

# The forgotten form of stratification: Sexual Orientation in large social survey research

Dr Scott Oatley

[scott.oatley@manchester.ac.uk](mailto:scott.oatley@manchester.ac.uk)

University of Manchester

# SUMMER SESSIONS



---

# Social Stratification

---

- Typically dominated by discussions of social class, gender, and ethnicity (Big Three)
- Sexual Orientation is often left out of models of social stratification

---

# Forgotten?

---

- Left out or forgotten entirely?
- There are a limited number of large social surveys in the UK that collect sexual orientation data
  - Even fewer that collect it well

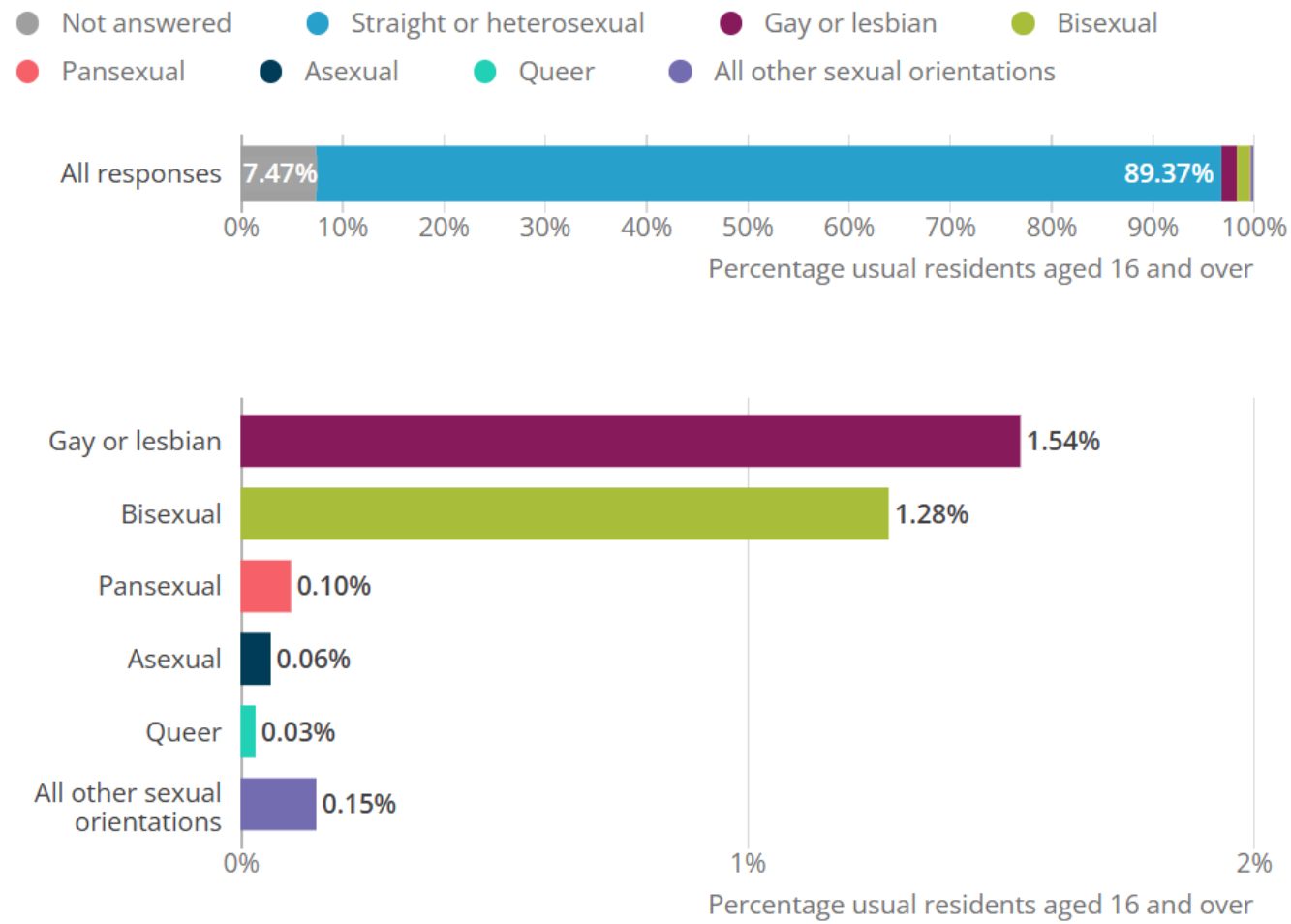
---

# Issues

---

- Data Collection strategies
  - Pretty hard to ask
  - Pretty obvious candidate for high levels of missingness
  - 'Moral panic' over asking young people
- Real world issues
  - The 'Queer' population is small

Figure 1: Sexual orientation, 2021, England and Wales



Source: Office for National Statistics – Census 2021

---

# Issues

---

- Real world issues
  - The 'Queer' population is small
- Very small sample size
- To my knowledge NO large social surveys in UK that collect sexual orientation construct appropriate weights for this population

---

# Issues

---

- Bad Collection and Recording
  - Couple data
  - Adults only
  - Only asking binary questions
  - Special License Locked



# Making the most out of what is left

---

- Enter the UKHLS
  - Contemporary large social survey in the UK
  - Has a robust question on sexual orientation that is collected from wave 3 and asked every other wave since then
- Small, but nationally representative sub-populations
- Unfortunately, no weights include sexual orientation – this was indicated on UKHLS forums at some point but seems to be forgotten
  - (Wouldn't that be a lovely postdoc...)

# Sexuality Pay Gap

---

Working Paper

---

# Sexuality Pay Gap

---

- Using UKHLS data to study the Sexuality Pay Gap
  - Very few papers on this topic worldwide
  - Almost exclusively US led
  - A few papers from the UK
  - Consensus is far from settled
- UKHLS also allows us to go beyond Pay gap analysis and look at growth curves
  - This has never been done for the study of sexual orientation pay gaps before

# UK Sexuality Pay Gap

---

- Aksoy (2018)
- Arabsheibani (2005)
- Arabsheibani (2007)
- Booth and Frank (2008)
- Bridges and Mann (2019)
- Bryson (2017)
- Frank (2008)

# UK Sexuality Pay Gap

---

- 7 studies
- 2 not nationally representative (Teacher survey + Academic survey)
- 5 nationally representative studies between 2005-2018 (13 years)
- Seminal research on this topic began in 1995 (Badgett 1995)
- If we add (Klawitter 2015; Drydakis 2022) 22% of the papers written on UK sexuality pay gaps are meta-analyses on the topic...

# Meta-analysis (Klawitter, 2015)

---

- From the earliest paper on sexuality income gaps (Badgett 1995) to 2015
- Consistently shows gay men earn less than straight men
- Lesbian women are sometimes more likely to outearn straight women
- Non-US studies report smaller earnings gaps
- Sexual orientation measured through self-identity rather than couple status or sexual behaviour reports smaller gaps
- Annual earnings rather than hourly reports larger earnings gap %
- Limiting to full-time workers increases earnings gap

---

# Meta- analysis (Klawitter, 2015)

---

- Earnings penalty of 11% for gay men
- Earnings premium of 9% for lesbian women
- No information for bisexuals...

---

# Working Data

---

- Sample consists of Waves 3-14 of UKHLS
- Sample includes those aged 16-66 AND those in some form of employment AND not in full time training or education
- Leaves us with a N=294,377 over 11 waves of data



# Non-starter Analysis?

---

Sexuality	Frequency (%)
Heterosexual	42,258 (14.36%)
Homosexual	738 (0.25%)
Bisexual	676 (0.23%)
Don't Know	29 (0.01%)
Inapplicable	75,041 (25.49%)
Missing	174,579 (59.30%)
Refused	1,056 (0.36%)

---

# Across Waves

UKHLS wave															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
N	27,340 (9.3%)	27,244 (9.3%)	24,783 (8.4%)	23,108 (7.8%)	22,284 (7.6%)	22,728 (7.7%)	21,670 (7.4%)	20,732 (7.0%)	19,197 (6.5%)	18,262 (6.2%)	17,100 (5.8%)	15,594 (5.3%)	15,196 (5.2%)	19,139 (6.5%)	294,377 (100.0%)
Sexuality															
Heterosexual	0 (.%)	0 (.%)	21,546 (97.5%)	0 (.%)	1,034 (94.8%)	0 (.%)	998 (94.3%)	0 (.%)	17,414 (96.7%)	0 (.%)	719 (91.2%)	0 (.%)	547 (88.1%)	0 (.%)	42,258 (96.8%)
Homosexual	0 (.%)	0 (.%)	341 (1.5%)	0 (.%)	31 (2.8%)	0 (.%)	25 (2.4%)	0 (.%)	318 (1.8%)	0 (.%)	10 (1.3%)	0 (.%)	13 (2.1%)	0 (.%)	738 (1.7%)
Bisexual	0 (.%)	0 (.%)	215 (1.0%)	0 (.%)	26 (2.4%)	0 (.%)	35 (3.3%)	0 (.%)	280 (1.6%)	0 (.%)	59 (7.5%)	0 (.%)	61 (9.8%)	0 (.%)	676 (1.5%)

---

# Non- starter Analysis?

---

- From base data sexuality is collected every other wave from wave 3
- Not a whole lot to work with
- Using Last Observation Carried Forward can fill in a lot of blanks here
- Last Observation Carried Backwards to get to wave 1-2 data is dangerous when dealing with sexual orientation data
  - We have no way of knowing when someone has 'come out'

