in George R. Feiwel, ed. (1987)
Arrow and the Ascent of Modern Economic Theory

Arrow and the Theory of Discrimination

Henry Y. Wan Jr.*

1 INTRODUCTION

Many years have passed, since the appearance of Arrow's (1973) study of discrimination – a study that is definitive in two senses:

First, it is the coda for the neoclassical literature on this topic, over the half century bracketed between Edgeworth (1922) and Becker (1971). Discrimination involves the efficiency and equity of an entire economic system, under *laissez-faire* as well as under government legislations. To reach general and conclusive results, there is no alternative to the theory of general equilibrium. As one of the founders of the abstract theory of general equilibrium, Arrow sets a high standard for all theorists. By his personal example, he demonstrates that the insights and perspective of abstract studies are ultimately justified by their services for social concerns. Rigorously and exhaustively, Arrow proves for all ages, that under the usually-made assumptions, those who indulge in discrimination must suffer reduced income in the short run. Additionally, only the least-discriminating firms may survive in the long run. In short, within the competitive system, the 'virtuous' always outcompetes the 'vicious'.

Second, the paper is also a clarion call for action. Always the complete scientist and never an ideologue, Arrow does not allow his appreciation of the power of neoclassical analysis to becloud his sense of realism. He passes his judgement with finality, 'since discrimination

* It is always an inspiring experience to work on themes pioneered by Professor Arrow and to use the methodology he fashioned. In 1970 Simone Clemhout and I had the pleasure of applying the theory of learning-by-doing to the area of infant industry protection, following the trail Arrow blazed. This time, my pleasure is two-fold for I apply an analysis that grew out of his theory of moral hazard to the topic of the economics of discrimination – an area where he made the pivotal contribution. We all wish that in the years to come we shall be able to repeat our fruitful exercises, following Arrow into the realm of many unresolved economic issues.

survives, ... the model must have ... limitations'. After pinpointing imperfect information as the fly in the ointment (and incidentally as the wave of the future in economics), Arrow then proceeds with the theory he and Phelps (1972) independently founded – the theory of 'statistical discrimination', the established theory, ever since (see also Aigner and Cain, 1977). This is the theory of the vicious cycle, with two components: (a) the discouraged group of the disadvantaged who would less frequently invest in their own human capital, and (b) the imperfectly-informed employers who perpetuate the expectation that those disadvantaged are less likely qualified to take on responsible functions. It does not rule out the need for government policies that may elevate society from a discriminatory second best to the non-discriminatory first best (see, for example, Lundberg and Startz, 1983). None the less, it is a theory with plenty of victims, but no genuine villain.

It would have been perfect if the above theory explained all there is, within our daily experience. Arrow's own example does not allow us to rest on such complacent thoughts, however. A nagging doubt remains: Despite more than a decade of government efforts – let alone in the absence of them – unequal treatments are meted out to equally qualified persons, based on race, sex and creed, with perfect information about productivity, and unpunished as well as unabated under our competitive system. One might dispute, case by case, the charge of discrimination on the grounds of malice. Or one might entertain the hope that the 'Mill of God' grinds slowly but surely, so that the day of reckoning lies ahead for the miscreant. None the less, following in Arrow's footsteps, we must search for some alternative, but internally consistent, explanations for what we have long suspected. There may be cases where, if the public does not act, there will be discrimination, but never punishment. Such a quest may lead us down novel and untrodden paths, but here again, Arrow (and Hahn, 1971) has set a precedent by exploring various alternative models, including a model of disequilibrium.

Many years have passed. In economic analysis too generations of newly-cast artillery have arrived at the breach. The new equipment *par excellence* within our arsenal pertains to the incentive compatibility constraints. Its lineage traces back to Arrow's (1963) theory of moral hazard (see Radner, 1982, for the 'roots' of the principal-agent model). The question then is: With new tools is the profession ready for new discoveries? The challenge then is: to prove analytically that competitive forces do not right all wrongs and that affirmative legislation is needed to end discrimination. And, is there much promise in our quest?

Our answer is positive but somewhat tentative. An example is provided where employer discrimination can carry on with impunity, under competive conditions. In fact, discrimination can go on in two alternative forms.

