THÈSE de DOCTORAT



de l'UNIVERSITÉ TOULOUSE CAPITOLE

Présentée et soutenue par

Madame Julia HOEFER MARTI

Le 16 mai 2024

Essays on the Economics of Gender and Norms

École doctorale : **TSE Toulouse Sciences Economiques** Spécialité : **Sciences Economiques - Toulouse** Unité de recherche : **TSE-R - Toulouse School of Economics -Recherche**

Thèse dirigée par Monsieur Paul SEABRIGHT

Composition du jury Rapporteur : M. Guido FRIEBEL Rapporteure : Mme Amma PANIN Examinatrice : Mme Josepa MIQUEL-FLORENSA Présidente Mme Emmanuelle AURIOL



Essays in the Economics of Gender and Norms

Ph.D Thesis

Julia Hoefer $Martí^1$

¹ Toulouse School of Economics

March 13, 2024

Contents

A	cknowledgments	7
Su	ımmary	8
1	Sexual Orientation and the Motherhood Earnings Penalty	10
	1 Introduction	12
	2 Literature Review	13
	3 Data and Methodology	16
	4 Descriptive Statistics	20
	5 Estimation and Results	24
	5.1 Main results	24
	5.2 Primary vs. Secondary Earners	27
	5.3 Primary vs. Secondary House-workers	30
	5.4 Effects on Employment	32
	6 Conclusions and Further Research	34
A	ppendices	41
	I.A Extra Tables	42
	1.B Primary vs. Secondary Workers	43
2	Alcohol, Behavioral Norms and Sexual Violence on U.S. College	
	Campuses	46
	$1 \qquad \text{Introduction} \dots \dots$	48
	2 Literature Review	51
	2.1 The link between alcohol and violence	51
	2.2 Other factors affecting the incidence of sexual violence on col-	
	lege campuses	52
	2.3 Reporting Sexual Violence to Police	53
	3 The Model	55

	$3.1 \qquad \text{Basic set-up} \dots \dots$	55
	3.2 The student's choices	57
	3.3 The incidence of assault	63
4	Data and Descriptive Statistics	67
	4.1 Data sources and choice of variables	67
	4.2 Using the model to derive an empirical specification	69
	4.3 Incidents by month and campus type	73
	4.4 Regional characteristics	77
	4.5 Characteristics of perpetrators and victims	78
5	Regression Results	78
	5.1 The effect of technical factors excluding social and psycholog-	
	ical costs	78
	5.2 Main findings: including social and psychological costs	80
	5.3 Robustness checks	82
6	HERI attitudinal data	92
	6.1 How do religious students differ from atheist ones?	93
	6.2 How do students on campuses with a nearby Planned Parent-	
	hood clinic differ from those without?	94
7	Conclusions and policy implications	97
A		104
Appen	Supplementary tables	104
2.A	Other support of U.S. data service and	104
2.B	Other sources of U.S. data considered	114
3 Rel	igion and State	116
1	Introduction	117
2	Literature Review	118
3	Data and Methodology	123
4	Descriptive Statistics	129
5	Results	132
	5.1 Preferred Estimation	133
	5.2 Interaction with Gender	135
	5.3 Extensive margin	137
6	Conclusion	139
Appen		147
$3.\mathrm{A}$	Current government support	147

$3.\mathrm{B}$	Funding vs. regulatory support	151
$3.\mathrm{C}$	Robustness test: Primary enrolment Gender Parity Index	154
3.D	Gini imputation	157

List of Figures

3.1	Timing of events and decisions in the model: objective payoffs .	60
3.2	Predicted effect of g on drunken assault for various values of γ .	66
4.1	Mean number of incidents by month	74
4.2	Mean incidents per month by campus alcohol policy	74
4.3	Incidents per capita by university size	75
4.4	Mean incidents per month by Planned Parenthood availability	76
4.5	Mean incidents per capita per month by campus religion	77
5.1	Property damage and PP clinics	84
5.2	Assault and PP clinics	84
5.3	Property damage and religious affiliation	84
5.4	Assault and religious affiliation	84
5.5	Black individuals' representation among offenders compared to	
	general and student populations	92
6.1	Beer consumption by religion	93
6.2	Wine consumption by religion	93
6.3	Attitudes towards abortion	93
6.4	Attitudes towards hookups	93
6.5	Hrs/week on homework	94
6.6	Hrs/week partying	94
6.7	Attitudes to abortion by campus	95
6.8	Attitudes towards hookups by campus	95
6.9	Beer consumption by campus	95
6.10	Wine consumption by campus	95
31	State support percentiles for regression sample countries	128

List of Tables

1.1	Main variables - definitions and availability	17
1.2	Descriptive means by motherhood status	21
1.3	Probit - Lesbian self-identification	23
1.4	Income effects of motherhood – EU-SILC data	25
1.5	Income effects of motherhood – PSID data	26
1.6	Primary vs. secondary earners – EU-SILC data	28
1.7	Primary vs. secondary earners – PSID data	29
1.8	Primary vs. secondary house-workers – PSID data	31
1.9	LFP effects of motherhood – EU-SILC data	32
1.10	LFP effects of motherhood – PSID data	33
1.A.1	Full list of EU-SILC variables	42
1.A.2	Full list of PSID variables	43
1.B.1	Primary vs. secondary workers – EU-SILC data	44
1.B.2	Primary vs. secondary workers – PSID data	45
		
4.1	Main variables and definitions	68
$\frac{4.1}{4.2}$	Main variables and definitions Summary means	68 69
$\frac{4.1}{4.2}$	Main variables and definitions Summary means Alcohol use by campus alcohol policy	68 69 75
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 $	Main variables and definitions	68 69 75 77
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type	68 69 75 77 78
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs	68 69 75 77 78 79
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data	68 69 75 77 78 79 81
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ 5.3 \\ $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data Regressions with other crimes as controls	68 69 75 77 78 79 81 86
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ 5.3 \\ 5.4 \\ $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data Regressions with other crimes as controls Regressions with excess female/male students	68 69 75 77 78 79 81 86 87
$ 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ 5.3 \\ 5.4 \\ 5.5 \\ $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data Regressions with other crimes as controls Regressions with excess female/male students Summary statistics for sexual assault (NCVS) - proportions of	68 69 75 77 78 79 81 86 87
$ \begin{array}{c} 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ 5.3 \\ 5.4 \\ 5.5 \\ \end{array} $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data Regressions with other crimes as controls Regressions with excess female/male students Summary statistics for sexual assault (NCVS) - proportions of total cases	68 69 75 77 78 79 81 86 87 89
$ \begin{array}{c} 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ 5.3 \\ 5.4 \\ 5.5 \\ \hline 5.6 \\ \end{array} $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data Regressions with other crimes as controls Regressions with excess female/male students Summary statistics for sexual assault (NCVS) - proportions of total cases Reporting rates by crime (NCVS)	68 69 75 77 78 79 81 86 87 89 89
$ \begin{array}{c} 4.1 \\ 4.2 \\ 4.3 \\ 4.4 \\ 4.5 \\ 5.1 \\ 5.2 \\ 5.3 \\ 5.4 \\ 5.5 \\ \hline 5.6 \\ 5.7 \\ \end{array} $	Main variables and definitions Summary means Alcohol use by campus alcohol policy Alcohol use by red county/blue county Alcohol policy by county type Alcohol policy by county type Regression excluding social and psychological costs Main regressions - incidents of sexual assault, monthly data Regressions with other crimes as controls Regressions with excess female/male students Summary statistics for sexual assault (NCVS) - proportions of total cases Reporting rates by crime (NCVS) Probit marginal effects - cases of sexual assault	68 69 75 77 78 79 81 86 87 89 89 89

2.A.1	NCVS variable definitions	104
2.A.2	NCVS probit marginal effects - cases of rape	105
2.A.3	IV Negative binomial - incidents of sexual assault	106
2.A.4	Main regressions - incidents of rape, monthly data	107
2.A.5	Regressions using Planned Parenthood availability	108
2.A.6	Main regressions - incidents of sexual assault, daily data	109
2.A.7	Main regressions - incidents of sexual assault per capita	110
2.A.8	Regressions with HERI attitudinal data (II)	111
2.A.9	Main regressions - restricted dataset	112
2.A.10	Regressions controlling for student demographics	113
2.1	Predicted effects of government support of religion	122
3.1	Outcome variables	124
3.2	Variables of interest	125
4.1	Religious attitudes and behaviors	129
4.2	Denomination membership (regression sample)	130
4.3	Religious governance descriptive statistics – RAS data	131
4.4	Government religious preference – GRP data	131
4.5	In what kinds of countries do governments support religions?	132
5.1	Main results - multi-level mixed-effects model, lagged support	134
5.2	Results by gender - multi-level mixed-effects model, lagged support	136
5.3	Denomination membership and religious person - multi-level mixed-	
	effects model, lagged support	138
3.A.1	Main results - multi-level mixed-effects model, current support .	148
3.A.2	Results by gender - multi-level mixed-effects model, current support	149
3.A.3	Denomination membership and religious person - multi-level mixed-	
	effects model, current support	150
3.B.1	Funding vs. regulatory support - current values	152
3.B.2	Funding vs. regulatory support - lagged values	153
3.C.1	Primary enrolment GPI - multi-level mixed-effects model, current	
	support	155
$3.\mathrm{C.2}$	Primary enrolment GPI - multi-level mixed-effects model, lagged	
	support	156
3.D.1	Gini imputation comparison	159
3.D.2	Imputed Gini vs. World Bank average	160
$3.\mathrm{D.3}$	Confidence in government: Mean in raw vs. imputed sample	162

Acknowledgments

I am deeply grateful to my advisor, Paul Seabright, for his insightful advice, thoughtprovoking debate, and indefatigable support, which have meant the world to me these past years.