---

	Male	Female	Total
Sexuality	Frequency	Frequency	Frequency
	(%)	(%)	(%)
Heterosexual	82,312	96,547	178,859
	(59.51%)	(61.87%)	(60.76%)
Homosexual	1,831	1,395	3,226
	(1.32%)	(0.89%)	(1.10%)
Bisexual	865	1,527	2,392
	(0.63%)	(0.98%)	(0.81%)
Missing	53,310	56,580	109,890
	(38.54%)	(36.62%)	(37.33%)
Total	138,318	156,049	294,367
	(100%)	(100%)	(100%)

---

---

	Male	Female	Total
Sexuality	Frequency	Frequency	Frequency
	(%)	(%)	(%)
Heterosexual	82,312	96,547	178,859
	(96.83%)	(97.06%)	(96.95%)
Homosexual	1,831	1,395	3,226
	(2.15%)	(1.40%)	(1.75%)
Bisexual	865	1,527	2,392
	(1.02%)	(1.54%)	(1.30%)
Total	85,008	99,469	184,477
	(100%)	(100%)	(100%)

---

- 
- Getting close to 2021 census data
  - Still large amounts of missingness
  - Could derive orientation from couple data
    - Bad idea
    - Erasure of bisexuality often occurs

	UKHLS wave														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	Total
<b>N</b>	27,340 (9.3%)	27,244 (9.3%)	24,783 (8.4%)	23,108 (7.8%)	22,284 (7.6%)	22,728 (7.7%)	21,670 (7.4%)	20,732 (7.0%)	19,197 (6.5%)	18,262 (6.2%)	17,100 (5.8%)	15,594 (5.3%)	15,196 (5.2%)	19,139 (6.5%)	294,377 (100.0%)
<b>Sexuality</b>															
<b>Heterosexual</b>	0 (.%)	0 (.%)	21,546 (97.5%)	17,035 (97.7%)	16,325 (97.4%)	13,974 (97.5%)	13,697 (97.3%)	12,444 (97.3%)	17,802 (96.6%)	15,345 (96.6%)	14,345 (96.4%)	12,633 (96.4%)	12,239 (96.0%)	11,475 (96.0%)	178,860 (97.0%)
<b>Homosexual</b>	0 (.%)	0 (.%)	341 (1.5%)	265 (1.5%)	270 (1.6%)	231 (1.6%)	231 (1.6%)	211 (1.7%)	332 (1.8%)	298 (1.9%)	277 (1.9%)	259 (2.0%)	261 (2.0%)	250 (2.1%)	3,226 (1.7%)
<b>Bisexual</b>	0 (.%)	0 (.%)	215 (1.0%)	144 (0.8%)	160 (1.0%)	126 (0.9%)	152 (1.1%)	128 (1.0%)	294 (1.6%)	234 (1.5%)	256 (1.7%)	206 (1.6%)	250 (2.0%)	227 (1.9%)	2,392 (1.3%)

- 
- Pretty healthy numbers from wave 3-14
  - Remember the longitudinal context
    - Is sexual orientation fixed?
      - Common question that pops its head every so often with supposedly ‘fixed’ variables such as ethnicity as well
  - IFF orientation is not fixed it would be inappropriate to use LOCF to fill in item missingness across waves for individual pidp units



	MALE	FEMALE	TOTAL
Total Sexuality Switchers	Frequency (%)	Frequency (%)	Frequency (%)
Bisexual -> Homosexual	3 (4.41%)	4 (3.57%)	7 (3.89%)
Bisexual -> Heterosexual	26 (38.24%)	31 (27.68%)	57 (31.67%)
Homosexual -> Bisexual	2 (2.94%)	5 (4.46%)	7 (3.89%)
Homosexual -> Heterosexual	3 (4.41%)	4 (3.57%)	7 (3.89%)
Heterosexual -> Bisexual	24 (35.29%)	58 (51.79%)	82 (45.56%)
Heterosexual -> Homosexual	10 (14.71%)	10 (8.93%)	20 (11.11%)
Total	68 (100%)	112 (100%)	180 (100%)

- 
- Out of 184,477 cases of sexual orientation only 180 people across 11 waves of UKHLS data have 'switched' their sexual orientation
  - Out of these 180 people 178 of them have only switched once
    - This is capturing people 'coming out'
      - The lower mean age of switchers at 35.5 compared to sample average of 42 seems to corroborate this
    - There are a handful of sexual minority -> heterosexuality switchers however
      - These people appear to have 'dipped their toes'
      - This is dominated by the bisexual -> heterosexual and vice versa categories, which are themselves dominated by women

- 
- Out of these 180 people 2 of them switched twice
  - Person A is recorded as heterosexual at wave 3-4, at wave 5 recorded as homosexual until wave 6 then from wave 8-14 heterosexual again
  - Person B is recorded at waves 3-4 as bisexual then from wave 5-8 recorded as homosexual and then from waves 9-14 bisexual again
  - Both individuals are women

# Switchers

---

- Given the very low number of people that appear to switch sexual orientations over the life course it is fair in my view to consider sexuality as 'fixed'

# Occupational Sorting

---

- The occupations that individuals sort into has a knock on impact when assessing wages
- Beyond wages, occupational sorting is a clear indicator of societal and cultural expectations, norms, and influences on individual behaviours that are constituted via larger unit groups
- Brief look at largest occupational titles ( $\Rightarrow 2\%$ ) sorted into by sexuality, separated by sex
  - Using SOC 2000 codes

# Men

Heterosexual Men	%		Homosexual Men	%		Bisexual Men	%
1121 Production, works and maintenance managers	2.4		1136 Information and communication technology managers	2.04		2132 Software professionals	6.97
1132 Marketing and sales managers	2.6		1135 Personnel, training and industrial relations managers	2.27		2211 Medical practitioners	2.66
1239 Managers and proprietors in other services n.e.c.	2.08		2132 Software professionals	2.5		2319 Teaching professionals n.e.c.	2.53
2132 Software professionals	2.36		2314 Secondary education teaching professionals	2.45		5434 Chefs, cooks	2.28
7111 Sales and retail assistants	2.11		2315 Primary and nursery education teaching professionals	2.39		6115 Care assistants and home carers	3.8
9259 Elementary sales occupations n.e.c.	2		7111 Sales and retail assistants	4.02		7111 Sales and retail assistants	2.92
						8211 Heavy goods vehicle drivers	2.15
						8214 Taxi, cab drivers and chauffeurs	2.79
						9223 Kitchen and catering assistants	2.15
						9241 Security guards and related occupations	3.93

# Women

Heterosexual Women			Homosexual Women			Bisexual Women	
2314 Secondary education teaching professionals	2.36		1152 Office managers	2.27		3211 Nurses	2.13
2315 Primary and nursery education teaching professionals	3.12		2315 Primary and nursery education teaching professionals	3.11		3539 Business and related associate professionals n.e.c.	2.99
3211 Nurses	4.69		3211 Nurses	4.93		4150 General office assistants/clerks	2.49
4122 Accounts and wages clerks, book-keepers, other financial clerks	3.11		3312 Police officers (sergeant and below)	2.43		6111 Nursing auxiliaries and assistants	2.28
4150 General office assistants/clerks	4.26		6111 Nursing auxiliaries and assistants	2.73		6115 Care assistants and home carers	4.48
6115 Care assistants and home carers	5.29		6115 Care assistants and home carers	3.64		7111 Sales and retail assistants	5.9
6124 Educational assistants	3.93		7111 Sales and retail assistants	2.96		7212 Customer care occupations	2.35
7111 Sales and retail assistants	4.66		7212 Customer care occupations	2.05		9223 Kitchen and catering assistants	3.63
9233 Cleaners, domestics	2.4					9233 Cleaners, domestics	2

# Occupational Sorting

---

- Women from all sexual orientations and bi men appear to be more heavily concentrated in a few occupational titles compared to straight and gay men
- Bi men are concentrated in <4000 occupations. Possible driver of a wage gap?
  - 7% of bi men are software professionals...
- Women across sexuality appear to focus on soft-skilled labour
  - Though gay and bi women appear to concentrate in authority-related occupations. Possible driver of a pay premium?
- Gay men concentrate in similarly soft-skilled labour positions (teachers)
  - Possible driver of a pay penalty?



