‘Phelpsian’ statistical discrimination: A brief history of thought

Ludvig Sinander
University of Oxford

drawing (toward the end) on work with

Matteo Escudé, Paula Onuchic, Quitzé Valenzuela-Stookey

‘Perspectives on Economic Theory’ conference
LSE, 4 June 2025

Economists on labour-market discrimination

Theory:

- first contribution, it seems: Edgeworth (1922)
- very influential: Becker (1957)
- surveys: many, recently Onuchic (2023)

Empirics: large lit.

Statistical discrimination

Two quite distinct strands of thought:

- equilibrium theories following Arrow (1973)
- pure inference theories following Phelps (1972a, 1972b)

Both called 'statistical discrimination'.

Today: the latter.

‘CliffsNotes’

Plot:

1	Idea (vaguely)	Phelps, 1972a, 1972b
2	Clarification (uncharitably)	Aigner–Cain, 1977
3	Modernisation (mathematically)	Chambers–Echenique, 2021
4	Revision (Blackwellly)	Blackwell, 1951, 1953

Some themes:

noisy signals	$\rightsquigarrow$	random beliefs
parametric models	$\rightsquigarrow$	‘flexible’ models
economies & games	$\rightsquigarrow$	decision problems
worry about / maximise $\mathbb{E}$	$\rightsquigarrow$	worry about / maximise min
econ with formalisation	$\rightsquigarrow$	maths with applications

Plot

Introduction	
Idea (vaguely)	Phelps (1972a, 1972b)
Clarification (uncharitably)	Aigner–Cain (1977)
Modernisation (mathematically)	Chambers–Echenique (2021)
Revision (Blackwellly)	Blackwell (1951, 1953)
Conclusion	

Setup I: workers

Lotta workers. Each worker has

– a skill type $\in \Theta$

– a social identity $\in \{A, B\}$. Speak of 'group A' & 'group B'.

Use 'probability / $\mathbf{P}$' as shorthand for 'fraction of workers'.

Assumption: groups have same skill distribution:

$$\mathbf{P}(\text{skill} = \theta | \text{identity} = A) = \mathbf{P}(\text{skill} = \theta | \text{identity} = B) \quad \forall \theta \in \Theta.$$

Assumption: firms care about skill, not identity.

$\implies$ if firms observe skill, then HR decisions $\perp$ identity.

'HR decisions': hiring, task assignment, pay, ...

No claim that assumptions are realistic. A thought experiment.

Setup II: information

Assumption: firms do not observe skill. Only observe

- identity
- a (possibly multi-dimensional) covariate $\in \mathcal{C}$
  (e.g. CV, test scores, ...)

Describe identity, skill & covariate as 'random variables' with some joint (cross-sectional) dist'n.

To inform HR decisions, firms must guess skill based on observables.

Assumption: firms are correctly-specified Bayesians. That is, for worker with observables $(\text{identity}, \text{covariate}) = (i, c)$, firm's (subjective) probability $p(\theta|c, i)$ that this worker has skill $= \theta$ is

$$p(\theta|c, i) = \mathbf{P}(\text{skill} = \theta|\text{identity} = i, \text{covariate} = c).$$

Setup III: firm homogeneity

In Phelps, firms homogenous: same pref's over skill types.

- all care about expectation of $f(\text{skill})$, where $f : \Theta \to \mathbf{R}$
- idea: single-task economy, skill = 'productivity',
  f = identity function.
- implication: workers vertically differentiated

Later (Chambers–Echenique): firms (extremely) heterogeneous

$\simeq$ workers horizontally different'd.

Phelps's idea

Basic point: typically, for any given covariate value $c \in \mathcal{C}$,

$$\begin{aligned}&\mathbf{E}(f(\text{skill})|\text{identity} = A, \text{covariate} = c)\\ \neq\ &\mathbf{E}(f(\text{skill})|\text{identity} = B, \text{covariate} = c),\end{aligned}$$

so HR decisions depend on identity (not only covariate).

Why? identity $\perp$ skill, but identity helps interpret covariate.

Example 1: $f(\text{skill}) \equiv \text{skill} \sim U([0,1])$,

$$\text{covariate} = \begin{cases} \text{skill} & \text{if identity} = A \\ 1 - \text{skill} & \text{if identity} = B. \end{cases}$$

Implies discrimination, says Phelps. Details left to imagination.

Discrimination in Phelps's model

Phelps says his model predicts discrimination.

