

# Wage & Employment Dynamics

## THE WED PROJECT



# Revisiting ethnic wage differences in England and Wales: New Payroll-Census data compared with the Annual Population Survey

Van Phan, Felix Ritchie & Damian Whittard (University of the West of England)

Alex Bryson (University College London)

John Forth (City, University of London)

Carl Singleton (University of Reading)

Lucy Stokes (National Institute of Economic and Social Research)

18<sup>th</sup> October 2021

*"This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates."*

## Key findings

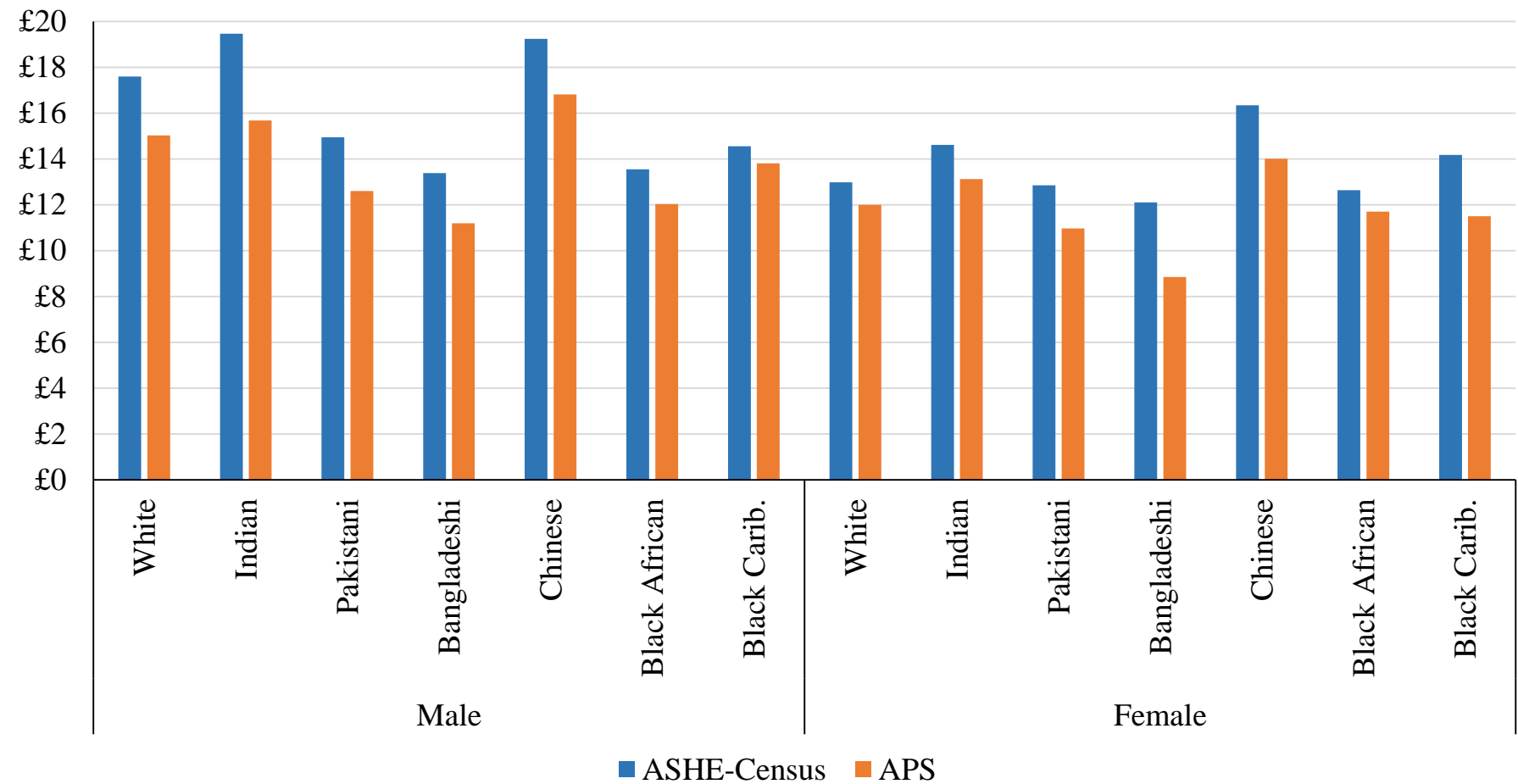
- The mean hourly pay of employees is higher in the payroll data from ASHE-Census than in the Annual Population Survey (APS), across all ethnic groups, in 2011
  - This is not a function of the survey/sample weights
  - Differences are much smaller at the median and earnings distributions confirm longer right-hand tails in ASHE-Census data
    - May be due to the different treatment of hours as the denominator across the two data sets: self-reported vs payroll based
    - Small N's imply there is value in focusing on the median rather than the mean
- Gender-ethnicity wage gaps at the median are similar in ASHE-Census and APS, and are as expected from the literature; e.g., highest median hourly earnings is for Indian and Chinese men
- Hourly pay distributions by ethnicity and gender are similar across ASHE-Census and APS (apart from longer right-hand tails in ASHE-Census and higher density of very low pay in APS)
  - Strong indicator of data quality for the ASHE-Census
  - The distributions are 'twin peaked' in both data sets in a number of cases, e.g., White women, Black African women, Bangladeshi men, Chinese men
- Adjusted wage gaps are all negative relative to White men (controlling for education, age, family, and other characteristics)
  - Very large differences across non-white ethnic groups and by gender
  - Gaps tend to conform in size to what we know from the literature, but there are some substantial differences between comparable estimates from ASHE-Census and APS.

# Contents

1. **Mean and median wages of ethnic minority employees, by gender:** *What were the average or typical ethnic pay gaps according to payroll and household survey data in 2011?* (Slides 4-8)
2. **Distributions of hourly earnings by ethnicity and gender:** *Were high or low wages more common in England and Wales for some groups of employees than others?* (Slides 9-15)
3. **Estimated wage penalties associated with ethnicity-gender:** *What were the average wage differences associated with ethnicity for employees with the same levels of education, experience, occupation, and other characteristics?* (slides 16-17)
4. **Appendix:** Details on data sources, methods and additional results (slides 18+)

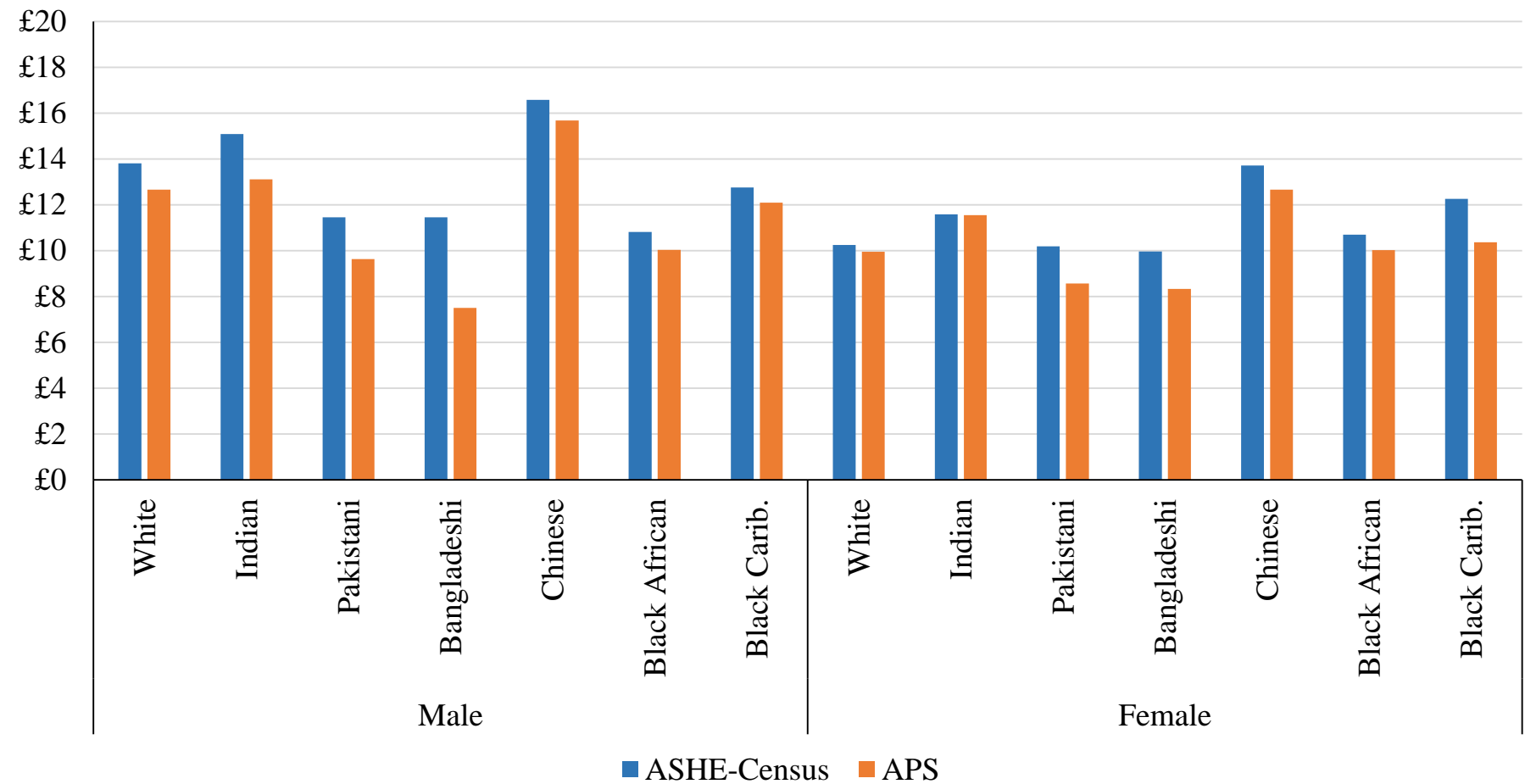
**1. Mean and median wages of ethnic minority employees, by gender:**  
*What were the average or typical ethnic pay gaps according to payroll and household survey data in 2011?*

