Wage Discrimination when Identity is Subjective: Evidence from Changes in Employer-Reported Race

Christopher Cornwell
Jason Rivera
and
Ian M. Schmutte

Department of Economics
University of Georgia

Applied Microeconomics Workshop
The Ohio State University
5 March 2015

What we do

- Estimate the effect of race on labor market earnings
- Using differences in the race
  - reported for the same worker
  - by different employers
- **Punchline:** 20-40 percent of cross-section wage gap between white and non-white workers

The Promise

- Do the impossible – panel data estimate of the racial earnings gap;
- exploiting variation in something malleable – employer 'perception' of race;
- changing racial identity is a rational response to discrimination

The Challenge

- Are changes in reported race 'real'?
- ... or are they classification errors?

Descriptive statistics, individual characteristics

			By Race History		
	All Workers (1)	Job Changers (2)	'11' (3)	'10' (4)	'01' (5)
Race History					
'11': White/White	n/a	0.485	1	0	0
'10': White/Non-White	n/a	0.139	0	1	0
'01': Non-White/White	n/a	0.132	0	0	1
'00': Non-White/Non-White	n/a	0.244	0	0	0
White					
Orig. Job	0.644	0.624	1	1	0
Dest. Job	n/a	0.618	1	0	1
Log Wage					
Orig. Job	6.536	6.404	6.462	6.390	6.376
Dest. Job	n/a	6.460	6.517	6.452	6.431
Male					
Orig. Job	0.649	0.717	0.658	0.745	0.742
Dest. Job	n/a	0.717	0.659	0.745	0.743
Age					
Orig. Job	35.010	31.4	31.1	31.4	31.3
Dest. Job	n/a	31.4	31.1	31.4	31.2
Education					
LTHS	0.446	0.461	0.409	0.461	0.477
High School	0.421	0.436	0.451	0.451	0.443
Some College	0.041	0.040	0.052	0.035	0.033
Bachelor's (+)	0.092	0.063	0.088	0.053	0.047
Num.Obs.	26,512,018	3,000,688	1,443,893	420,759	397,030

Descriptive statistics, plant characteristics

			By Race History		
	All Workers (1)	Job Changers (2)	'11' (3)	'10' (4)	'01' (5)
Plant Mean Log Wage					
Orig. Job	6.528	6.459	6.503	6.445	6.449
Dest. Job	n/a	6.510	6.556	6.510	6.493
Plant White Share					
Orig. Job	0.626	0.614	0.822	0.749	0.363
Dest. Job	n/a	0.613	0.816	0.374	0.750
Plant Employment					
Orig. Job	755.437	662.532	551.536	549.636	703.130
Dest. Job	n/a	757.640	654.183	800.152	620.993
Plant Separation Rate					
Orig. Job	0.633	1.150	1.139	1.197	1.121
Dest. Job	n/a	1.466	1.503	1.360	1.693
Num.Obs.	26,512,018	3,000,688	1,443,893	420,759	397,030

Race in Brazil

Brazil vs US

- Historical Similarities
  - Colonial repression of indigenous population
  - Import of African slaves in large numbers
- Historical Differences
  - Portuguese colonists encouraged to populate with natives
  - No "race science" in Brazil
  - No history of segregation, "one-drop" rules, or anti-miscegenation laws in Brazil

It's skin color

Open-ended query about race elicits 136 color descriptions (PNAD, 1976)

Portuguese	English
Acastanhada	Somewhat chestnut-coloured
Alva rosada	Pinkish white
Azul	Blue
Branca	White
Canela	Cinnamon
Cor-de-café	Coffee-coloured
Meio-branca	Half-white
Morena	Dark-skinned, brunette
Rosada	Rosy
Sapecada	Singed
Turva	Murky

Official race categories and population shares

Portuguese	English	Share
Branca	"White"	55.71
Pardo	"Brown"	36.05
Preto	"Black"	7.54
Amarelo	"Yellow"	0.50
Indigeno	"Indigenous"	0.21