– Question 1 (next): discrimination in <u>which</u> HR decisions?

– Question 2 (later): definition of 'discrimination'?

Definition: random conditional mean

Useful: define random variable M^i by

$$M^i := \mathbf{E}(f(\text{skill})|\text{identity} = i, \underbrace{\text{covariate}}_{\text{random}}).$$

Describes within-group-i heterogeneity ('randomness')
of covariate-based estimate (= expectation) of $f(\text{skill})$.

Charitable reading of Phelps: hiring

Consider hiring. Simplest version:
worker hired iff expectation of her $f(\text{skill})$ exceeds a threshold.

$\implies$ fraction of group i hired $= \mathbf{P}\big(M^i \geq \text{threshold}\big)$

where $M^i = \mathbf{E}(f(\text{skill})|\text{identity} = i, \underbrace{\text{covariate}}_{\text{random}})$

<u>Example 2:</u> $f(\Theta) = \{1, 2\}$, covariate $= \begin{cases} \text{skill} & \text{if identity} = A \\ \varnothing & \text{if identity} = B \end{cases}$
$1 < \text{threshold} < 2$.

- if $\mathbf{E}(f(\text{skill})) < \text{threshold}$:
  fraction A hired $= \mathbf{P}(f(\text{skill}) = 2) > 0 =$ fraction B hired

- if $\mathbf{E}(f(\text{skill})) \geq \text{threshold}$:
  fraction A hired $= \mathbf{P}(f(\text{skill}) = 2) < 1 =$ fraction B hired.

So Phelps's model predicts discrimination in hiring.

(jump to slide 17 / slide 32 / slide 35)

Charitable reading of Phelps: minimum wage

Following variant is closest to what's actually in Phelps (1972a).

Pay in competitive market with minimum wage:

- worker paid expectation of her $f(\text{skill})$ if it's $\geq \text{min\_wage}$
- otherwise worker paid zero (not hired)

<u>Example 2 again:</u> assume $1 < \text{min\_wage} < 2$.

- if $\mathbf{E}(f(\text{skill})) < \text{min\_wage}$:
  As' avg. pay $= 2\mathbf{P}(f(\text{skill}) = 2) > 0 = B$s' avg. pay
- if $\mathbf{E}(f(\text{skill})) \geq \text{min\_wage}$:
  As' avg. pay $= 2\mathbf{P}(f(\text{skill}) = 2) < \mathbf{E}(f(\text{skill})) = B$s' avg. pay

So Phelps's model predicts discrimination in pay.

Plot

Uncharitable reading of Phelps: pay

Consider pay in a frictionless competitive market:
worker paid expectation of her $f(\text{skill})$.

Average pay in group i: $\quad \mathbf{E}(M^i)$.

Law of iterated expectations + equal skill distributions:

$$
\begin{aligned}
\mathbf{E}\Big(M^A\Big) &= \mathbf{E}\big(\mathbf{E}(f(\text{skill}) \mid \text{identity} = A, \text{covariate})\big) \\
&= \quad \mathbf{E}(f(\text{skill}) \mid \text{identity} = A) \\
&= \quad \mathbf{E}(f(\text{skill}) \mid \text{identity} = B) \\
&= \mathbf{E}\big(\mathbf{E}(f(\text{skill}) \mid \text{identity} = B, \text{covariate})\big) = \mathbf{E}\Big(M^B\Big).
\end{aligned}
$$

So Phelps's model predicts <u>no</u> discrimination in pay.

Aigner and Cain (1977)...

– claim that Phelps claimed otherwise,
– 'prove him wrong' as above.

The critique in full

Fully, Aigner–Cain complain

(1) that Phelps's model predicts no pay discrimination
   - upshot (next slide): need non-linearity

(2) that 'identity helps interpret covariate' is a red herring
   - indeed (slide after next)

Pay discrimination requires non-linearity

Upshot: to have statistical discrimination in pay
in frictionless competitive model,
pay cannot be expectation of f(skill).

Expectation ≡ linear function(al) of skill dist'n (Riesz repres'n theorem)

$\Longrightarrow$ pay must be non-linear f'n of skill dist'n.

One story: firms dislike variance of f(skill)
$\Longrightarrow$ if covariate more informative about skill for A than for B,
then A paid more than B on average.

Aigner–Cain seem quite wedded to this story.

It's special, though. In other natural stories,
more info not always better. Recall Example 2 on slide 12!