Mean Hourly Earnings in England and Wales in 2011  
– Comparison of ASHE-Census and APS



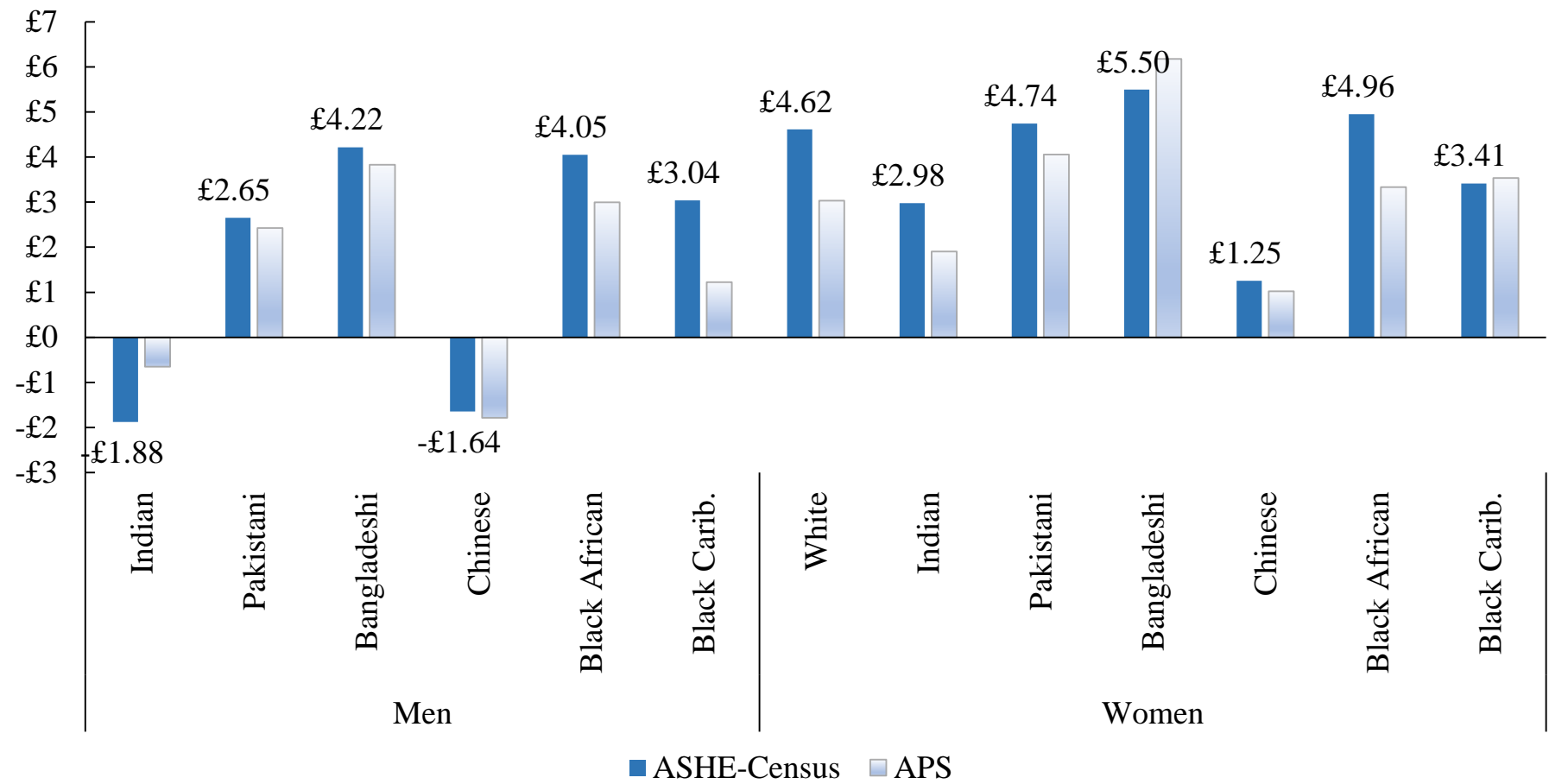
Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here use survey weights. See slide 23 for a table of these wage statistics and slide 25 for the equivalents without using survey weights.

Median Hourly Earnings in England and Wales in 2011  
– Comparison of ASHE-Census and APS



Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here use survey weights. See slide 24 for a table of these wage statistics and slide 26 for the equivalents without using survey weights.

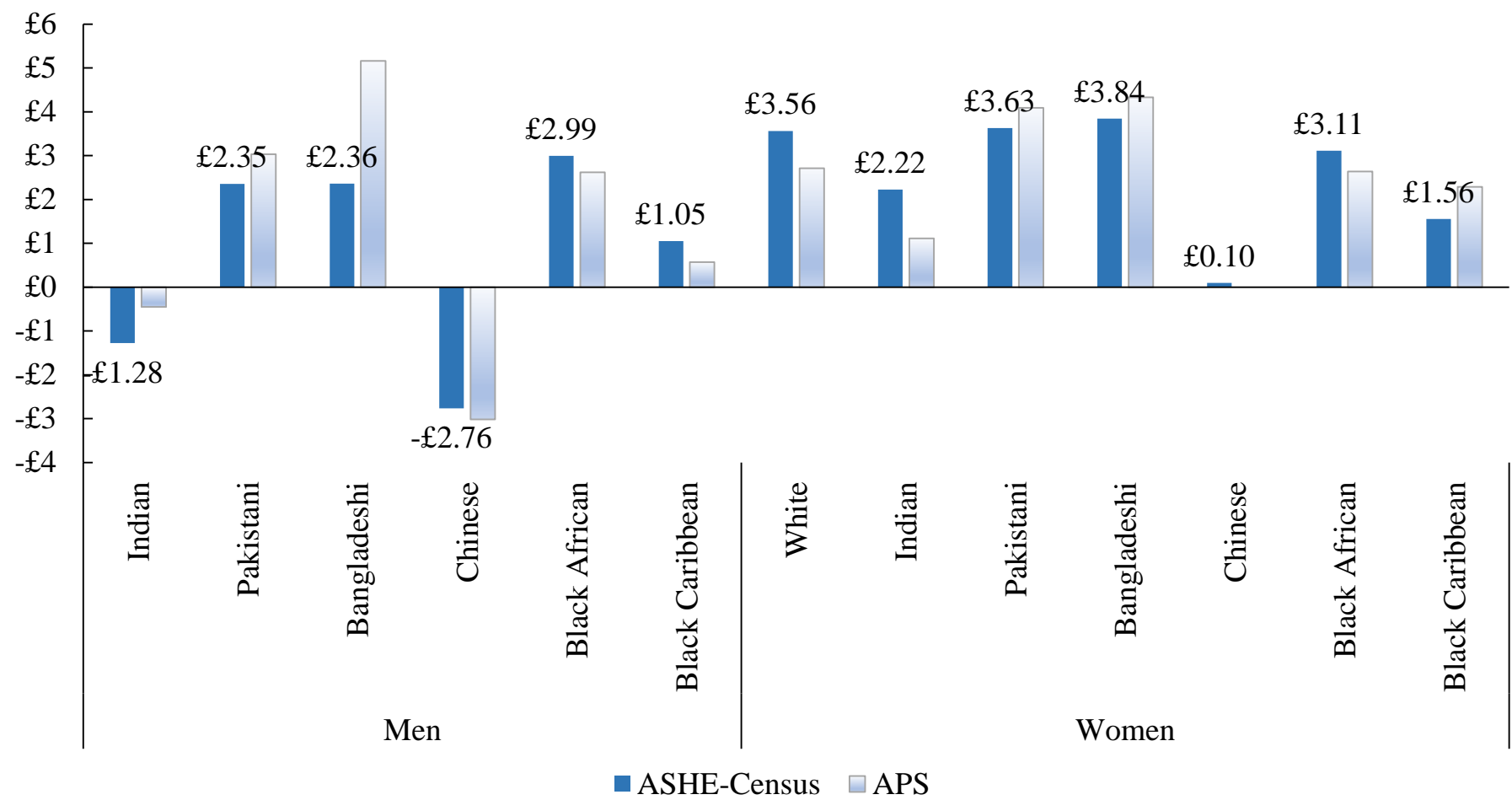
Mean Absolute Hourly Earnings wage gap in England and Wales in 2011  
- Difference from white men - Comparison of ASHE-Census and APS



Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here use survey weights. See slide 23 for sample wage statistics.