The basis of our example is information asymmetry in the principal-agent model, as in Hurwicz and Shapiro (1978), Harris and Townsend (1981), and Foster and Wan (1984a,b). That model is slightly modified in Section 2 and discrimination will take the 'first-to-fire, last-to-hire' form towards the disadvantaged group. Unlike in the statistical theory of discrimination in this model, the employers are always perfectly informed that there is no behavioural difference between the disadvantaged and the privileged groups and the employers who discriminate do not fare any worse in whichever possible way relative to those who do not discriminate.

In Section 3, we vary the previous example to allow for two types of jobs, each producing a particular 'productive service' and the combination of the latter yields the final output. One type of job admits perfect monitoring; the other type is shirking-prone so that superior performance is encouraged with bonus payments that include a 'bribe'. At equilibrium, the former job yields less utility to workers than does the latter, yet the incentive compatibility constraint bars 'arbitrage', since any pay reduction on the latter job is against the employer's interest. Which worker is assigned to what job is in the power of the employer – a power that may be abused to satisfy his taste to discriminate.

Finally, in Section 4, to provide a deeper perspective, we explore the relationship between our model and the conventional model of competitive equilibrium.

2 THE 'BASIC' MODEL AND DISCRIMINATORY EMPLOYMENT

Foster and Wan (1984a) shows that in an economy with M identical firms and L units of homogeneous labour, information asymmetry may cause each firm to hire N units only, leaving $L\text{-}MN$ jobless. While the jobless fare worse than those working, they cannot get hired by wage concessions. This happens because firms, accepting their own inability to monitor, pay voluntarily a bonus to deter shirking. As it is in the interest of the employer to reward more, it is futile for the jobless to offer to be paid less. In our context, any employer who rations coveted jobs among workers with identical productivity, can also discriminate with impunity, on grounds totally irrelevant to production.

The following summary of the gist of the example in Foster and Wan (1984a) helps make our discussion self-contained, and prepares the ground for Section 3:

We assume that the output of a worker depends on the effort and status of that worker and the number of workers in the firm. Specifically, the output is proportional to effort. The worker's status may be t (for tired) or h (for healthy), with the output under h higher than the output under t by a factor $a > 1$, other things being equal. Finally, congestion reduces exponentially the output per worker, as the number of workers increases. The (net) utility of the worker is the utility of reward minus the disutility of work; specifically, it is the reward in output units minus the square of effort. Firms know only the probability of workers' statuses, distributed identically and independently over all workers, but not the status of any particular worker. They know the workers' preferences, not any individual's effort. Workers know their own efforts, and their own statuses just before they decide their own efforts. Unable to deduce exactly the worker's effort, firms reward workers by output. They select the size of their labour force and the (possibly non-linear) reward schedule to maximize their expected profit, which is output minus reward. Since workers have two statuses only, in setting the reward schedule a firm focuses on output and reward targets, for each of the statuses, making sure that workers would neither 'quit' nor 'shirk'. Quitting means zero effort, hence zero output, so that even by paying zero reward, the firm nets zero profit. For prevention, the reward schedule must be sufficiently generous in absolute terms, for each status s. Shirking means the concealing of the true status s, and producing the target output for status $s' \neq s$, so that one may become better off under the reward schedule. This subverts the firm's plan and generally reduces its expected profit. For prevention, the reward schedule must be sufficiently generous, in relative terms, for each true status s, *vis-à-vis* any other $s' \neq s$, corresponding to all (s,s') pairs. These form the individual rationality (*IR*) and self-selection (*SS*) constraints for the target rewards and outputs.

Here, we come to the crux of the matter for Foster and Wan (1984a). Congestion means diminishing returns for employment. Firms hire no more than their equilibrium size of labour force, unless workers make concessions on contract terms. The jobless may be ready to make concessions to modify the equilibrium reward schedule. Yet the schedule is determined by *IR* and *SS* constraints. Concessions violating *IR* will not be made: or else working is worse than quitting. Concessions violating *SS* will not be accepted: such concessions have no way to be enforced, given information asymmetry favouring the worker. Hence, 'involuntary' unemployment may exist in an equilibrium, and the coveted jobs must be rationed among a larger number of equally qualified applicants.

Such a situation is avoidable, in principle, either by a heavy application fee, exacted from the workers, or by permitting negative reward for low outputs. Both work by end-running the *IR* constraint. In reality, the first is financially infeasible for the workers, and the second is legally unenforceable due to the prohibition of human bondage.

