

SUPPLEMENTARY MATERIAL TO:

Does intergenerational educational mobility vary by sexual identity?

A comparative analysis of five OECD countries

Contents

Online Appendix A. Overview of findings from previous research on sexual orientation and education	2
Online Appendix B. Results using only last available wave with full information for each individual.....	5
Online Appendix C. Results using logistic regression models and multiple imputation	7
Online Appendix D. Further considerations on use of sexual-identity questions	11
Online Appendix E. Results for other response categories to the sexual-identity question.....	13
Online Appendix F. Results using alternative specifications of education	16
Online Appendix G. Descriptive statistics stratified by sexual identity.....	22
Online Appendix H. Results using alternative methods.....	25
Online Appendix I. Additional models explaining educational attainment	31

Online Appendix A. Overview of findings from previous research on sexual orientation and education

Study	Country	Data	Sample	Sexual Orientation Measure	Sexual Orientation Categories	Educational Outcome	Key Findings
Mollborn and Everett (2015)	United States	Add Health (Waves I and IV)	middle and high school cohort (grades 7 to 12 in 1994-1995)	self-reported	completely heterosexual, mostly heterosexual, bisexual, mostly gay (homosexual), or completely gay (homosexual).	educational attainment (high school/GED, college attendance, college completion)	mostly gay/100% gay/mostly heterosexual men > heterosexual men and women mostly gay/100% gay and bisexual (especially) women < heterosexual/mostly heterosexual men and women bisexual men < heterosexual men and women (but conditional on college enrollment > heterosexual men and mostly heterosexual women)
Mittleman (2022)	United States	HSLs (2009-17)	high school cohort		“lesbian or gay,” “straight, that is, not lesbian or gay,” “bisexual,” “I don’t know the answer,” and “refused.”	bachelor’s degree enrollment and high-school academic achievement/behavior/attitudes	gay men > straight men and women lesbian and bisexual (especially) women < straight men and women bisexual boys ~ straights boys until high school, but less likely to enroll in college
		NHIS (2013 to 2019)	adult population (25+)	self-reported		educational attainment (high school, college, professional/graduate)	gay men > straight men and women
		NSDUH (2015-19)	adult population (25+)				lesbian women > straight women (primarily older cohorts of white lesbian women)
		NCVS (2017-19)	adult population (25+)				bisexual men, and especially women < straight men and women

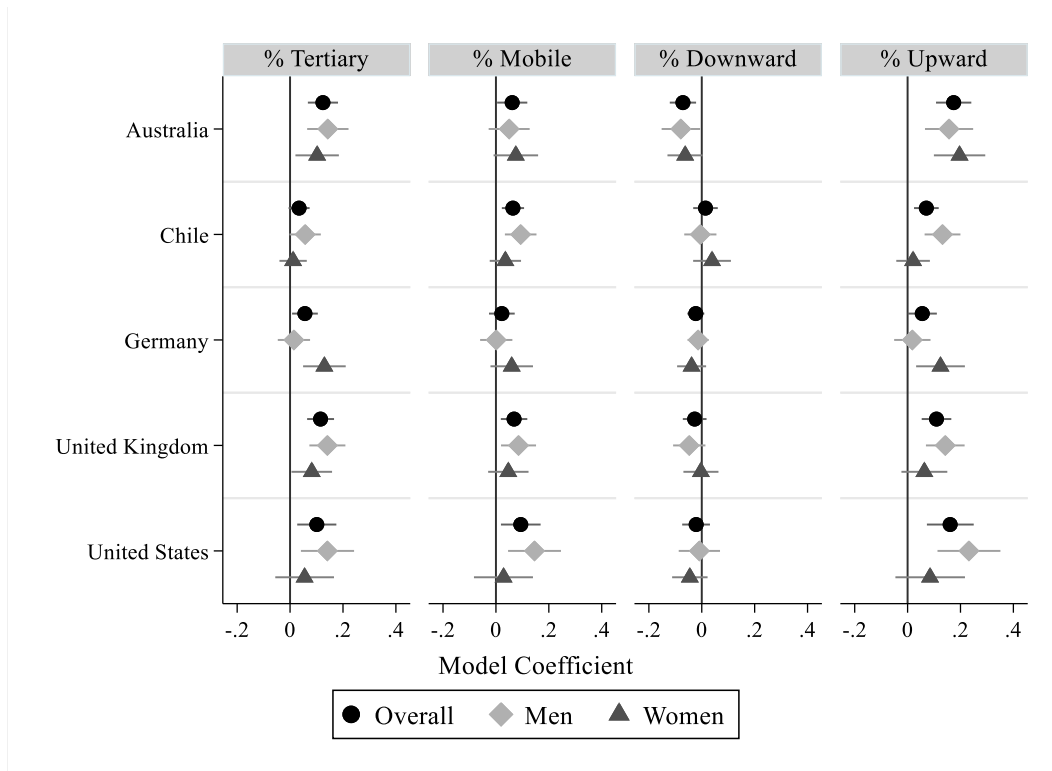
Verbakel (2013)	Netherlands	Dutch Labor Force Surveys (1994 to 2007)	adult population in coresidential partnerships (25-54)	gender composition of coresidential partnership	men and women in either same-sex or different-sex coresidential partnership	occupational status, education field, educational attainment	same-sex partnership men > different-sex partnership men (occ + educ level) same-sex partnership women > different-sex partnership women (occ + educ level) same-sex partnership men + caring field and - technical field (wrt different-sex men) same-sex partnership women + caring field and - general field (wrt different-sex partnership women)
Pascoe (2007)	United States	Ethnography	"River High" high school/suburban California	self-reported, observed behaviors	gay, lesbian, straight, gender non-conforming, masculine/feminine	high school achievement/ behavior/ attitudes	Ricky (one of the few gay boys at River High) suffers extensive abuse from both peers and educators, and drops out of high school during the ethnography; Rebeca (basketball girls) and Jessie are openly lesbian and enjoy high social status at school both from peers and educators; Genevieve and Lacy (GSA girls) present an ambiguous gender expression and sexuality - they suffer from administrative negligence and peer harassment.
Black, Makar, Sanders, and Taylor (2003)	United States	General Social Surveys (1989-96)	adult population	past sexual behavior	gay/bisexual men and women (having at least as many same-sex partners as opposite-sex partners)	focus on earnings, but years of education descriptively reported as well as father's years of education	gay/bisexual men > heterosexual men (years of education) gay/bisexual women > heterosexual women (years of education) gay/bisexual men = heterosexual men (father's years of education)

Black, Gates, Sanders, Taylor (2000)	United States	5% and 1% Public Use Microdata Samples of the 1990 census	adult population in coresidential partnerships (18+)	gender composition of coresidential partnership	men and women in either same-sex or different-sex coresidential partnership	educational attainment (some high school, high school diploma, some college, college degree, postcollege)	same-sex partnership men > different-sex partnership men (all education levels) same-sex partnership women > different-sex partnership women (all education levels)
Gates (2014)	United States	Gallup (2014)	adult population (25+)	self-reported	lesbian, gay, bisexual, transgender, non-LGB/T	college degree	LGB people ~ non-LGB
		GSS (2008-12)	adult population (25+)	self-reported	lesbian, gay, bisexual, non-LGB		LGB people > non-LGB
		NHIS (2013)	adult population (25+)	self-reported	lesbian, gay, bisexual, non-LGB		LGB people > non-LGB
		NSFG (2006-10)	adult population (25-44)	self-reported	lesbian, gay, bisexual, non-LGB		LGB people ~ non-LGB
Badgett (1995)	United States	GSS (1989-91)		past sexual behavior	(1) having had one or more same-sex sexual partners ; (2) having had more than one same-sex sexual partner; (3) having had at least as many same-sex sexual partners as opposite- sex sexual partners; and (4) having had either more than one same-sex sexual partner or at least as many same-sex sexual partners as opposite-sex sexual partners	focus on wages, but education years descriptively	LGB people ~ non-LGB

Notes: wrt = with respect to; LGB = lesbian, gay, and bisexual; educ = education; ~ = similar outcome; > = higher outcome; < = lower outcome.

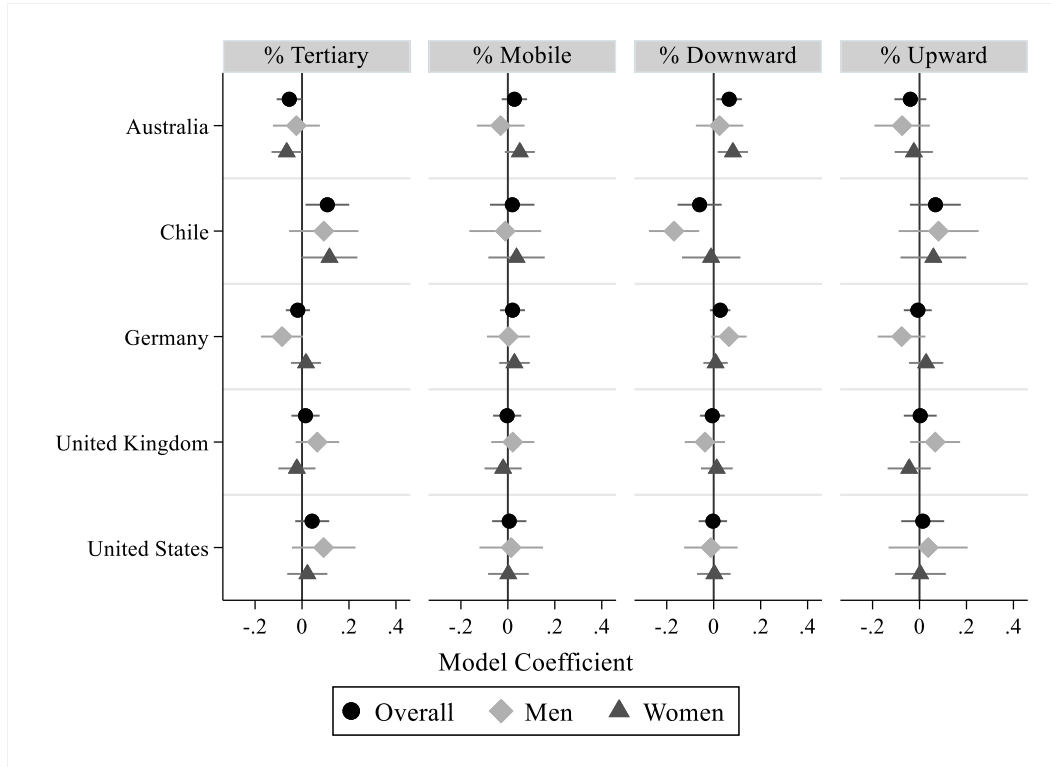
Online Appendix B. Results using only last available wave with full information for each individual (this applies to HILDA, SOEP and UK-HLS).