Median Absolute Hourly Earnings wage gap in England and Wales -  
Difference from white men - Comparison of ASHE-Census and APS



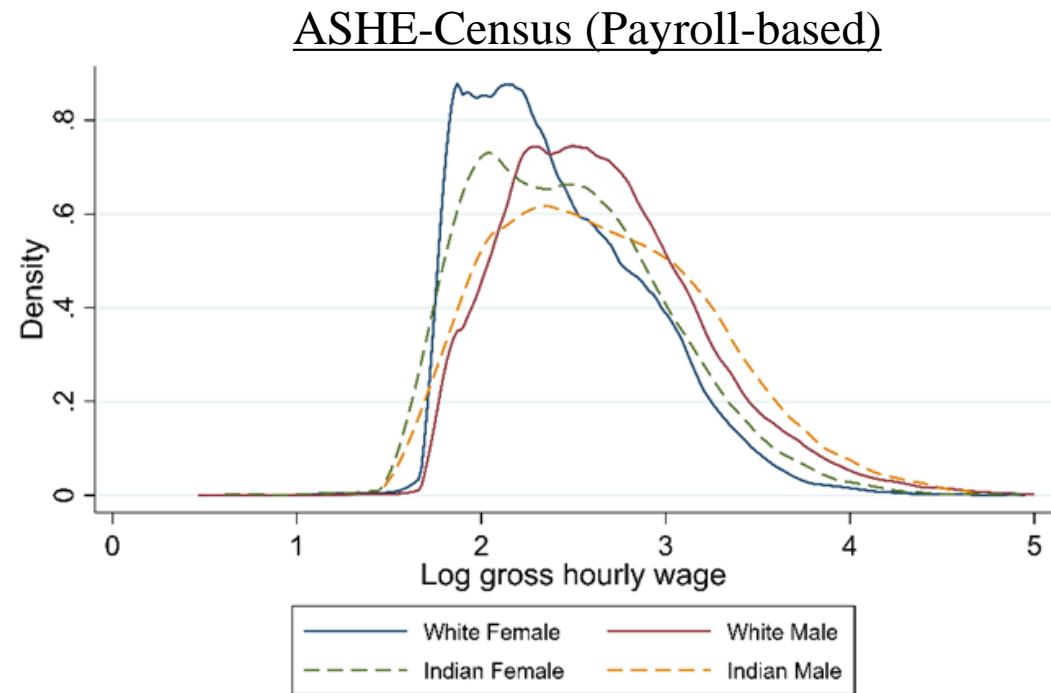
Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here use survey weights. See slide 25 for sample wage statistics.

## 2. Distributions of hourly earnings by ethnicity and gender:

*Were high or low wages more common in England and Wales for some groups of employees than others?*

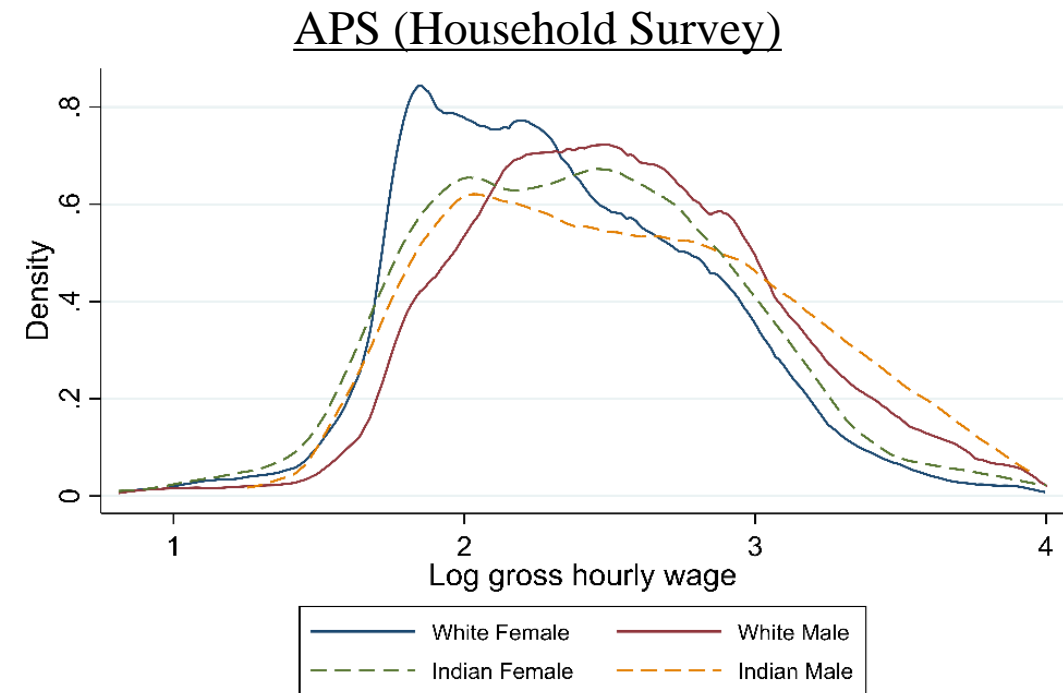
Distributions of hourly earnings in England and Wales in 2011

White vs Indian: Comparison of ASHE-Census and APS



ASHE-Census 2011

Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Indian



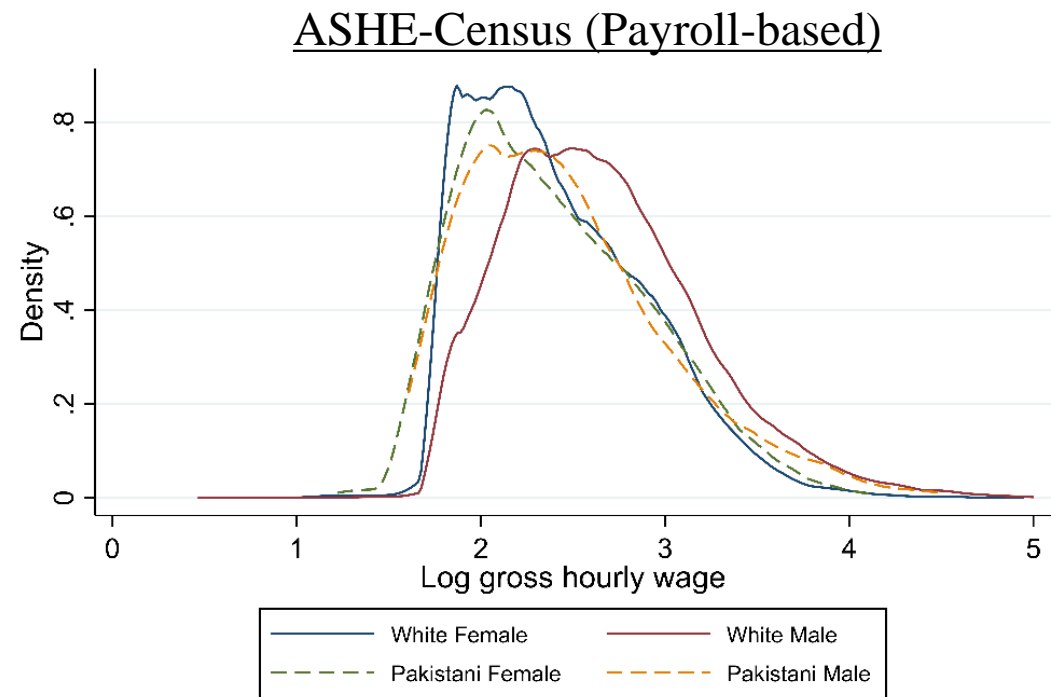
APS 2011

Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Indian

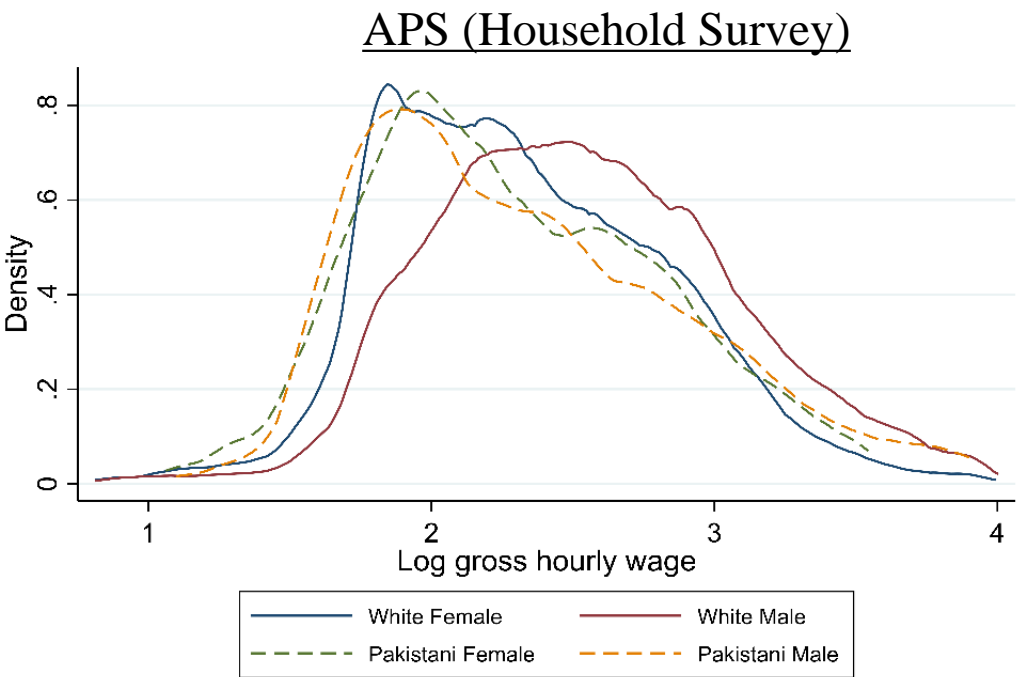
Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions. See slides 30+ for the same kernel density estimates but overlapping ASHE-Census and APS on the same sets of axes.