Source: PNAD, 2009

Malleability of race

- Individual manipulation of identity
  - Affirmative action in education (Francis and Tannuri-Pianto 2013)
- Variation in Other's Perception of Racial Classification
  - Survey numerators and respondents (Telles 2002)
  - Parents and children (Schwartzman 2007)

Evidence of racial inequality in the labor market

- Qualitative evidence of workplace discrimination (Telles 2002)
- Disparities in labor-market earnings
- Workplace segregation

The RAIS data and employer-reported race

Relação Anual de Informações Sociais (RAIS)

- Collected *from employers* to administer *Abono Salarial* ("Thirteenth Salary")
- Covers the population of formal-sector jobs (∼40 million per year)
- Data items include
  - job characteristics: wage, hours, occupation, tenure ...
  - plant characteristics: industry, size, location ...
  - worker characteristics: education, race, sex ...

We use RAIS under an agreement with the Brazilian Ministry of Labor and Employment (MTE).

How employers collect race data

- Worker presents "Worker Record Booklet" at date of hire
  - Includes usual identification information and a photograph
  - It does *not* report race
- Worker must also provide a photograph and proof of education for the position
- Employer makes entry in an "Employer Registration Book"
  - Legal requirement to collect worker's name, date of hire and other information related to the job
  - Not required to collect information on race and gender, but they routinely do
  - Information provided by worker and verified by administrative staff
- No affirmative-action or equal-opportunity laws in Brazil

Carteira de Trabalho e Prevêdencia Social

Carteira de Trabalho e Prevêdencia Social

Registro De Empregado

RANIERI
CONSULTORIA E TREINAMENTO LTDA.

FICHA DE REGISTRO DE EMPREGADO

EMPREGADOR:	ENDEREÇO:

NÚMERO DE ORDEM	NOME DO EMPREGADO	Nº DE MATRÍCULA

	NOME DO PAI	NACIONALIDADE
	NOME DA MÃE	NACIONALIDADE

DATA DE NASCIMENTO	IDADE	NACIONALIDADE	ESTADO CIVIL	LOCAL DE NASCIMENTO	U.F.	CARTEIRA DE IDENTIDADE Nº

CTPS Nº	SÉRIE Nº	CERTIF. DE RESERVISTA Nº	CATEGORIA	CPF / CIC	TÍTULO DE ELEITOR Nº	CARTEIRA DE SAÚDE Nº

C.B. O. Nº	CARTEIRA MODELO 19 Nº	É CASADO(A) COM BRASILEIRO(A)?	É NATURALIZADO(A)?	TEM FILHOS BRASILEIROS?
DATA EM QUE CHEGOU AO BRASIL	Nº DO REGISTRO GERAL	NOME DO(A) CONJUGE		QUANTOS FILHOS?

AUTENTICAÇÃO DO MTB

ENDEREÇO	(DDD) TELEFONE
MUDANÇA DE ENDEREÇO E TELEFONE	

COR	ALTURA	PESO	CABELOS	OLHOS	SINAIS

NOMES DOS DEPENDENTES / BENEFICIÁRIOS	PARENTESCO	DATA DE NASCIMENTO

PROGRAMA DE INTEGRAÇÃO SOCIAL – PIS

CADASTRADO EM:
SOB O Nº
ENDEREÇO:
BANCO: AGÊNCIA:

DATA DE ADMISSÃO	DATA DO REGISTRO	CARGO	SEÇÃO	SALÁRIO INICIAL R$	COMISSÕES	TAREFA	FORMA DE PAGAMENTO

SITUAÇÃO PERANTE O FUNDO DE GARANTIA POR TEMPO DE SERVIÇO

É OPTANTE?	DATA DE OPÇÃO	DATA DE RETRATAÇÃO

BANCO DEPOSITÁRIO
CAIXA ECONÔMICA FEDERAL

HORÁRIO DE TRABALHO

ENTRADA	REPOUSO / ALIMENTAÇÃO	SAÍDA	DESCANSO SEMANAL

Declaro que estou de pleno acordo com as informações acima e que exprimem a verdade.