Figure B1. Coefficients of linear probability models based on matched samples explaining mobility among gay/lesbian individuals (Reference category: “heterosexual”)



Notes: Coefficients from linear probability models (95% confidence intervals) based on CEM. Matching variables: age, ethno-migrant or racial background, parental education, and gender (in the pooled models). Each coefficient comes from a separate model for a specific country and mobility type; for the estimates for men and women separate models by gender are run. Results using only last available wave with full information for each individual (this applies to HILDA, SOEP and UK-HLS; results for CASEN and GSS remain unchanged).

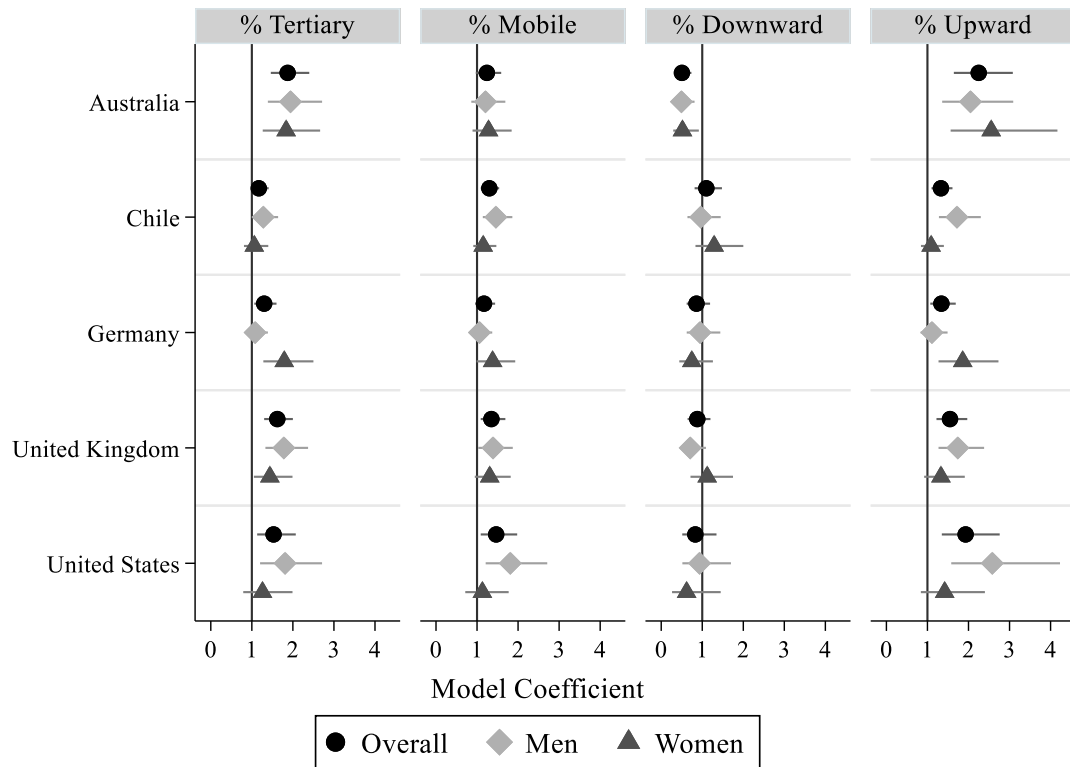
Figure B2. Coefficients of linear probability models based on matched samples explaining mobility among bisexual individuals (Reference category: “heterosexual”)



Notes: Coefficients from linear probability models (95% confidence intervals) based on CEM. Matching variables: age, ethno-migrant or racial background, parental education, and gender (in the pooled models). Each coefficient comes from a separate model for a specific country and mobility type; for the estimates for men and women separate models by gender are run. Results using only last available wave with full information for each individual (this applies to HILDA, SOEP and UK-HLS; results for CASEN and GSS remain unchanged).

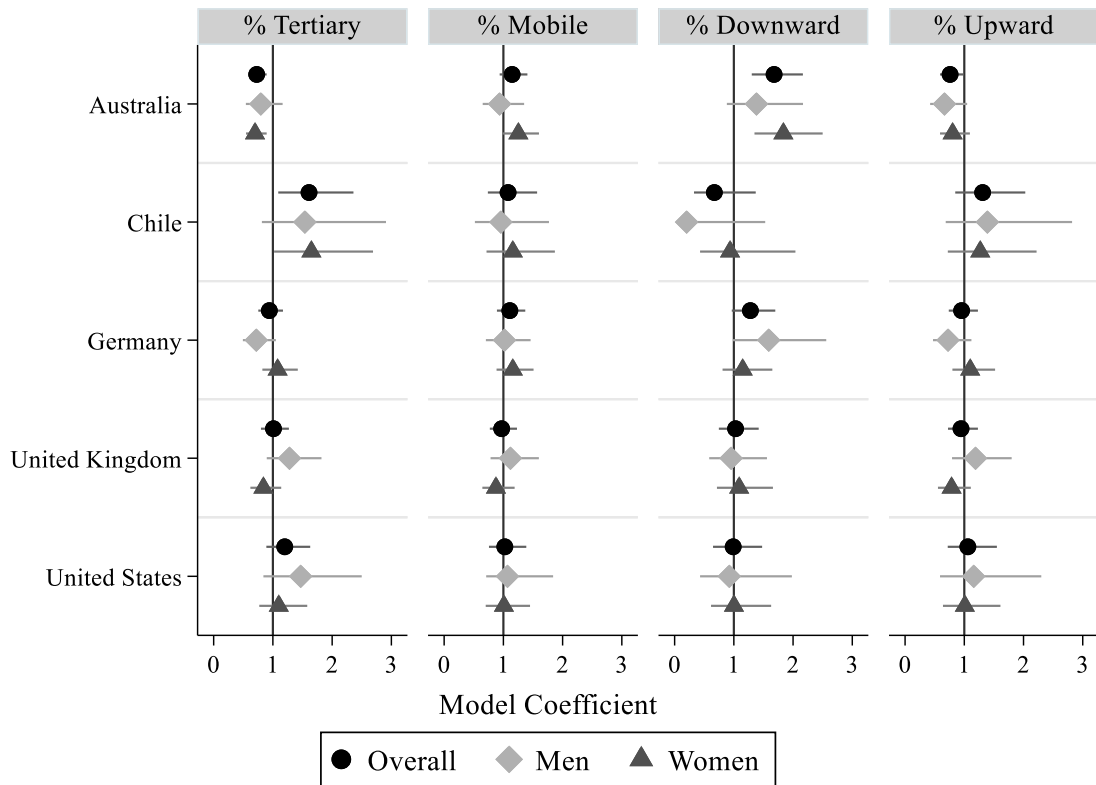
Online Appendix C. Results using logistic regression models and multiple imputation

Figure C1. Odds ratios from logistic regression models based on matched samples explaining mobility among gay/lesbian individuals (Reference category: *Heterosexual*)



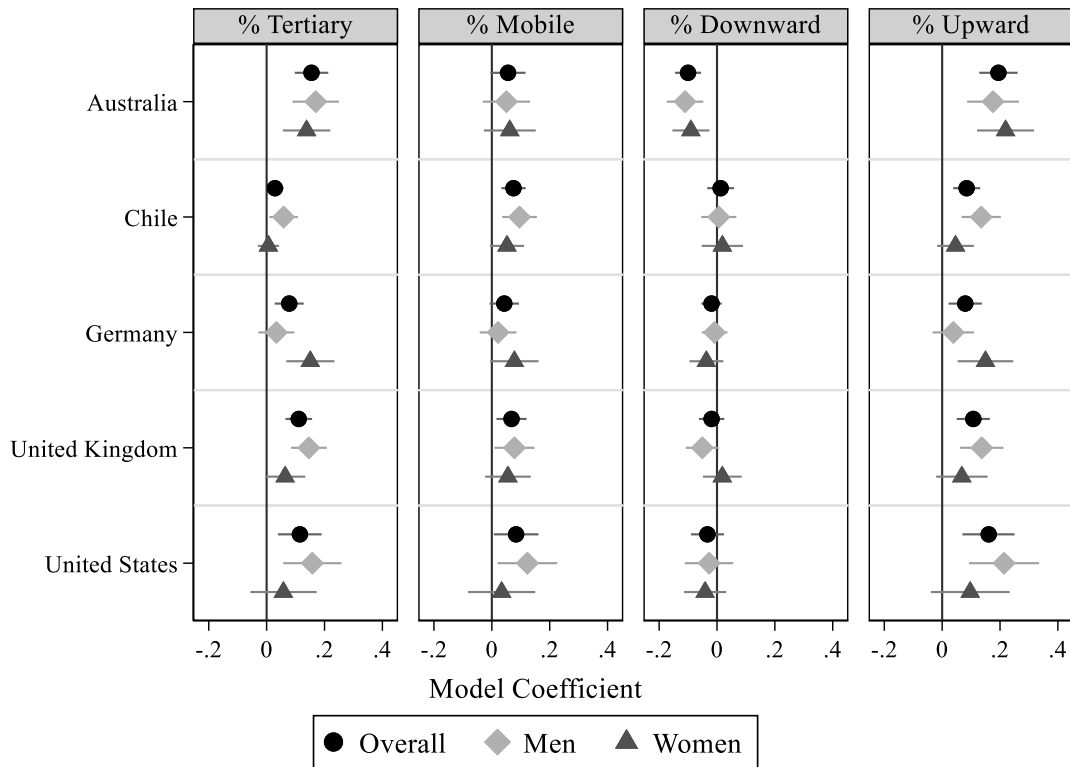
Notes: Odds ratios from logistic regression models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each odds ratio comes from a separate model for a specific country, mobility type, and gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US)