Distributions of hourly earnings in England and Wales in 2011  
White vs Pakistani: Comparison of ASHE-Census and APS



ASHE-Census 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Pakistani



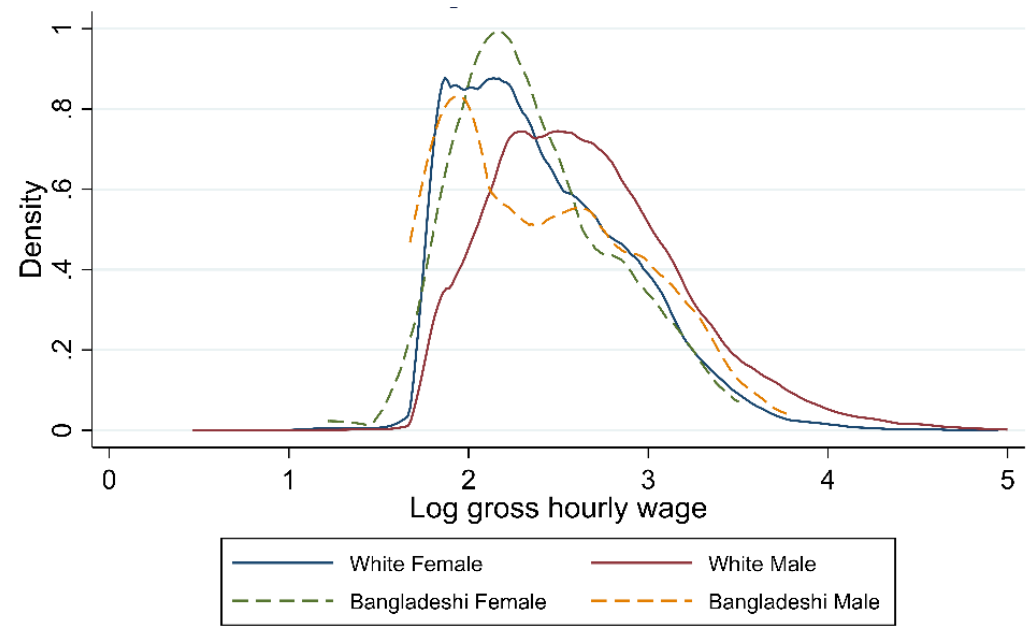
APS 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Pakistani

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions. See slides 30+ for the same kernel density estimates but overlapping ASHE-Census and APS on the same sets of axes.

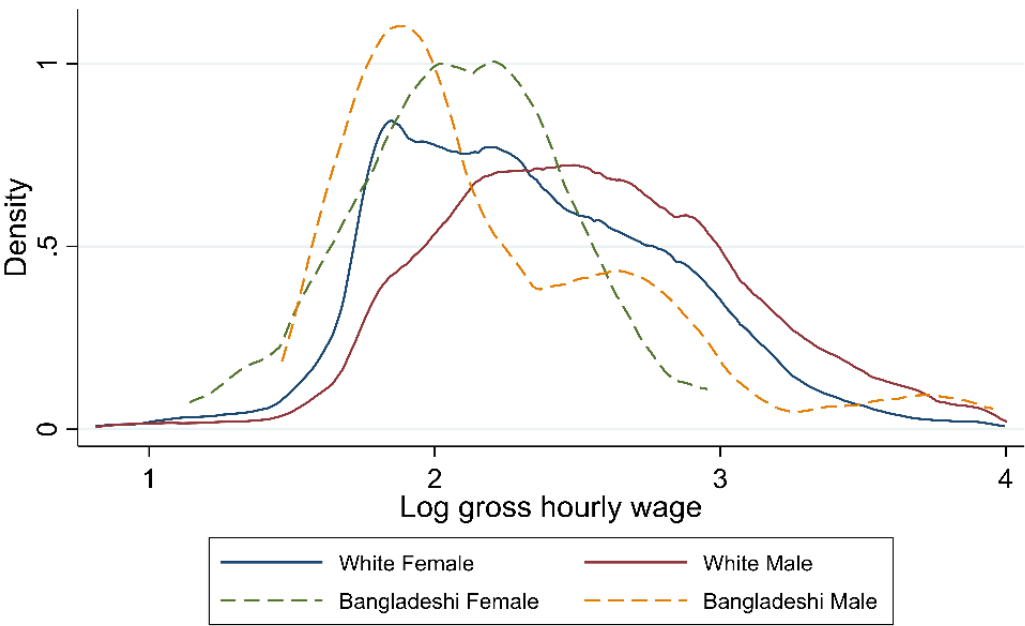
Distributions of hourly earnings in England and Wales in 2011  
White vs Bangladeshi: Comparison of ASHE-Census and APS

ASHE-Census (Payroll-based)



ASHE-Census 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Bangladeshi

APS (Household Survey)

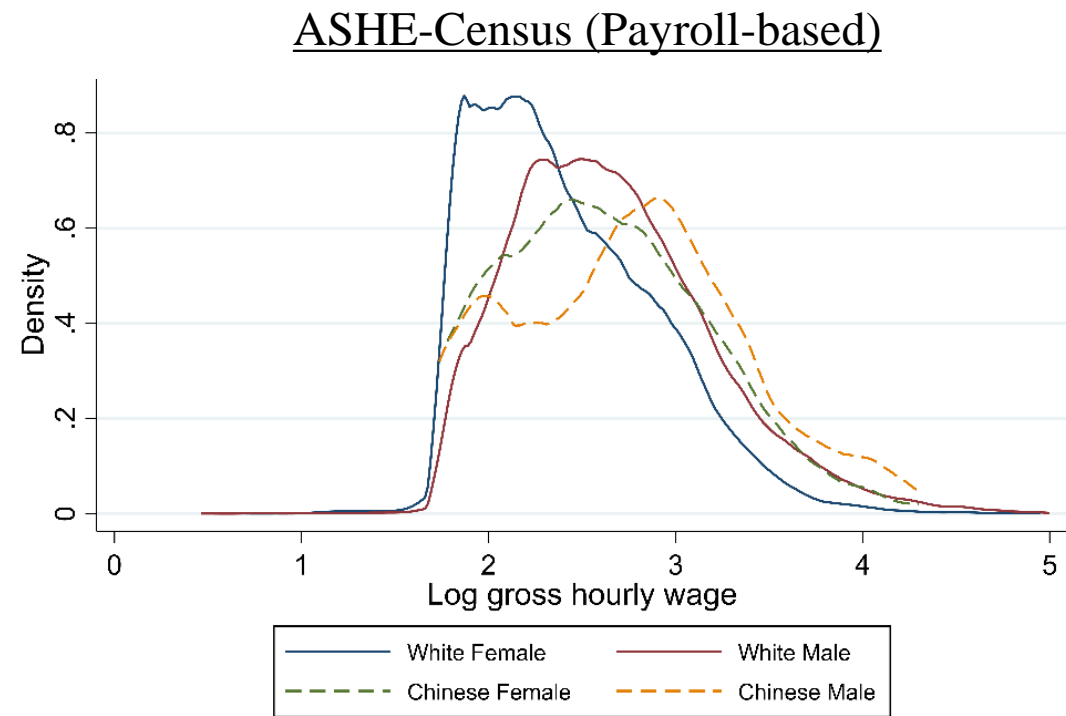


APS 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Bangladeshi

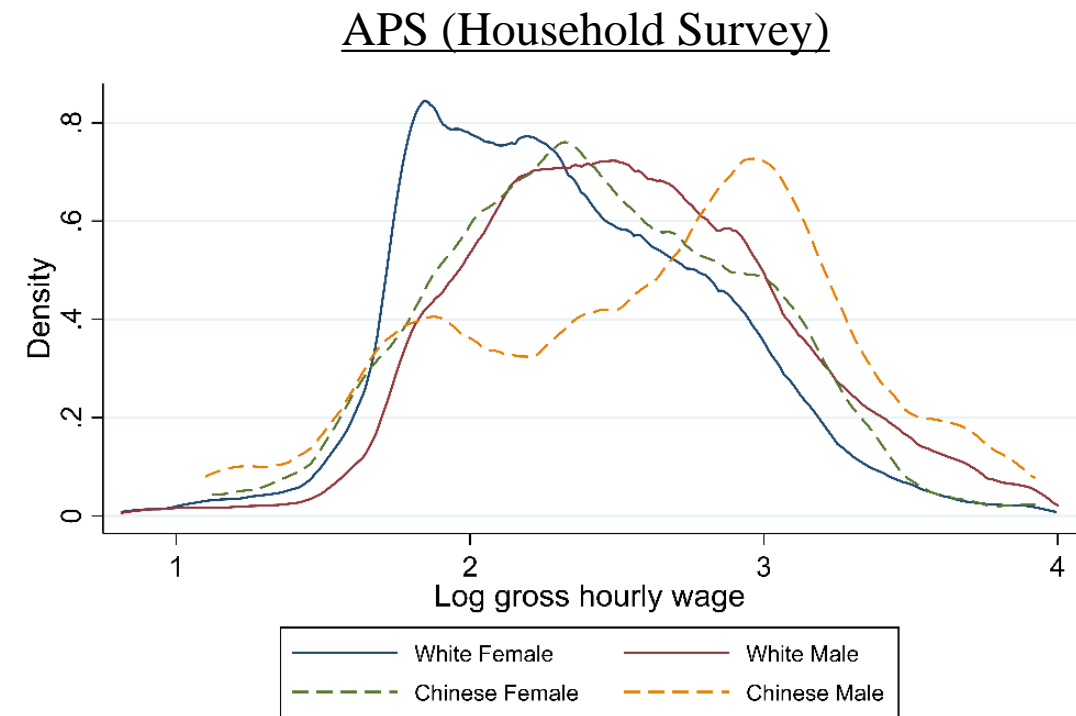
Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions. See slides 30+ for the same kernel density estimates but overlapping ASHE-Census and APS on the same sets of axes.