POLEGAR DIREITO

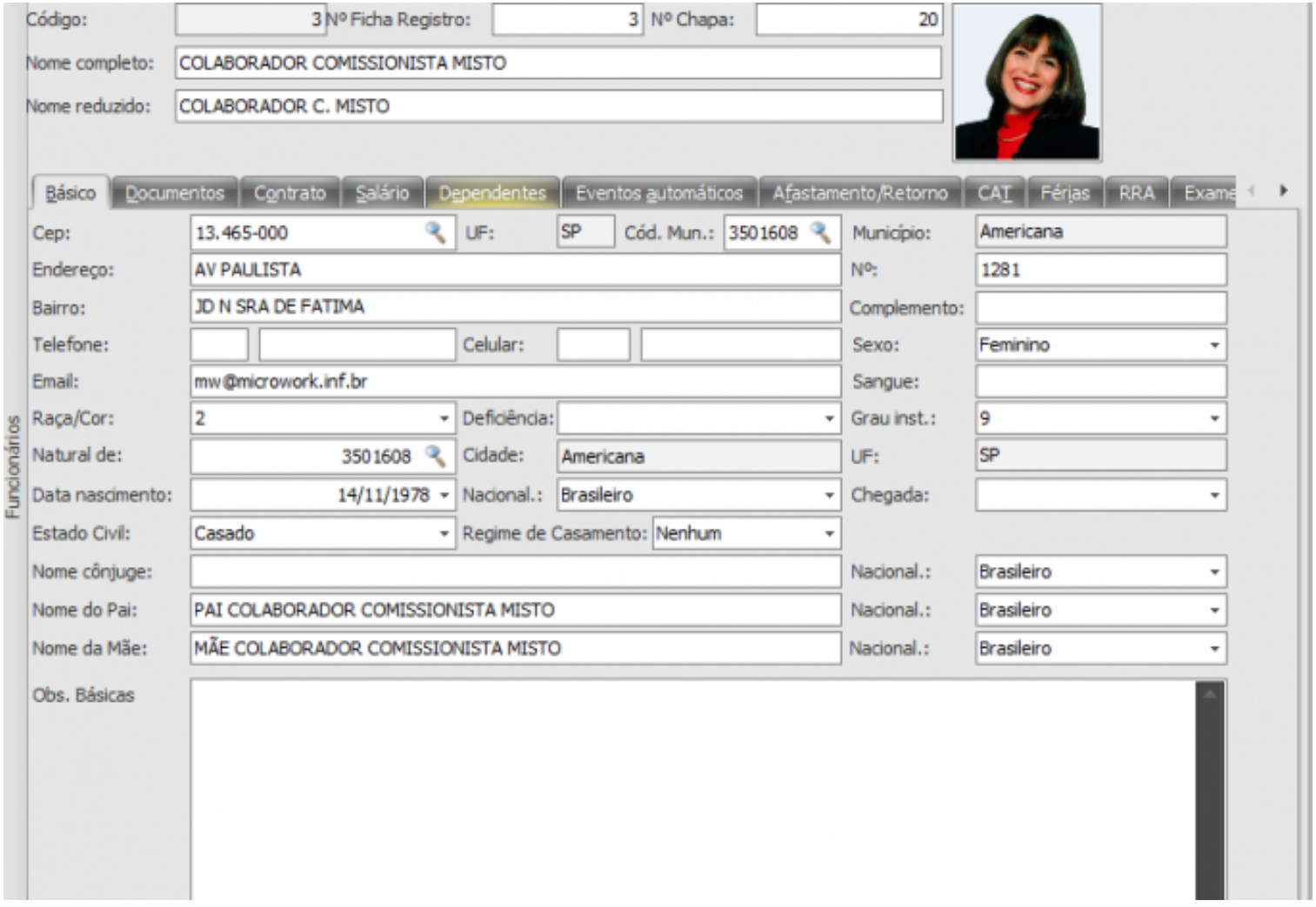
CARIMBO E VISTO DO EMPREGADOR

Assinatura do Empregado ne Admissão

Data de Demissão: _____/_____/_____

Assinatura do Empregado na Rescisão

Registro De Empregado



Registro De Empregado

REGISTRO DE EMPREG

Cor Parda

Cabelo Preto

Barba Grande

Bigode Enorme

Olhos Pretos

Altura 1,74 m

Peso 59 kg

VISTO DA F

O Sr. Jaca Robertino

Carteira profissional nº 666 série 000 Carteira de Trabalho de mer

série foi admitido em 31 de Fevereiro de 1987 para exercer a

com o salário de Cr$ 5.000,00 (cinco m

Job changers and race change

Construction of the analysis sample

From the 2010 wave of RAIS

- Choose workers with an ongoing full-time job at the start of the year
- ...who start another full-time job in 2010
- ...and assemble their employer-reported information from both jobs
- Limit to white, brown and black workers