Figure C2. Odds ratios from logistic regression models based on matched samples explaining mobility among bisexual individuals (Reference category: *Heterosexual*)



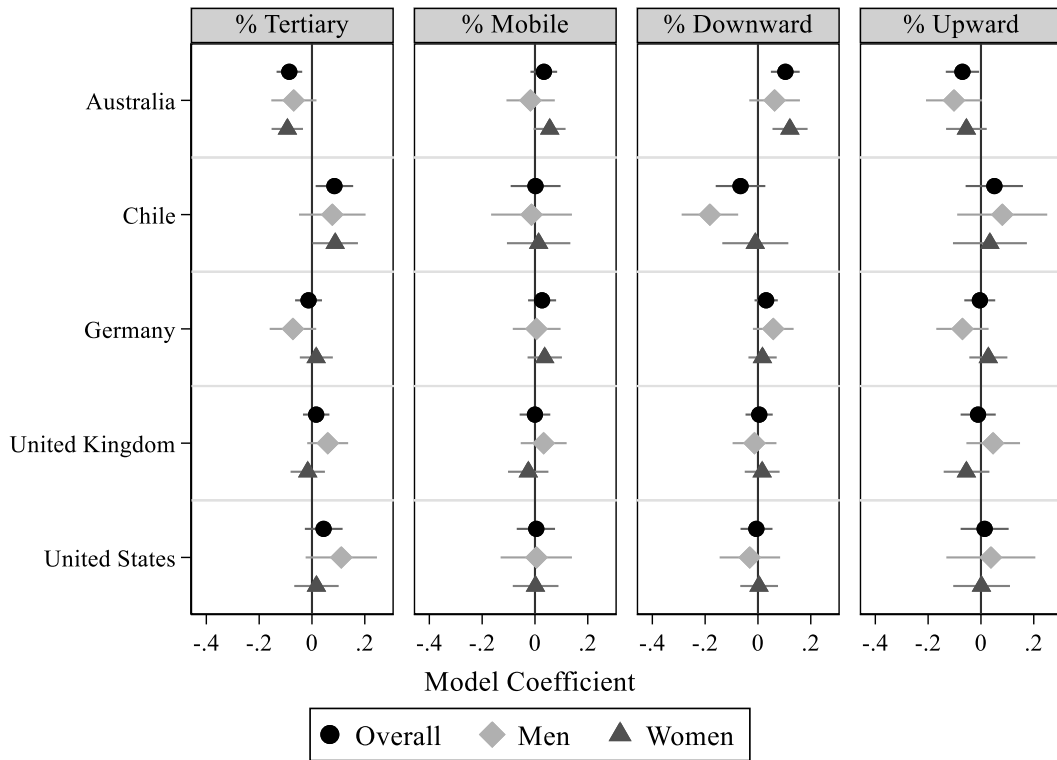
Notes: Odds ratios from logistic regression models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each odds ratio comes from a separate model for a specific country, mobility type, and gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US)

Figure C3. Coefficients of linear probability models based on matched samples explaining mobility among gay/lesbian individuals (Reference category: “heterosexual”) using multiple imputation



Notes: Coefficients from linear probability models (95% confidence intervals) based on CEM. Matching variables: age, ethno-migrant or racial background, parental education, and gender (in the pooled models). Each coefficient comes from a separate model for a specific country and mobility type; for the estimates for men and women separate models by gender are run. Results based on 20 imputations of missing information on own and parental education using chained imputation and ordered logit models and age, foreign-born status, wave, gender/sex and sexual orientation as predictors in the imputation model. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US)

Figure C4. Coefficients of linear probability models based on matched samples explaining mobility among bisexual individuals (Reference category: “heterosexual”) using multiple imputation



Notes: Coefficients from linear probability models (95% confidence intervals) based on CEM. Matching variables: age, ethno-migrant or racial background, parental education, and gender (in the pooled models). Each coefficient comes from a separate model for a specific country and mobility type; for the estimates for men and women separate models by gender are run. Results based on 20 imputations of missing information on own and parental education using chained imputation and ordered logit models and age, foreign-born status, wave, gender/sex and sexual orientation as predictors in the imputation model. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US)

Online Appendix D. Further considerations on use of sexual-identity questions

In our main analyses, we excluded individuals who responded *Other* to the sexual-identity question. This course of action aligns with that followed in previous studies (see e.g., Aksoy et al., 2018; Mittleman, 2022; Mize, 2016). In this section, we provide some more information on this group and our justification to consider it separately from both the *Heterosexual* and *Gay/Lesbian* groups. We also present selected results for this group of respondents.

Other can be an option chosen by individuals with non-heterosexual identities other than gay/lesbian or bisexual, including pansexual, queer, and asexual (Goldberg et al., 2020). However, some of these non-heterosexual individuals might also opt for one of the LGB labels when presented with the traditional format of surveys questions on sexual-identity, or refuse to respond (Ridolfo et al., 2012). Previous research has shown that the *Other* category also comprises heterosexual individuals who do not understand the question or the terms used in the response options (Elliott et al., 2019; Wooden, 2014). The *Other* group is therefore difficult to handle and interpret.

In our data, 7% of the sample answered *Other* in Germany, compared to just 0.8% in Australia and the UK. This might hint at a larger share of non-heterosexual individuals not identifying as LGB in Germany, but could also be related to the lack of a word for “straight” in the German answer option *Heterosexual* (Kühne et al., 2019). Inclusion of the word “straight” as part of that response option has been shown to reduce the share of individuals answering *Other* in English-based surveys (Ridolfo et al., 2012).

In Tables D1 to D3 we show how sexual identities changed across waves. Of individuals reporting an *Other* identity in the first wave of the panel datasets, 52% (HILDA), 75% (UK-HLS) and 80% (SOEP) identified as heterosexual in the second wave. Only a small share identified as *Other* in both waves (9% to 20) and an even smaller share identified as LGB in the second wave (1% to 4%).

Table D1. Sexual identity in Wave 16 of HILDA by sexual identity in Wave 12 (%)

	Sexual identity in Wave 16					Frequency
	Heterosexual	Gay/Lesbian	Bisexual	Other	Refused / Unsure	
Sexual identity in Wave 12						
Heterosexual	96.2	0.1	0.8	0.5	2.4	12,168
Gay/Lesbian	2.4	89.3	3.0	1.2	4.2	177
Bisexual	36.5	5.4	45.5	3.6	9.0	176
Other	51.8	2.3	1.2	20.0	24.7	92
Refused/Unsure	43.8	2.6	5.0	4.2	44.4	448

Notes: *Refused/Unsure* includes Refused to answer/not stated, don't know/unsure and prefer not to say. Data from 2012 and 2016.

Table D2. Sexual identity in Wave 35 of SOEP based on identity in Wave 32 (%)

	Sexual identity in Wave 35					Frequency
	Heterosexual	Gay/Lesbian	Bisexual	Other	Refused / Unsure	
Sexual identity in Wave 32						
Heterosexual	96.0	0.2	0.6	N/A	3.3	14,748
Gay/Lesbian	7.7	82.1	6.0	N/A	4.3	117
Bisexual	34.4	3.2	53.8	N/A	8.6	93
Other	80.1	0.3	0.7	N/A	19.0	1,108
Refused/Unsure	69.0	0.7	1.6	N/A	28.8	858

Notes: Refused/Unsure includes refused to answer/not stated, don't know/unsure and prefer not to say. Data from 2016 and 2019.

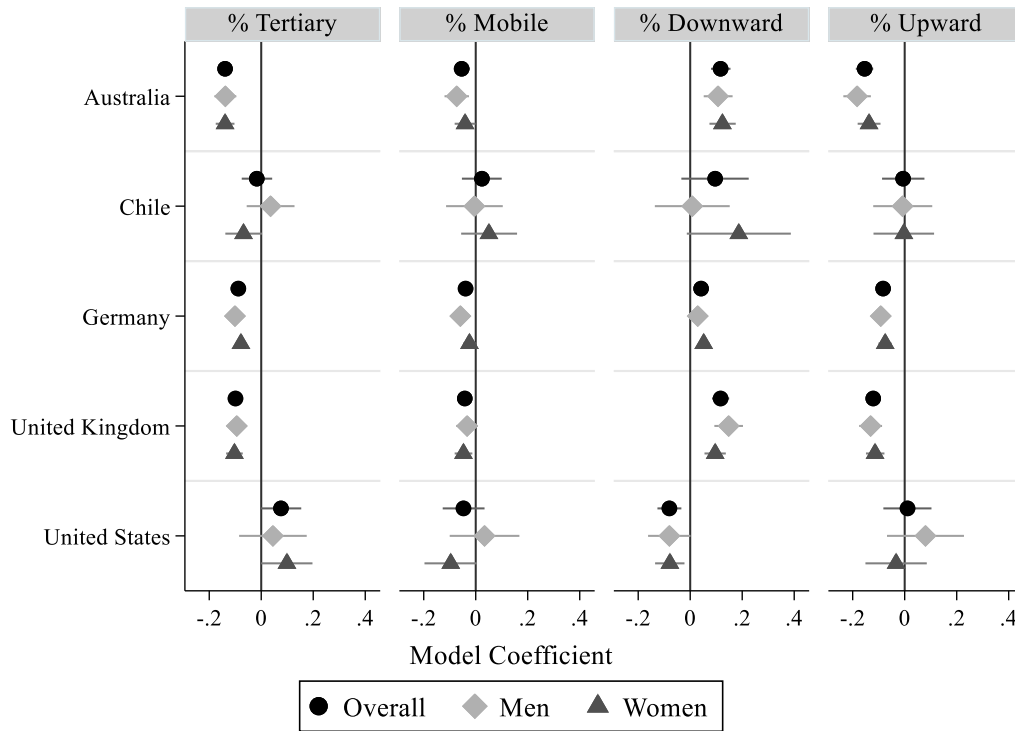
Table D3. Sexual identity in Wave 9 of UK-HLS by sexual identity in Wave 3 (%)