Distributions of hourly earnings in England and Wales in 2011  
White vs Chinese: Comparison of ASHE-Census and APS



ASHE-Census 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Chinese



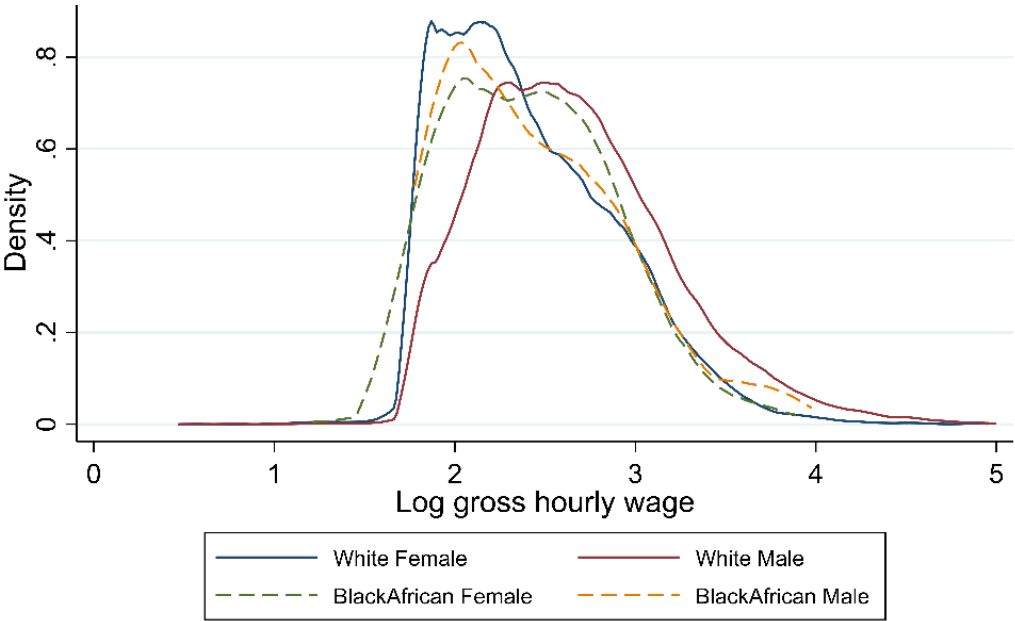
APS 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for Chinese

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions. See slides 30+ for the same kernel density estimates but overlapping ASHE-Census and APS on the same sets of axes.

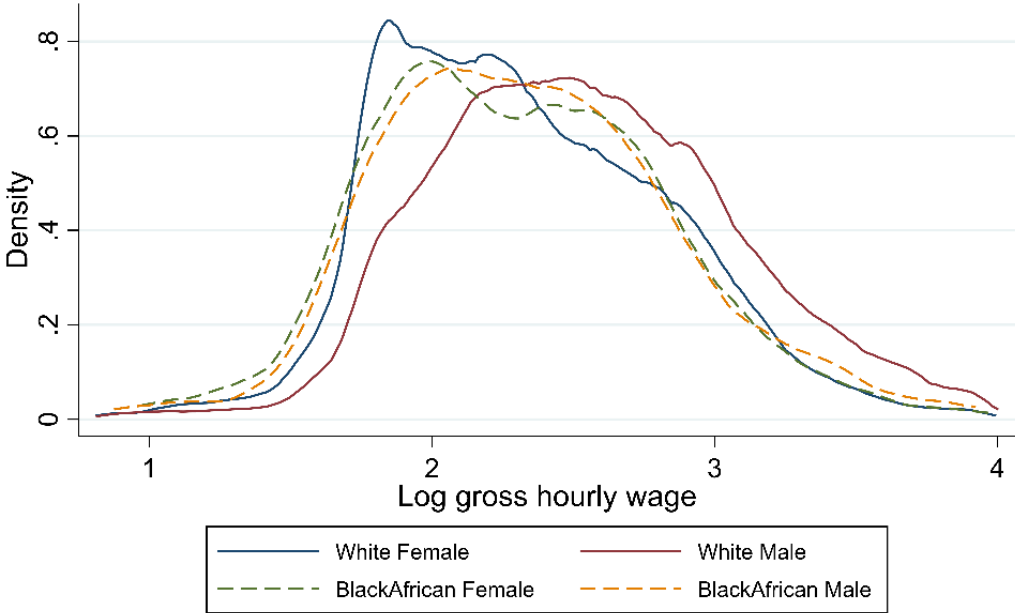
Distributions of hourly earnings in England and Wales in 2011  
White vs Black African: Comparison of ASHE-Census and APS

ASHE-Census (Payroll-based)



ASHE-Census 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for BlackAfrican

APS (Household Survey)

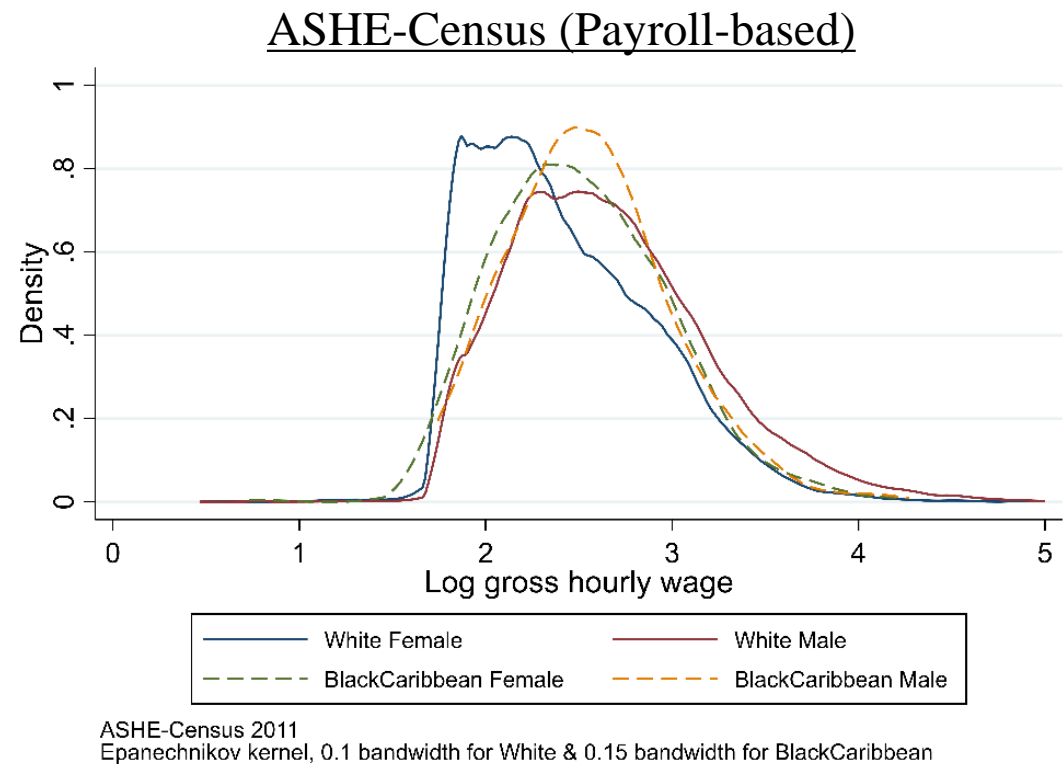


APS 2011  
Epanechnikov kernel, 0.1 bandwidth for White & 0.15 bandwidth for BlackAfrican

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.  
Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions. See slides 30+ for the same kernel density estimates but overlapping ASHE-Census and APS on the same sets of axes.



Distributions of hourly earnings in England and Wales in 2011  
White vs Black Caribbean: Comparison of ASHE-Census and APS