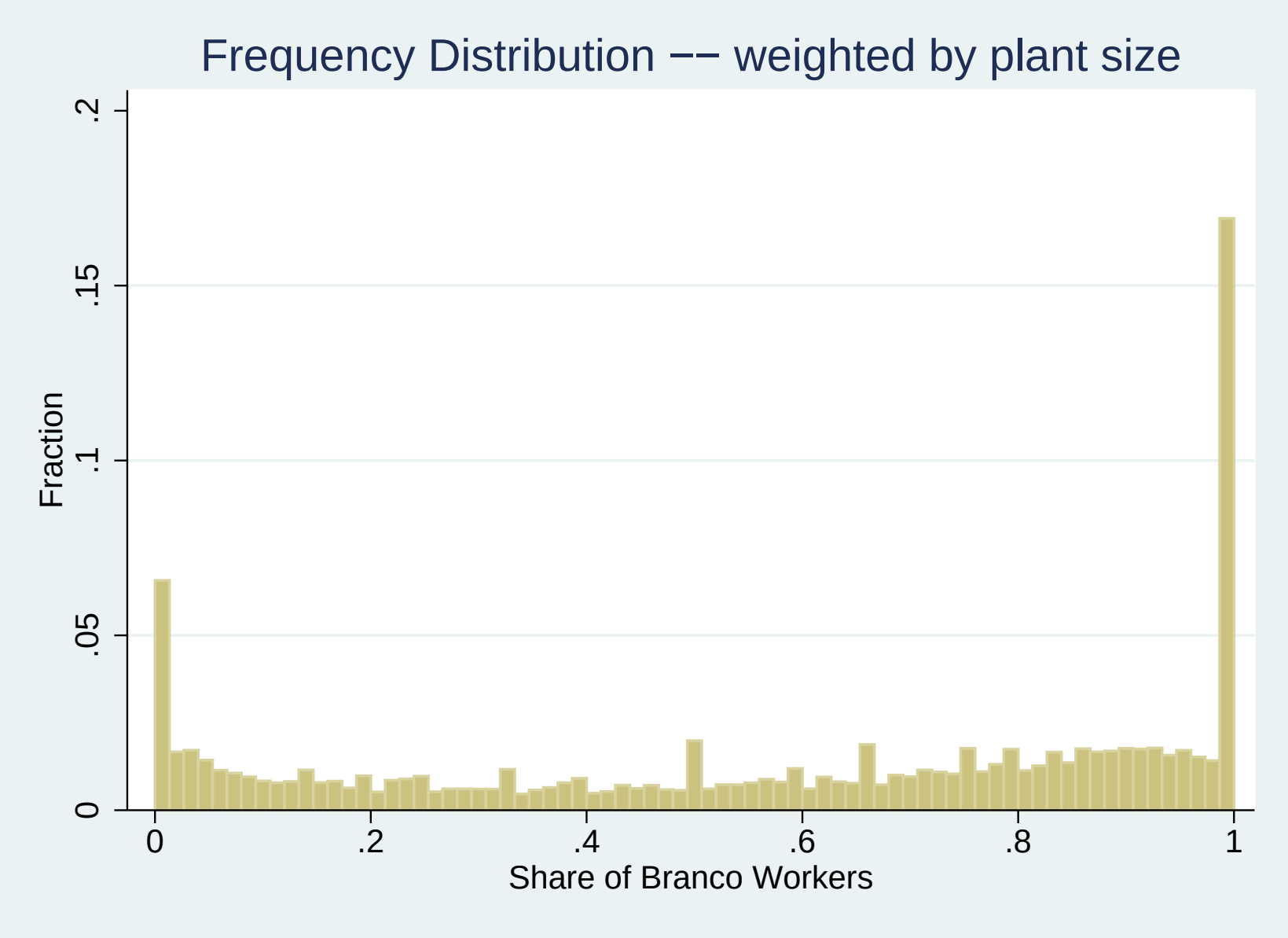
Cross-section racial wage gaps

	All Workers		Job Changers	
			Orig. Job Wage	Dest. Job Wage
	(1)	(2)	(3)	(4)
White	0.132	0.078	0.065	0.048
	(0.0002)	(0.001)	(0.001)	(0.001)
Plant Characteristics?	N	Y	Y	Y
N	26,512,018	26,512,018	3,000,688	3,000,688
R^2	0.3621	0.6804	0.5515	0.5276

Control variables

- Individual: gender, education, quadratic in age
- Plant: industry, state, employment level, share white, mean log wage, separation rate

Racial distribution across plants



Source: RAIS, 2010

Race change is not pure misclassification

Basic elements of the misclassification model

Adapt correlated random effects model of Card (1996)

- Two notions of race
  - "Market" race (r^*) – worker's wage depends on this
  - Employer-reported race (r^M) – what is observed?
- Reject: r^* is immutable
- Cannot reject: $r^M = r^*$

▸ Model Details

Effects of race history on wages

Reduced-form wage equations

$$\begin{aligned} w_{i1} &= a_1' + b_1 x_i + d_1 R_i + e_{i1} \\ w_{i2} &= a_2' + b_2 x_i + d_2 R_i + e_{i2} \end{aligned}$$

Notation:

- R_{ih}: indicator for the hth employer race history
- $h \in \{00, 01, 10, 11\}$
- Concerned with elements of d_1 and d_2
- Specifically, $d_1 - d_2$.

Estimated race-history effects

	Orig. Job Log Wage (1)	Dest. Job Log Wage (2)	Dest.–Orig. (3)
Race History			
'11': White/White	0.072	0.069	−0.003
	(0.001)	(0.001)	(0.001)
'10': White/Non-White	0.046	0.025	−0.021