	Sexual identity in Wave 9					Frequency
	Heterosexual	Gay/Lesbian	Bisexual	Other	Refused / Unsure	
Sexual identity in Wave 3						
Heterosexual	97.0	0.1	0.6	0.4	1.9	21,448
Gay/Lesbian	5.4	87.1	2.2	1.8	3.6	279
Bisexual	40.6	6.1	45.3	3.8	4.3	212
Other	75.1	2.1	1.6	8.8	12.4	193
Refused/Unsure	65.6	2.8	3.4	2.0	26.2	642

Notes: Refused/Unsure includes refused to answer/not stated, don't know and prefer not to say. Data from 2011/2012 and 2017/2018.

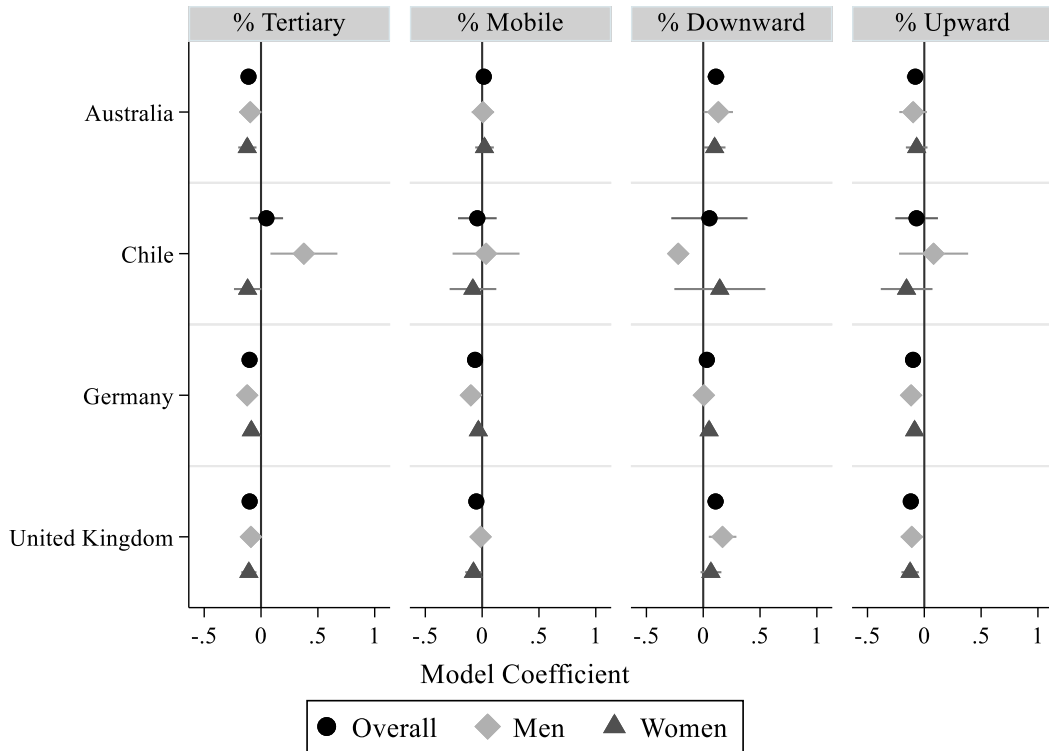
Online Appendix E. Results for other response categories to the sexual-identity question

Figure E1. Coefficients of linear probability models based on matched samples explaining mobility among individuals who answered *Other*, *Don't Know*, *Prefer Not to Say* or *Refused to answer* to the sexual-identity question (Reference category: *Heterosexual*)



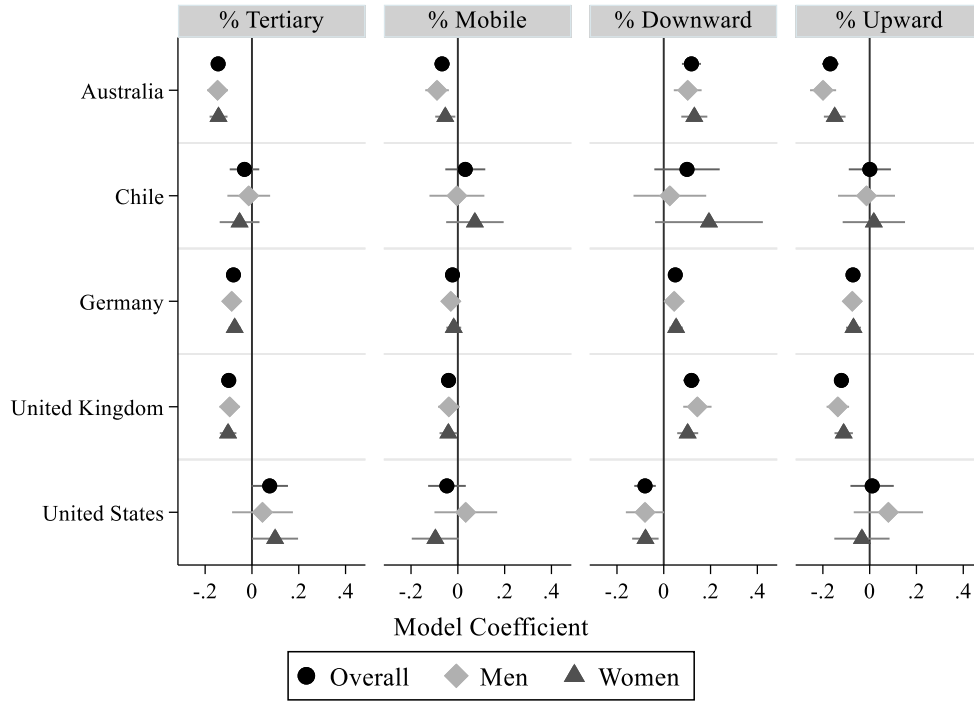
Notes: Coefficients from linear probability models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each coefficient comes from a separate model for a specific country, mobility type, and gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US).

Figure E2. Coefficients of linear probability models based on matched samples explaining mobility among individuals who answered *Other* to the sexual-identity question (Reference category: *Heterosexual*)



Notes: Coefficients from linear probability models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each coefficient comes from a separate model for a specific country, mobility type, and gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US).

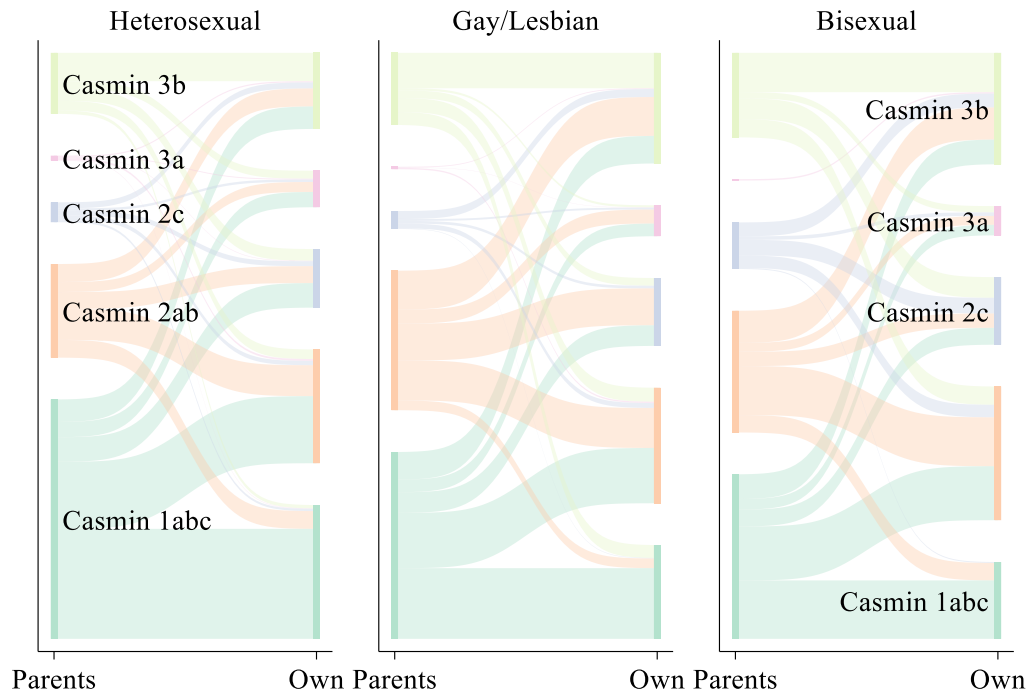
Figure E3. Coefficients of linear probability models based on matched samples explaining mobility among individuals who answered *Don't Know*, *Prefer Not to Say* or *Refused to answer* to the sexual-identity question (Reference category: *Heterosexual*)



Notes: Coefficients from linear probability models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each coefficient comes from a separate model for a specific country, mobility type, and gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US).