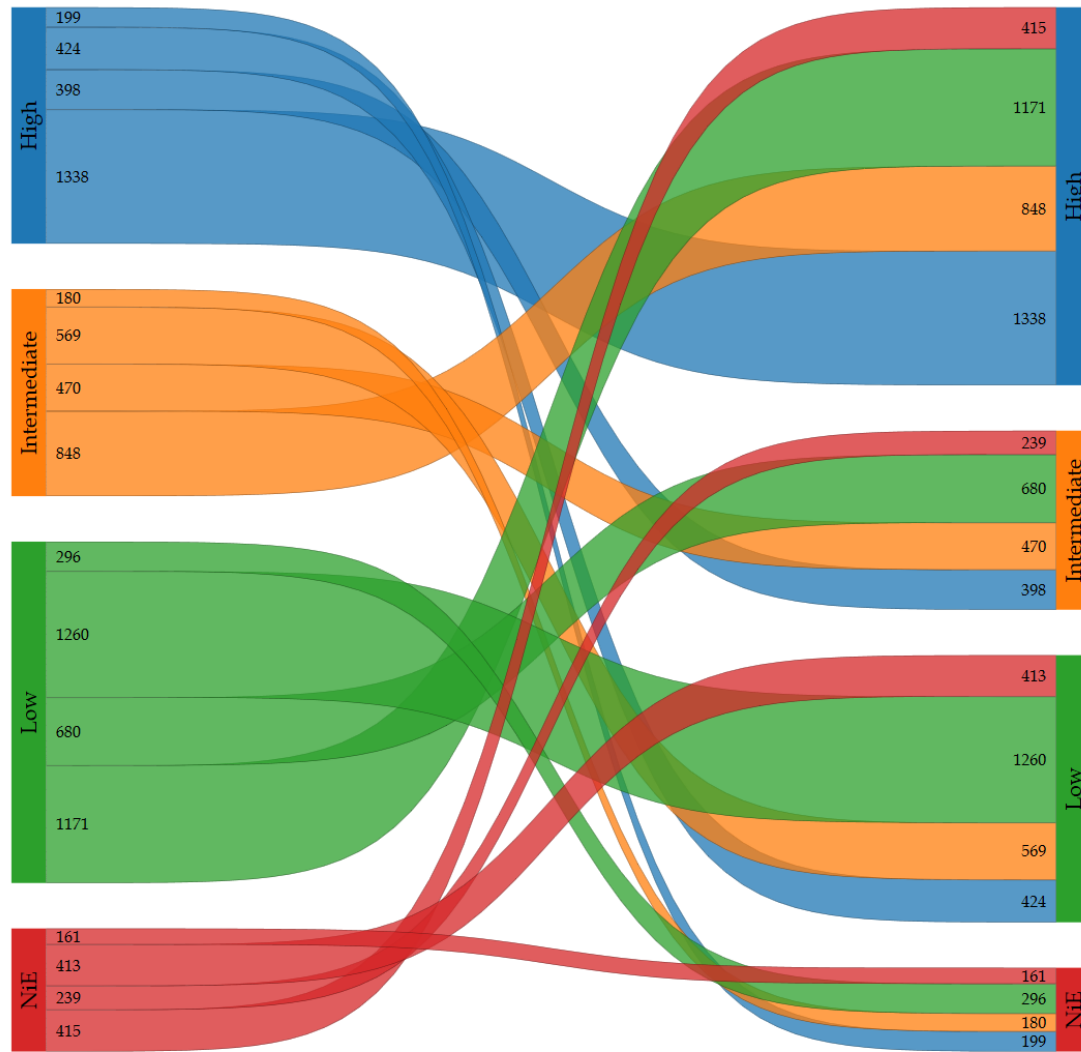
# Social Mobility

---

- Prior to modelling some introspection on the lives of sexual minorities is called for
- Social Mobility is a good starting point here
- How much do origins actually matter for current destinations

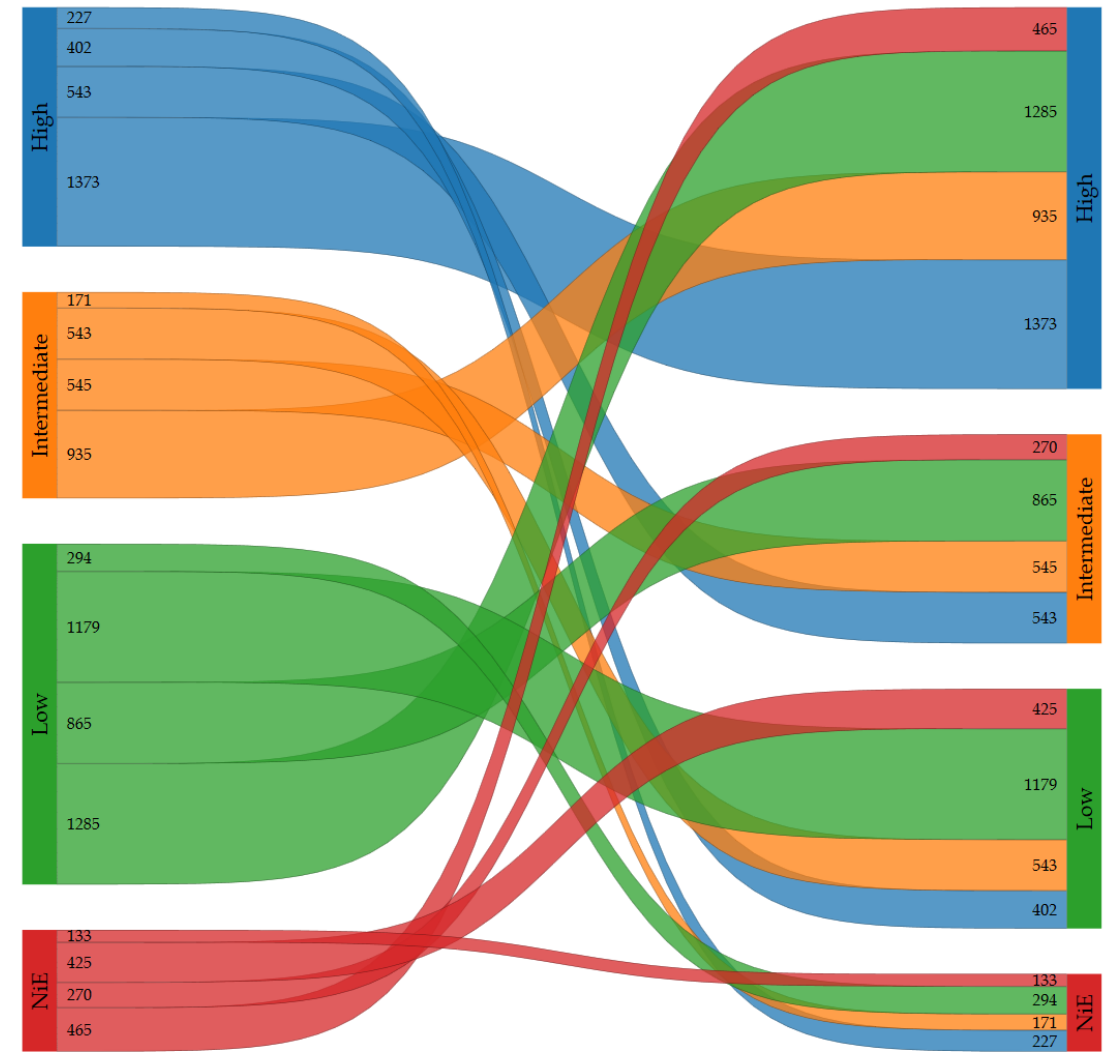
# Social Mobility: Origin to Latest Destination