	(0.001)	(0.001)	(0.001)
'01': Non-White/White	0.016	0.033	0.017
	(0.001)	(0.001)	(0.001)
N	3,000,688	3,000,688	3,000,688
R^2	0.565	0.599	0.195

Alternative mechanism – Plant-specific reporting behavior

	No Controls (1)	Full Contols (2)
Non-reporting share = 0 (Always report)	−0.031	−0.012
	(0.0006)	(0.0007)
Non-reporting share	−0.163	0.012
	(0.0031)	(0.0037)
N	3,000,009	3,000,009
R^2	0.0010	0.0709

Alternative mechanism – Plant-specific reporting behavior

	Benchmark (1)	Reporting Contols (2)	Always Report (3)	Not Always Report (4)	Plant Effects (5)
Race History					
'11': White/White	−0.003	−0.001	−0.002	0.009	0.001
	(0.0010)	(0.0010)	(0.0012)	(0.0031)	(0.001)
'10': White/Non-White	−0.021	−0.022	−0.021	−0.021	−0.010
	(0.0010)	(0.0010)	(0.0013)	(0.0035)	(0.001)
'01': Non-White/White	0.017	0.020	0.016	0.032	0.010
	(0.0010)	(0.0010)	(0.0013)	(0.0036)	(0.001)
Plant Effects	N	N	N	N	Y
N	3,000,688	3,000,009	1,864,636	250,447	3,000,688
R^2	0.195	0.1938	0.2111	0.1313	0.378

Alternative identification

$$w_{2i} = a + \zeta w_{1i} + bx_i + m \times OrigWhite_i + k_{10}R_{10} + k_{01}R_{01} + \psi_{J(2i)} + e_{2i}$$

	ΔLog Wage (1)	Dest. Wage (2)
Race History		
'11': White/White	−0.003	
	(0.001)	
'10': White/Non-White	−0.021	−0.034
	(0.001)	(0.001)
'01': Non-White/White	0.017	0.022
	(0.001)	(0.001)
Log Wage (Origin Job)		0.307
		(0.001)
White (Origin Job)		0.043
		(0.001)
Plant Effects	N	Y
N	3,000,688	3,000,688
R^2	0.1948	0.7450