We now supply the details of the example. Let $y(s)$, $r(s)$, $z(s)$ and $Z(s)$ be the target-output, target-reward, target-effort and the disutility associated with such effort, for a worker in status $s=t$ or h. Let $Q>0$ (with $L-MQ>0$) be a constant, then, the production function is:

$$y(t)=z\exp(-N/Q)$$
$$y(h)=az\exp(-N/Q),$$

the utility index for the worker is:

$$\begin{aligned} u&=r-Z\\ &=r-z^2, \end{aligned}$$

the effort requirement per output is:

$$k(s,N)\equiv\begin{cases} \exp(N/Q) & s=t\\ a^{-1}\exp(N/Q) & s=h \end{cases}$$

the worker's utility in fulfilling the status s' target, when the actual status is s, is:

$$u(s',s,N)=r(s')-y^2(s')k^2(s,N),$$

so that the problem for the firm is:

$$\max_{N\geqslant 0}\quad \max_{r(s),y(s)}\quad \mathrm{E}\{y(s)-r(s)\}=P$$

$$
\text{s.t.}\quad
\begin{array}{lllll}
u(t,t,N) & & & & \geqslant 0 \\
 & u(h,h,N) & & & \geqslant 0 \\
u(t,t,N) & & -u(h,t,N) & & \geqslant 0 \\
 & u(h,h,N) & & -u(t,h,N) & \geqslant 0
\end{array}
\quad
\begin{array}{l} IR \\ \\ SS \end{array}
$$

$$N, y(s), r(s) \geqslant 0$$

where $E(\cdot)$ is the expected value operator.

Alternatively, one can transform the above problem to:

$$
\underset{N \geqslant 0}{\text{Max}} \quad \underset{r(s),\, y(s)}{\text{Max}} \quad E\{-r(s) + Z^{1/2}(s)/k(s,N)\} \tag{1}
$$

$$
\text{s.t.} \quad
\begin{pmatrix}
1 & -1 & 0 & 0 \\
0 & 0 & 1 & -1 \\
1 & -1 & -1 & a^2 \\
-1 & a^{-2} & 1 & -1
\end{pmatrix}
\begin{pmatrix}
r(t) \\ Z(t) \\ r(h) \\ Z(h)
\end{pmatrix}
\geqslant 0
$$

$$N, r(s), Z(s) \geqslant 0.$$

Given N, the above is a concave programming problem with a polyhedral constraint set. Its optimal solution can be shown as:

$$
\begin{aligned}
r(t) &= a^4(1-p)^2 \exp(-N/Q)/4(a^2-p)^2 \\
r(h) &= a^2((a^2-p)^2 + (a^2-1)(1-p)^2)\exp(-N/Q)/4(a^2-p)^2 \\
Z(t) &= a^4(1-p)^2 \exp(-N/Q)/4(a^2-p)^2 \\
Z(h) &= a^2 \exp(-N/Q)/4
\end{aligned}
$$

where p is the probability of any worker being in the status h, and $0 < p < 1$. The dual vector is: $(N, 0, 0, Np)$. The optimality property may be verified by the saddle-point criterion for the Lagrangian (see Foster and Wan, 1984b).

We can then show that the optimal value of N is Q.

The essence of this exercise is that, at the equilibrium:

(i) Unemployment exists.

(ii) Such unemployment is involuntary, since for the jobless, the expected utility is zero. For those working, it is:

$$u_0 = p[r(h) - Z(h)] = pa^2(a^2-1)(1-p)^2/4(a^2-p)^2\exp(1) > 0.$$

3 THE EXTENDED MODEL AND DISCRIMINATORY ASSIGNMENT

We now turn to the case where between two equally qualified persons, one privileged and one disadvantaged, the latter is not left jobless, but assigned a job yielding a lower level of expected utility than the one assigned to his privileged counterpart. This appears similar to the job segregation phenomenon of Bergman (1971). For such a situation to persist, there must be some barrier against the disadvantaged to take the better job. Thus, it seems this is a case close to our example in the last section. Call the less desirable job X and the better job Y. The same mechanism which prevents the wage structure to fall in Y may resemble what prevails in Section 2.

Seeking insights and not generality at this stage, we shall expand the example in Section 2 in two ways: First, change the nature of the product in Section 2 from final good to intermediate good, and call it Y-good. Secondly, introduce job X to absorb all workers not working on Y-good. Their output will be called X-good. X-good and Y-good are combined to form the universal consumption good, G, by some production function. This is assumed to be a constant returns, Cobb-Douglas form: $G = C\sqrt{XY}$, which is probably as simple as one can get.