## Heterosexual Men



n = 9,061

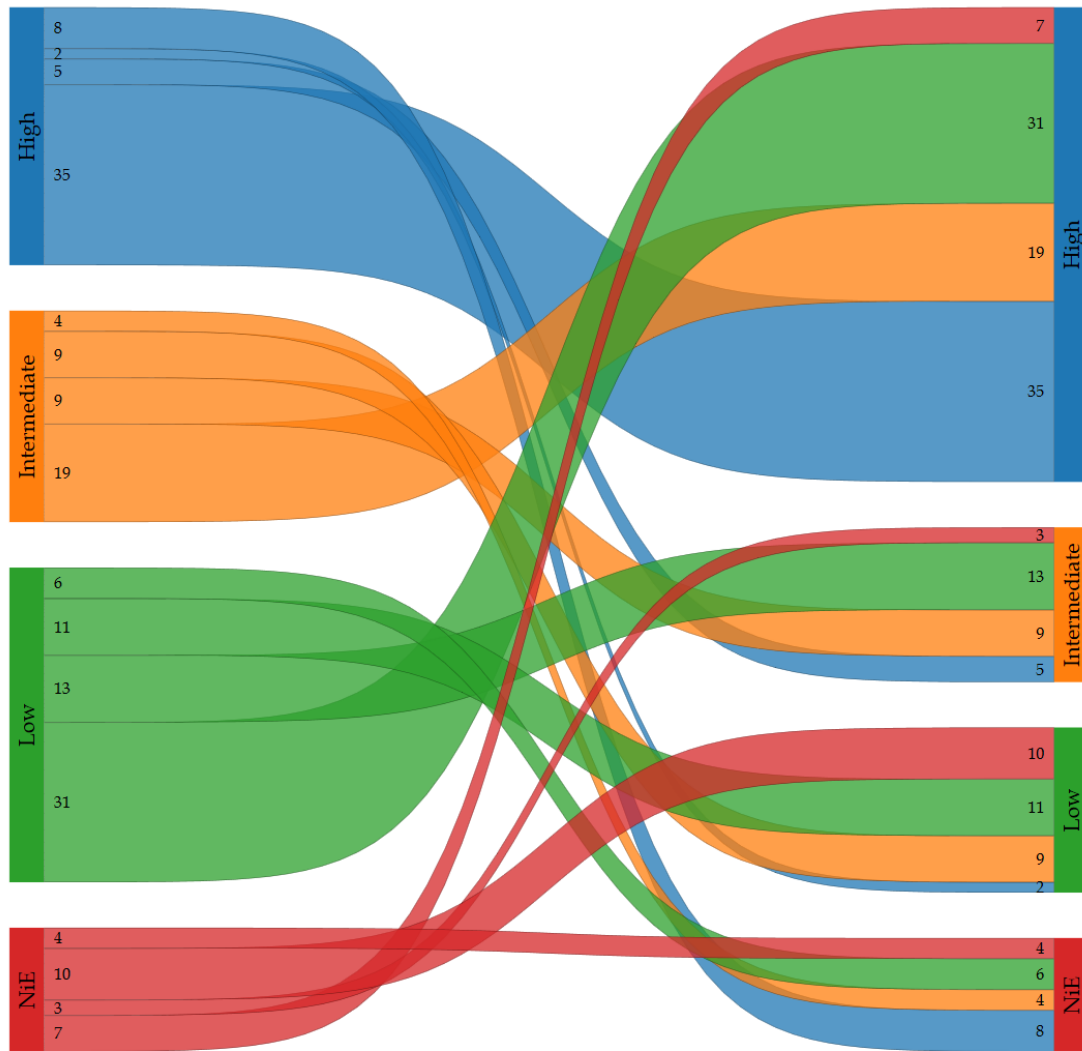
## Heterosexual Women



n = 9,655

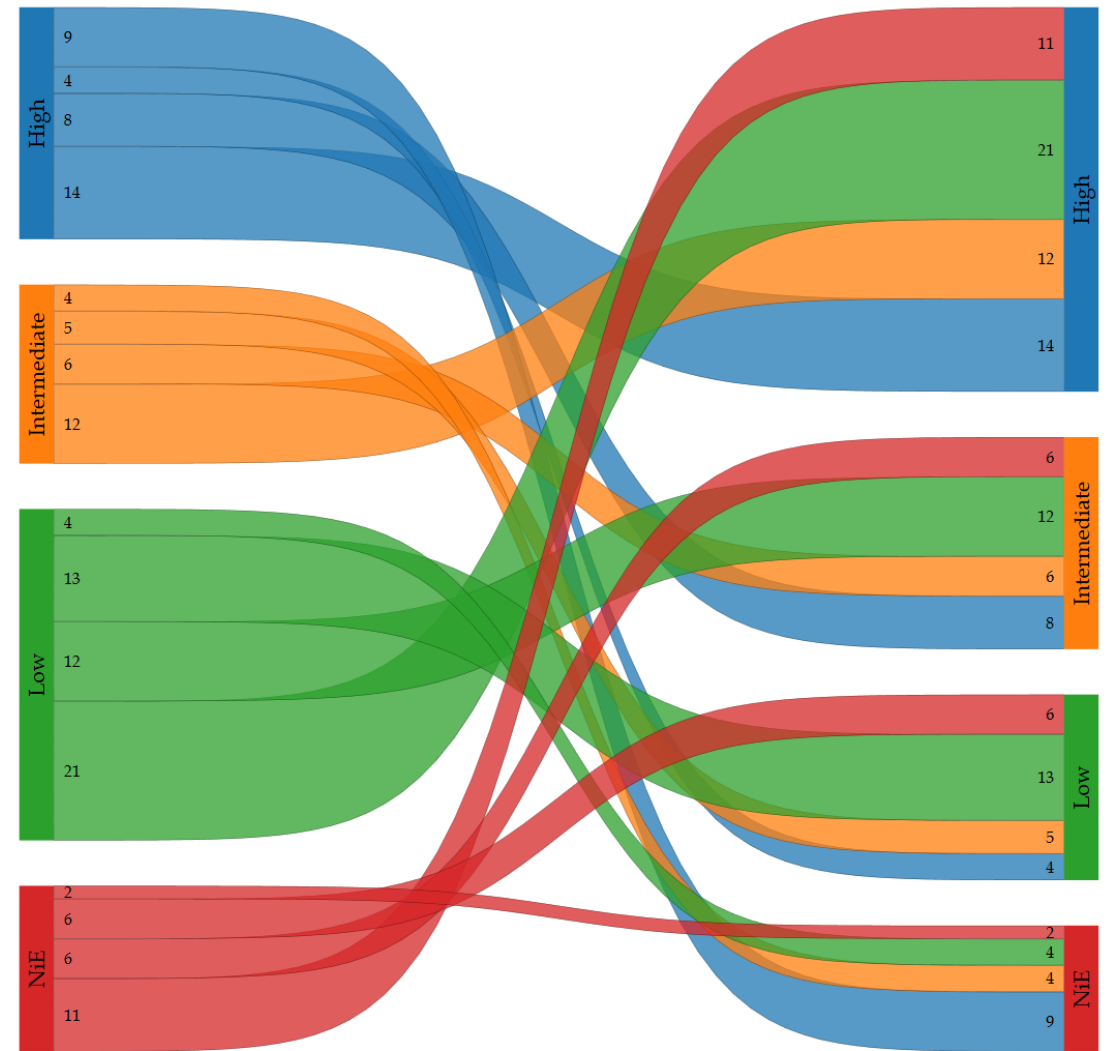
# Social Mobility: Origin to Latest Destination