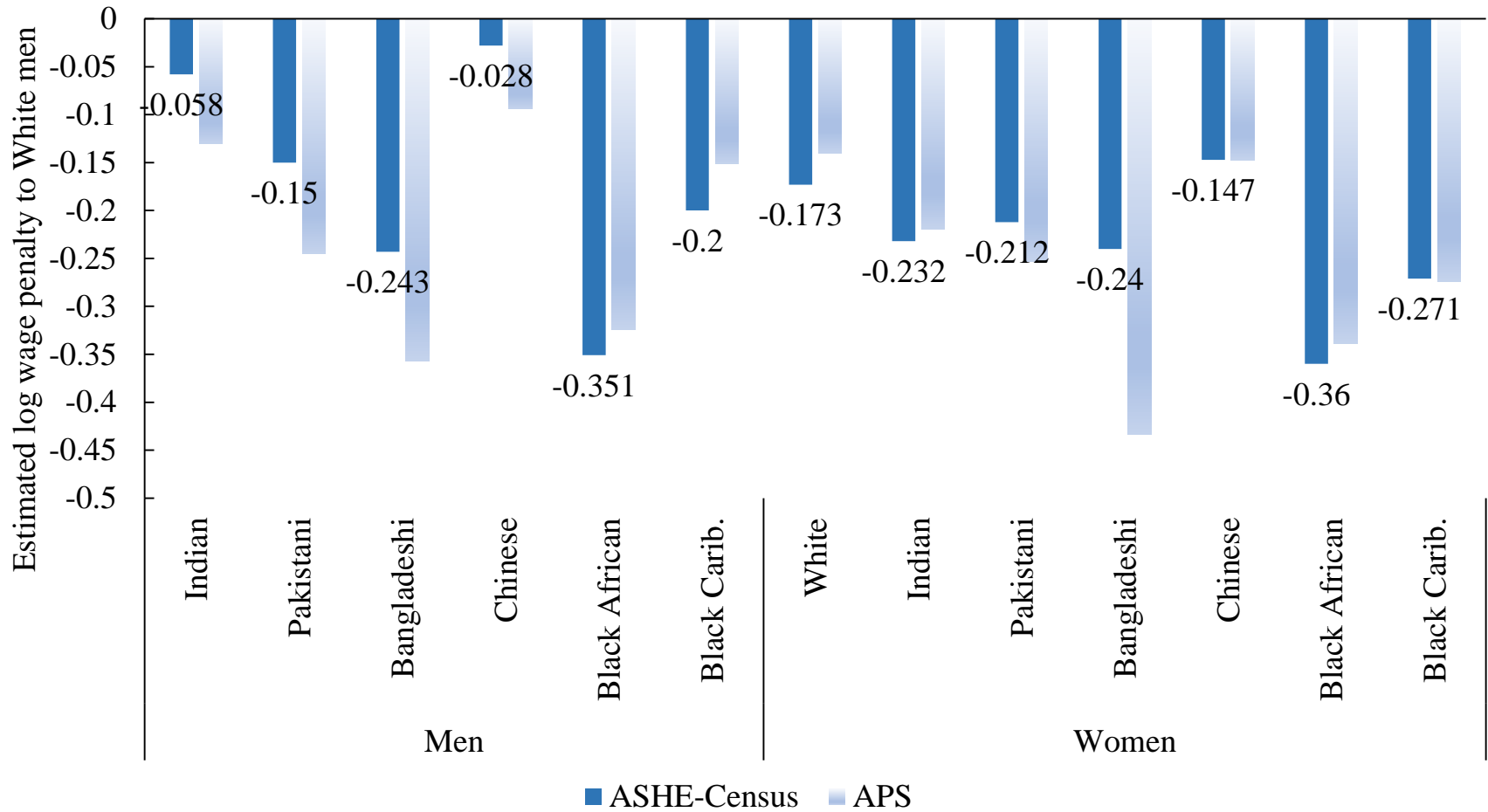




### 3. Estimated wage penalties associated with ethnicity-gender:

*What were the average wage differences associated with ethnicity for employees with the same levels of education, experience, occupation, and other characteristics?*

Adjusted Hourly Earnings wage penalty in England & Wales in 2011 -  
Difference from white men - Comparison of ASHE-Census and APS



Notes: author calculations using ASHE-Census & Annual Population Survey. Value labels are for ASHE-Census. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here use survey weights. See slide 27 for full estimates and measures of precision/statistical significance. Control variables are functions of age, education, marital status, work region, tenure, language, migrant status, health, part-time, and family characteristics. See slides 28-29 for more details.

## 4. Appendix

*Details on data sources, methods and additional results*

The ASHE-Census dataset – A payroll-based earnings dataset with detailed employee and job characteristics

Dataset: ASHE-Census 2011, those in England and Wales

Can be accessed: Soon, via Office for National Statistics Secure Research Service – Contact WED project team for advice. ASHE itself can be accessed as: Office for National Statistics. (2021). Annual Survey of Hours and Earnings, 1997-2020: Secure Access. [data collection]. 18th Edition. UK Data Service. SN: 6689, <http://doi.org/10.5255/UKDA-SN-6689-17>

Sample: those in the age group 25-64, main jobs only, on an adult rate of pay, not on an apprentice pay rate, and without a loss of pay for any reason in the reference period. We also drop those whose recorded hours of work were >99 or equal to 0.

The measure of “Hourly Earnings” used in ASHE is ‘he’. This is calculated from gross weekly pay divided by total hours worked, including any overtime. The sample is trimmed for those in the bottom 0.5 percentile, and in the top 0.5 percentile before all analysis and statistics are computed.

The weights used, where mentioned, were designed to address the ASHE-Census linking patterns, through a an adjustment of the standard cross-section survey weights available in the ASHE dataset.

## The Annual Population Survey – A household survey dataset

Dataset: APS individual 2011, observations working in England and Wales

Can be accessed as: Office for National Statistics, Social Survey Division. (2021). Annual Population Survey, 2004-2021: Secure Access. [data collection]. 20th Edition. UK Data Service. SN: 6721, <http://doi.org/10.5255/UKDA-SN-6721-19>

Sample: Employees not earning an apprentice rate of pay, aged 25+. We also drop those whose hourly earnings were >£99 or equal to £0.

The measure of “Hourly Earnings” used in APS is ‘hourpay’, which is calculated by  $\text{grsswk}/\text{bacthr}$  (gross weekly pay divided by reported basic actual hours). The sample is trimmed for those in the bottom 0.5 percentile, and in the top 0.5 percentile before all analysis and statistics are computed.

The weights used, where mentioned, are “pwta14” in the publicly available version of the dataset.

# Unweighted samples sizes for the analysis

	<u>ASHE-Census</u>		<u>APS</u>	
	Male	Female	Male	Female
White	38,130	37,961	21,754	24,374
Indian	1,189	1,176	686	597
Pakistani	514	314	321	152
Bangladeshi	171	98	119	55
Chinese	172	208	79	98
Black African	487	557	312	347
Black Caribbean	396	599	153	266
Total	41,059	40,913	23,424	25,889

# Mean Hourly Earnings in England and Wales in 2011 – Comparison of ASHE-Census and APS (weighted)

		ASHE-Census	APS
<u>Male</u>	White	£17.60	£15.03
	Indian	£19.47	£15.68
	Pakistani	£14.94	£12.60
	Bangladeshi	£13.38	£11.20
	Chinese	£19.24	£16.82
	Black African	£13.55	£12.04
	Black Caribbean	£14.56	£13.81
<u>Female</u>	White	£12.98	£12.00
	Indian	£14.62	£13.13
	Pakistani	£12.85	£10.97
	Bangladeshi	£12.10	£8.85
	Chinese	£16.34	£14.01
	Black African	£12.64	£11.70
	Black Caribbean	£14.18	£11.50

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-21 for details on these datasets, and slide 22 for unweighted sample sizes.

# Median Hourly Earnings in England and Wales in 2011 – Comparison of ASHE-Census and APS (weighted)

		ASHE-Census	APS
<u>Male</u>	White	£13.81	£12.66
	Indian	£15.09	£13.11
	Pakistani	£11.46	£9.63
	Bangladeshi	£11.45	£7.50
	Chinese	£16.58	£15.68
	Black African	£10.82	£10.04
	Black Caribbean	£12.76	£12.09
<u>Female</u>	White	£10.25	£9.95
	Indian	£11.59	£11.55
	Pakistani	£10.19	£8.57
	Bangladeshi	£9.97	£8.33
	Chinese	£13.72	£12.66
	Black African	£10.70	£10.02
	Black Caribbean	£12.26	£10.37

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-21 for details on these datasets, and slide 22 for unweighted sample sizes.



# Mean Hourly Earnings in England and Wales in 2011 – Comparison of ASHE-Census and APS (unweighted)

		ASHE-Census	APS
<u>Male</u>	White	£16.99	£14.59
	Indian	£18.19	£15.15
	Pakistani	£13.97	£12.31
	Bangladeshi	£12.92	£11.04
	Chinese	£18.72	£16.40
	Black African	£13.05	£11.96
	Black Caribbean	£14.22	£13.67
-			
<u>Female</u>	White	£12.87	£11.68
	Indian	£14.05	£12.66
	Pakistani	£12.77	£11.22
	Bangladeshi	£11.83	£8.68
	Chinese	£16.03	£12.99
	Black African	£12.57	£11.49
	Black Caribbean	£13.89	£11.58

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-21 for details on these datasets, and slide 22 for unweighted sample sizes.

# Median Hourly Earnings in England and Wales in 2011 – Comparison of ASHE-Census and APS (unweighted)

		ASHE-Census	APS
<u>Male</u>	White	£13.46	£12.28
	Indian	£13.63	£12.36
	Pakistani	£10.59	£9.53
	Bangladeshi	£11.01	£7.50
	Chinese	£15.74	£15.38
	Black African	£10.34	£10.04
	Black Caribbean	£12.58	£11.83
<u>Female</u>	White	£10.27	£9.63
	Indian	£11.40	£10.92
	Pakistani	£10.20	£8.74
	Bangladeshi	£9.94	£8.33
	Chinese	£13.12	£11.10
	Black African	£10.89	£9.74
	Black Caribbean	£11.81	£10.44

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-21 for details on these datasets, and slide 22 for unweighted sample sizes.

# Adjusted wage gap estimates

## Log wage regression model

## Comparison of ASHE-Census and APS

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-21 for details on these datasets, samples and survey weights.

Notes: Other control variables are functions of age, education, marital status, work region, tenure, English national identity, long-standing health problem, part-time, and family characteristics. See slides 28-29 for more details.

\*, \*\*, \*\*\* indicate statistical significance from zero at 10%, 5%, 1% levels, two-sided, robust standard errors in parentheses.

Interpretation: The log wage penalty associated with a characteristic (approx. %), compared to excluded category. E.g., the adjusted penalty for an Indian man compared to a white man, in the first column, is  $\{-0.073 - 0.001 = -0.072$ , or 7.2 log points, or approx. 7%.