We assume that Y is produced in a way similar to the final good in Section 2, with one exception: v, the unit value of Y, is no longer a constant.

Next we assume that the production of X causes no disutility, and depends on neither the worker's status, nor effort; presence alone is what matters. We can select unit for good X, so that the output of a worker is unity. Given the form of the worker's utility function, the utility yielded by job X is equal to the wage rate of job X, which is also the same as the unit value of the service of X. This value will be called w.

We now utilize the property of the specific Cobb-Douglas function, that the total value of X, always equals the total value of Y:

$$wX = vY.$$

or,

$$\begin{aligned} w(L - MQ) &= vMQEy\ (s), \\ &= vMQy_0 \end{aligned} \tag{2}$$

say.

The last equation is derived under the assumption, that like in Section 2, we have that particular type of equilibrium where each of the M firms will hire Q units of labour. In such an equilibrium, X-jobs should yield less utility than Y-jobs. Our intuition indicates this will be the case for large value of L/M. So we test it below.

From Section 2 the expected utility for a worker on a Y-job is:

$$u_0 v \tag{3}$$

The fact that y_0 and u_0 in Section 2 should be replaced by vy_0 and $u_0 V$ in (2) and (3) can be verified by replacing $1/k(s, N)$ with $v/k(s, N)$ in (1) Section 2. Our previous calculation states that the utility for a worker on an X-job is;

$$w = \frac{MQy_0}{L - MQ} v \qquad \text{(by (2))}$$

Hence,

$$w \leqslant u_0 v$$

if, and only if:

$$\frac{Q}{(L/M) - Q} \leqslant \frac{u_0}{y_0}$$

which is equivalent to:

$$L/M \geqslant \frac{1 + (u_0/y_0)}{(u_0/y_0)} Q\text{'}$$

which clearly substantiates our intuition.

In other words, when the supply of labour is abundant relative to the 'means of production' (as represented by the outfit of plant and equipment owned by the firms), and when some jobs (by the information structure) tend to promise higher utility levels for the worker, firms can discriminate with impunity.

The 'technocratic infrastructure' ushered in by the rise of our quintessential market economy seems to generate Y-jobs in various parts of the economy. The monitoring of the effort intensity in mental endeavours is difficult, both in actual and fictional form, (even in 1984!) Consequently, one cannot trust the market mechanism alone to deal with discriminatory job assignments.

4 FINAL REMARKS

We now relate our examples within the context of Arrow and Debreu (1954) in three remarks:

First, the discriminatory equilibria here *are not* the competitive equilibria of the literature. Discrimination takes various forms, two of which are considered in this paper: In one, a job is denied to the disadvantaged, but is open to members of the privileged group, with exactly the same qualifications. In the other, two equally qualified persons hold two jobs which differ in pay and amenities, with the disadvantaged being worse off than the privileged. Both seem to be quite common. The latter case corresponds to the notion that the powerful favours his favourite. The former case is also frequent grist for the journalists' mill: newspapers often highlight the high unemployment rates for the disadvantaged, and such unemployment must be involuntary. Were the choices between working and not working indifferent to the marginal worker, such news would never be newsworthy. Both cases imply that the labour market is not cleared, hence, not in the equilibrium of either Marshall or Walras. However, they are equilibrium positions in a principal–agent model, and, therefore, in a game theoretic model. (Recall that one period principal–agent models are games in two moves, as Radner (1981) noted.)