## Homosexual Men



n = 176

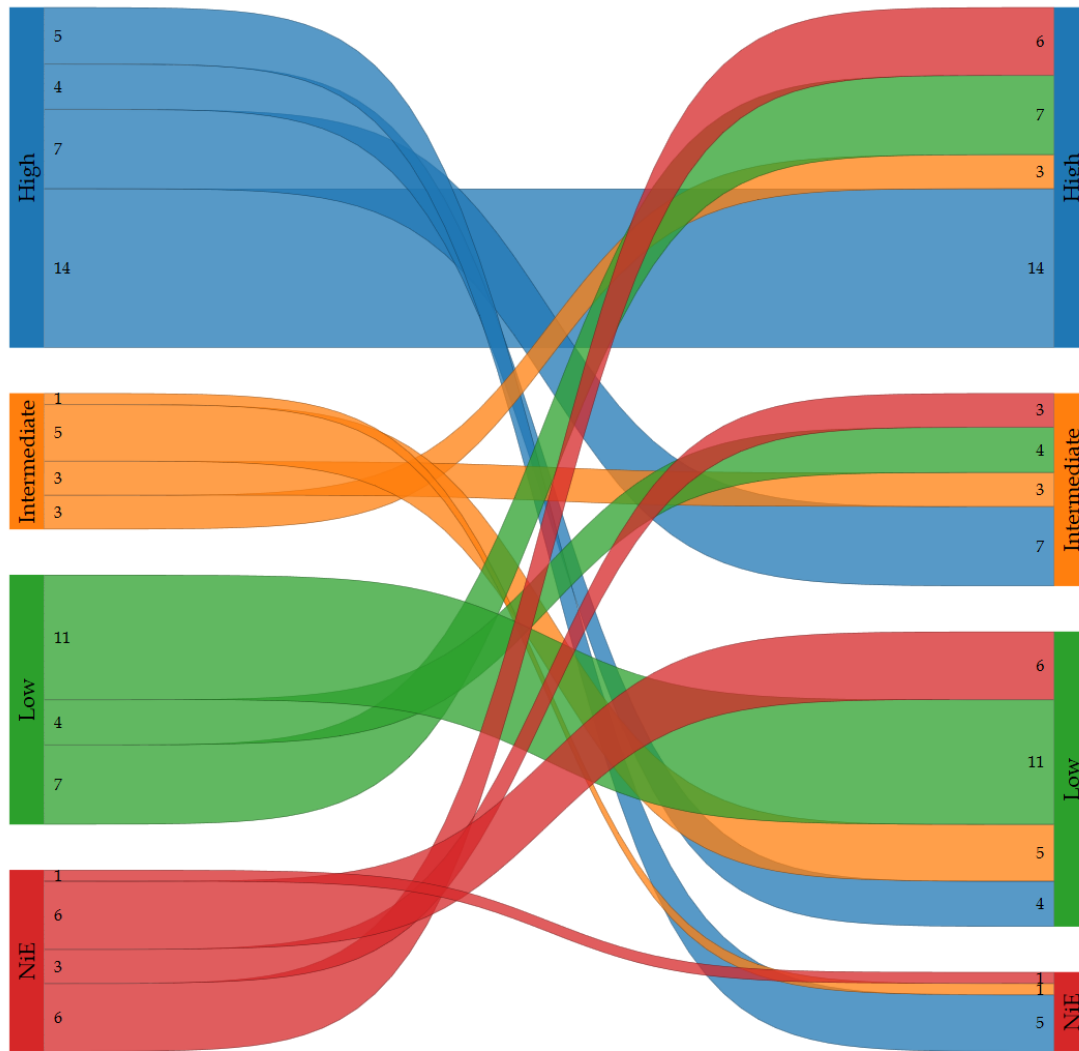
## Homosexual Women



n = 137

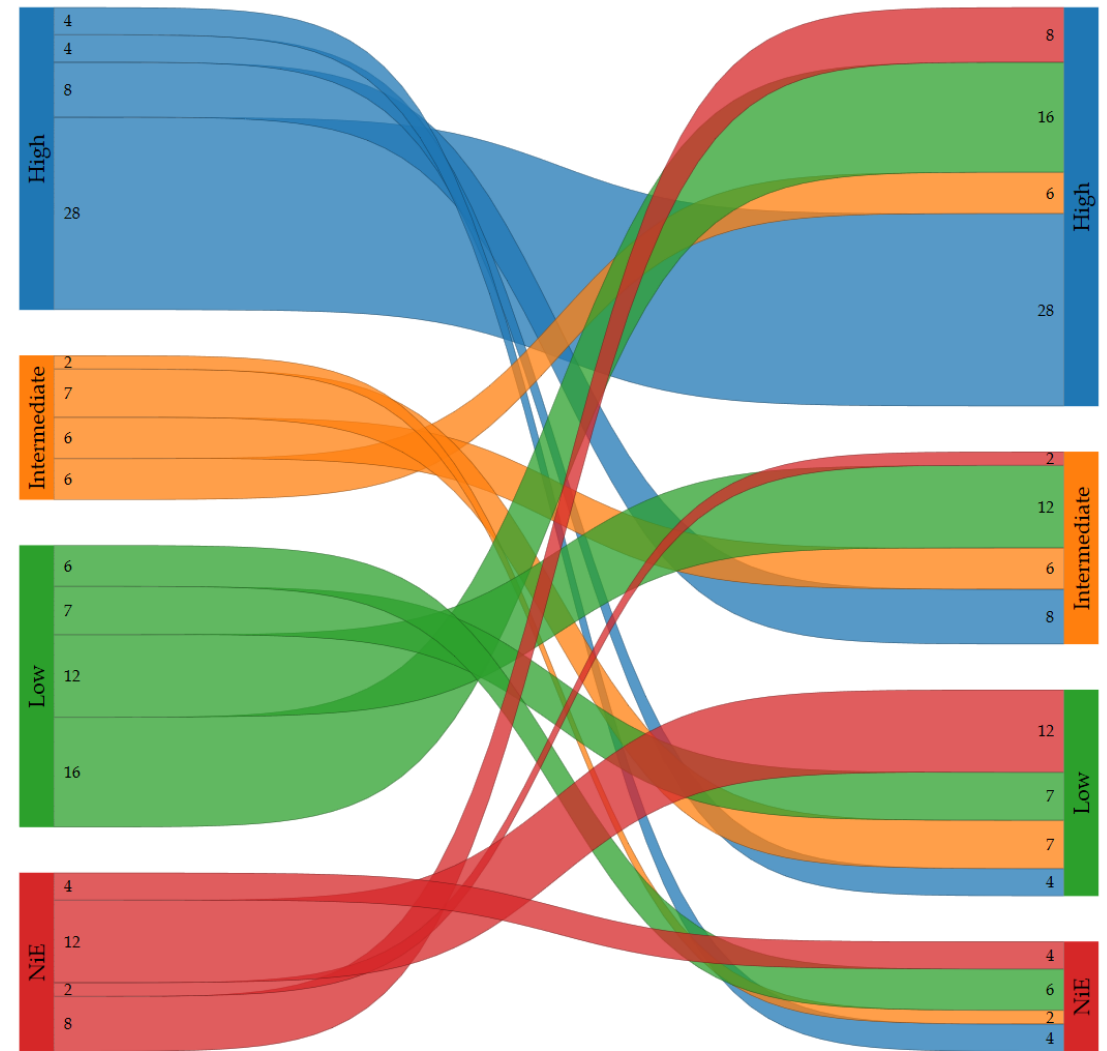
# Social Mobility: Origin to Latest Destination

## Bisexual Men



n = 80

## Bisexual Women



n = 132

- 
- Slightly larger proportions of gay male individuals 'leapfrogging' in terms of upward social mobility
  - Is this an origins effect or an 'in-spite of origins' effect?
  - Supportive households are presumably more likely to have a net positive origin effect on individuals

# Modelling the sexuality pay gap

---

- Four components to this modelling strategy
- 1: Demographics
- 2: Human Capital
- 3: Geography
- 4: Time/Growth

# Demographics

---

- Social Origins
- Sexuality
- Age
- Ethnicity
- Sex
- Housing Tenure
- Marital Status
- Parental Status
- Long Term Illness

# Human Capital

---

- Education
- Work Hours
- Current NS-SEC
- Sector
- Industry
- Size of Firm



# Geography

---

- Urban
- Region

# Time/Growth

---

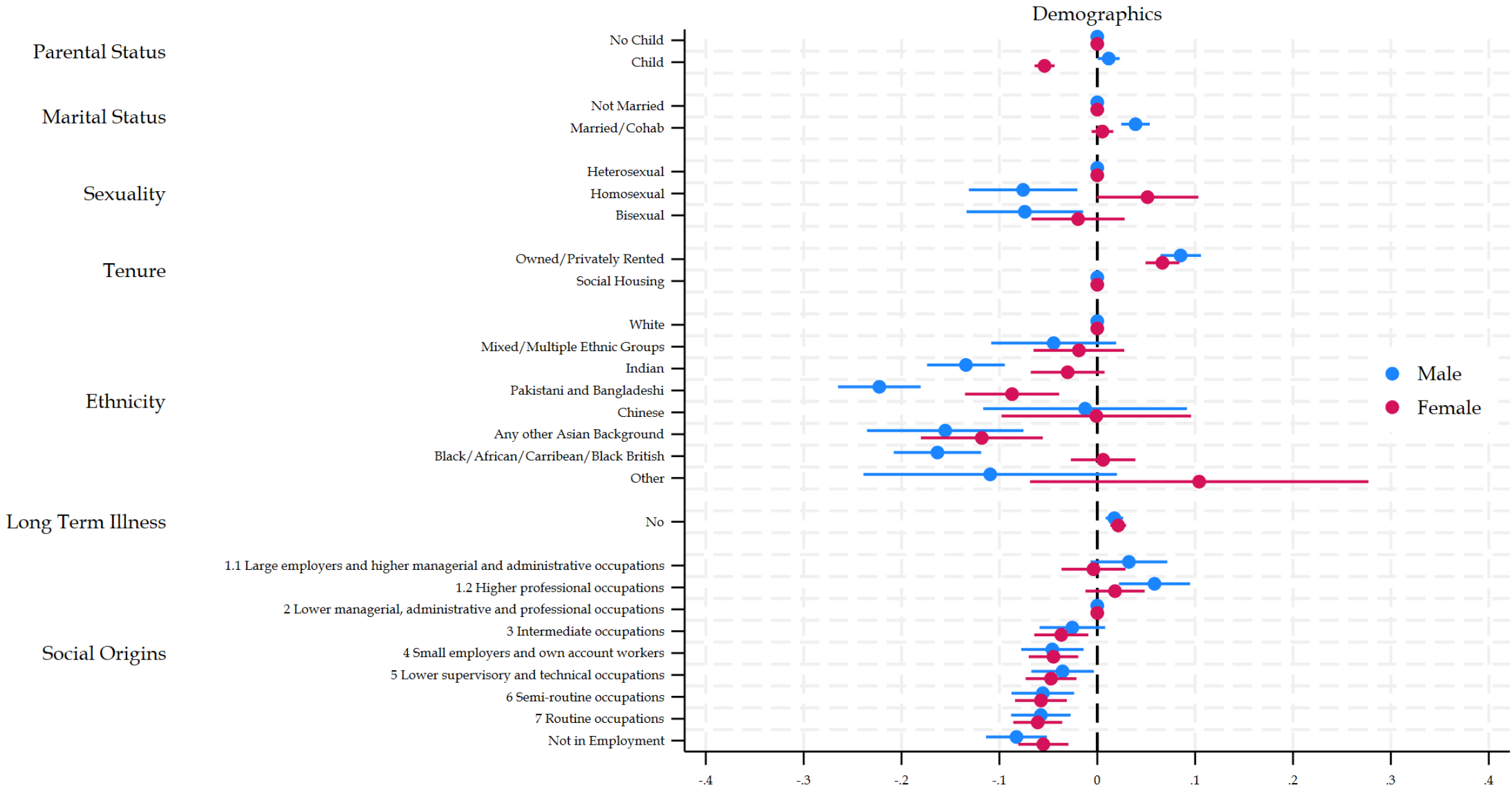
- Wave as a function of time
- Modelling through an appropriate panel set up via a unique individual pidp
- Measure change over time