'Identity helps interpret covariate' is red herring

Example 2: $f(\Theta) = \{0, 1\}$, covariate $= \begin{cases} \text{skill} & \text{if identity} = A \\ \varnothing & \text{if identity} = B \end{cases}$

- recall discrimination occurs
- but identity doesn't help interpret covariate: covariate perfectly reveals identity.

This is very general:

- group i's average outcome (avg. pay, fraction hired, etc.) is a function of the dist'n of $f(\text{skill})$ conditional on ' identity $= i$, $\underbrace{\text{covariate}}_{\text{random}}$ '
- this dist'n obviously doesn't change if replace covariate by covariate$^\star :=$ (covariate, identity), & obviously identity doesn't help interpret covariate*.

What really matters: what info covariate conveys about skill.

Some more uncharitable reading

To make point on previous slide, Aigner–Cain invent terms:

(i) 'individual-level discrimination': for some $c \in \mathcal{C}$,

$$\mathbf{E}(f(\text{skill})|\text{identity} = A, \text{covariate} = c)$$
$$\neq \mathbf{E}(f(\text{skill})|\text{identity} = B, \text{covariate} = c).$$

(ii) 'group-level discrimination':
different average outcomes for groups A & B.

Phelps employs neither definition;
instead leaves meaning of 'discrimination' vague.

Aigner and Cain (1977)...
- claim that Phelps called (i) 'discrimination'
- note that (ii) is a better definition of 'discrimination'.

Plot

Introduction	
Idea (vaguely)	Phelps (1972a, 1972b)
Clarification (uncharitably)	Aigner–Cain (1977)
Modernisation (mathematically)	Chambers–Echenique (2021)
Revision (Blackwellly)	Blackwell (1951, 1953)
Conclusion	

FanFic origin story

maths Phelps:

Phelps, R. R. (2000). *Lectures on Choquet's theorem* (2nd). Springer. https://doi.org/10.1007/b76887

econ Phelps:

Phelps, E. S. (1972b). The statistical theory of racism and sexism. *American Economic Review*, *62*(4), 659–661

Chambers and Echenique (2021):
apply Phelps to Phelps!

Chambers–Echenique setup I: firm heterogeneity

Stick with Aigner–Cain story:

- discrimination in pay
- competitive market, no frictions (e.g. minimum wage)
- requisite non-linearity: convexity $\iff$ info good for avg. pay.

But formalise the story 'non-parametrically' / 'flexibly'
$\iff$ consider (extremely) heterogeneous firms

- a task is a vector $\in \mathbf{R}^{\Theta}$ (Θ finite)
  = surplus as f'n of skill of worker performing the task

- a firm is a finite set of tasks

Assumption: consider all firms.

Firms (very) heterogeneous ('consider all firms')
$\iff$ workers horizontally differentiated
(different firms value different skills)

Chambers–Echenique setup II: production, pay

Production = task assignment.

Given firm $\subseteq \mathbf{R}^{\Theta}$ & belief $\in \Delta(\Theta)$ about worker,

$$\text{pay} = \text{expected surplus} = \max_{\text{task} \in \text{firm}} (\text{belief} \cdot \text{task}).$$

A firm's exp. surplus f'n $\text{belief} \mapsto \max_{\text{task} \in \text{firm}} (\text{belief} \cdot \text{task})$
is a convex f'n $\Delta(\Theta) \to \mathbf{R}$.

– all firms $\simeq$ all convex f'ns $\Delta(\Theta) \to \mathbf{R}$ (formally: up to uniform closure)

– 'special case': f'n = mean − k × variance

Summary

	Aigner–Cain	Chambers–Echenique
workers	vertically differentiated	horizontally diff'ed
firms	homogeneous	(very) heterogeneous
surplus	'parametric'	'non-parametric' / 'flexible'
	(mean $-$ $k \times$ variance)	(arbitrary convex f'n)

Definition: random conditional distribution

Let P^i be random vector $\in \Delta(\Theta)$ defined by

$$P^i_\theta \coloneqq \mathbf{P}\big(\text{skill} = \theta \;\big|\; \text{identity} = i,\ \underbrace{\text{covariate}}_{\text{random}}\big) \quad \forall \theta \in \Theta.$$

Describes within-group-i heterogeneity ('randomness')
of covariate-based estimate of (= belief about) skill dist'n.

Random belief. 'Belief-based approach' $\begin{cases} \text{Blackwell,} \\ \text{Aumann–Maschler,} \\ \text{Kamenica–Gentzkow} \end{cases}$

CE go as far as to <u>identify</u> covariate with P^i! <u>Very</u> modern.