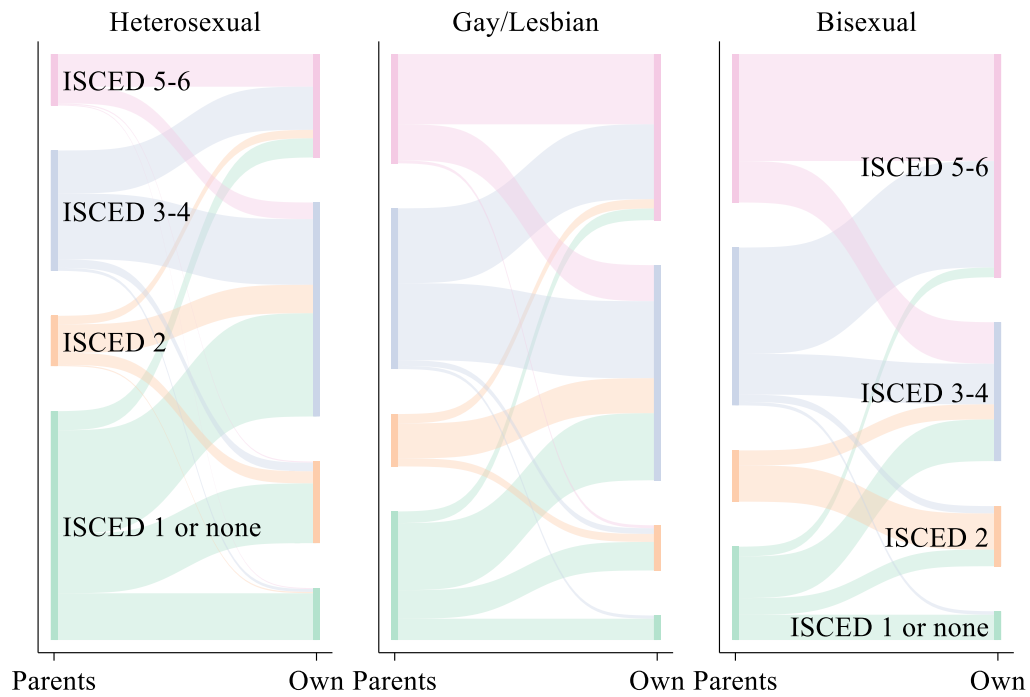
Online Appendix F. Results using alternative specifications of education

Figure F1. Share of respondents attaining *Hauptschule* (Casmin 1abc), *Realschule* (Casmin 2ab), *Abitur* (Casmin 2c), lower tertiary education (Casmin 3a) and higher tertiary education (Casmin 3b) by parental education (Germany)



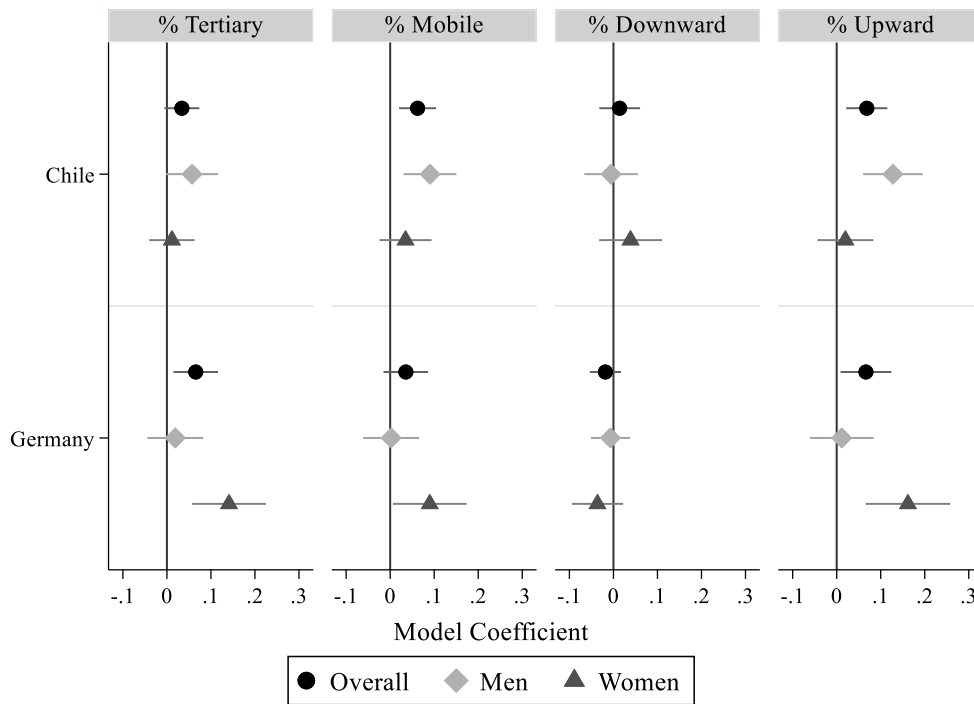
Notes: Flowcharts based on weighted descriptive statistics (Table 1). Data: SOEP 2016 & 2019.

Figure F2. Share of respondents attaining primary (ISCED 1), lower secondary (ISCED 2), upper secondary (ISCED 3-4) and tertiary (ISCED 5-6) education by parental education (Chile)



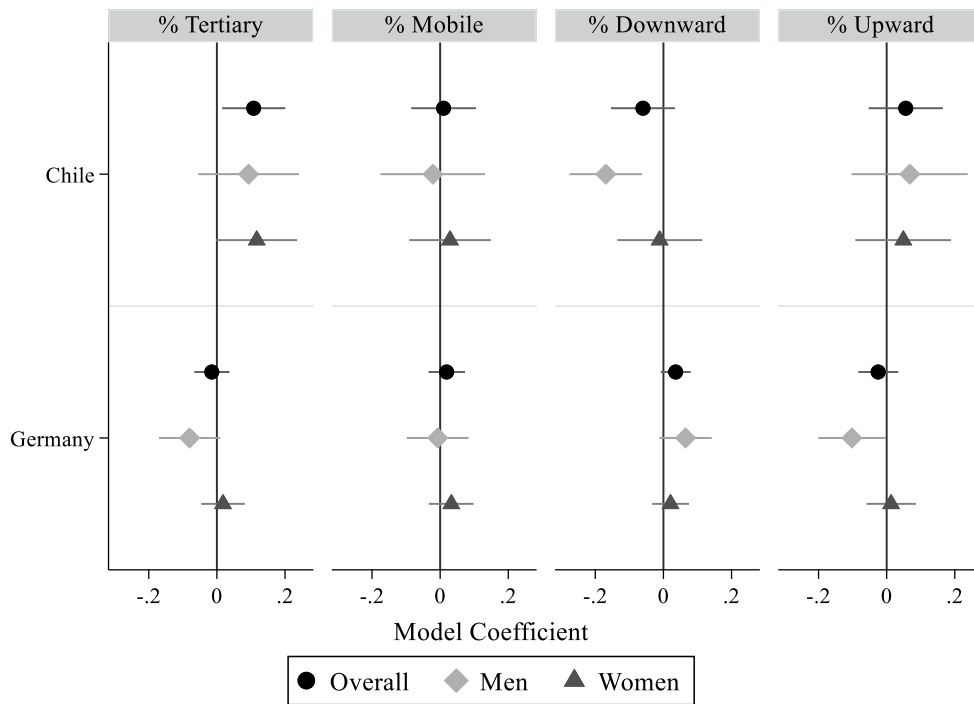
Notes: Flowcharts based on weighted descriptive statistics (Table 1). Data: CASEN 2017.

Figure F3. Attainment and mobility differences between gay/lesbian and heterosexual individuals (matched samples) using detailed educational categories (CASMIN for Germany, primary, lower secondary, upper secondary, tertiary for Chile)



Notes: Coefficients from linear probability models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each coefficient comes from a separate model for a specific country, mobility type, and gender. Data from 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany).

Figure F4. Attainment and mobility differences between bisexual and heterosexual individuals (matched samples) using detailed educational categories (CASMIN for Germany, primary, lower secondary, upper secondary, tertiary for Chile)



Notes: Coefficients from linear probability models (95% confidence intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant or racial background, parental education, and gender. Each coefficient comes from a separate model for a specific country, mobility type, and gender. Data from 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany).

Tables F1a-e. Share of respondents attaining low, middle, and high education according to highest level of parental education and parents' combined level of education.

<i>Australia</i>	Own education (ISCED; %)								
	Bisexual			Gay/Lesbian			Heterosexual		
	5-6	3-4	1-2	5-6	3-4	1-2	5-6	3-4	1-2
<i>Parental Education (Highest)</i>									
ISCED 5-6 (High)	61	31	8	75	22	3	68	26	6
ISCED 3-4 (Mid)	33	47	20	59	34	7	40	41	18
ISCED 1-2 (Low)	45	33	21	35	53	12	30	36	34
<i>Parental Education (Combination of both parents)</i>									
Both High	86	12	2	85	15	0	82	16	2
High/Mid	47	43	9	78	19	2	66	29	6
High/Low	39	42	19	64	31	5	56	31	12
Both Mid	24	59	17	64	30	7	46	43	12
Mid/Low	34	43	23	54	38	8	38	39	22
Both Low	49	32	20	30	57	13	32	35	33

<i>Chile</i>	Own education (ISCED; %)								
	Bisexual			Gay/Lesbian			Heterosexual		
	5-6	3-4	1-2	5-6	3-4	1-2	5-6	3-4	1-2
<i>Parental Education (Highest)</i>									
ISCED 5-6 (High)	72	28	0	64	33	3	64	33	4
ISCED 3-4 (Mid)	67	25	7	46	48	5	35	55	10
ISCED 1-2 (Low)	7	39	54	12	56	32	10	47	43
<i>Parental Education (Combination of both parents)</i>									
Both High	76	24	0	69	31	0	71	27	3
High/Mid	65	35	0	71	29	0	69	30	1
High/Low	100	0	0	42	44	14	42	47	11
Both Mid	63	33	4	55	43	2	44	50	6
Mid/Low	72	18	10	34	56	10	28	59	13
Both Low	7	38	56	9	59	32	10	47	43