# Regression Models

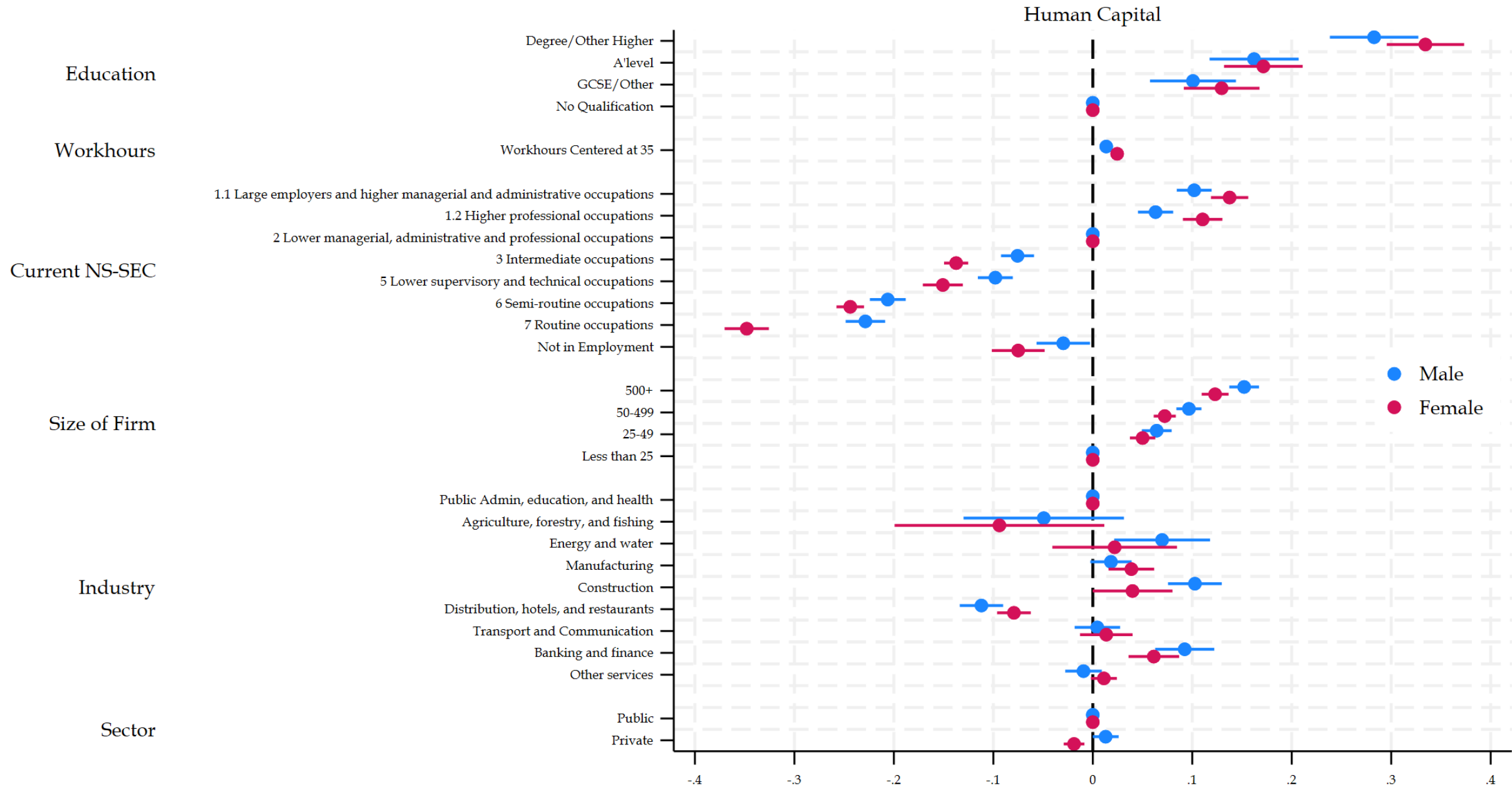
---

- Big model
- Shown in the form of coefficient plots and predictive marginal effects
- Shown in 'clusters'
- Not interpreting all effects (Shout at me if you want me to go back or look at a specific effect)