I would like to thank Emmanuelle Auriol, François Poinas, Karine Van Der Straeten, Ingela Alger, Isis Durrmeyer, Ian Hoefer, and participants at TSE Student Workshops, the ERINN Workshop on social norms, and UCL ENTER seminars for insightful feedback and discussion, as well as the faculty of the Toulouse School of Economics in general for providing a stimulating learning environment that I have been truly privileged to benefit from.

I am likewise deeply grateful to Nour Meddahi, Laurence Delorme, Ludmila Namolovan and Louise Strachan for their aid and advice in navigating the system on multiple occasions.

I would like to thank my friends in the Ph.D program in Toulouse, in particular Joana Duran, Jacint Enrich, Filippo Maria D'Arcangelo, Ling Zhou, Jacopo Bregolin, Miguel Zerecero, Chris Sandmann, Yuting Yang, and Olga Bernard, for the many years of fond memories. I would also like to thank my non-economist friends, particularly Emmalie Hall-Skank, Cyrus Fisher, and An Kelley, for their truly outstanding emotional support (and patience in indulging my jargon filled rants).

Finally, the biggest of thank-yous to my family, without whom none of this would have been possible. I love you always.

Summary

This thesis contains three essays studying the various ways in which outcomes are shaped by the interaction of economic factors, social norms, and institutional constraints, with a focus on women.

In the first chapter, I study the earnings effects of motherhood, and how they interact with women's sexual orientation. Motherhood carries an earnings penalty for women that is both well known and persistent; however, the share attributable to differences in human capital or productivity has decreased, while the unexplained share has increased over the decades. This has led to a renewed focus on household labor specialization, and the part played by gendered norms in determining it. Lesbian couples are not subject to the same social norms about housework that straight couples are, and indeed, preliminary evidence seemed to suggest that lesbian women see an earnings premium of motherhood.

I investigate empirically, using data from the EU Standards of Income and Living Conditions and US Panel Study of Income Dynamics. I confirm the earnings penalty for straight women but reject a general earnings premium for lesbian mothers; instead, I find suggestive evidence of a motherhood wage premium specifically for lesbian mothers who do not specialize in housework. In contrast, earning more, working more, or doing less housework than their partner are not enough to eliminate the motherhood earnings penalty for straight women, though all reduce it.

The second chapter is co-authored with Paul Seabright; in it, we investigate the idea that social norms against casual sex may inadvertently result in higher rates of sexual assault by incentivizing individuals to consume alcohol as a "disinhibitor" before attempting to find a sexual partner. We construct a decision-theoretic model in which a student may make the decision to consume alcohol as a way of strategically weakening the pressure from social norms against casual sex; the consumption of alcohol then drives incidents of sexual violence.

We take this model to data from the US National Incident-Based Reporting System, using the presence of Planned Parenthood in the university's county as a proxy for the strength of social norms against casual sex. Campuses in counties with no Planned Parenthood present do see higher rates of sexual violence where alcohol is involved. Finally, we examine attitudinal data from the Higher Education Research Institute and find limited evidence that incidents alcohol-fueled sexual violence are related to attitudes against consensual sex.

In the third chapter I study the effects of state intervention in the market for religion. Government support can deliver a dominant market position to a religious denomination, but it may also distort the institutions incentives and reduce its responsiveness to its own adherents. Previous research has attempted to measure the effects of state regulation and/or support of religion only at the population level, and often with only binary measures of state support.

I construct a standardized index of the strength of states' support of religion and study its effects on individuals' external and internal religiosity, and the legitimacy of religious institutions. I find that stronger government support increases external religiosity, but at the cost of decreasing individuals' internal religiosity and trust in religious institutions. These negative effects are felt mostly over the long-term, and are concentrated on members of those religions that receive state support. Women increase their external religiosity less, and the decrease in internal religiosity and trust in religious institutions is more keenly felt. Ultimately, government support has a negative effect on religious affiliation at the population level, but individuals abandon their denomination far more readily than their faith.

Chapter 1

Sexual Orientation and the Motherhood Earnings Penalty

Julia Hoefer Martí¹

Abstract

The effect of motherhood on women's earnings has spawned a large body of literature, and garnered increasing focus among researchers as the motherhood penalty has become an ever larger proportion of the overall gender earnings gap. Attention has been drawn to the role of gender norms and household labor in opposite-sex couples, making same-sex couples a useful comparison. While established literature has found a significant motherhood wage penalty for heterosexual women, preliminary evidence appeared to suggest no effect, or even a motherhood earnings premium, for homosexual ones.

I use longitudinal data from the EU Standards of Income and Living Conditions (EU-SILC) and the US Panel Study of Income Dynamics (PSID) to investigate the interaction of gender, sexual orientation, and parenthood. I confirm that heterosexual women see a significant earnings penalty, but homosexual women do not experience any significant income effects of motherhood at a group level. I further find suggestive evidence that the penalty for heterosexual women is driven by specialization in household labor, and that homosexual women who do less housework than their partners may indeed see an earnings premium of motherhood. ${\bf Keywords:}\ {\bf motherhood},\ {\bf sexual orientation},\ {\bf norms}$

JEL codes: J13, J31, J71

1 Introduction

Economists have been tackling the issue of unequal labor market outcomes almost for as long as it has been part of the public conscience. From S. Becker, 1957's first attempts at creating a parsimonious model for discrimination all the way up to the present day, economists the world over have worked to document and explain workplace inequalities in all their forms.

One of the most persistent forms of inequality in labor market outcomes is the gender gap in earnings, to which the motherhood earnings penalty is a large – and proportionally increasing – contributor. The motherhood earnings penalty has persisted despite attempts at correction through policies such as parental leave or equal pay legislation, and attempts at accounting for the gap through controlling for human capital differences, employment characteristics, or worker productivity have failed to close it completely.

Concurrently, as LGBT+ individuals have become a more visible part of society over the past few decades, economists have naturally begun to look at the possibility of a sexual orientation effect on income. Results here have been somewhat surprising, however: while gay men do experience an earnings penalty relative to straight men, lesbian women appear instead to experience an earnings premium relative to straight women. While the idea that historical anti-LGBT+ attitudes might have translated into labor market discrimination seems sensible prima facie, the effect of sexual orientation on income interacts in unexpected ways with gender and gender roles. It therefore becomes logical to ask whether gender and sexual orientation might also interact in unexpected ways when it comes to motherhood. If lesbian women do indeed experience an earnings premium over straight women, is this despite also suffering similar wage effects of motherhood – or might it be because they do *not*?

Section 2 covers the existing literature on the motherhood wage gap and the earnings effects of sexual orientation, and motivates in greater depth the idea that gender, orientation, and parenthood may interact in unexpected ways. Section 3 describes the data and lays out the methodological approach, Section 4 presents a selection of descriptive statistics, and Section 5 presents the results of the estimations. Section 6 concludes.

2 Literature Review

The gap in earnings between women who have children and those who do not, termed the "motherhood wage penalty", has been the subject of substantial (and ongoing) analysis. An earnings penalty of motherhood has been extensively documented in countries such as the U.S. (Waldfogel, 1998; Benard et al., 2007; Budig and Hodges, 2010; Gough and Noonan, 2013), the U.K. (Harkness, 2016), Germany (Gangl and Ziefle, 2009), Denmark (Kleven et al., 2019), Spain (de Quinto et al., 2021) – to name just a few. It is significant in both the statistical and colloquial sense: depending on the country and study, estimates of the immediate negative effect of motherhood on women's income tend to range between 5 and 10%, and is further compounded over time (Angelov et al., 2016). The motherhood penalty has proven remarkably difficult to combat, persisting despite notable convergence in human capital between mothers and non-mothers over the past three or four decades (Jee et al., 2019). The proportion of the overall *gender* gap in earnings attributable human capital differences has diminished over the past three to four decades (Blau and Kahn, 2017). At the same time, the proportion attributable to the motherhood penalty has increased from 40% in 1980 to 80% in 2013 (Kleven et al., 2019), making this issue ever more important to address properly.