<i>Germany</i>	Own education (ISCED; %)								
	Bisexual			Gay/Lesbian			Heterosexual		
	5-6	3-4	1-2	5-6	3-4	1-2	5-6	3-4	1-2
<i>Parental Education (Highest)</i>									
ISCED 5-6 (High)	56	39	4	52	29	18	63	34	3
ISCED 3-4 (Mid)	33	58	9	39	57	5	30	61	9
ISCED 1-2 (Low)	5	65	30	16	60	24	17	58	25
<i>Parental Education (Combination of both parents)</i>									
Both High	83	11	6	74	22	4	74	25	1
High/Mid	46	49	4	52	21	27	63	35	2
High/Low	48	52	0	52	48	0	46	44	9
Both Mid	33	58	9	38	56	5	35	59	6
Mid/Low	34	56	9	39	55	5	23	63	13
Both Low	5	65	30	16	60	24	17	58	25

UK	Own education (ISCED; %)								
	Bisexual			Gay/Lesbian			Heterosexual		
	5-6	3-4	1-2	5-6	3-4	1-2	5-6	3-4	1-2
<i>Parental Education (Highest)</i>									
ISCED 5-6 (High)	79	17	4	76	25	0	73	21	6
ISCED 3-4 (Mid)	38	46	15	55	32	12	43	40	17
ISCED 1-2 (Low)	37	36	27	45	27	28	22	32	47
<i>Parental Education (Combination of both parents)</i>									
Both High	77	19	4	79	21	0	82	12	6
High/Mid	85	10	5	72	28	0	72	23	5
High/Low	60	40	0	85	15	0	56	32	12
Both Mid	46	39	16	57	32	11	48	40	12
Mid/Low	35	52	14	56	29	15	37	40	23
Both Low	40	33	26	46	25	29	22	31	47

US	Own education (ISCED; %)								
	Bisexual			Gay/Lesbian			Heterosexual		
	5-6	3-4	1-2	5-6	3-4	1-2	5-6	3-4	1-2
<i>Parental Education (Highest)</i>									
ISCED 5-6 (High)	67	32	2	66	34	0	63	36	1
ISCED 3-4 (Mid)	29	69	3	48	51	1	29	68	3
ISCED 1-2 (Low)	27	54	19	13	77	10	13	65	22
<i>Parental Education (Combination of both parents)</i>									
Both High	73	27	0	84	16	0	70	29	1
High/Mid	55	41	4	53	47	0	58	42	0
High/Low	87	13	0	100	0	0	41	48	11
Both Mid	34	62	3	48	51	1	31	67	2
Mid/Low	0	100	0	51	49	0	24	72	4
Both Low	27	54	19	13	77	10	13	65	22

Notes: Figures based on weighted descriptive statistics (Table 1).

Online Appendix G. Descriptive statistics stratified by sexual identity

Table G1. Descriptive statistics, Australia

	Heterosexual %	Gay/Lesbian %	Bisexual %	Other %	Refused %
<i>Gender</i>					
Woman	51.5	34.6	62.6	56.5	57.9
<i>Age Group</i>					
25-34 years	21.0	34.2	45.2	31.1	26.5
35-44 years	21.4	23.4	21.9	12.6	14.7
45-54 years	20.4	22.6	11.6	15.5	17.7
55-64 years	17.8	13.4	13.0	21.8	15.8
65-74 years	13.0	5.4	6.4	10.5	15.0
75-84 years	6.4	1.0	1.9	8.5	10.2
<i>Ethno-migrant background</i>					
Native born, Indigenous	1.6	1.7	4.5	3.1	3.5
Native born, Non-Indigenous	67.7	68.0	67.4	54.0	39.8
Foreign born, major English-speaking country	12.3	14.6	8.9	2.5	4.9
Foreign born, other country	18.3	15.7	19.3	40.4	51.8
<i>n (observations)</i>	32,969	574	563	292	1,442

Notes: Weighted descriptive statistics. 2012, 2016, 2020 data from HILDA.

Table G2. Descriptive statistics, Chile

	Heterosexual %	Gay/Lesbian %	Bisexual %	Other %	Refused %
<i>Gender</i>					
Woman	43.8	40.1	57.5	71.2	42.9
<i>Age Group</i>					
25-34 years	14.2	37.8	45.1	9.6	14.9
35-44 years	17.4	22.1	18.3	15.0	9.2
45-54 years	20.8	18.6	10.6	15.4	30.1
55-64 years	21.2	12.4	11.3	12.0	23.0
65-74 years	17.2	7.4	11.2	10.6	17.1
75-84 years	9.3	1.8	3.5	37.4	5.7
<i>Migration-background</i>					
Native born	94.8	89.0	93.6	99.0	97.5
Foreign born	5.2	10.9	6.4	0.1	2.5
Missing					
<i>n (observations)</i>	47,534	546	108	33	126

Notes: Weighted descriptive statistics. 2017 CASEN

Table G3. Descriptive statistics, Germany

	Heterosexual %	Gay/Lesbian %	Bisexual %	Other %	Refused %
<i>Gender</i>					
Woman	50.4	40.8	65.7	56.7	58.3
<i>Age Group</i>					
25-34 years	17.8	28.8	39.3	9.9	8.2
35-44 years	17.7	19.5	22.4	9.6	9.6
45-54 years	22.2	26.5	16.9	16.1	17.9
55-64 years	19.0	14.7	8.3	21.6	21.7
65-74 years	13.5	8.0	9.8	21.7	19.9
75-84 years	9.8	2.6	3.3	21.0	22.7
<i>Migration-background</i>					
Native born	84.6	89.9	91.0	88.0	85.7
Foreign born	15.4	10.1	9.0	12.0	14.3
<i>Missing</i>					
<i>n (observations)</i>	37,285	509	395	1,444	2,069

Notes: Weighted descriptive statistics. 2016 and 2019 data from SOEP.

Table G4. Descriptive statistics, UK

	Heterosexual %	Gay/Lesbian %	Bisexual %	Other %	Refused %
<i>Gender</i>					
Woman	53.9	40.2	55.0	53.8	54.0
<i>Age Group</i>					
25-34 years	16.3	21.6	37.4	23.7	14.8
35-44 years	19.8	23.8	22.9	18.5	18.0
45-54 years	21.4	30.4	21.6	21.2	19.7
55-64 years	19.4	15.1	9.7	15.4	17.4
65-74 years	15.4	6.9	6.8	15.0	18.3
75-84 years	7.8	2.3	1.6	6.2	11.8
<i>Migration-background</i>					
Native born	89.8	89.6	85.6	73.6	71.9
Foreign born	10.2	10.4	14.4	26.4	28.1
<i>n (observations)</i>	35,552	495	337	293	1,168

Notes: Weighted descriptive statistics. 2011/2012 and 2017/2018 data from UK-HLS.

Table G5. Descriptive statistics, US

	Heterosexual %	Gay/Lesbian %	Bisexual %	Refused %
<i>Gender</i>				
Woman	54.2	47.6	72.5	62.5
<i>Age Group</i>				
25-34 years	21.5	34.3	46.6	17.7
35-44 years	21.6	21.9	25.2	22.4
45-54 years	21.2	21.5	15.5	15.4
55-64 years	19.4	15.5	8.4	25.0
65-74 years	11.3	6.1	2.6	13.6
75-84 years	5.0	0.6	1.5	5.8
<i>Migration-background</i>				
Native born	86.5	92.3	84.9	63.8
Foreign born	13.5	7.7	15.1	36.2
<i>Ethnicity</i>				
Black	13.5	13.3	14.7	14.1
Other	9.9	6.8	12.9	23.8
White	76.5	79.9	72.4	62.1
<i>n (observations)</i>	8,949	183	206	170

Notes: Weighted descriptive statistics. 2008-2018 data from GSS.

Online Appendix H. Results using alternative methods

Table H1. Unidiff models testing for differences in intergenerational association of education across groups

	Overall		Men		Women	
	<i>Coef.</i>	<i>SE.</i>	<i>Coef.</i>	<i>SE.</i>	<i>Coef.</i>	<i>SE.</i>
<i>Australia</i>						
Bisexual	-0.30	0.18	-0.39	0.21	-0.27	0.20
Gay/Lesbian	0.01	0.16	0.13	0.36	-0.17	0.27
<i>Chile</i>						
Bisexual	0.10	0.21	0.76*	0.30	-0.23	0.33
Gay/Lesbian	-0.13	0.11	-0.16	0.16	-0.14	0.17
<i>Germany</i>						
Bisexual	-0.02	0.15	-0.23	0.23	0.01	0.16
Gay/Lesbian	-0.38*	0.19	-0.26	0.31	-0.50	0.31
<i>UK</i>						
Bisexual	-0.11	0.63	-0.35	0.27	0.03	0.20
Gay/Lesbian	-0.56*	0.23	-0.29	0.33	-0.88*	0.44
<i>US</i>						
Bisexual	-0.29	0.21	-0.75	0.56	-0.11	0.21
Gay/Lesbian	-0.24	0.24	-0.34	0.32	-0.03	0.35

Notes: * $p < 0.05$. Unidiff models test whether the association between three-category parental education and three-category own education differs by sexual identity (Erikson & Goldthorpe, 1992; Jann & Seiler, 2020). We ran three models for each country separately: one for the whole sample, one for men, and one for women. These models are not based on matched samples. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US).