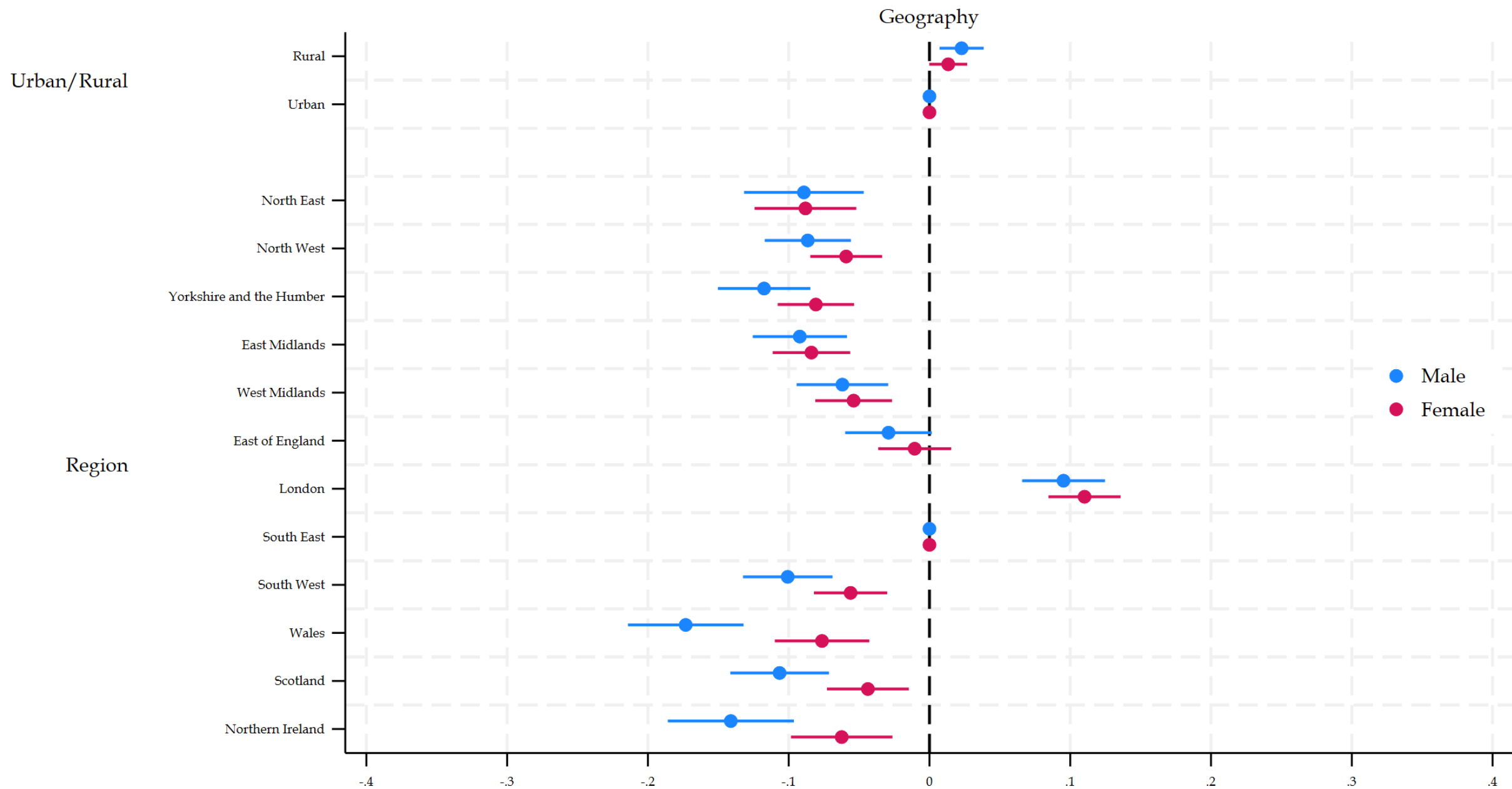
# Coefficient Plot of Sexuality Pay Gap



# Coefficient Plot of Sexuality Pay Gap



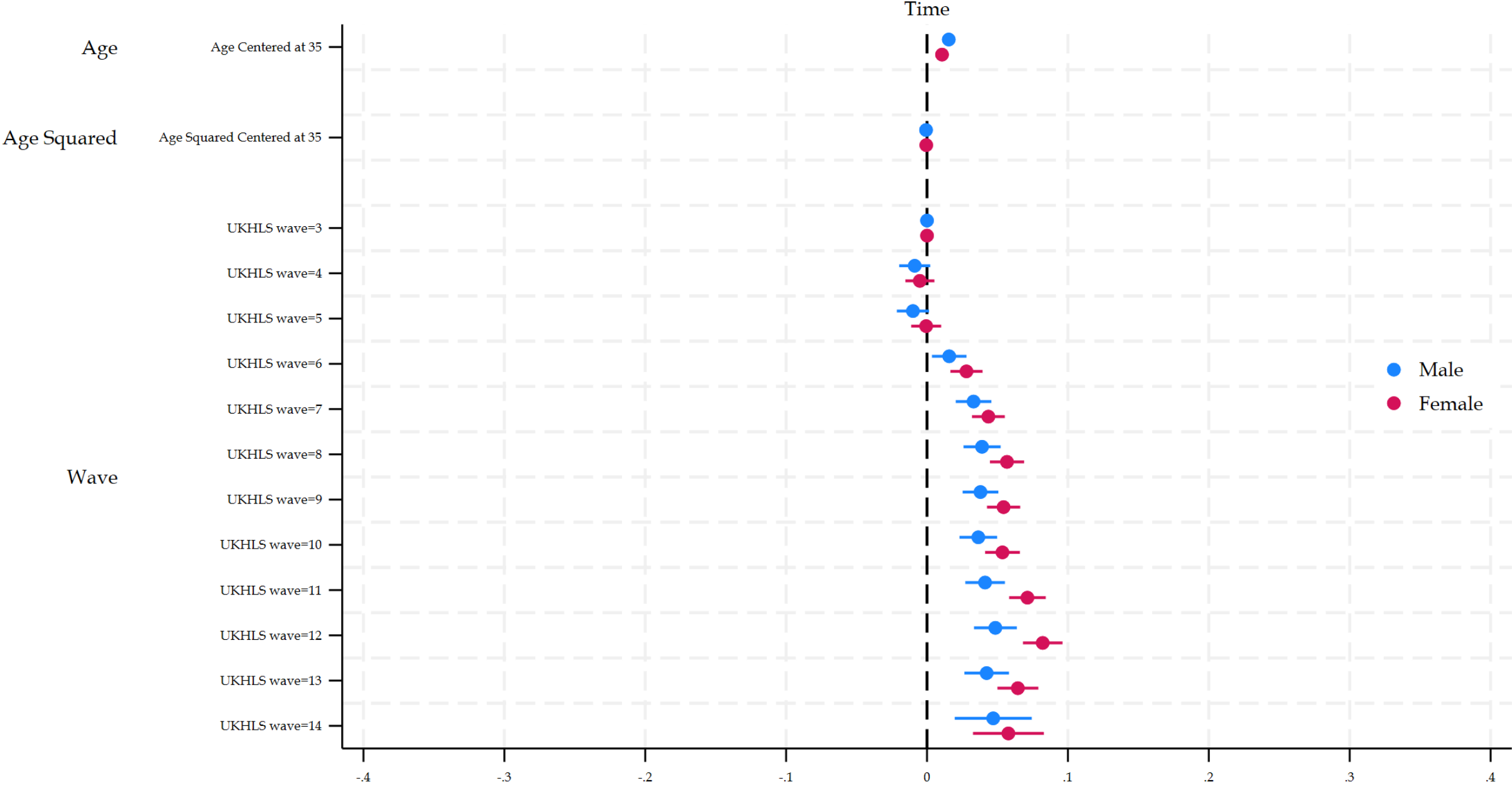
# Coefficient Plot of Sexuality Pay Gap



Data Source: UKHLS Waves 3-14

Male Model n=40,135 & Female Model n=48,917

# Coefficient Plot of Sexuality Pay Gap



# Sexual Orientation

---

- Gay men earn 8% less than straight men annually and bi men 7% less controlling for all other variables
- Gay women and bi women not statistically significant



# Sexual Orientation

---

- Remember the meta-analysis?
  - Gay men 11% penalty + Gay women 9% premium
- My model:
  - Gay men 8% penalty
  - Bi men 7% penalty
- Reduction in penalty

# Sexual Orientation

---

- Converting this into real £s would also be beneficial
- Set a new constant:
  - Straight white men with no children that are married age 35 and work 35 hours a week and own their own home with a degree in an urban region of London with NS-SEC 2 parents in an NS-SEC 2 job for a size 500+ company in the private banking industry at wave 3.
  - New constant = 10.28 OR £30,000
- Straight men = £30,000 gross OR £23,303 net (£1,942 a month)
- Gay men = £26,900 gross OR £21,501 net (£1,792 a month)
- Bi men = £27,173 gross OR £21,660 net (£1,805 a month)

# Sexual Orientation

---

- Straight men = £30,000 gross OR £23,303 net (£1,942 a month)
- Gay men = £26,900 gross OR £21,501 net (£1,792 a month)
- Bi men = £27,173 gross OR £21,660 net (£1,805 a month)
- Compared to women?
  - Women earn less than men of all orientations across the board

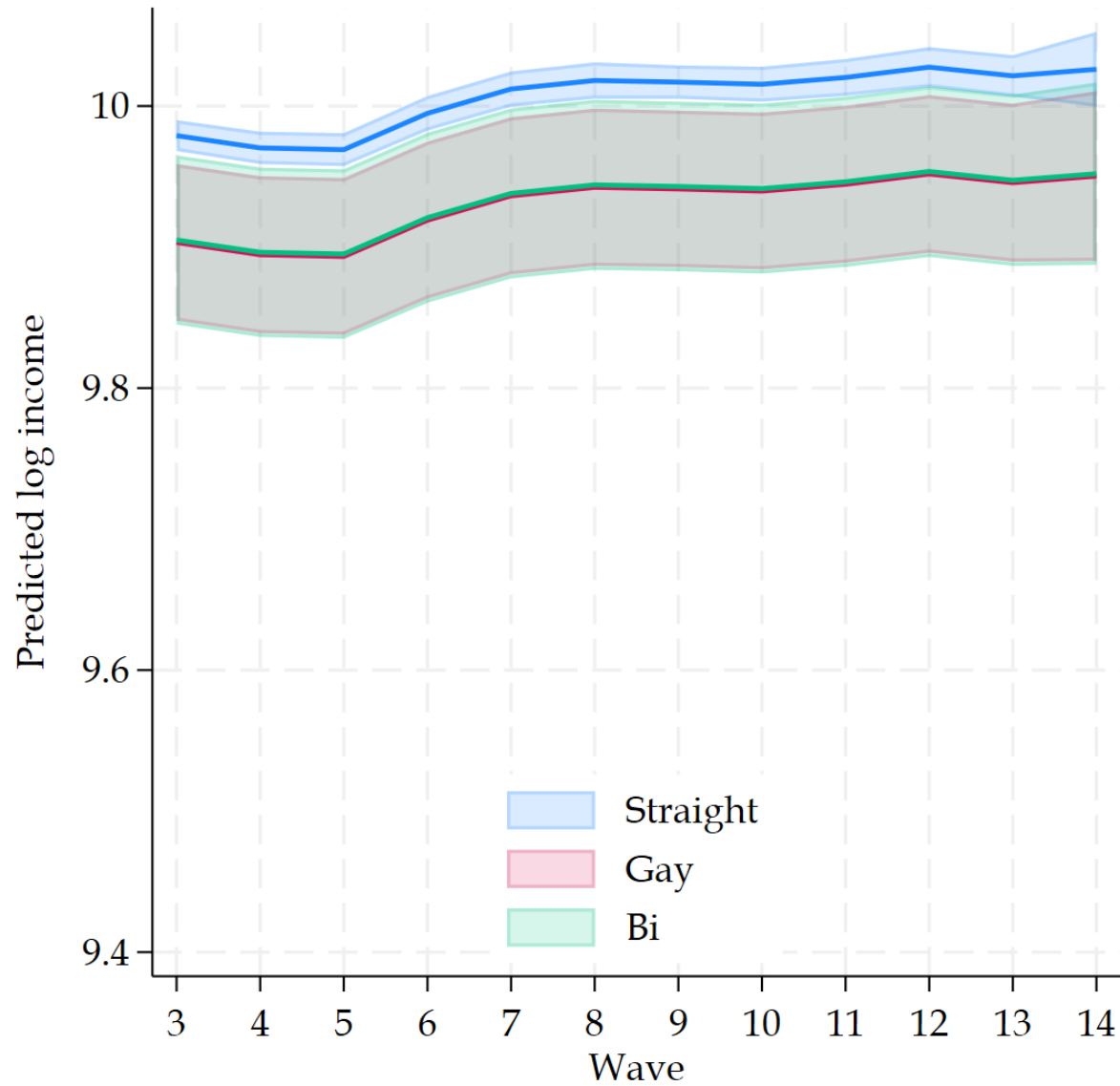
# Growth Curves

---

- Using wave as a function of time
- Possibility to look at age and age2 for curvilinear affects
- Careful of age=period-cohort effects
- Model may need tweaking
- Synthetic cohorts + age instead of waves?

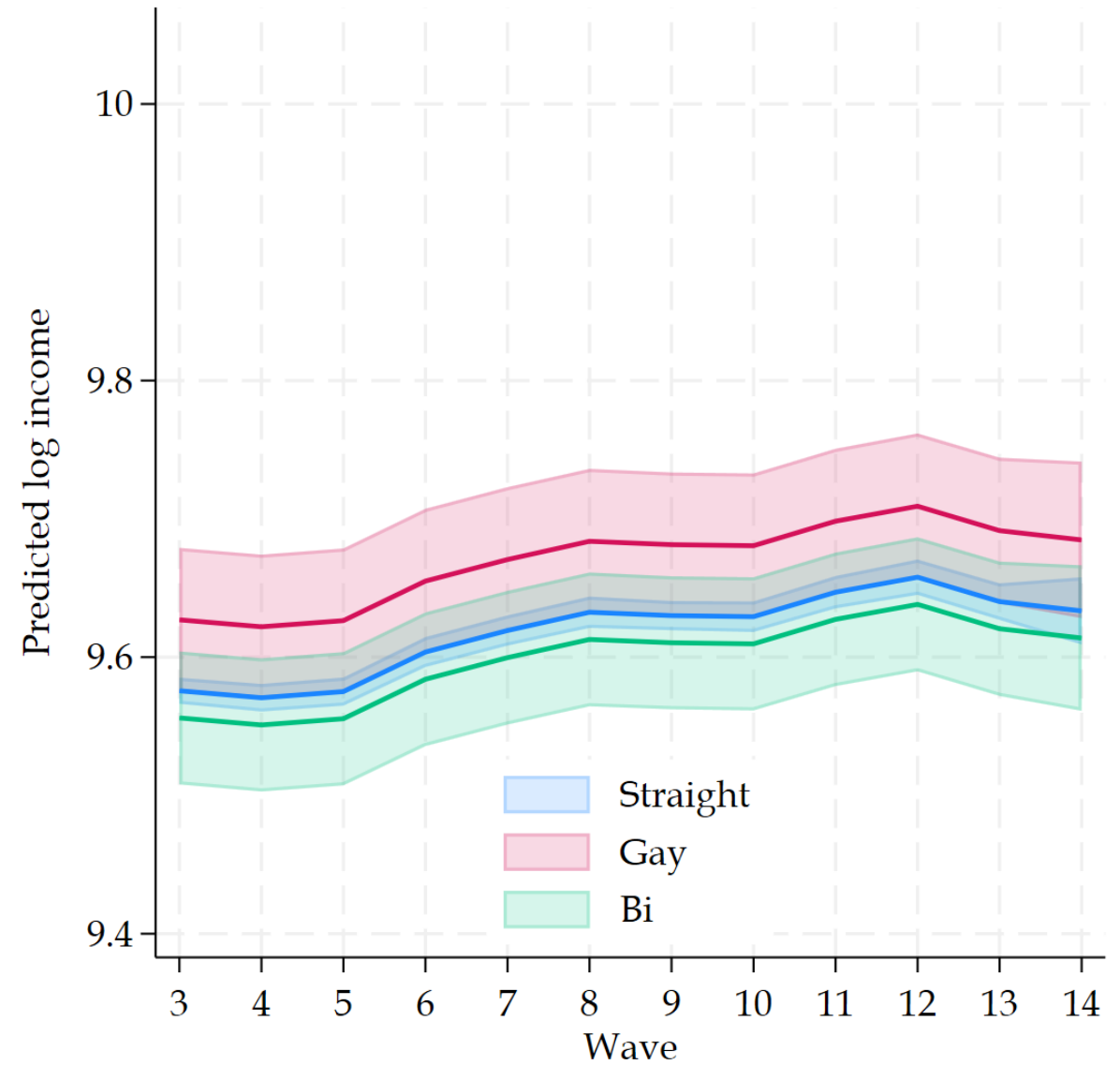
# Predicted log income growth curves by sexuality