Robustness to Endogenous Mobility

	Benchmark (1)	JUJ (2)	Educ. Same (3)	Educ. Down (4)
Race History				
'11': White/White	−0.003	−0.007	−0.002	−0.007
	(0.0010)	(0.0022)	(0.0013)	(0.0023)
'10': White/Non-White	−0.021	−0.021	−0.022	−0.019
	(0.0010)	(0.0024)	(0.0014)	(0.0024)
'01': Non-White/White	0.017	0.019	0.017	0.013
	(0.0010)	(0.0024)	(0.0014)	(0.0024)
N	3,000,688	513,335	1,657,397	551,214
R^2	0.1948	0.2544	0.1791	0.2287

Final thoughts

- Rhetoric of 'post-racial' US is probably like Brazil's 'racial democracy'
- The need to understand racial inequalities will persist
- Race may become increasingly difficult to measure and model
  - Saperstein and Penner (2012)
  - Liebler et al. (2014)

Thank You

Ian M. Schmutte
schmutte@uga.edu

Bonus slides: misclassification model

Return to Presentation

Modeling Framework

Three different concepts of race

- The 'market race' (unobserved) (r^*)
- The 'employer race' (observed) (r^M)
- The 'self-race' (unobserved) (r^S)

Set up a Chamberlain-style correlated random effects model with misclassification of market race (Card 1996).

Wages (Structural Model):

$$w_{ij} = a_j + \beta_j x_i + \delta r^*_{ij} + \varepsilon_{ij}$$

- w_{ij} is the log wage reported for worker i by employer $j \in \{1, 2\}$
- x_i includes both stationary characteristics and the complete history of time-varying observables
- r^*_{ij} indicates the market race of worker i with employer j
- δ is the coefficient of discrimination
- $\varepsilon_{it} = \alpha_i + \varepsilon'_{it}$

Notation:

- R^*_{ih}: indicator for the hth market race history (*unobserved*)
- R^M_{ih}: indicator for the hth employer race history (*observed*)
- $h \in \{00, 01, 10, 11\}$

Project person effect onto unobservable $R*_i$ and observable x_i

$$\alpha_i = \phi_1 + \sum_{h \neq 00} R^*_{ih}\phi_h + \lambda x_i + \xi_i$$

With two employers, of data, wages are

$$w_{i1} = a_1 + \phi_1 + (\beta_1 + \lambda)x_i + (\delta + \phi_{10})R^*_{i10} + \phi_{01} \quad R^*_{01} + (\phi_{11} + \delta)R^*_{i11} + \xi_i + \varepsilon'_{i1}$$

$$w_{i2} = a_2 + \phi_1 + (\beta_2 + \lambda)x_i + \quad \phi_{10} \; R^*_{i10} + (\phi_{01} + \delta)R^*_{01} + (\phi_{11} + \delta)R^*_{i11} + \xi_i + \varepsilon'_{i2}$$

Problem: R^*_{ih} is unobservable.

Work with projection of R^*_{ih} onto observed race histories:

$$R^*_{ih} = \gamma_{0h} + \gamma_h R^M_i + \gamma_{xh} x_i + \eta_{ih}$$

- $\gamma_h = [\gamma_{h,11}, \gamma_{h,10}, \gamma_{h,01}]$
- $\gamma_{h,k}$ measures the conditional correlation between observed history k and market race history h

Reduced form:

$$
\begin{aligned}
w_{i1} &= a_1' + b_1 x_i + d_1 R_i + e_{i1} \\
w_{i2} &= a_2' + b_2 x_i + d_2 R_i + e_{i2}
\end{aligned}
$$

Estimating equations:

$$d_1 = [(\delta + \phi_{10})\gamma_{10} + \quad \phi_{01} \;\; \gamma_{01} + (\delta + \phi_{11})\gamma_{11}]$$
$$d_2 = [\quad \phi_{10} \;\; \gamma_{10} + (\delta + \phi_{01})\gamma_{01} + (\delta + \phi_{11})\gamma_{11}]$$