Several explanations have been put forward to explain this persistent earnings penalty associated with motherhood. Firstly, motherhood affects women's commitment to the labor market: women who have or plan to have children may seek out jobs with more flexibility in balancing work and home life, or which require fewer hours in total. While there is clear occupational sorting, with mothers far more likely to work reduced hours or part-time jobs, flexible or "mother-friendly" job characteristics have in fact been shown to have relatively little explanatory power (Budig and England, 2001, Weeden, 2005). Consistent with this, policy interventions attempting to "flexibilize" women's attachment to the labor force through paid maternal leave have seen mixed results (Budig et al., 2012; Mandel and Semyonov, 2006).

A second possibility is that motherhood may cause women to be temporarily less productive at their jobs - an explanation known as the "work effort" hypothesis. Attempts to test the validity of the work effort hypothesis have consistently found results that are inconsistent with it, however (Crittenden, 2002; Anderson et al., 2003): women who return to work soon after becoming parents, when the effects on their focus, energy, and commitment should be strongest, in fact see the smallest effect on their earnings. In addition, both of the explanations laid out above above fail to account for the root causes of why having a child would cause one parent to seek a lower-intensity job, or muster less effort in their existing employment, but not the other.

This has led to a resurgence in attention on the role of intra-household specialization. The idea that specialization is a driver of income inequalities both within couples and in society as a whole is not new (G. S. Becker, 1965; G. S. Becker, 1991), but specialization predictions based on pre-parenthood earnings clash with the fact that women in opposite-sex couples take on a majority of household labor, regardless of whether both members of the couple work or not, and even when they are the primary earner (Apter, 1993; Bittman et al., 2003; Bertrand et al., 2015). It also fails to explain why marriage – which itself increases household specialization (Jepsen and Jepsen, 2015) – yields an earnings premium for both members of an opposite-sex couple, but parenthood has differential effects based on gender (Killewald and Gough, 2013).

Consequently, there has been a marked increase in proposed explanations that explicitly account for gender norms, bias, and discrimination. The idea that gendered specialization in paid labor versus household labor among opposite-sex couples lies at the root of the motherhood wage penalty - and the wider gender pay gap - has received empirical support (Montag, 2015; Downs et al., 2023). Given the increased focus on gender-based norms and their effect on women's labor supply and income, it makes sense to look at women in couples that do not, and cannot have differently gendered roles for each member: same-sex couples.

Studies looking at the earnings effects of LGBT+ status have found, since the beginning, a consistent earnings penalty for gay men relative to straight men (Badgett, 1995; Klawitter and Flatt, 1998; Blandford, 2003). Research has tended to find also that lesbian women experience an earnings premium over straight women in the U.S. (Berg and Lien, 2002; Black et al., 2003; Antecol et al., 2008; Aksoy et al., 2018), the U.K. (Arabsheibani et al., 2004), the Netherlands (Plug and Berkhout, 2004), and Sweden (Ahmed et al., 2011; Ahmed et al., 2013), among others – although not all studies into the issue agree (Carpenter, 2005; Del Río and Alonso-Villar, 2019).

Some evidence has suggested lesbian women may experience an advantage in hiring relative to straight women (Baert, 2014), be less penalized for motherhood (Peplau and Fingerhut, 2004; Jepsen, 2007), and be more likely to hold workplace authority

(Aksoy et al., 2019; de Vries and Steinmetz, 2023). There is also evidence to suggest lesbian women are considered more competent than straight women, *and* that this difference hinges on whether and how gender-stereotypical behavior is performed (Niedlich et al., 2015). On the other hand, there is also ample evidence of discrimination against lesbians as well as gay men (Coffman et al., 2017; Fasoli et al., 2017), and some studies have found a hiring disadvantage rather than an advantage (Drydakis, 2015; Mourelatos, 2023).

Time use and household specialization likewise differ in same-sex couples as compared to opposite-sex couples, and often in ways that challenge the idea that specialization is based on comparative advantage within the couple (Tebaldi and Elmslie, 2006; Martell and Roncolato, 2016). Given that the motherhood wage penalty is a significant and persistent component of the gender wage gap, it is logical to ask whether straight and lesbian women experience motherhood differently in terms of its effects on their income, household specialization, and labor market status. Initial forays into the question appeared to suggest lesbian women might experience an earnings premium of motherhood (Baumle, 2009). Subsequent studies have not confirmed this result, but overall indicate that lesbian mothers suffer a much smaller earnings penalty than straight mothers, where it exists at all (Andresen and Nix, 2022; Downs et al., 2023). Lesbian mothers are also less likely to reduce their labor market participation than straight mothers (Leppel, 2009; Antecol and Steinberger, 2013).

It should be noted, of course, that the evidence for the income and labor market effects of being LGBT+ is still largely very new, and research on the topic has faced additional challenges beyond those inherent in studying small minority populations. Household surveys often have not collected information on sexual orientation, forcing researchers to assign LGBT+ status based on available information, such as presence of a same-sex partner in the household – which risks mis-classifying unpartnered LGBT+ individuals (Martell and Eschelbach Hansen, 2017; Badgett et al., 2021). This matters, because while some studies find an earnings premium, not all do, and results often change significantly depending on whether LGBT+ status is self-reported or assigned – see e.g. Carpenter, 2005 or Badgett, 2018, which use self-reported orientation and find that only bisexual individuals face a significant earnings penalty and increased chance of poverty relative to straight individuals, respectively. There is also some evidence that inferring sexual orientation based on cohabitation or self-reported homosexual behavior may lead to inflated estimates of

lesbian women's earnings advantage over straight women (Martell, 2021). Attempts to account for this variation in the estimated earnings effects of sexual orientation have largely succeeded for the gay penalty, but have failed to find concrete answers for the lesbian premium in particular (Klawitter, 2015). If any one fact has been made clear by the research conducted thus far on gender, sexual orientation, and parenthood, it is how much more research still remains to be done.

This paper contributes to the as-yet scant literature on the income effects of parenthood for lesbian women. At the start of time of writing, and up until 2022, the literature indicated a motherhood earnings premium for lesbian women. Building on the approach first employed in Baumle, 2009, I improve on the cross-sectional analysis in several key ways: firstly, I obtain panel data for the EU as well as the US, and conduct a longitudinal regression analysis.¹ I add a variety of relevant individual-level and macroeconomic control variables drawn from the literature on the income effects of parenthood and the gender pay gap, and restrict the sample to partnered women to improve comparability between groups. My results confirm the continued presence of an earnings penalty of motherhood for straight women, but resoundingly reject a motherhood premium for lesbian women as a whole. Like Andresen and Nix, 2022, I also investigate various explanations for why motherhood income effects might differ for lesbian and straight women, including labor market disadvantage prior to parenthood and gender norm-driven household specialization, but using a more diverse dataset which features multiple Western countries. I obtain suggestive evidence that lesbian mothers who do not specialize in household labor may experience a motherhood earnings premium, corroborating the results from both Andresen and Nix, 2022 and Downs et al., 2023.

3 Data and Methodology

I use data from two primary sources in order to investigate motherhood earnings effects among straight and lesbian women: firstly, I use the EU Standards of Income and Living Conditions (EU-SILC) survey; and secondly, the US Panel Study of Income Dynamics (PSID) survey. Both are longitudinal surveys which collect information from all members of a particular household or family unit. A key difference between the two is that EU-SILC interviews a given household for four to

¹As is detailed in Section 3, I use both random- and fixed-effects approaches.

six consecutive years, after which the household exits the sample. In contrast, PSID began in 1968 and has followed the initial households, and their offshoots, to the present day where possible.²

Variable	Definition	EU-SILC	PSID
Income	Yearly income from labor	Y	Y
Has child	Dummy, $=1$ if individual has at least one child	Υ	Υ
Secondary education	Dummy, $=1$ if individual has completed secondary	Υ	Υ
Post-secondary education	Dummy, =1 if individual has completed post- secondary education such as vocational school	Y	-
Tertiary education	Dummy, =1 if individual has an undergraduate degree	Υ	Υ
Post-graduate education	Dummy, $=1$ if individual has a post-graduate degree	-	Υ
Black	Dummy, $=1$ if individual is Black	-	Υ
Asian	Dummy, $=1$ if individual is Asian or Pacific Islander	-	Υ
Age	Individual's age in years	Υ	Υ
Urban residence	Dummy, $=1$ if individual resides in an urban area	Υ	Υ
Hours worked/week	Average hours individual spends working each week	Υ	Υ
Hours housework/week	Average hours individual spends on housework each week	-	Υ
Occupation	Dummies for individual's type of occupation	Υ	-
Industry	Dummies for individual's industry of occupation	-	Υ

Table 1.1: Main variables - definitions and availability

Both datasets contain information on individual characteristics such as age, gender, educational attainment, occupational status, etc. Most importantly, both contain information which makes it possible to identify couples comprised of two individuals of the same gender. The EU-SILC data contains a Spouse/Partner ID variable, while the PSID survey records the relationship of all household members to the Reference Person³, as well as assigning each discrete couple a sequence number.