## Men



n=40,135

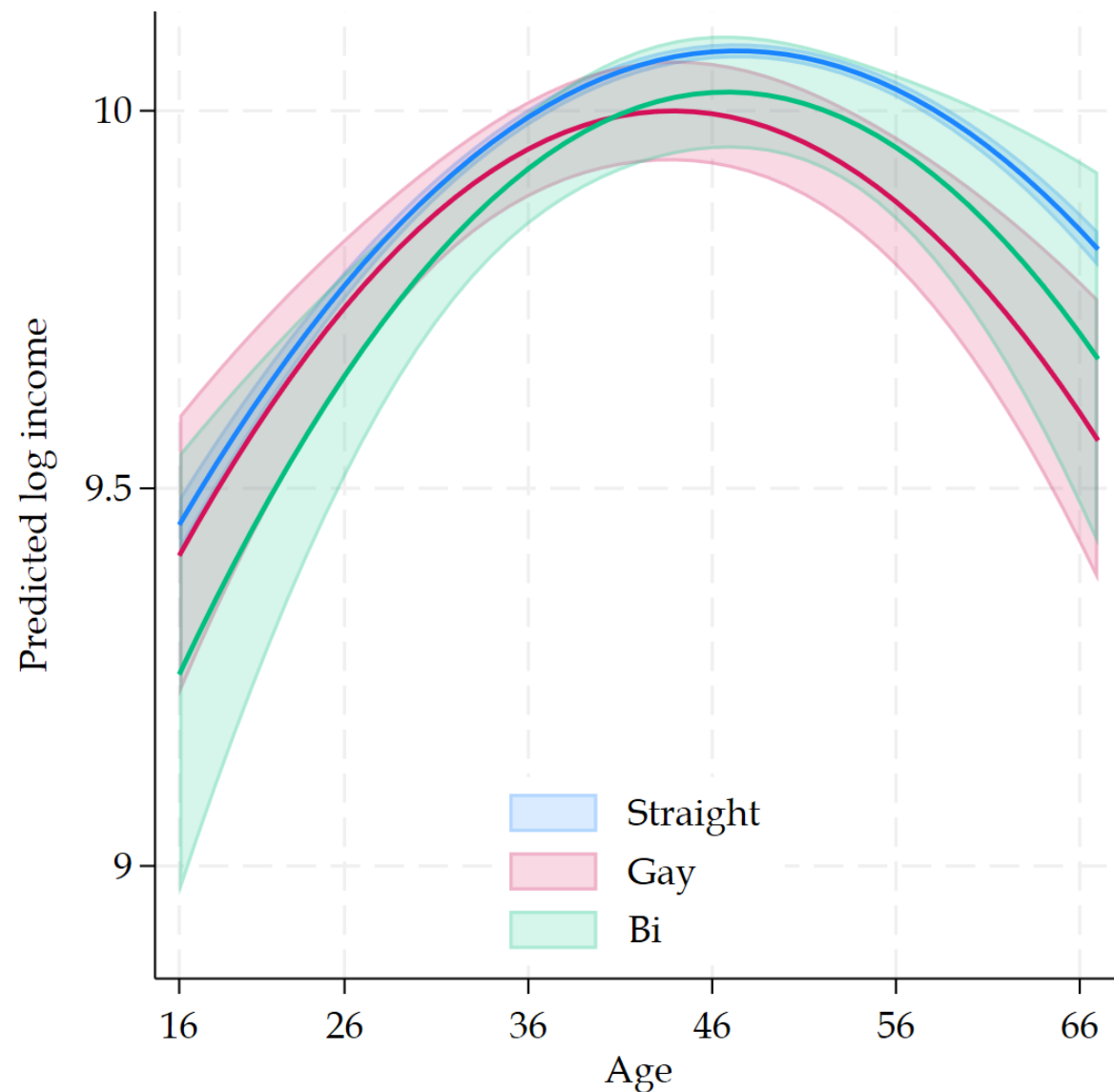
## Women



n=48,917

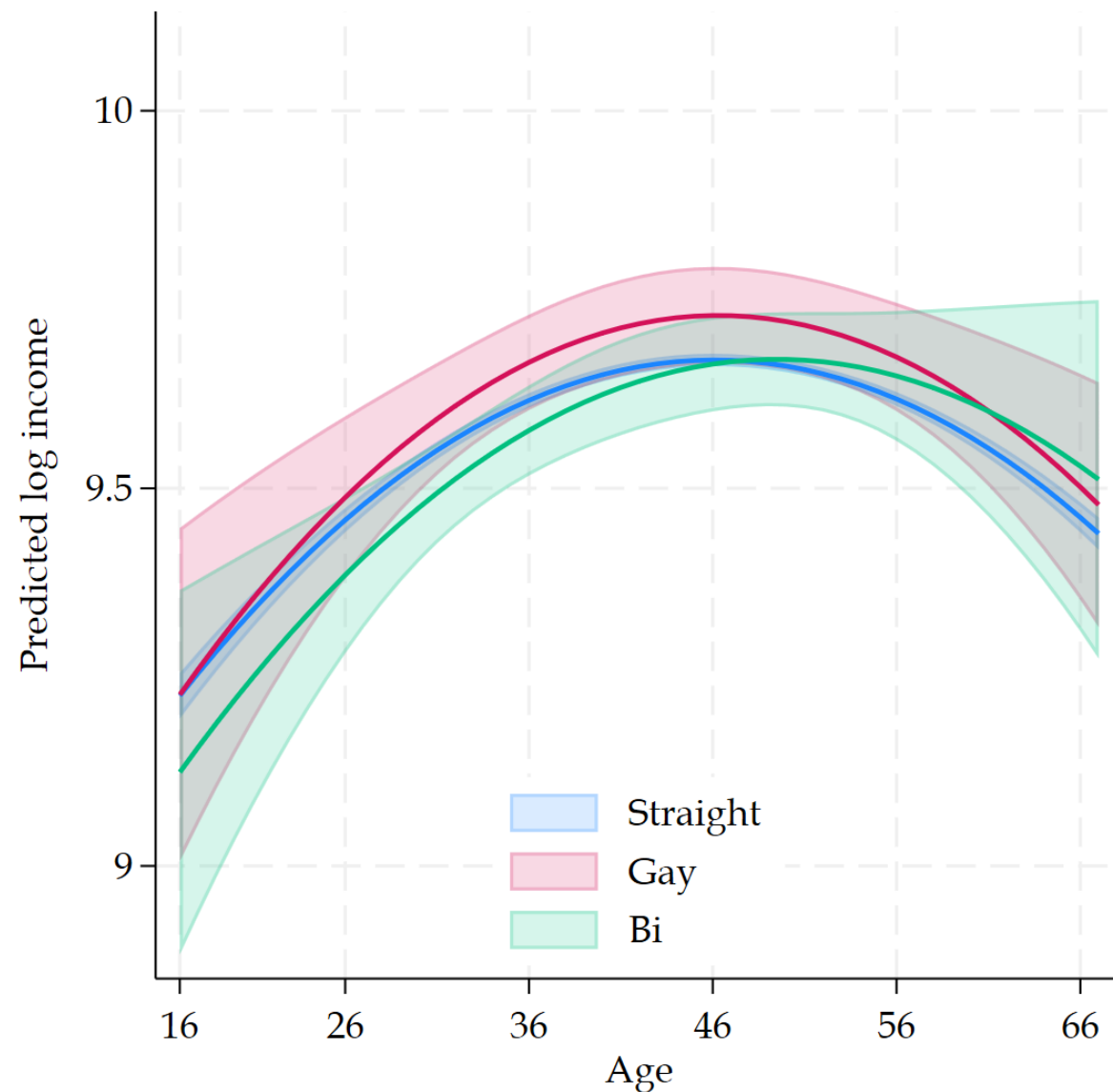
# Predicted log income growth curves by sexuality

Men



n=40,135

Women



n=48,917

# Future Plans

---

- Growth curve needs tweaking a little
  - Stay with xtreg or move over to mixed
    - Possibility to random slopes
- Complex survey design adjustments
- Handling missing data
- Experimental longitudinal decomposition techniques
  - Used to explain possible 'discrimination'

# Questions

---