	Unweighted		Weighted	
	ASHE-Census	APS	ASHE-Census	APS
Man (excl. Female)	0.158*** (0.003)	0.140*** (0.004)	0.173*** (0.004)	0.140*** (0.005)
Indian (excl. White)	-0.073*** (0.013)	-0.083*** (0.019)	-0.059*** (0.017)	-0.079*** (0.022)
Pakistani	-0.038 (0.024)	-0.109*** (0.034)	-0.039 (0.025)	-0.113*** (0.039)
Bangladeshi	-0.065* (0.034)	-0.275*** (0.047)	-0.067* (0.036)	-0.294*** (0.048)
Chinese	0.027 (0.031)	-0.047 (0.045)	0.026 (0.032)	-0.007 (0.058)
Black African	-0.189*** (0.016)	-0.194*** (0.024)	-0.187*** (0.017)	-0.199*** (0.027)
Black Caribbean	-0.093*** (0.016)	-0.100*** (0.022)	-0.098*** (0.018)	-0.134*** (0.027)
Indian & Male	-0.001 (0.020)	-0.042 (0.026)	0.001 (0.025)	-0.051* (0.030)
Pakistani & Male	-0.130*** (0.031)	-0.107** (0.042)	-0.111*** (0.034)	-0.132*** (0.051)
Bangladeshi & Male	-0.177*** (0.046)	-0.060 (0.063)	-0.176*** (0.051)	-0.063 (0.068)
Chinese & Male	-0.062 (0.049)	-0.045 (0.074)	-0.054 (0.051)	-0.087 (0.093)
Black African & Male	-0.150*** (0.026)	-0.110*** (0.036)	-0.164*** (0.030)	-0.125*** (0.041)
Black Caribbean & Male	-0.096*** (0.025)	-0.036 (0.039)	-0.102*** (0.027)	-0.017 (0.045)
Disability	-0.062*** (0.006)	-0.049*** (0.005)	-0.061*** (0.007)	-0.046*** (0.006)
R-squared	0.425	0.385	0.439	0.383
Observations	75,234	49,030	75,234	49,018

# Variables in the wage regression models

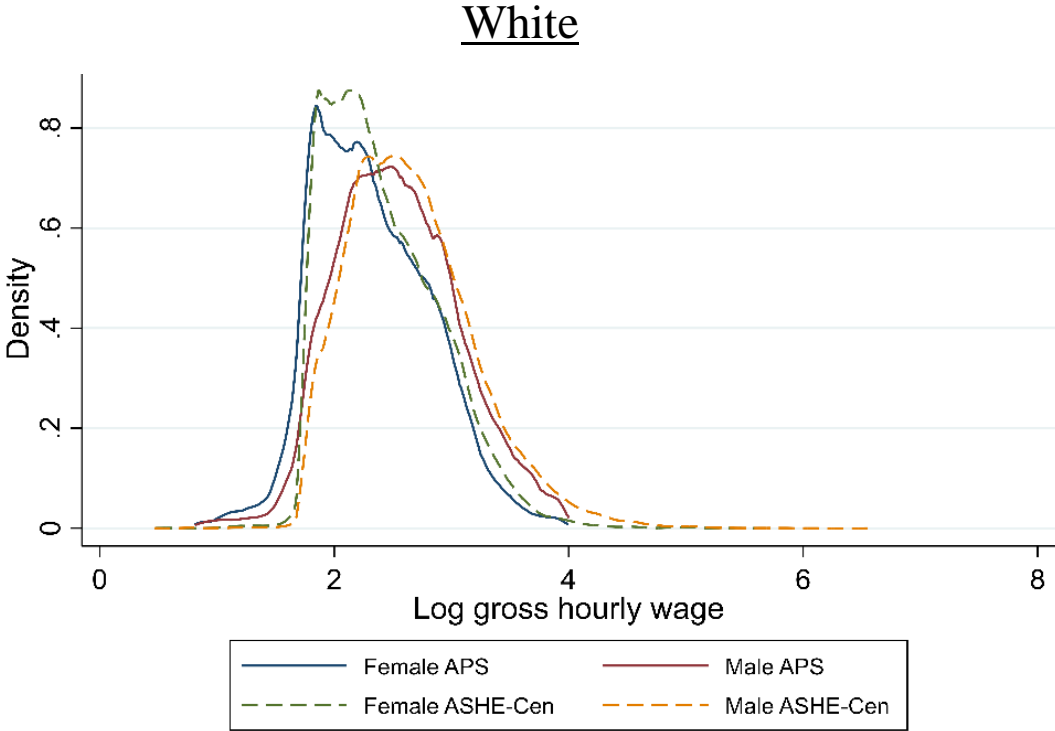
Variable	Categories	ASHE-Census	APS
Hourly pay/earnings	Continuous variable	he	hourpay
Male	Dummy variable	sex	sex
Ethnicity	Categorical variable (White, Indian, Pakistani, Bangladeshi, Chinese, Black Caribbean, Black Africa; observations in the Mixed and Other categories are omitted from all the results)	ethpuk11	Ethew18
Age	Categorical variable for APS Continuous variable for ASHE-Census, with squared value	age	ages
Education	Categorical variable (ASHE-Census: No qualification, GCSEs, apprenticeship, A-level, Degree, Other/vocational qualification) (APS: NQF level 4 and above, NQF level 3, Trade apprenticeships, NQF level 2, below NQF level 2, Other qualification, No qualification)	hlqpuk11	levqul11
Marital status	Dummy variable (1 for those married or in a registered same-sex civil partnership)	marstat	marsta
Disability	ASHE-Census: Long-term sick or disabled. APS: Dummy variable for those who report a disability that restricts or limits their work	disability	discurr
Work region	Categorical variable (NUTS1: North East, North West, Yorkshire, East Midlands, West Midlands, South West, East, London, South East, Wales)	wgor	gorwkr
Tenure	Continuous variable. This is derived from when an employee started working for their employer. The squared value is included	empsta	conmpy

# Variables in the wage regression models

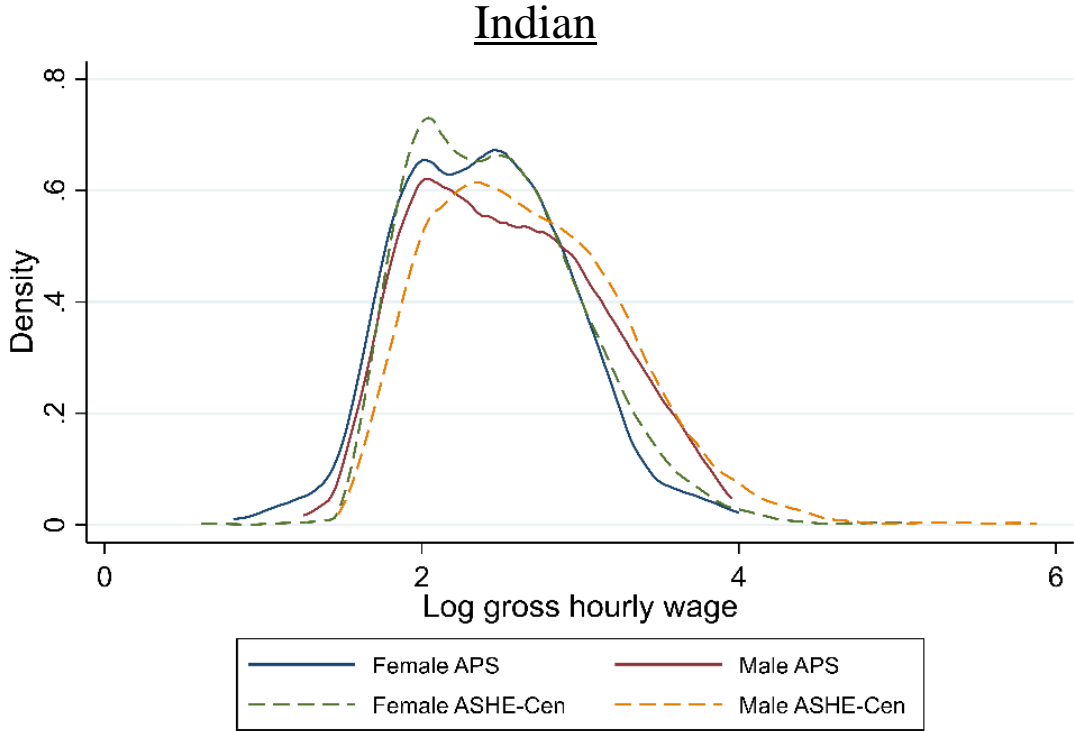
Variable	Categories	ASHE-Census	APS
English nationality	Dummy variable for English or non-English	lrespuk11	natide11
English language	Dummy variable (1: English is the main language)	mainlang	-
Health problem	Dummy variable whether the person has a long-standing and work-limiting health problem	ehlthm	health
Part-time job	Dummy variable. In ASHE-Census, it is derived from basic working hours (<30)	h_basic	ftptwk
Number of children	Continuous variable – number in family	dpcfamuk11	fdpch19
Age of the youngest child	Categorical variable. This is a derived variable from the number of children in each age group in the family	dpcfamuk11	fdpch2, fdpch4, fdpch9, fdpch15, fdpch16, fdpch19
Carer duty	Dummy variable for people who report a responsibility as a carer for somebody with a disability or old-age	carer	-
Industry Sector of Employer	Categorical variable (1 digit classification)	sic07	inde07m
Occupation	Categorical variable (1digit classification)	occ10	nsecmj10

# Distributions of hourly earnings in England and Wales in 2011

## Comparison of ASHE-Census and APS – White & Indian



Epanechnikov kernel, 0.05 bandwidth.