Each dataset has strengths and weaknesses: EU-SILC is a much larger dataset, and the survey questionnaire allows individuals to designate their spouse or long-term partner, resulting in a higher rate of identification of same-sex couples. On the other hand, the PSID survey collects information about time spent working *and* time spent on housework – useful when analysing a topic where household specialization may

²To ensure general comparability both between EU-SILC countries in the sample, and between EU and US data, I exclude Eastern European and ex-USSR countries. Of the remainder, gross yearly income is available for Austria, Belgium, Denmark, Finland, France, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, and the United Kingdom. The final EU-SILC sample spans from 2003 to 2013.

³ "Head of Household" prior to 2017.

play a crucial part – while EU-SILC data contains information only about time spent at work. If, as we will discuss in greater detail in Section 5, the persistence of the motherhood wage penalty is due to the fact that (straight) women take on the lion's share of household labor and child-rearing, then seeing who does the most household labor within a particular couple is useful indeed. The PSID survey also tracks households long-term, with some family units having been in the survey since its inception in 1968, which leads to a much more complete picture of individuals' lives and the trajectory of their careers.

Unfortunately, the PSID data has a number of crucial drawbacks: firstly, and most importantly, it has very few LGBT+ observations. Until 1987, the PSID survey classified same-sex partners of the Reference Persons as "other individuals not in family unit" (a general category which also included, for instance, friends of children of the family unit who resided in the household). Additionally, while European nations mostly legalized marriage equality in the early 2000s, barring a few holdouts such as France and the UK, marriage equality came later for most US states, and the country as a whole only legalized it in 2015. Marriage equality is highly correlated with people being willing to identify a same-sex partner, and the relatively late arrival of marriage equality laws in the US is reflected in the data: LGBT women in the regression sample have an average age of 32 as opposed to 40 for straight women in the PSID sample, while in the EU-SILC sample, the average ages of straight and lesbian women are almost identical (just over 43 years for both groups). A second drawback of the PSID data is that, for budgetary reasons, the survey switched to a biennial format in 1997, and some attempt has been made to compensate for this by asking for the last two years of income, other important household variables get the same treatment. As a result, it was necessary to impute the average hours per week spent working and spent on household labor, for "off" years.

As a final note, the approach to identifying same-sex couples I employ here has several shortcomings that must be acknowledged. Most obviously, it does not allow us to identify lesbian women who were single during the entire survey period, and instead assumes all women who are single for the duration are heterosexual - which is clearly not the case. For this reason, I restrict the sample to straight and lesbian women in a couple. Secondly, this approach does not distinguish between bisexual and lesbian women, but it is perfectly possible that bisexual women might behave differently from both straight and lesbian women, whatever their partner's gender; they may be subject to different incentives and constraints, or may be perceived differently. There is some evidence that this is the case: Carpenter, 2005, for instance, finds a wage penalty associated with bisexuality for women but no penalty for homosexuality. However, as the dataset is not suited to discuss this problem, we must ignore the possibility.

To estimate the combined effect of sexual orientation and motherhood on women's income, I use both random-effects and fixed-effects regressions. The random-effects estimation tells us how mothers differ from non-mothers, while fixed-effects tells us what happens to a given woman's salary when she becomes a mother. As we will see in Section [5] both are necessary for a complete picture. If the difference between the random- and fixed-effects regressions itself depends on sexual orientation – as we will indeed see is the case – then this tells us something about how women attempt to compensate (or don't) for motherhood's effect on their careers and income. People, after all, are generally rational and forward-looking; if they anticipate an earnings penalty of motherhood (*whatever* the reason), women who intend to become mothers may, for instance, opt for a high-paying job at the beginning of their working lives, and then plan to move to less demanding ones or quit the labor market entirely after motherhood. This is a difference between women who are mothers and women who are not, but nevertheless remains an earnings effect of motherhood.

A number of different specifications are used. The preferred estimations use all lesbian and straight women over the age of 16 who are in a couple. A second specification splits couples into primary and secondary earners, depending on which individual earned more in the first year of the survey in which both members of the couple appear. Given that men and women in heterosexual couples often specialize in household labor or market labor, treating members of a homosexual couple as being identical to each other because they are the same gender may hide important differences in behavior. I address this problem in two ways: first, I split couples into primary and secondary earners, depending on which partner reported the higher income in the first year both members of a couple were present in the survey. Lastly, a PSID-only regression splits couples into primary and secondary house workers.

I control for the individual's level of education, usual weekly work hours, and include a dummy variable indicating whether same-sex marriage was legal in the individual's country and year. I also include a number of occupational variables, to account for the fact that lesbian women tend to sort themselves into different fields of work than heterosexual women, and in particular have higher propensity to enter traditionally male-dominated fields (Del Río and Alonso-Villar, 2019). Occupational variables for the EU-SILC sample are based on the European Commission's occupational classification systems, ISCO-88 and ISCO-08. Occupational variables for the PSID sample are drawn from the three- and four-digit US Census codes. It should be noted that, as these classification system have changed multiple times, some categories are fairly general to ensure compatibility across years and classifications. Finally, to account for labor market conditions women and mothers face in the various European countries – which often differ quite sharply – I include a number of macroeconomic variables: the ratio of female to male labor force participation; the number of weeks of paid maternity and paternity leave mandated by a country's laws; and the enrolment in childcare⁴ for children 0-3 years old. While, again, it would be preferable to account for these macroeconomic variables in my regressions using PSID data as well, their inclusion reduces the size of an already small sample by around a third. However, the inclusion or omission of these variables affects the significance of coefficients only for regressions looking at probability of employment.

4 Descriptive Statistics

I now turn to a selection of descriptive statistics. The regression sample consists of women age 16 and over, with a positive income ('income' being defined in both the EU-SILC and PSID surveys as cash or near-cash income resulting from labor), adjusted for inflation, to control for the tendencies for a) nominal incomes to rise and b) more countries to introduce LGBT+ marriage rights as time goes on. The first three columns in Table 1.2 therefore refer only to women with a positive income. The exception is *In labor market*, for which it makes little sense to look only at women with an income, given that income almost always requires the individual to be employed (and thus in the labor market).

⁴That is, nursery and preschool.

		(N)	Yearly income $(\$ \neq)$	Hrs worked /week	Hrs housework /week	In labor market (%)
	Straight non-mothers	(58,927)	23,964	34.3	-	63
EU-SILC	Straight mothers	(144, 689)	$22,\!245$	32.4	-	72
	Lesbian non-mothers	(1,213)	25,581	36.5	-	73
	Lesbian mothers	(1,731)	$14,\!310$	37.5	-	76
	Straight non-mothers	(11,072)	22,363	19.2	12.1	78
PSID	Straight mothers	(52,743)	$18,\!158$	18.7	17.6	67
	Lesbian non-mothers	(94)	15,753	8.9	10.1	90
	Lesbian mothers	(90)	17,003	15.6	17	81

Table 1.2: Descriptive means by motherhood status

These descriptive statistics highlight both similarities and differences. In the EU-SILC sample, mothers have lower income than non-mothers, and this is true for both straight and lesbian women. In contrast, in the PSID sample, LGBT+ women differ sharply from straight women in one obvious respect: lesbian mothers have higher average yearly labor income than non-mothers. In both EU-SILC and PSID data, straight mothers reduce their working hours relative to non-mothers, while lesbian mothers work longer hours than lesbian non-mothers, suggesting a fundamental difference in how these two groups of women react to motherhood that is reflected in the results of the empirical estimation.

In the PSID data, both straight and lesbian mothers dedicate more time to household labor than non-mothers – but there is one key difference, which is that straight women have a male partner. While fatherhood is not the focus of this paper, it may be illuminating to note that straight fathers work about an hour more per week than non-fathers, but dedicate almost identical amounts of time to housework (specifically, straight fathers dedicate an average of 24 minutes more per week to housework than non-fathers, in contrast with an over five hour difference between straight mothers and non-mothers). Finally, in the EU-SILC data, mothers are slightly more likely to be in the labor market, but in PSID data, mothers are *less* attached to the labor market – though this is perhaps unsurprising for a country with no federally mandated maternity leave and weaker labor protection.

The EU-SILC sample contains a total of 67,145 observations corresponding to country-

years with marriage equality, and 139,415 observations without. Indeed, there is some evidence that lesbian women's willingness to openly identify is contingent on the state of LGBT+ marriage laws: although observations with marriage equality make up just just 33% of the EU-SILC sample in total, 45% of lesbian women observations are from country-years with marriage equality. Lesbian women in a relationship make up around 2% of the EU-SILC sample in country-years with marriage equality, and 1% of the sample in country-years without. This figure is consistent with that found in most recent studies of LGBT+ demographics in European countries, which place the proportion of the population that is lesbian at around $1-2\%^{5}$. For the PSID sample, the ratios are even more extreme: just 5% of the observations in the sample are for states and years with marriage equality (2,967 observations, compared to 59,824), but over 12% of lesbian women observations are found here.

