firm know their types for certain, and that this type reveals without uncertainty the exit option values U^h, V^h with $h = B, W$. Suppose, as described above, that $U_B > U_W$ and $V_B > V_W$. Suppose also that both type of workers receive the same wage $\widetilde{w}$ when jobs are not productive. Defining e^h and i^h as the investment levels chosen respectively by workers and firms of type h and firms employing a worker of type h one must simply find the values solving separately the following best response curves systems :

$$\Phi^e(e^h, i^h) = \lambda_{e^h}(w^h - \widetilde{w}) - (r + s + \lambda(e^h, i^h))^2 = 0$$

$$\Phi^i(e^h, i^h) = \lambda_{i^h}(y - w^h + \widetilde{w}) - (r + s + \lambda(e^h, i^h))^2 = 0$$

with

$$w^h = (1 - \beta) r U^h + \beta(y - r V^h)$$

Now, the average rate for both types of populations is simply :

$$\overline{w}^h = w^h + \frac{1}{1 + \Lambda/s}(\widetilde{w} - w^h)$$

since all workers within a group and all firms show identical preferences.

Simulations evidence the impact of taste discrimination for various values of the transition technology parameters ρ and σ. Using the same values for all other parameters as in table 7, except for the exit option value of firms employing type B workers ($rV_B = 1.3$), one gets :

ρ	α	e_W	i_W	λ_W	$\overline{w}_W$	e_B	i_B	λ_B	$\overline{w}_B$
−3.0	1.0	0.62	0.74	0.67	1.37	0.58	0.73	0.64	1.23
−1.0	0.6	0.37	0.52	0.61	1.36	0.33	0.53	0.59	1.22
+0.5	1.0	0.32	1.27	0.72	1.38	0.23	1.47	0.72	1.25

It is remarkable to see that the lower λ for black workers is mostly due to a decline in their own effort, while the lower investment by the firm is proportionally much more limited. This is because wages incorporate the outside option of workers, and that the marginal returns to investment is increased by a larger V for black workers, while the marginal return to black workers'effort is strongly reduced.

1.2.3 Multiple types, incomplete observability of types

We build an incomplete information framework to analyze the impact of statistical discrimination on wage differentials between two types of workers indexed, as above, by $h = B, W$. But now, let us assume that workers of type B can be of two kinds once training investments have been made: the low productive BL, and the high productive BH. Let us assume that firms have a prior distribution of the two kinds of B workers : a proportion p is low productive and a proportion $1 - p$ is high productive. This prior distribution is common knowledge among firms and workers. Given the structure of the investment game, no revisions of these priors will take place after investment levels are revealed.

As for W workers, the setup is unchanged. The W investment Nash solutions (e^W, i^W) solve the following equations:

$$\Phi^e(e^W, i^W) = \lambda_{e^W}(w^h - \widetilde{w}) - (r + s + \lambda(e^W, i^W))^2 = 0$$

$$\Phi^i(e^W, i^W) = \lambda_{i^W}(y - w^W + \widetilde{w}) - (r + s + \lambda(e^W, i^W))^2 = 0$$

As for B workers, firms have to decide what level i^B of training to invest given the distribution of high and low productivity workers:

$$i^B = \arg\max \quad p\left(\widetilde{J}^{BL} - i^B\right) + (1 - p)\left(\widetilde{J}^{BH} - i^B\right)$$

with

$$\begin{aligned} r\widetilde{J}^h &= \widetilde{y} - \widetilde{w} + \lambda(e^h, i^B)(J^h - \widetilde{J}^h) + s(V - \widetilde{J}^h) \\ rJ^h &= y^h - w^h + s(V - J^h) \end{aligned} \tag{11}$$

for $h = BL, BH$. The first order condition gives

$$p\frac{\lambda\left(e^{BL}, i^B\right)(y^{BL} - \Delta w^{BL})}{\left(r + s + \lambda\left(e^{BL}, i^B\right)\right)^2} + (1 - p)\frac{\lambda\left(e^{BH}, i^B\right)(y^{BL} - \Delta w^{BH})}{\left(r + s + \lambda\left(e^{BH}, i^B\right)\right)^2} = 1$$

High and low productivity workers have to decide separately of their effort levels, which requires as before the following condition to be satisfied:

$$\frac{\lambda\left(e^h, i^B\right)(w^h - \widetilde{w})}{(r + s + \lambda(e^h, i^B))^2} = 1$$

for $h = BL, BH$. The Bayesian Nash solution in $\left(e^{BL}, e^{BH}, i^B\right)$ given p, solves the three previous equations.

Finally, intra-group average wages are

$$\overline{w}^W = w^W + \frac{1}{1 + \Lambda^W/s}(\widetilde{w} - w^W)$$

and

$$\overline{w}^B = (\widetilde{N}^B/N^B)\widetilde{w} + \left(\frac{N^{BL} - \widetilde{N}^{BL}}{N^B}\right)w^{BL} + \left(\frac{N^{BH} - \widetilde{N}^{BH}}{N^B}\right)w^{BH}$$

$$= (\widetilde{N}^B/N^B)\widetilde{w} + \left(\frac{pN^B - \widetilde{N}^{BL}}{N^B}\right)w^{BL} + \left(\frac{(1-p)N - \widetilde{N}^{BH}}{N^B}\right)w^{BH}$$

$$= \frac{1}{1+\Lambda/s}\widetilde{w} + p\frac{\lambda\left(e^{BL},i^B\right)}{s+\lambda\left(e^{BL},i^B\right)}w^{BL} + (1-p)\,\frac{\lambda\left(e^{BH},i^B\right)}{s+\lambda\left(e^{BH},i^B\right)}w^{BH}$$

Given parameters in table 7, except with $y^W = y^{BH} = 3$, $y^{BL} = 2.5$ and with $p = 0.5$

ρ	α	e_W	i_W	λ_W	$\overline{w}_W$	e_{BL}	e_{BH}	λ_{BL}	λ_{BH}	i_B	$\overline{w}_B$
−3.0	1.0	0.62	0.74	0.67	1.37	0.54	0.61	0.60	0.65	0.70	1.25