Secondly, the competitive equilibrium *is* a particular type of market game equilibrium. We now clarify the nature of the noeclassical competitive equilibrium from some new vantage point. Consider the market for the labour service provided by one particular kind of household. For simplicity, assume there is a continuum of such workers, none of them working part time. Moreover, assume that leisure generates no utility, but work generates disutility, the intensity of which varies from job to job. The first implication of these is that the labour supply is always constant. The second implication is that the market co-ordinates labour allocation by a signal of labour scarcity other than a single wage rate.[1] The presence of the compensating wage differential means that workers consider alternative job offers according to the utility levels these jobs respectively promise. Hence, for the derived demand schedule of labour for any firm, employment quantity may be plotted against the expected utility this firm is willing to promise its employee, and not the wage rate, in the more general case. Consider now the contours of the maximum (expected) profit over the first quadrant for (N, r) pairs, where N stands for the size of labour force and r stands for the actual utility promised to the worker. Such contours in the neclassical context are concentric loci of the horseshoe shape, opening at the bottom. This shape means that profit declines when the employment deviates from its optimal value in either direction, but profit always improves if the firm can promise less utility to the worker. Since the market signal is the minimum utility, R, which a firm must promise, the second implication above assures that the firm will promise exactly that utility level, no more and no less. No less because otherwise no employee will be forthcoming; no more because otherwise less profit will be earned. When the disutility of work is not job-specific, the wage serves well as the market signal and we return to the familiar case: competitive firms pay market wages. What is usually not realized is that this result follows two separate rules of not paying more, and not paying less, each with its own different reasoning. The two – as one probably will expect – do not always hold true together as we shall see later. Presently, the above discussion is illustrated with Figure 14.1. The line joining the apex points of all iso-profit loci is the derived demand curve.

Finally, our solution concept in the example of Sections 2 and 3 *may be regarded as* a generalization of the concept of competitive equilibrium.[2] We shall show that it is not always in the interest of the employer to promise the employee a barely adequate utility level to attract him. Under competitive equilibrium, employers always promise only the minimum. In Figure 14.1, the employer's optimal (N,r) pair is always found at the boundary of the feasible set. For our examples, say, the one in Section 3, an interior optimum may arise. But then, a boundary optimum may also happen. To make our point, introduce an additional equality constraint on the employer's programming problem in Section 2:

$$E[r(s) - Z(s)] = r,$$

and define the maximized value of P as $P(N,r)$. It can be shown (see Foster and Wan, 1984b) that for each r, there exists a unique N where $P(N,r)$ is a maximum. The iso-profit contours are nested 'simple closed