Table H2a-e. Log-linear models testing for statistically significant differences in upward and downward mobility across groups

Table H2a. Log-linear models explaining the frequency of upward and downward mobility (Australia)

	Overall				Men				Women			
	Bisexual		Gay/Lesbian		Bisexual		Gay/Lesbian		Bisexual		Gay/Lesbian	
	IRR	SE	IRR	SE	IRR	SE	IRR	SE	IRR	SE	IRR	SE
Parental Education (Ref. ISCED 5-6)												
ISCED 3-4	1.22***	0.02	1.21***	0.02	1.21***	0.03	1.21***	0.03	1.22***	0.02	1.22***	0.02
ISCED 1-2	0.69***	0.01	0.69***	0.01	0.68***	0.02	0.68***	0.02	0.71***	0.02	0.71***	0.02
Bisexual (Ref. Heterosexual)	0.02***	0.00			0.01***	0.00						
Gay/Lesbian (Ref. Heterosexual)			0.03***	0.00			0.03***	0.00	0.03***	0.00	0.03***	0.00
Bisexual × Parental ISCED 3-4	0.84***	0.09			1.14	0.22			0.73*	0.10		
Bisexual × Parental ISCED 1-2	0.65***	0.09			0.84	0.21			0.57**	0.09		
Gay/Lesbian × Parental ISCED 3-4			0.44***	0.05			0.53***	0.09			0.36*	0.07
Gay/Lesbian × Parental ISCED 1-2			0.29***	0.04			0.49***	0.09			0.15***	0.03
<i>Mobility (Ref. Immobile)</i>												
Downwardly mobile	0.34***	0.01	0.34***	0.01	0.34***	0.01	0.34***	0.01	0.34***	0.01	0.34***	0.01
Upwardly mobile	0.88***	0.01	0.88***	0.01	0.89***	0.02	0.89***	0.02	0.88***	0.02	0.88***	0.02
Bisexual × Downwardly mobile	1.32*	0.15			1.24	0.25			1.37*	0.18		
Bisexual × Upwardly mobile	0.95***	0.11			0.66	0.13			1.12	0.15		
Gay/Lesbian × Downwardly mobile			0.48***	0.07			0.52**	0.11			0.44***	0.09
Gay/Lesbian × Upwardly mobile			2.31***	0.26			2.00***	0.30			2.79***	0.49
BIC Model	2985.88		2884.31		1340.7		1365.6		1411.40		1363.5	
BIC Model without interaction	2986.08		2979.67		1341.7		1399.9		1411.26		1422.7	

Table H2b. Log-linear models explaining the frequency of upward and downward mobility (Chile)

	Overall				Men		Women					
	Bisexual		Gay/Lesbian		Bisexual	SE	Gay/Lesbian		Bisexual		Gay/Lesbian	
	IRR	SE	IRR	SE			IRR	SE	IRR	SE	IRR	SE
Parental Education (Ref. ISCED 5-6)												
ISCED 3-4	2.15***	0.04	2.15***	0.04	2.02***	0.05	2.02***	0.05	2.29***	0.06	2.29***	0.06
ISCED 1-2	4.92***	0.08	4.92***	0.08	4.43***	0.11	4.43***	0.11	5.48***	0.14	5.48***	0.14
Bisexual (Ref. Heterosexual)	0.01***	0.00			0.00***	0.00						
Gay/Lesbian (Ref. Heterosexual)			0.03***	0.00			0.03***	0.00	0.01***	0.00	0.02***	0.00
Bisexual × Parental ISCED 3-4	0.35***	0.10			0.23*	0.15			0.33**	0.11		
Bisexual × Parental ISCED 1-2	0.19***	0.05			0.23**	0.11			0.14***	0.05		
Gay/Lesbian × Parental ISCED 3-4			0.48***	0.06			0.34***	0.06			0.70	0.14
Gay/Lesbian × Parental ISCED 1-2			0.24***	0.04			0.14***	0.03			0.41***	0.09
<i>Mobility (Ref. Immobile)</i>												
Downwardly mobile	0.23***	0.00	0.23***	0.01	0.21***	0.01	0.21***	0.01	0.26***	0.01	0.26***	0.01
Upwardly mobile	0.57***	0.01	0.57***	0.01	0.60***	0.01	0.60***	0.01	0.55***	0.01	0.55***	0.01
Bisexual × Downwardly mobile	1.37	0.51			0.61	0.65			1.30*	0.52		
Bisexual × Upwardly mobile	1.44	0.32			1.49	0.54			1.48	0.44		
Gay/Lesbian × Downwardly mobile			1.24	0.19			1.03	0.23			1.49***	0.34
Gay/Lesbian × Upwardly mobile			1.71***	0.17			2.17***	0.33			1.44**	0.20
BIC Model	9792.93		9925.66		4621.3		4687.0		5116.05		5167.8	
BIC Model without interaction	9788.97		9948.16		4617.2		4708.2		5112.34		5157.2	

Table H2c. Log-linear models explaining the frequency of upward and downward mobility (Germany)

	Overall				Men		Women					
	Bisexual		Gay/Lesbian		Bisexual	Gay/Lesbian	Bisexual		Gay/Lesbian			
	IRR	SE	IRR	SE	IRR	SE	IRR	SE	IRR	SE		
Parental Education (Ref. ISCED 5-6)												
ISCED 3-4	2.54***	0.04	2.54***	0.04	2.24***	0.05	2.24***	0.05	2.79***	0.05	2.79***	0.05
ISCED 1-2	0.40***	0.01	0.40***	0.01	0.35***	0.01	0.35***	0.01	0.45***	0.01	0.45***	0.01
Bisexual (Ref. Heterosexual)	0.01***	0.00			0.01***	0.00						
Gay/Lesbian (Ref. Heterosexual)			0.02***	0.00			0.02***	0.01	0.02***	0.00	0.01***	0.00
Bisexual × Parental ISCED 3-4	0.83	0.11			0.90	0.19						
Bisexual × Parental ISCED 1-2	0.57**	0.12			0.49	0.20						
Gay/Lesbian × Parental ISCED 3-4			0.73*	0.09			0.91	0.14	0.77	0.12	0.56**	0.11
Gay/Lesbian × Parental ISCED 1-2			0.41**	0.08			0.44**	0.11	0.58*	0.18	0.40**	0.12
<i>Mobility (Ref. Immobile)</i>												
Downwardly mobile	0.18***	0.00	0.18***	0.00	0.16***	0.00	0.16***	0.00	0.20***	0.00	0.20***	0.00
Upwardly mobile	0.67***	0.01	0.67***	0.01	0.86***	0.02	0.86***	0.01	0.52***	0.01	0.52***	0.01
Bisexual × Downwardly mobile	1.38*	0.19			1.70*	0.40			1.20	0.21		
Bisexual × Upwardly mobile	0.91	0.11			0.76	0.17			1.11	0.17		
Gay/Lesbian × Downwardly mobile			0.91	0.13			1.02	0.19			0.83	0.20
Gay/Lesbian × Upwardly mobile			1.42**	0.15			1.06	0.14			1.91***	0.33
BIC Model	5347.1		5508.6		2034.1		2045.54		2785.05		2766.37	
BIC Model without interaction	5531.0		5514.4		2035.8		2039.98		2780.62		2776.82	

Table H2d. Log-linear models explaining the frequency of upward and downward mobility (UK)

	Overall				Men				Women			
	Bisexual		Gay/Lesbian		Bisexual		Gay/Lesbian		Bisexual		Gay/Lesbian	
	IRR	SE	IRR	SE	IRR	SE	IRR	SE	IRR	SE	IRR	SE
Parental Education (Ref. ISCED 5-6)												
ISCED 3-4	3.17***	0.05	3.17***	0.05	3.21***	0.08	3.21***	0.08	3.14***	0.07	3.13***	0.07
ISCED 1-2	1.67***	0.03	1.67***	0.03	1.80***	0.05	1.80***	0.05	1.58***	0.04	1.58***	0.04
Bisexual (Ref. Heterosexual)	0.02***	0.00			0.02***	0.00			0.02***	0.00		
Gay/Lesbian (Ref. Heterosexual)			0.02***	0.00			0.02***	0.00			0.02***	0.00
Bisexual × Parental ISCED 3-4	0.46***	0.06			0.46***	0.10			0.46***	0.08		
Bisexual × Parental ISCED 1-2	0.26***	0.05			0.31***	0.08			0.22***	0.05		
Gay/Lesbian × Parental ISCED 3-4			0.55***	0.07			0.65***	0.12			0.46***	0.08
Gay/Lesbian × Parental ISCED 1-2			0.32***	0.05			0.42***	0.09			0.21***	0.05
<i>Mobility (Ref. Immobile)</i>												
Downwardly mobile	0.33***	0.01	0.33***	0.01	0.36***	0.01	0.36***	0.01	0.31***	0.01	0.31***	0.01
Upwardly mobile	0.84***	0.01	0.84***	0.01	0.79***	0.01	0.79***	0.01	0.87***	0.01	0.87***	0.01
Bisexual × Downwardly mobile	0.86	0.13			0.89	0.21			0.85	0.17		
Bisexual × Upwardly mobile	1.17	0.14			1.41	0.27			1.00	0.17		
Gay/Lesbian × Downwardly mobile			0.87	0.12			0.77	0.16			0.97	0.19
Gay/Lesbian × Upwardly mobile			1.79***	0.18			1.92***	0.26			1.68***	0.26
BIC Model	3001.01		3025.20		879.35		885.73		1211.15		1225.56	
BIC Model without interaction	2996.93		3056.86		877.68		910.23		1206.56		1232.13	

Table H2e. Log-linear models explaining the frequency of upward and downward mobility (US)