For all true histories, h,

$$d_{2,h} - d_{1,h} = \delta(\gamma_{01,h} - \gamma_{10,h})$$

Closing the Model:

- Still need the attenuation parameters (elements of γ)
- And, a specification for the misclassification process

Define

- **False negative:** $P(r_{it}^M = 0 | r_{it}^* = 1) = 1 - q_1$
- **False positive:** $P(r_{it}^M = 1 | r_{it}^* = 0) = q_0$

Assume

$$P(r_{i1}^M, r_{i2}^M | r_{i1}^*, r_{12}^*, x_i) = P(r_{i1}^M | r_{i1}^*) \cdot P(r_{i2}^M | r_{i2}^*)$$

Misclassification Matrix:

- π_k: the share of workers with $R^*_{ik} = 1$ (*unobserved*)
- p_j: the share of workers with $R^M_{ij} = 1$ (*observed*)

Then

$$p = \mathrm{E}(R_i) = \mathrm{E}(R^*_i T) = \pi T$$

T is a 4×4 matrix whose (j, k) entry is the misclassification probability $\tau_{j,k} = P(R^M_{ij} = 1 | R^*_{ik} = 1)$.

Project market and employer race histories onto observables:

$$
\begin{aligned}
R^*_{ih} =& \pi_h + (x_i - \bar{x})c_h + \nu_{ih} \\
R_{ih} =& p_h + (x_i - \bar{x})\zeta_h + \nu'_{ih}
\end{aligned}
$$

Finally, a model for γ falls out of partitioned regression:

$$\gamma_h = \left[\text{var}(R) - \Omega c^{\text{T}} V_{xx} c \Omega^{\text{T}}\right]^{-1} \cdot \left\{\text{cov}(R, R_h^*) - \Omega c^{\text{T}} V_{xx} c_h\right\}$$

where V_{xx} is the covariance matrix of x_i

Estimation:

- **Step 1:** Estimate the reduced-form models for wages and observed race histories

- **Step 2:** Use minimum distance estimator to fit
  - nine unrestricted sample moments
    $(d_{11}, d_{12}, d_{13}, d_{21}, d_{22}, d_{23}, p_{11}, p_{10}, p_{01})$
  - to nine parameters
    $(q_1, q_0, \pi_{11}, \pi_{10}, \pi_{01}, \phi_{11}, \phi_{10}, \phi_{01}, \delta)$

Model 1: Market Race does not Change

Testable Restriction: No person has true history R^*_{10} or R^*_{01}

- $\pi_{10} = \pi_{01} = 0$
- ϕ_{10} and ϕ_{01} are not identified

Model 2: No Measurement Error

Testable Restrictions: Employer report identical to market race ($r_j^* = r_j^M$)

- $q_1 = 1$ (no false negatives)
- $q_0 = 0$ (no false positives)

Summary of Structural Tests: RAIS 2010

	Model	
	No Race Change (1)	No Meas. Error (2)
Obj. Fcn Value	0.0005	$1.049e^{-5}$
Test Statistic	$1,588$	0.5313

Summary of Structural Tests: RAIS 2010

Panel A: Structural Parameter Estimates		
Parameter	Model	
	No Race Change (1)	No Meas. Error (2)
$\kappa = (\delta + \phi_{11})$	0.283	0.071
	(0.0030)	(0.0001)
δ	–	0.019
		$(2.7e^{-5})$
ϕ_{11}	–	0.052
		$(9.9e^{-5})$
ϕ_{10}	–	0.026
		$(7.6e^{-5})$
ϕ_{01}	–	0.015
		$(8.8e^{-5})$
q_1	0.884	$--$
	(0.0002)	
q_0	0.236	$--$
	(0.0002)	
π_{11}	0.583	0.481
	(0.0004)	(0.0003)
π_{10}	–	0.141
		(0.0002)
π_{01}	–	0.132
		(0.0002)
Panel C: Model Fit		
Obj. Fcn Value	0.0005	$1.049e^{-5}$
Test Statistic	$1,588$	0.5313