Self-selection both into motherhood and into an LGBT+ relationship undoubtedly play a large role in determining wages in and of themselves, and LGBT+ identity may also have an effect on the evolution of a given woman's income on becoming a parent. For this reason, I include a simple descriptive probit looking at the association between LGBT+ identification and marriage equality laws, income, education, etc. As entering a relationship with a person of the same gender is the mechanism by which I identify queer women, for this descriptive regression I use the entire sample of women in the EU-SILC and PSID, single or not.

 $^{^5 \}mathrm{See}$ e.g. "Integrated Household Survey (Experimental statistics): January to December 2014". ons.gov.uk.

	EU-SILC		\mathbf{PS}	ID	
	Coef.	Std. Err.	Coef.	Std. Err.	
Marriage equality	0.303***	(0.012)	0.193**	(0.067)	
Employed	-0.076**	(0.031)	-0.174^{**}	(0.064)	
$\log(\text{Income})$	-0.070***	(0.006)	0.007	(0.016)	
High school	-0.411***	(0.016)	0.347^{**}	(0.111)	
Post-secondary education	-0.487***	(0.048)	-	-	
Tertiary education	-0.243***	(0.015)	0.436^{***}	(0.109)	
Post-Graduate education	-	-	0.658^{***}	(0.113)	
Black	-	-	-0.197***	(0.045)	
Asian	-	-	-0.612***	(0.312)	
Age	-0.001	(0.005)	-0.050***	(0.009)	
Age2	-0.0000	(0.000)	0.0004^{**}	(0.000)	
Urban	0.117***	(0.013)	0.116^{***}	(0.044)	
Constant	-1.337***	(0.113)	-1.672^{***}	(0.223)	
N	313,729		67,914		
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001					

Table 1.3: Probit - Lesbian self-identification

Marriage equality, unsurprisingly, carries a positive and significant coefficient for both EU and US samples. The reason for this positive association is twofold: firstly, as marriage equality laws stem from changing attitudes towards homosexuality, the presence of these laws serves as a proxy for an overall cultural environment that is more accepting of LGBT+ individuals. Secondly, marriage equality enables samesex couples to access a raft of privileges, including tax benefits, next-of-kin status, etc. – a powerful inducement to formally identify a same-sex partner where one exists.

Most other coefficients are of the expected sign: urban residence is associated with a slightly increased willingness to identify a same-sex partner both in Europe and the U.S., while employment and income are overall negatively associated. Curiously, education is associated with a higher propensity to self-identify as queer in the US sample, but a lower propensity in the EU one – though the size of the coefficient decreases with level of education. This may reflect different attitudes towards (and acquisition rates of) tertiary education: not only are these rates lower across the board in the EU-SILC sample, there is also a much more pronounced difference between straight and lesbian women in the US than there is in the EU. While in the EU there's just a two-percentage-point difference (35%) of lesbian women have a college degree, as opposed to 37% of straight women), in the US the gap is a whopping eleven percentage points: 55% of lesbian women have a college degree, as compared to 44% of straight women (though this gap is then reversed for postgraduate education, where straight women gain the lead again).

5 Estimation and Results

5.1 Main results

Table 1.4 contains the main results for EU-SILC data; Table 1.5 presents the main results for PSID data.

In both the EU and US, straight mothers earn less on average than non-mothers, though the size of the gap differs significantly: in the EU, mothers earn on average 14% less than non-mothers; in the US this gap is twice as large, at 28%. Another commonality between the two datasets is that the fixed-effects coefficient for straight mothers is far larger than the random-effects coefficient – that is to say, the effects on a given *individual* straight woman's earnings and career path are far larger than the earnings differential between who are mothers and women who are not, when evaluated as a group.

The fact that the fixed-effects motherhood coefficient for straight women is almost twice the size of the random-effects coefficient suggests that women are selecting into motherhood in a way that is strategic: (straight) women who go

	Random Effects		Fixed	Effects	
	Coef.	Std. Err.	Coef.	Std. Err.	
Straight mother	-0.142***	(0.00)	-0.264***	(0.01)	
Lesbian mother	-0.132***	(0.02)	-0.099+	(0.06)	
Secondary education	0.212^{***}	(0.01)	0.024^{*}	(0.01)	
Post-secondary education	0.273^{***}	(0.01)	0.037^{*}	(0.02)	
Tertiary eduction	0.471^{***}	(0.01)	0.065^{***}	(0.02)	
Urban	0.057^{***}	(0.00)	0.003	(0.01)	
Age	0.085^{***}	(0.00)	0.172^{***}	(0.01)	
Age^2	-0.001***	(0.00)	-0.001***	(0.00)	
Labor hours/week	0.019^{***}	(0.00)	0.009^{***}	(0.00)	
Marriage equality	0.012 +	(0.01)	0.019^{*}	(0.01)	
Maternity leave, weeks	-0.006***	(0.00)	-0.003***	(0.00)	
Paternity leave, weeks	0.007^{***}	(0.00)	0.003***	(0.00)	
Ratio of female to male LFP	0.005^{***}	(0.00)	0.010***	(0.00)	
Childcare enrolment 0-3	0.002^{***}	(0.00)	0.001	(0.00)	
Constant	6.641^{***}	(0.08)	3.787^{***}	(0.13)	
Occupational dummies	Y	ES	NO		
Country FE	Y	ES	Ν	10	
Year FE	YES NO		10		
No. of obs.	280,537		$280{,}537$		
R^2	0.395		0.082		
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001					

Table 1.4: Income effects of motherhood – EU-SILC data

on to become mothers are richer on average than those who do not, even though mothers as a group are poorer than non-mothers. This makes intuitive sense – children are, after all, expensive – and is also consistent with explanations that point to household specialization based on relative income as the main cause of the remaining unexplained motherhood earnings penalty. If the amount of household labor each partner does is driven by bargaining power, relatively richer women may be ceteris paribus more likely to select into motherhood, expecting a more equal division of the ensuing labor. We will turn to this explanation in the following subsection.

	Random Effects		Fixed	Effects	
	Coef.	Std. Err.	Coef.	Std. Err.	
Straight mother	-0.282***	(0.02)	-0.435***	(0.03)	
Lesbian mother	-0.116	(0.18)	0.046	(0.21)	
Black	-0.003	(0.02)			
American-Indian	-0.022	(0.06)			
Asian	0.072	(0.07)			
High school	0.312***	(0.03)	0.032	(0.07)	
College	0.529***	(0.03)	0.136 +	(0.08)	
Post-grad	0.785***	(0.04)	0.263^{**}	(0.09)	
Age	0.101***	(0.01)	0.142^{***}	(0.01)	
Age 2	-0.001***	(0.00)	-0.001***	(0.00)	
Labor hours/week	0.014***	(0.00)	0.007^{***}	(0.00)	
Housework hours/week	-0.013***	(0.00)	-0.013***	(0.00)	
Marriage equality	-0.033	(0.03)	-0.065*	(0.03)	
Constant	5.874***	(0.13)	6.474^{***}	(0.14)	
Occupational dummies	Y	ES	NO		
State FE	Y	ES	NO		
Year FE	Y	ES	Ν	10	
No. of obs.	63	,660	63,999		
R^2	0.319		0.	080	
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001					

Table 1.5: Income effects of motherhood – PSID data

Lesbian women differ substantially from their straight counterparts in both the EU and US data as well: while in EU data the earnings gap between lesbian mothers and non-mothers is about the same as the earnings gap between straight mothers and non-mothers, the fixed effects coefficient on motherhood for lesbian women is *smaller* than the random-effects coefficient, and is no longer significant at the 5% level. In the US data, both random-effects and fixed-effects coefficients are insignificant. Motherhood itself appears to have no effect on individual lesbians' incomes in either dataset, and in the US, lesbian mothers do not earn less than lesbian non-mothers on average. This lack of a significant (and large) earnings penalty of motherhood for lesbian women may contribute to the overall lesbian earnings premium found by some researchers, but there is no evidence here of a specific lesbian *motherhood* premium, as some early research had suggested (Baumle, 2009).

5.2 Primary vs. Secondary Earners

In her 1993 book "Working Women Don't Have Wives", Terri Apter interviewed over one hundred working women and found that the vast majority, even those who fit the stereotype of a "career woman", dedicated significant time and energy to their families. Apter argued that the main cause of the persistent gap in labor market outcomes between men and women is that working women largely do not have anyone to take care of domestic affairs while they work: much the opposite, they take on the lion's share of household work in addition to whatever career they may have (see e.g. Bittman et al., 2003). This, then, has important implications for lesbian couples: if female same-sex couples specialize in household labor, and if specialization into housework is a part of the remaining unexplained motherhood earnings penalty for straight women, then simply looking at the motherhood earnings effects of lesbian women as a group will obfuscate the full picture.