	Overall				Men		Women					
	Bisexual		Gay/Lesbian		Bisexual	SE	Gay/Lesbian		Bisexual		Gay/Lesbian	
	IRR	SE	IRR	SE			IRR	SE	IRR	SE	IRR	SE
Parental Education (Ref. ISCED 5-6)												
ISCED 3-4	1.76***	0.05	1.76***	0.05	1.71***	0.06	1.71***	0.06	1.80***	0.06	1.80***	0.06
ISCED 1-2	0.43***	0.01	0.43***	0.01	0.39***	0.02	0.39***	0.02	0.47***	0.02	0.47***	0.02
Bisexual (Ref. Heterosexual)	0.04***	0.00			0.02***	0.00			0.05***	0.01		
Gay/Lesbian (Ref. Heterosexual)			0.03***	0.00			0.03***	0.01			0.02***	0.01
Bisexual × Parental ISCED 3-4	0.62***	0.09			0.64	0.18			0.63***	0.11		
Bisexual × Parental ISCED 1-2	0.48***	0.11			0.68	0.27			0.40***	0.11		
Gay/Lesbian × Parental ISCED 3-4			0.59***	0.10			0.47**	0.11			0.77**	0.20
Gay/Lesbian × Parental ISCED 1-2			0.40***	0.11			0.32**	0.12			0.60**	0.29
<i>Mobility (Ref. Immobile)</i>												
Downwardly mobile	0.17***	0.01	0.17***	0.01	0.19***	0.01	0.19***	0.01	0.16***	0.01	0.16***	0.01
Upwardly mobile	0.63***	0.02	0.63***	0.02	0.60***	0.02	0.60***	0.02	0.65***	0.02	0.65***	0.02
Bisexual × Downwardly mobile	1.66**	0.28			1.68	0.56			1.57*	0.30		
Bisexual × Upwardly mobile	0.90	0.15			1.07	0.33			0.81	0.16		
Gay/Lesbian × Downwardly mobile			1.69*	0.36			1.98*	0.55			1.33	0.44
Gay/Lesbian × Upwardly mobile			1.79**	0.32			2.59***	0.63			1.14	0.32
BIC Model	3788.20		3683.21		1605.56		1622.16		2043.32		1961.79	
BIC Model without interaction	3790.85		3689.72		1602.11		1634.35		2044.49		1957.03	

Notes: * $p < 0.05$. These analyses are based on Poisson models predicting frequencies of cells in three-way tabulation of sexual identity, own and parental education. We ran three models for each country separately: one for the whole sample, one for men, and one for women. These models are not based on matched samples. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US). IRR: = Incidence Rate Ratio; SE = Standard Error; ISCED = International Standard Classification of Education; BIC = Bayesian Information Criterion.

Online Appendix I. Additional models explaining educational attainment

Examining educational mobility may help us understand the educational advantage of gay men and lesbian women observed in Figures 2 and 3 and documented in earlier research for the US (Mittleman, 2022). To generate these insights, Table I1 shows several models comparing college attainment between gay/lesbian and heterosexual individuals in the five countries under consideration.

Two key observations can be made based on this analysis. First, the estimated effect of parental education on respondent's educational attainment is weaker among gay/lesbian than heterosexual people. Although interaction effects are occasionally not statistically significant, the overall pattern of results lends support to the proposition that the intergenerational transmission of educational advantage is weaker among gay/lesbian individuals (i.e., that mobility is greater for this group).

Second, the educational advantage observed for gay/lesbian people is primarily driven by their higher levels of upward mobility. Gay and lesbian individuals' educational premium is greater among those with lower-educated parents compared to those with higher-educated parents (in both cases, this is relative to heterosexual individuals with similarly educated parents). For instance, in Chile, Germany and the UK, the educational advantage of gay/lesbian individuals fades entirely for individuals with highly educated parents.

Additional analyses indicated that the estimated effect of identifying as gay/lesbian on educational attainment for individuals with highly educated parents was not statistically significant in Chile, Germany, the UK or the US (effects were larger and significant in Australia). In short, we only find firm evidence that gay/lesbian people are more highly educated than heterosexual people for individuals who come from lower educational backgrounds. This suggests that explanations for the observed educational differences between heterosexual and gay/lesbian individuals should focus on explaining why gay/lesbian individuals with lower-educated parents are more likely to be upwardly mobile, rather than on factors that may explain an overall educational premium for gay/lesbian people across the board.

Table 11. Results from linear probability models based on matched samples explaining college attainment

	Australia		Chile		Germany		UK		US	
	<i>Coef.</i>	SE	<i>Coef.</i>	SE	<i>Coef.</i>	SE	<i>Coef.</i>	SE	<i>Coef.</i>	SE
Panel 1: Pooled sample										
Gay/Lesbian	0.18***	0.04	0.04*	0.02	0.09**	0.03	0.15***	0.03	0.12**	0.04
Parent(s) college educated	0.33***	0.01	0.46***	0.01	0.35***	0.01	0.34***	0.01	0.39***	0.02
Gay/Lesbian × Parent(s) college ed.	-0.07	0.06	-0.05	0.05	-0.09	0.06	-0.14*	0.06	-0.06	0.08
Constant	0.32***	0.01	0.20***	0.00	0.34***	0.00	0.38***	0.00	0.22	0.01
Panel 2: Men										
Gay/Lesbian	0.17**	0.05	0.08*	0.03	0.03	0.04	0.15***	0.04	0.20**	0.06
Parent(s) college educated	0.30***	0.02	0.44***	0.01	0.31***	0.01	0.34***	0.02	0.38***	0.03
Gay/Lesbian × Parent(s) college ed.	-0.01	0.08	-0.08	0.07	-0.04	0.08	-0.08	0.09	-0.21†	0.11
Constant	0.31***	0.01	0.22***	0.00	0.38***	0.01	0.36***	0.00	0.22***	0.01
Panel 3: Women										
Gay/Lesbian	0.19**	0.06	0.02	0.03	0.17***	0.05	0.14**	0.05	0.02	0.06
Parent(s) college educated	0.34***	0.01	0.46***	0.01	0.42***	0.01	0.34***	0.01	0.40***	0.03
Gay/Lesbian × Parent(s) college ed.	-0.13	0.08	-0.04	0.09	-0.16†	0.09	-0.19*	0.08	0.14	0.11
Constant	0.40***	0.01	0.18***	0.00	0.26***	0.01	0.42***	0.01	0.25***	0.01

Notes: Linear probability model coefficients (95% Confidence Intervals) based on coarsened exact matching. Matching variables: age, ethno-migrant background, parental education, gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US). SE: Standard error. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$.

Table I2. Results from linear probability models based on matched samples explaining college attainment

	Australia		Chile		Germany		UK		US	
	<i>Coef.</i>	SE	<i>Coef.</i>	SE	<i>Coef.</i>	SE	<i>Coef.</i>	SE	<i>Coef.</i>	SE
Men										
Bisexual	-0.04	0.06	0.04	0.07	-0.07	0.05	0.05	0.05	0.10	0.08
Parent(s) college educated	0.28***	0.02	0.44***	0.02	0.32***	0.01	0.32***	0.02	0.29***	0.05
Bisexual × Parent(s) college ed.	-0.06	0.10	0.23 [†]	0.13	-0.06	0.11	0.04	0.09	0.02	0.14
Constant	0.33***	0.01	0.18***	0.00	0.39***	0.01	0.37***	0.01	0.26***	0.03
Women										
Bisexual	-0.07 [†]	0.04	0.15*	0.07	0.01	0.03	-0.06	0.04	0.01	0.05
Parent(s) college educated	0.31***	0.02	0.45***	0.02	0.41***	0.01	0.33***	0.01	0.38***	0.03
Bisexual × Parent(s) college ed.	-0.06	0.06	-0.13	0.13	0.13	0.07	0.06	0.08	0.03	0.09
Constant	0.40***	0.01	0.19***	0.00	0.26***	0.01	0.43***	0.01	0.23***	0.02

Notes: Linear probability model coefficients (95% Confidence Intervals) based on coarsened exact matching. Matching variables: age, country of birth/ethnicity, parental education, gender. Data from 2012, 2016 and 2020 for HILDA (Australia), 2017 for CASEN (Chile), 2016 and 2019 for SOEP (Germany), 2011/2012 and 2017/2018 for UK-HLS (UK), and 2008-2018 for the GSS (US). SE: Standard error. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; [†] $p < 0.10$.

References

- Aksoy, C. G., Carpenter, C. S., & Frank, J. (2018). Sexual orientation and earnings: new evidence from the United Kingdom. *ILR Review*, 71(1), 242-272.
- Elliott, M. N., Dahlhamer, J. M., MacCarthy, S., Beckett, M. K., Orr, N., Guerino, P., ... & Martino, S. C. (2019). Using ancillary sociodemographic data to identify sexual minority adults among those responding “Something Else” or “Don’t Know” to sexual orientation questions. *Medical Care*, 57(12), e87.
- Goldberg, S. K., Rothblum, E. D., Russell, S. T., & Meyer, I. H. (2020). Exploring the Q in LGBTQ: Demographic characteristic and sexuality of queer people in a US representative sample of sexual minorities. *Psychology of Sexual Orientation and Gender Diversity*, 7(1), 101.
- Kühne, S., Kroh, M., & Richter, D. (2019). Comparing self-reported and partnership-inferred sexual orientation in household surveys. *Journal of Official Statistics*, 35(4), 777-805.
- Mittleman, J. (2022). Intersecting the academic gender gap: The education of lesbian, gay, and bisexual America. *American Sociological Review*, 87(2), 303-335.
- Mize, T. D. (2016). Sexual orientation in the labor market. *American Sociological Review*, 81(6), 1132-1160.
- Ridolfo, H., Miller, K., & Maitland, A. (2012). Measuring sexual identity using survey questionnaires: How valid are our measures?. *Sexuality Research and Social Policy*, 9(2), 113-124.
- Wooden, M. (2014). The measurement of sexual Identity in Wave 12 of the HILDA survey (and associations with mental health and earnings). *HILDA Project Discussion Paper Series*, No. 1/14, February 2014.