Epanechnikov kernel, 0.15 bandwidth.

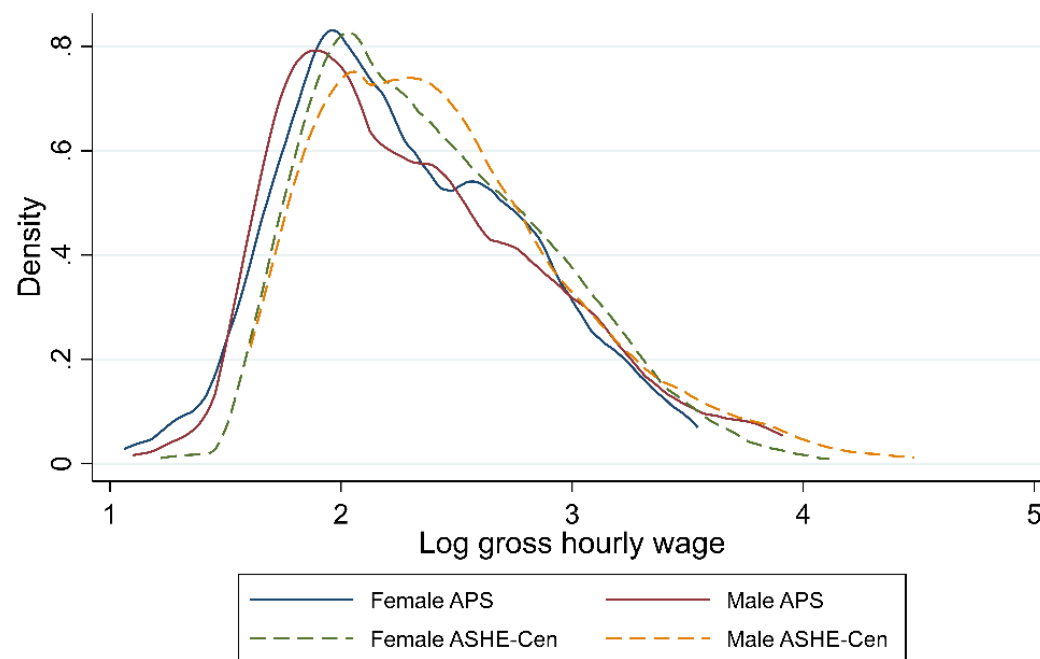
Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions.

# Distributions of hourly earnings in England and Wales in 2011

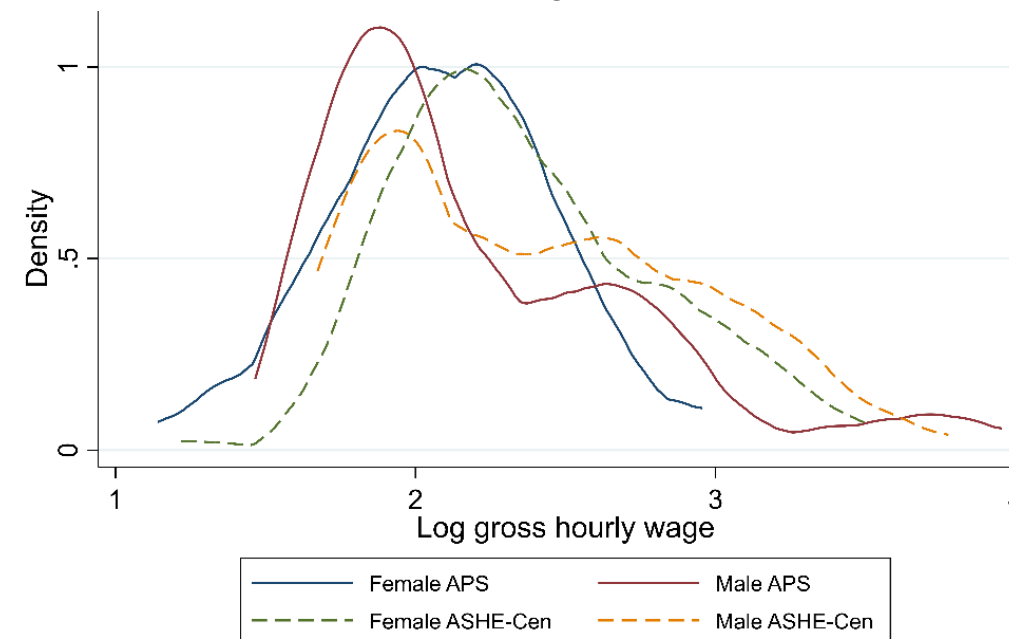
## Comparison of ASHE-Census and APS – Pakistani and Bangladeshi

Pakistani



Epanechnikov kernel, 0.15 bandwidth.

Bangladeshi



Epanechnikov kernel, 0.15 bandwidth.

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

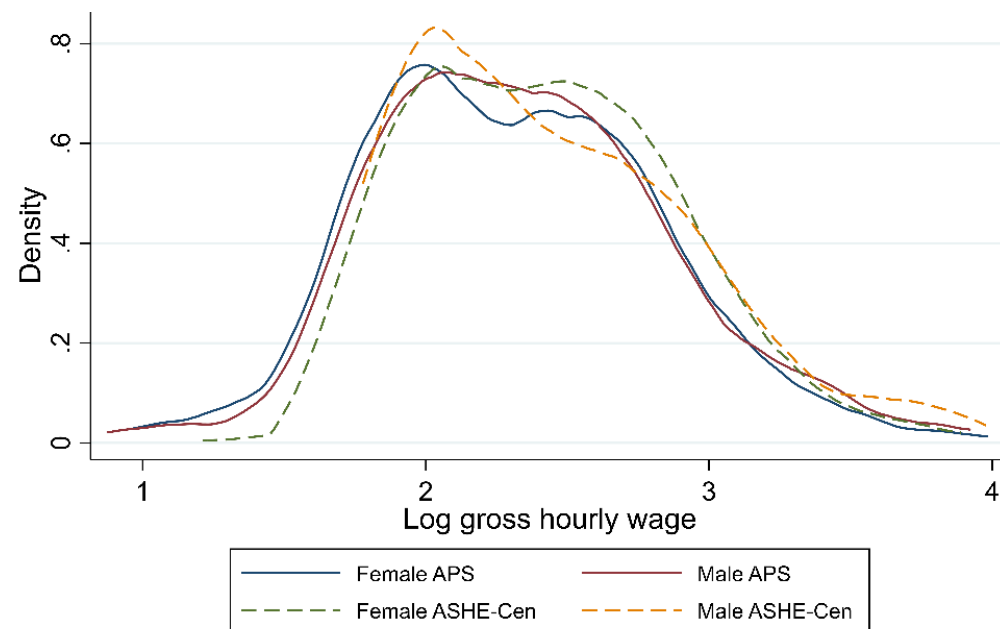
Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions.



# Distributions of hourly earnings in England and Wales in 2011

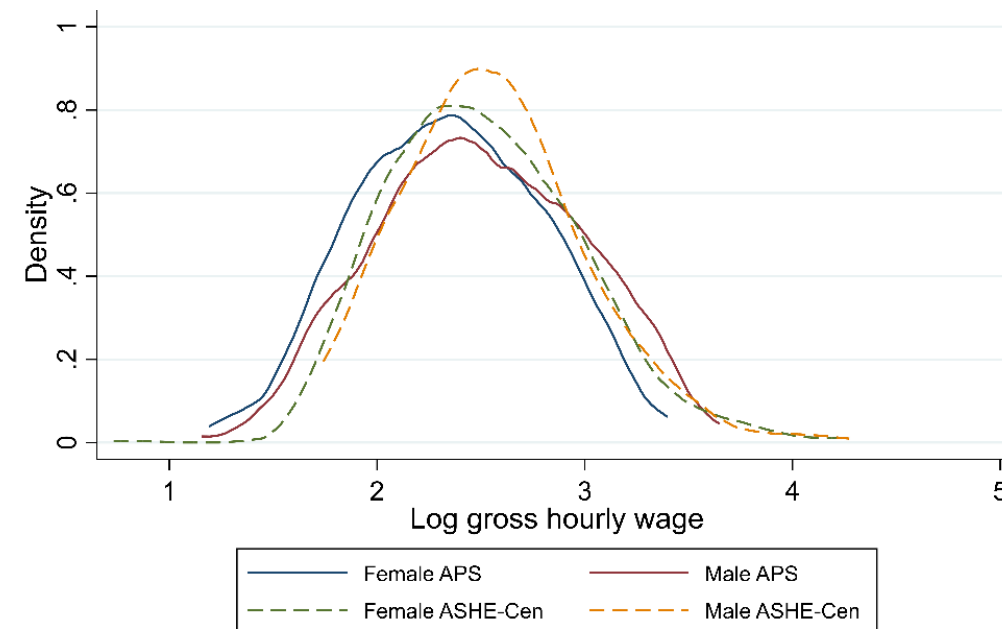
## Comparison of ASHE-Census and APS – Black African and Black Caribbean

### Black African



Epanechnikov kernel, 0.15 bandwidth.

### Black Caribbean



Epanechnikov kernel, 0.15 bandwidth.

Source: author calculations using ASHE-Census & Annual Population Survey. See slides 19-22 for details on these datasets and samples, the wage measure, and unweighted sample sizes. Statistics here do not use survey weights Constructed as kernel density estimates, with chosen bandwidth displayed above.

Interpretation: shows the tendency (density) of employees in each group to earn relatively high or low wages, with the log scale on the X-axis compressing the displayed distributions.