For this reason, in this and the following subsection, I examine possible household specialization from two angles. First I evaluate the hypothesis that household specialization is driven by relative bargaining power, and that the lower-earning partner – statistically likely to be female in an opposite-sex couple – decreases their attachment to the labor market in order to take on the additional household labor that having a child entails. In Subsection 5.3, I use data on actual self-reported time spent on housework, to evaluate the possibility that the income effects of mother-hood *are* driven by household specialization, but that the specialization itself is not driven by relative income.

I define as "primary earners" those individuals who earned more than their spouse or partner in the first year of the dataset in which both members of the couple appear. I also define as "primary earners" individuals with an income whose partner's initial income is either 0 or missing. Individuals whose partner earns more than them are defined as "secondary earners", whether they have a positive income or not. I use initial income rather than the average over all years in which the household appears, as the latter would pick up the income effect of individuals having children during the years covered by the survey.

I then interact the primary- and secondary-earner dummies with the motherhood dummy, resulting in four dummy variables total. Column (I) contains the random effects coefficients; column (II) contains the fixed effects coefficients. The abbreviated results are presented below.

	Random Effects		Fixed Effects		
	Coef.	Std. Err.	Coef.	Std. Err.	
Primary earner straight mothers	-0.031***	(0.01)	-0.287***	(0.02)	
Secondary earner straight mothers	-0.217***	(0.00)	-0.260***	(0.01)	
Primary earner lesbian mothers	-0.389***	(0.04)	-0.129	(0.13)	
Secondary earner lesbian mothers	-0.403***	(0.04)	-0.129	(0.13)	
Secondary education	0.211***	(0.01)	0.024^{*}	(0.01)	
Post-secondary education	0.267***	(0.01)	0.038*	(0.02)	
Tertiary education	0.463^{***}	(0.01)	0.066^{***}	(0.02)	
Urban	0.060***	(0.00)	0.003	(0.01)	
Age	0.086***	(0.00)	0.171^{***}	(0.01)	
Age^2	-0.001***	(0.00)	-0.001***	(0.00)	
Labor hours/week	0.019***	(0.00)	0.009^{***}	(0.00)	
Marriage equality	0.011	(0.01)	0.019^{*}	(0.01)	
Maternity leave, weeks	-0.006***	(0.00)	-0.003***	(0.00)	
Paternity leave, weeks	0.007***	(0.00)	0.003^{***}	(0.00)	
Ratio of female to male LFP	0.005***	(0.00)	0.010^{***}	(0.00)	
Childcare enrolment	0.002***	(0.00)	0.001	(0.00)	
Constant	6.670***	(0.08)	3.788^{***}	(0.13)	
Occupation dummies	Y	ES	Ν	IO OI	
Country FE	YES		NO		
Year FE	Y	ES	Ν	10	
No. of obs	280),537	280	,537	
R^2	0.	434	0.	082	
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001					

Table 1.6: Primary vs. secondary earners – EU-SILC data

It is important to note that primary- and secondary-earner status is only a proxy for division of household labor – a variable not available in the main EU-SILC dataset. For heterosexual couples, research has indicated that, while women's share of household labor declines as she nears parity of income with her husband, it increases again as her income surpasses his (Bertrand et al., 2015). It is therefore not particularly surprising that the coefficients for primary- and secondary-earner straight mothers are similar, even if motherhood coefficients are (at least in part) a function of individual's allocation of household labor.

	Random Effects		Fixed Effects	
	Coef.	Std. Err.	Coef.	Std. Err.
Primary earner straight mothers	-0.093***	(0.02)	-0.278***	(0.04)
Secondary earner straight mothers	-0.354***	(0.02)	-0.485***	(0.04)
Primary earner lesbian mothers	-0.360	(0.22)	-0.148	(0.23)
Secondary earner lesbian mothers	0.133	(0.24)	0.232	(0.25)
Black	-0.021	(0.02)		
American-Indian	-0.023	(0.05)		
Asian	0.063	(0.07)		
High school	0.316^{***}	(0.03)	0.029	(0.07)
College	0.527^{***}	(0.03)	0.129 +	(0.08)
Post-grad	0.778^{***}	(0.04)	0.256^{**}	(0.09)
Age	0.101^{***}	(0.01)	0.141^{***}	(0.01)
Age 2	-0.001***	(0.00)	-0.001***	(0.00)
Labor hours/week	0.014^{***}	(0.00)	0.007^{***}	(0.00)
Housework hours/week	-0.012***	(0.00)	-0.013***	(0.00)
Marriage equality	-0.032	(0.03)	-0.065*	(0.03)
Constant	5.880^{***}	(0.13)	6.505^{***}	(0.14)
Occupation dummies	YES		NO	
State FE	YES		NO	
Year FE	YES		NO	
No. of obs	63,660		63,999	
R^2	0.434		0.081	
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001				

Table 1.7: Primary vs. secondary earners – PSID data

Interestingly, in both samples – but especially the EU data – primary-earner straight mothers earn almost the same as primary-earner non-mothers when we look at group averages. However, when we turn to look at the fixed-effects regression we see that, in the EU, primary- and secondary-earner women experience an almost identical motherhood earnings penalty – but in the US, secondary earner straight women experience a motherhood earnings penalty almost twice as large as primary earner straight women. These results would appear to suggest that "breadwinner" status does have some sort of effect on the size of the motherhood penalty, but reject the idea that specialization in work on the labor market or in the household is entirely, or even primarily, driven by relative income or bargaining power.

Unfortunately, splitting the sample of lesbians twice results in a very small sample

size in US data. This makes inference a challenge in and of itself, but the use of income to split sub-samples generates further difficulties. One such difficulty is that assigning primary- and secondary-earner status in this way necessarily introduces noise: for instance, a partner that usually makes less money may have had a windfall in the first year in which the couple appears in the data, resulting in both members of the couple being misclassified. In addition to this sort of stochastic noise, which member of a couple earns more in any given year is overall much less stable for homosexual couples than it is for heterosexual ones, making the small sample size an even bigger issue.

5.3 Primary vs. Secondary House-workers

This regression is, as stated previously, unfortunately only possible with the PSID data. I assign the categories of primary and secondary house-worker depending on which partner reports spending more weekly hours on average on household labor. As time spent on household labor is itself affected by the presence of a child in the household, I include only those years prior to the birth of the couple's first child; however, as time spent on household labor is far more variable than individual income, I instead use the average for the entire period where a couple cohabited without a child, rather than taking just the values for the first year in which a couple appears.

This is where we find the first evidence of a possible motherhood earnings premium for lesbian women – one that does indeed appear to be driven by specialization in housework. Lesbian women who took on more of the household labor prior to becoming mothers experience a negative coefficient of motherhood, though this coefficient is significant at the 5% level only in the random-effects regression. Lesbian women whose partners did a majority of the household labor – those working women who have "wives", in the sense of the word that Apter uses, *do* have a positive and significant fixed effects coefficient of motherhood. These findings are, however, only indicative, due to the small size of the sample.

	Random Effects		Fixed Effects	
	Coef.	Std. Err.	Coef.	Std. Err.
Primary house-worker straight mothers	-0.305***	(0.02)	-0.460***	(0.03)
Secondary house-worker straight mothers	-0.136***	(0.03)	-0.243***	(0.06)
Primary house-worker lesbian mothers	-0.390*	(0.16)	-0.219	(0.23)
Secondary house-worker lesbian mothers	0.416 +	(0.23)	0.430^{*}	(0.18)
Black	-0.010	(0.02)		
American-Indian	-0.024	(0.05)		
Asian	0.069	(0.07)		
High school	0.310***	(0.03)	0.029	(0.07)
College	0.526^{***}	(0.03)	0.131 +	(0.08)
Post-grad	0.780***	(0.04)	0.255^{**}	(0.09)
Age	0.101***	(0.01)	0.141^{***}	(0.01)
Age ²	-0.001***	(0.00)	-0.001***	(0.00)
Labor hours/week	0.014***	(0.00)	0.007^{***}	(0.00)
Housework hours/week	-0.012***	(0.00)	-0.013***	(0.00)
Marriage equality	-0.031	(0.03)	-0.064*	(0.03)
Constant	5.875^{***}	(0.13)	6.500^{***}	(0.14)
Occupation dummies	YES		NO	
State FE	YES		NO	
Year FE	YES		NO	
No. of obs	63,660		63,999	
R^2	0.424		0.080	
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001				

Table 1.8: Primary vs. secondary house-workers - PSID data

Straight women see a significant earnings penalty of motherhood, and accounting for primary or secondary house-worker status does not eliminate this negative coefficient. Nevertheless, being the person who does the least housework within a couple does decrease the size of this coefficient. Of course, this risks obfuscating an important fact discussed in Section 2 which is that straight women are far more likely to be primary house-workers than not. In this regression sample, 46,901 straight women are primary house-workers, and only 5,480 - 12% - do less housework than their male partners.

5.4 Effects on Employment

Finally, while the bulk of this paper has analysed the earnings penalty of motherhood, decreased income is far from the only effect motherhood may have on women's careers. It is worth taking a moment to discuss motherhood's effects on labor force participation (LFP), and not just on income: after all, women who quit their jobs to specialize entirely in household labor after the birth of a child are, in essence, taking a pay cut of 100%.