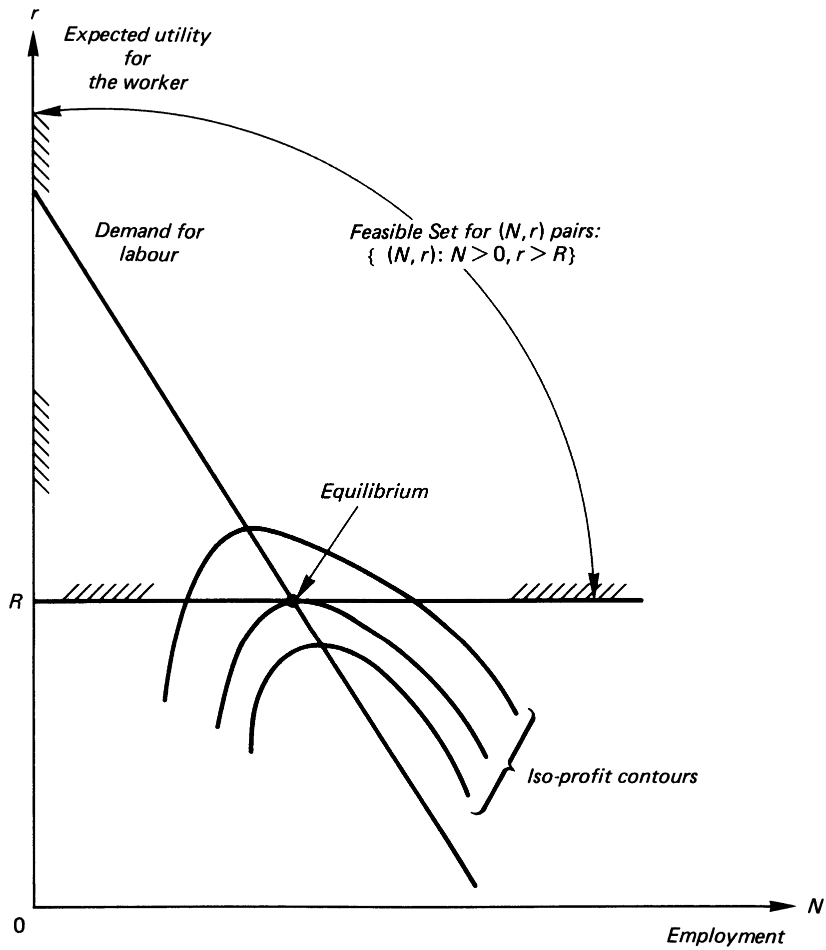

Figure 14.1 Labour market: conventional model of competitive equilibrium

curves', with an unconstrained maximum located at (N_0, r_0), for some $r_0 > 0$, as in Figures 14.2 and 14.3. In the example of Section 2, $R = 0$, showing that the unemployed workers have no alternative. In the example of Section 3, a loose labour market will have $0 < R < r_0$ as in Figure 14.2, and a tight market will have $r_0 \leqslant R$, as in Figure 14.3. The employer has an interior solution in the former, offering the opportunity to discriminate with impunity. He has a boundary solution in Figure 14.3, quite similar to the situation in Figure 14.1, under the

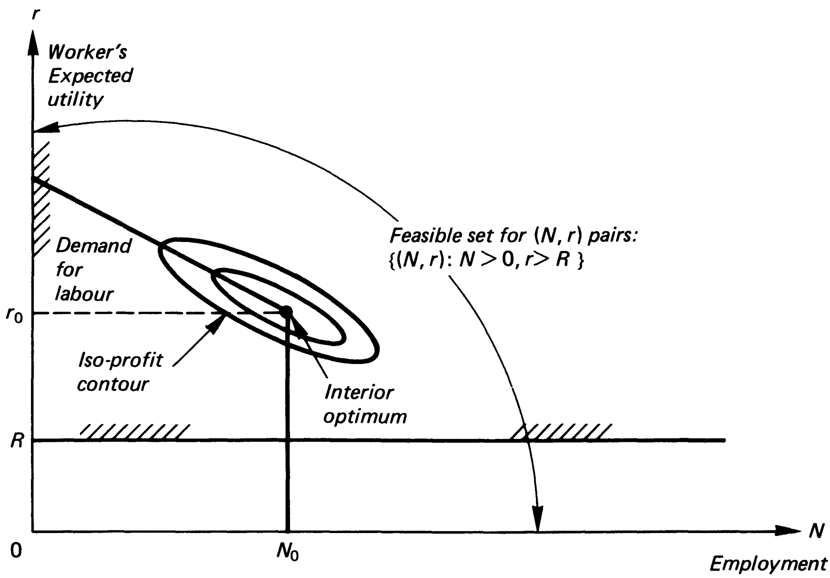

Figure 14.2 An extended model with loose labour market

conventional model of competitive equilibrium. *R* is the expected utility for workers in the *X*-sector in Section 3. Here, discrimination is a costly taste.

One should also note that information asymmetry is not the only reason for the employer to be interested in rewarding the employee at more than the market rate. This is so also because of *inter alia* (a) physical reasons; better pay means better health and better working potential, (b) social reasons; a certain wage structure is regarded as conducive to high productivity, for example, the rule-of-thumb restriction:

the foreman's pay/the subordinate's pay $\geqslant 1.2$,

which is gaining credibility among personnel managers (see, for example, Wan, 1973). Each and every such case implies that the wage structure affects productivity, that workers on certain jobs fare better than their peers, and that which one of a large number of equilibria[3] will prevail depends upon the employer's whim. Here lies the source of discrimination that the competitive force cannot redress.