CE's definition of '(statistical) discrimination'

CE's def'n: (statistical) discrimination against group B iff

some firm pays Bs strictly less on avg.: $\exists$ firm $\subseteq \mathbf{R}^{\Theta}$ s.t.

$$\mathbf{E}\left(\max_{\text{task} \in \text{firm}} P^A \cdot \text{task}\right) > \mathbf{E}\left(\max_{\text{task} \in \text{firm}} P^B \cdot \text{task}\right)$$

Results & interpretation

Note can view skill = 'state',
covariate = 'signal' = 'Blackwell experiment' = 'info structure'.

Question: when is there (CE-def'n) discrimination?

Answer: iff skill dist'n not identified off covariate iff XYZ.

Proved via Choquet theory from 'maths Phelps' book.

Big upshot from CE's introduction:

> *We show that the focus on informativeness in Phelps (1972b) and Aigner and Cain (1977) is misleading. There may be statistical discrimination even when the information structure of one [group] is not more informative than the other. [...] Aigner and Cain trace statistical discrimination to pure informativeness. We argue that the situation is more general.*

Plot

Comments on CE

CE model very natural. Comments on results / interpretation:

(1) CE's definition of 'discrimination' is weak.
Propose a better definition.

(2) Contrary to CE's claim, in CE's model,
discrimination <u>is precisely</u> about informativeness
(of covariate about skill).

(3) Relabelling Blackwell's theorem yields nice
characterisation of discrimination in CE's model.

Better definition of '(statistical) discrimination'

New def'n: (statistical) discrimination against group B iff both

(1) every firm pays Bs weakly less on avg.: $\forall$ firm $\subseteq \mathbf{R}^{\Theta}$,

$$\mathbf{E}\left(\max_{\text{task} \in \text{firm}} P^A \cdot \text{task} \right) \geq \mathbf{E}\left(\max_{\text{task} \in \text{firm}} P^B \cdot \text{task} \right)$$

(2) some firm pays Bs strictly less on avg.: $\exists$ firm $\subseteq \mathbf{R}^{\Theta}$ s.t.

$$\mathbf{E}\left(\max_{\text{task} \in \text{firm}} P^A \cdot \text{task} \right) > \mathbf{E}\left(\max_{\text{task} \in \text{firm}} P^B \cdot \text{task} \right)$$

Recall $P^i_\theta = \mathbf{P}\big(\text{skill} = \theta \,\big|\, \text{identity} = i,\ \underbrace{\text{covariate}}_{\text{random}}\big) \quad \forall \theta \in \Theta$

CE's def'n: (2) only. Can interpret as 'robustness concern': worry about 'worst-case' firm. ('maxmin')

Opinion: that's too weak to deserve name 'discrimination'.

Discrimination = informativeness I

CE model	$\rightsquigarrow$	Blackwell decision model
skill	$\rightsquigarrow$	state
covariate	$\rightsquigarrow$	signal / experiment / info struc.
task	$\rightsquigarrow$	action
firm	$\rightsquigarrow$	decision problem
(avg.) pay / surplus	$\rightsquigarrow$	(exp.) value

Recall def'n of Blackwell (strictly) less informative:

'weakly lower exp. value in every decision problem
(& strictly lower exp. value in some decision problem')

Obs'n: (new-definition) statistical discrimination against Bs

$\iff \begin{cases} B\text{s weakly lower avg. pay in every firm} \\ \&\ B\text{s strictly lower avg. pay in some firm} \end{cases}$

$\iff$ covariate str. less info'tive about skill for Bs than for As

Discrimination = informativeness I

CE model		Blackwell decision model
skill	$\rightsquigarrow$	state
covariate	$\rightsquigarrow$	signal / experiment / info struc.
task	$\rightsquigarrow$	action
firm	$\rightsquigarrow$	decision problem
(avg.) pay / surplus	$\rightsquigarrow$	(exp.) value

Recall def'n of Blackwell (strictly) less informative:

'weakly lower exp. value in every decision problem
(& strictly lower exp. value in some decision problem')

Obs'n: CE-definition statistical discrimination against Bs

$\iff$ Bs strictly lower avg. pay in some firm

$\iff$ not: Bs weakly higher avg. pay in every firm

$\iff$ covariate not more info'tive about skill for Bs than for As.