	Random Effects		Fixed Effects		
	Coef.	Std. Err.	Coef.	Std. Err.	
Straight mother	0.004***	(0.00)	-0.011***	(0.00)	
Lesbian mother	-0.000	(0.01)	-0.014	(0.02)	
Secondary education	0.040***	(0.00)	0.006 +	(0.00)	
Post-secondary education	0.051^{***}	(0.00)	0.011^{*}	(0.00)	
Tertiary eduction	0.068^{***}	(0.00)	0.004	(0.00)	
Urban residence	-0.004***	(0.00)	-0.002	(0.00)	
Age	0.002^{***}	(0.00)	-0.003*	(0.00)	
Age^2	-0.000*	(0.00)	-0.000+	(0.00)	
Marriage equality	-0.014***	(0.00)	0.007^{***}	(0.00)	
Maternity leave, weeks	-0.001***	(0.00)	0.000+	(0.00)	
Paternity leave, weeks	0.002***	(0.00)	-0.001***	(0.00)	
Ratio of female to male LFP	-0.008***	(0.00)	-0.003***	(0.00)	
Childcare enrolment	0.002***	(0.00)	0.001^{***}	(0.00)	
Constant	1.549^{***}	(0.02)	1.401^{***}	(0.04)	
Occupation dummies	YES		NO		
Country FE	YES		NO		
Year FE	YES		NO		
No. of obs	356,690		$356,\!690$		
R^2	0.408		0.082		
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001					

Table 1.9: LFP effects of motherhood – EU-SILC data

Here the differences between EU and US data are perhaps at their most pronounced. In the EU data, mothers are very slightly more attached to the labor force than nonmothers, and motherhood has a minimal effect on the likelihood of being in the labor force for a given straight woman, reducing her probability of being in the labor force by just over a single percentage point. For lesbian women both the random- and fixed-effects coefficients are both insignificant and near zero.

In the US data, on the other hand, the effect that motherhood has on women's

labor force participation is large and significant. The fixed-effects coefficients, in particular, are far larger in the US data than they are for the EU. Straight mothers are just over 2 percentage points less likely to be in the labor force, and becoming a mother reduces a given straight woman's likelihood of being in the labor force by 18 percentage points. While lesbian mothers are about equally attached to the labor force as lesbian non-mothers, as a group, motherhood decreases a given lesbian woman's likelihood by over 25 percentage points; notably, the fixed-effects straight mother coefficient is not statistically different from the fixed-effects lesbian mother coefficient. If motherhood reduces lesbian women's likelihood of being in the labor force, it appears to be only to the extent that it affects women as a whole.

	Random Effects		Fixed Effects	
	Coef.	Std. Err.	Coef.	Std. Err.
Straight mother	-0.022***	(0.00)	-0.180***	(0.01)
Lesbian mother	0.064	(0.05)	-0.257^{***}	(0.06)
Black	-0.008+	(0.00)		
American-Indian	-0.009	(0.01)		
Asian	0.005	(0.01)		
High school	0.027***	(0.01)	0.034	(0.02)
College	0.047^{***}	(0.01)	0.061^{*}	(0.03)
Post-grad	0.080***	(0.01)	0.101^{***}	(0.03)
Age	-0.001***	(0.00)	0.038^{***}	(0.00)
Age 2	0.000**	(0.00)	-0.000***	(0.00)
Housework hours/week	-0.003***	(0.00)	-0.007***	(0.00)
Marriage equality	-0.012+	(0.01)	-0.030***	(0.01)
Constant	0.173^{***}	(0.02)	0.222^{***}	(0.04)
Occupation dummies	YES		NO	
State FE	YES		NO	
Year FE	YES		NO	
No. of obs	103,980		104,435	
R^2	0.735		0.100	
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001				

Table 1.10: LFP effects of motherhood – PSID data

While delving into what drives women's decision to enter or leave the labor force is beyond the scope of this paper, it bears remarking that the EU-SILC results are broadly consistent with research that has cast doubt on the efficacy of maternity leave in increasing women's labor force participation. If these policies make having a child more attractive to women, more women may become mothers, resulting in more women taking breaks – however temporary – from the labor market, resulting in the negative coefficients we see on maternity and paternity leave.⁶ Childcare enrolment, in contrast, carries an unambiguously positive and significant coefficient in both random- and fixed-effects. Evaluated as a whole, the evidence in this section suggests that it is important not only to ensure women can leave the labor force for childbirth and later return in *theory*, but that they have the ability to do so in *practice* as well.

6 Conclusions and Further Research

The fact that mothers on average earn less money than women without children has been documented by a wide body of literature, but the causes and the evolution of this earnings gap over time have been a matter of some intrigue. With the initial gap in human capital falling away, and the unexplained portion of it increasing over the past three to four decades, specialization in household labor has emerged as one of the leading hypotheses explaining its persistence.

Using panel data on fourteen European countries from the EU Standards of Income and Living Conditions (EU-SILC), as well as US data from the Panel Study of Income Dynamics (PSID), and a variety of relevant outcome variables, I have attempted to delve into the question in greater detail. I reproduce the common finding in the literature of a significant earnings penalty for straight women on becoming mothers. This penalty is small but significant for the EU countries, and much larger in the US. Straight women appear to partially compensate for this penalty, as the group-level difference in incomes between straight mothers and non-mothers is smaller than the difference in a given straight woman's income before and after motherhood. Lesbian women, on the other hand, do not appear to experience any significant income effects of motherhood whatsoever, in either sample.

Being the primary earner in a couple – which prima facie should yield greater incentives to specialize in market labor, and translate to a greater bargaining power when dividing household labor – nevertheless does not remove the income penalty of motherhood experienced by straight women. For lesbian women, primary or sec-

⁶However, it should be noted that all European countries in this dataset have at least *some* mandated maternity leave, and the effects on labor force participation of increasing the duration of parental leave may not necessarily be the same as the effects of introducing parental leave where there previously was none.
ondary earner status appears largely irrelevant (and the signs of the coefficients run counter to the bargaining power hypothesis in the random-effects estimation for US data). These results would appear to confirm that specialization in household labor due to lower relative bargaining power is not the primary driver of the (straight) motherhood earnings penalty, as controlling for higher-earner status fails to eliminate the motherhood earnings penalty. Neither does controlling for primary or secondary house-worker status eliminate the straight motherhood earnings penalty, though both primary earner and secondary house-worker status reduce it; however, there is suggestive evidence that specialization in household versus market labor may drive the income effects of motherhood for lesbian women.

The impact that becoming a parent has on women's careers does not stop only at income, however: interestingly, motherhood appears to affect straight and lesbian women's probability of being in the labor force to a very similar degree. The greatest difference here is between samples: in the EU, the coefficients are significant but close to zero in all cases, while in the US, motherhood has a large and significant negative effect on women's labor force participation. Furthermore, the clearest positive effect is of enrolment in childcare, which speaks to the importance of providing women with accessible, practical options for returning to work after parenthood.

Much work, of course, remains to be done. The motherhood penalty is still a very real issue affecting women's careers, and is still not completely understood. Comparing straight and lesbian couples has begun to reveal the ways in which it is driven by the complex nexus of gender, sexual orientation, and the cultural context in which the decision to become a mother occurs. The results of this study serve to underline two important matters: first, that there is significant scope for improvement in data quality, and in particular that questions on gender and sexual identity must be included in household surveys; and second, that far more explicit attention must be paid to the role of social and gender norms.

Bibliography

- Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2011). Inter- and Intra-Household Earnings Differentials among Homosexual and Heterosexual Couples. British Journal of Industrial Relations, 49, s258–s278.
- Ahmed, A. M., Andersson, L., & Hammarstedt, M. (2013). Sexual orientation and full-time monthly earnings, by public and private sector: Evidence from swedish register data. *Review of Economics of the Household*, 11, 83–108.
- 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.
- Aksoy, C. G., Carpenter, C. S., Frank, J., & Huffman, M. L. (2019). Gay glass ceilings: Sexual orientation and workplace authority in the uk. *Journal of Economic Behavior & Organization*, 159, 167–180.
- Anderson, D. J., Binder, M., & Krause, K. (2003). The Motherhood Wage Penalty Revisited: Experience, Heterogeneity, Work Effort, and Work-Schedule Flexibility. *ILR Review*, 56(2), 273–294.
- Andresen, M. E., & Nix, E. (2022). What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples. Journal of Labor Economics, 40(4), 971– 1004.
- Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the Gender Gap in Pay. Journal of Labor Economics, 34(3), 545–579.
- Antecol, H., Jong, A., & Steinberger, M. D. (2008). The Sexual Orientation Wage Gap: The Role of Occupational Sorting and Human Capital. *ILR Review*, 61(4), 518–543.
- Antecol, H., & Steinberger, M. D. (2013). Labor supply differences between married heterosexual women and partnered lesbians: A semi-parametric decomposition approach. *Economic Inquiry*, 51(1), 783–805.
- Apter, T. (1993). Working Women Don't Have Wives: Professional Success in the 1990s. St. Martin's Press.