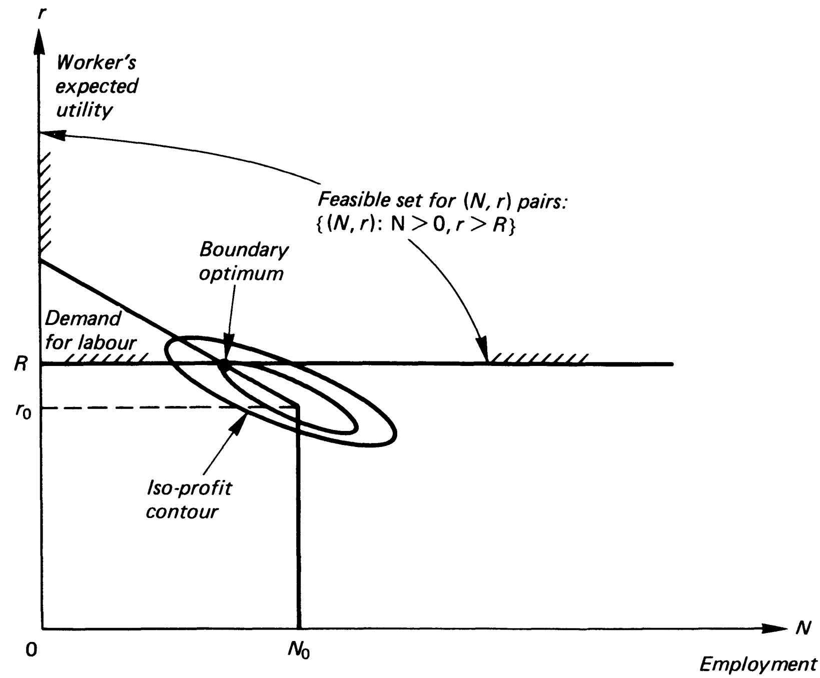

Figure 14.3 An extended model with tight labour market

NOTES

1. An alternative is to give up the concept of an aggregate market demand schedule for labour, which is summed up over the needs of all 'jobs'. Then, one can follow Arrow and Debreu (1954), treating the demands for the same type of labour to fill different jobs, as demands for different types of labour services, and a household may supply some labour for more than one job type. But if we treat the same work at different effort intensities as different labour types, we may have to deal with the added complexities of infinite-dimensional commodity spaces; the effort level can take any non-negative values. Thus, for expository purposes, our approach also has its advantage.
2. As Radner (1982) notes, most of the extant axiomatic theory of competitive equilibrium only deals with exogenous risks, and not endogenous moral hazard like 'shirking' in our model. Prescott and Townsend (1984) is an exception for the linear technology, zero profit case. Since their firms are indifferent about employment size, their contribution cannot be applied for our purpose in Sections 2 and 3 for analyzing discrimination.
3. Each distinguished by which worker fills what job, if any.

REFERENCES

Aigner, D. J. and G. G. Cain (1977) 'Statistical Theories of Discrimination in Labor Markets', *Industrial and Labor Relations Review*, 30: 175–87.

Arrow, K. J. (1973) 'The Theory of Discrimination', in O. Aschenfelter and A. Rees (eds) *Discrimination in Labor Markets* (Princeton: Princeton University Press).

Arrow K. J. (1963) 'Uncertainty and the Welfare Economics of Medical Care', *American Economic Review*, 53: 941–73.

Arrow, K. J. and G. Debreu (1954) 'Existence of an Equilibrium for a Competitive Economy', *Econometrica*, 22: 265–90.

Arrow, K. J. and F. H. Hahn (1971) *General Competitive Analysis* (San Francisco: Holden-Day).

Becker, G. S. (1971) *The Economics of Discrimination* (Chicago: University of Chicago Press).

Bergman, B. G. (1971) 'The Effect on White Incomes of Discrimination in Employment', *Journal of Political Economy*, 79: 194–313.

Edgeworth, F. Y. (1922) 'Equal Pay to Men and Women for Equal Work', *Economic Journal*, 32: 431–57.

Foster, J. E. and H. Y. Wan, Jr. (1984a), 'Involuntary Unemployment as a Principal–Agent Equilibrium', *American Economic Review*, 74: 476–84.

Foster, J. E. and H. Y. Wan, Jr. (1984b) *Competition Restores Efficiency*, mimeo.

Harris, M. and R. M. Townsend (1981) 'Resource Allocation Under Asymmetric Information', *Econometrica*, 49, 33–64.

Hurwicz, L. and L. Shapiro (1978). 'Incentive Structures Maximizing Residual Gain Under Incomplete Information', *Bell Journal of Economics*, 9: 180–91.

Lundberg, S. and R. Startz (1983) 'Private Discrimination and Social Intervention', *American Economic Review*, 73: 340–47.

Phelps, E. S. (1972) 'The Statistical Theory of Racism and Sexism', *American Economic Review*, 62: 659–61.

Prescott, E. C. and R. M. Townsend (1984) 'Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard', *Econometrica*, 52: 21–46.

Radner, R. (1981) 'Monitoring Cooperative Agreements in a Repeated Principal–Agent Relationship', *Econometrica*, 48: 1127–48.

Radner, R. (1982) 'Equilibrium Under Uncertainty' in K. J. Arrow and M. D. Intriligator (eds) *Handbook of Mathematical Economics*, vol. 2 (Amsterdam: North Holland).

Wan, H. Y. Jr. (1973) *A General Theory of Wages, Employment and Human Capital*, Department of Economics Working Paper, no. 51 (Cornell University).