Discrimination = informativeness II

Upshot: contrary to CE's claim, in their model,
discrimination is precisely about informativeness
(of covariate about skill).

However: ∃ other natural models
in which CE's claim is true
(recall Example 2 on slide 12).

Identification and inevitability

Recall Obs'n: CE-definition statistical discrimination

$\iff$ covariate not more info'tive for Bs than for As.

Corollary: 'CE-discrimination' against neither As nor Bs

$\iff$ covariate both more and less info'tive for Bs than for As

$\iff$ groups informationally identical. Extremely stringent.

Upshot: on CE's def'n, 'discrimination' is inevitable!

(Not shocking. Again, CE's def'n too weak.)

Modulo details, this is CE's 'identification' result,
re-stated in non-econometric language.

Characterising discrimination in CE's model

Informativeness well-understood, so can borrow insights. E.g.

Blackwell's theorem. The following are equivalent:

(i) (new-definition) statistical discrimination against Bs: covariate str. less info'tive about skill for Bs than for As

(ii) P^B strictly less variable than P^A in convex-order sense (a.k.a. 'mean-preserving spread')

(iii) B's covariate is a non-trivial garbling of A's

Recall $P^i_\theta = \mathbf{P}\big(\text{skill} = \theta \,\big|\, \text{identity} = i,\ \underbrace{\text{covariate}}_{\text{random}}\big) \quad \forall \theta \in \Theta$

Characterising discrimination in CE's model

Informativeness well-understood, so can borrow insights. E.g.

Blackwell's theorem v2. The following are equivalent:

(i) CE-definition statistical discrimination against Bs: covariate not more info'tive about skill for Bs than for As

(ii) P^B not more variable than P^A in convex-order sense (a.k.a. 'mean-preserving spread')

(iii) A's covariate is not a garbling of B's

Recall $P^i_\theta = \mathbf{P}\big(\text{skill} = \theta \,\big|\, \text{identity} = i,\ \underbrace{\text{covariate}}_{\text{random}}\big) \quad \forall \theta \in \Theta$

Suggestion for future work

One observation:

- Lit since Aigner–Cain very focussed on models
  in which more info $\iff$ higher avg. pay.
- But this is quite special. Recall Example 2 on slide 12.

Needed: analysis of statistical discrimination
in labour-market models beyond this special class.

Thanks!

$$b^2 - 4ac$$

References I

Aigner, D. J., & Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, *30*(2), 175–187. https://doi.org/10.1177/001979397703000204

Arrow, K. J. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in labor markets* (pp. 3–33). Princeton University Press.

Aumann, R. J., & Maschler, M. B. (1968/1995). *Repeated games with incomplete information* [circulated 1966–68, published 1995]. MIT Press.

Becker, G. (1957). *The economics of discrimination*. University of Chicago Press.

Blackwell, D. (1951). Comparison of experiments. In J. Neyman (Ed.), *Berkeley symposium on mathematical statistics and probability* (pp. 93–102, Vol. 2). University of California Press.

References II

Blackwell, D. (1953). Equivalent comparisons of experiments. *Annals of Mathematics and Statistics*, *24*(2), 265–272. https://doi.org/10.1214/aoms/1177729032

Chambers, C. P., & Echenique, F. (2021). A characterisation of 'Phelpsian' statistical discrimination. *Economic Journal*, *131*(637), 2018–2032. https://doi.org/10.1093/ej/ueaa107

Edgeworth, F. Y. (1922). Equal pay to men and women for equal work. *Economic Journal*, *32*(128), 431–457. https://doi.org/10.2307/2223426

Escudé, M., Onuchic, P., Sinander, L., & Valenzuela-Stookey, Q. (2022). *Statistical discrimination and statistical informativeness* [working paper, 31 May 2022]. https://doi.org/10.48550/arXiv.2205.07128

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, *101*(6), 2590–2615. https://doi.org/10.1257/aer.101.6.2590

References III

Onuchic, P. (2023). *Recent contributions to theories of discrimination* [working paper, May 2023]. https://doi.org/10.48550/arXiv.2205.05994

Phelps, E. S. (1972a). *Inflation policy and unemployment theory: The cost–benefit approach to monetary planning*. Macmillan.

Phelps, E. S. (1972b). The statistical theory of racism and sexism. *American Economic Review*, *62*(4), 659–661.

Phelps, R. R. (2000). *Lectures on Choquet's theorem* (2nd). Springer. https://doi.org/10.1007/b76887