- Arabsheibani, G. R., Marin, A., & Wadsworth, J. (2004). In the pink: Homosexualheterosexual wage differentials in the uk. *International Journal of Manpower*, 25(3-4), 343–354.
- Badgett, M. L. (1995). The Wage Effects of Sexual Orientation Discrimination. ILR Review, 48(4), 726–739.
- Badgett, M. L. (2018). Left Out? Lesbian, Gay, and Bisexual Poverty in the U.S. Population Research and Policy Review, 37(5), 667–702.
- Badgett, M. L., Carpenter, C. S., & Sansone, D. (2021). LGBTQ Economics. Journal of Economic Perspectives, 35(2), 141–170.
- Baert, S. (2014). Career lesbians. getting hired for not having kids? Industrial Relations Journal, 45(6), 543–561.
- Baumle, A. K. (2009). The Cost of Parenthood: Unraveling the Effects of Sexual Orientation and Gender on Income. Social Science Quarterly, 90(4), 983– 1002.
- Becker, G. S. (1965). A Theory of the Allocation of Time. The Economic Journal, 75(299), 493–517.
- Becker, G. S. (1991). A Treatise on the Family. Cambridge, MA: Harvard University Press.
- Becker, S. (1957). *The Economics of Discrimination*. The University of Chicago Press.
- Benard, S., Paik, I., & Correll, S. J. (2007). Cognitive Bias and the Motherhood Penalty. Hastings LJ, 59, 1359.
- Berg, N., & Lien, D. (2002). Measuring the effect of sexual orientation on income: Evidence of discrimination? Contemporary Economic Policy, 20(4), 394–414.
- Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender Identity and Relative Income within Households. The Quarterly Journal of Economics, 130(2), 571–614.
- Bittman, M., England, P., Sayer, L., Folbre, N., & Matheson, G. (2003). When Does Gender Trump Money? Bargaining and Time in Household Work. American Journal of Sociology, 109(1), 186–214.
- Black, D. A., Makar, H. R., Sanders, S. G., & Taylor, L. J. (2003). The Earnings Effects of Sexual Orientation. *ILR Review*, 56(3), 449–469.
- Blandford, J. M. (2003). The Nexus of Sexual Orientation and Gender in the Determination of Earnings. *ILR Review*, 56(4), 622–642.
- Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3), 789–865.

- Budig, M. J., & England, P. (2001). The Wage Penalty for Motherhood. American Sociological Review, 204–225.
- Budig, M. J., & Hodges, M. J. (2010). Differences in Disadvantage: Variation in the Motherhood Penalty across White Women's Earnings Distribution. American Sociological Review, 75(5), 705–728.
- Budig, M. J., Misra, J., & Boeckmann, I. (2012). The Motherhood Penalty in Cross-National Perspective: The Importance of Work–Family Policies and Cultural Attitudes. Social Politics, 19(2), 163–193.
- Carpenter, C. S. (2005). Self-Reported Sexual Orientation and Earnings: Evidence from California. *ILR Review*, 58(2), 258–273.
- Coffman, K. B., Coffman, L. C., & Marzilli Ericson, K. M. (2017). The Size of the LGBT Population and the Magnitude of Antigay Sentiment Are Substantially Underestimated. *Management Science*, 63(10), 3168–3186.
- Crittenden, A. (2002). The Price of Motherhood: Why the Most Important Job in the World is Still the Least Valued. Henry Holt; Company.
- Del Río, C., & Alonso-Villar, O. (2019). Occupational segregation by sexual orientation in the U.S.: exploring its economic effects on same-sex couples. *Review* of Economics of the Household, 17, 439–467.
- de Quinto, A., Hospido, L., & Sanz, C. (2021). The child penalty: Evidence from spain. SERIEs, 12, 585–606.
- de Vries, L., & Steinmetz, S. (2023). Sexual Orientation, Workplace Authority and Occupational Segregation: Evidence from Germany. Work, Employment and Society, 09500170231158513.
- Downs, B., Foster, L., Nesbit, R., & Sandler, D. H. (2023). Same-Sex Couples and the Child Earnings Penalty (Working Papers No. 23-25). Center for Economic Studies, U.S. Census Bureau. https://ideas.repec.org/p/cen/wpaper/23-25.html
- Drydakis, N. (2015). Sexual orientation discrimination in the united kingdom's labour market: A field experiment. *Human Relations*, 68(11), 1769–1796.
- Fasoli, F., Maass, A., Paladino, M. P., & Sulpizio, S. (2017). Gay- and Lesbian-Sounding Auditory Cues Elicit Stereotyping and Discrimination. Archives of Sexual Behavior, 46, 1261–1277.
- Gangl, M., & Ziefle, A. (2009). Motherhood, Labor Force Behavior, and Women's Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany, and the United States. *Demography*, 46, 341–369.

- Gough, M., & Noonan, M. (2013). A Review of the Motherhood Wage Penalty in the United States. Sociology Compass, 7(4), 328–342.
- Harkness, S. E. (2016). The Effect of Motherhood and Lone Motherhood on the Employment and Earnings of British Women: A Lifecycle Approach. European Sociological Review, 32(6), 850–863.
- Jee, E., Misra, J., & Murray-Close, M. (2019). Motherhood Penalties in the U.S., 1986–2014. Journal of Marriage and Family, 81(2), 434–449.
- Jepsen, C., & Jepsen, L. K. (2015). Labor-Market Specialization within Same-Sex and Difference-Sex Couples. *Industrial Relations: A Journal of Economy and Society*, 54(1), 109–130.
- Jepsen, L. K. (2007). Comparing the Earnings of Cohabiting Lesbians, Cohabiting Heterosexual Women, and Married Women: Evidence from the 2000 Census. Industrial Relations: A Journal of Economy and Society, 46(4), 699–727.
- Killewald, A., & Gough, M. (2013). Does Specialization Explain Marriage Penalties and Premiums? American Sociological Review, 78(3), 477–502. https://doi. org/10.1177/0003122413484151
- Klawitter, M. (2015). Meta-Analysis of the Effects of Sexual Orientation on Earnings[†]. Industrial Relations: A Journal of Economy and Society, 54(1), 4– 32.
- Klawitter, M. M., & Flatt, V. (1998). The effects of state and local antidiscrimination policies on earnings for gays and lesbians. *Journal of Policy Analysis and Management*, 17(4), 658–686.
- Kleven, H., Landais, C., & Søgaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. American Economic Journal: Applied Economics, 11(4), 181–209.
- Leppel, K. (2009). Labour Force Status and Sexual Orientation. *Economica*, 76(301), 197–207.
- Mandel, H., & Semyonov, M. (2006). A Welfare State Paradox: State Interventions and Women's Employment Opportunities in 22 Countries. American Journal of Sociology, 111(6), 1910–1949.
- Martell, M. E. (2021). Labor market differentials estimated with researcher-inferred and self-identified sexual orientation. *Economics Letters*, 205, 109959.
- Martell, M. E., & Eschelbach Hansen, M. (2017). Sexual identity and the lesbian earnings differential in the u.s. *Review of Social Economy*, 75(2), 159–180.
- Martell, M. E., & Roncolato, L. (2016). The homosexual lifestyle: Time use in samesex households. Journal of Demographic Economics, 82(4), 365–398.

- Montag, J. (2015). What Drives the Gender Gap? An Analysis Using Sexual Orientation. Kyklos, 68(4), 577–608.
- Mourelatos, E. (2023). Does mood affect sexual and gender discrimination in hiring choices? evidence from online experiments.
- Niedlich, C., Steffens, M. C., Krause, J., Settke, E., & Ebert, I. D. (2015). Ironic Effects of Sexual Minority Group Membership: Are Lesbians Less Susceptible to Invoking Negative Female Stereotypes than Heterosexual Women? Archives of Sexual Behavior, 44, 1439–1447.
- Peplau, L. A., & Fingerhut, A. (2004). The Paradox of the Lesbian Worker. Journal of Social Issues, 60(4), 719–735.
- Plug, E., & Berkhout, P. (2004). Effects of sexual preferences on earnings in the netherlands. Journal of Population Economics, 17, 117–131.
- Tebaldi, E., & Elmslie, B. (2006). Sexual orientation and labour supply. Applied Economics, 38(5), 549–562.
- Waldfogel, J. (1998). Understanding the "Family Gap" in Pay for Women with Children. Journal of Economic Perspectives, 12(1), 137–156.
- Weeden, K. A. (2005). Is there a flexiglass ceiling? Flexible work arrangements and wages in the United States. Social Science Research, 34(2), 454–482.

Appendix

1.A Extra Tables

Table 1.A.1:	Full list of EU-SILC variables	

Income	Yearly earnings from salary and wages
log(Income)	Logarithm of yearly earnings from salary and wages
Motherhood	Dummy, $= 1$ if individual has a child
Lesbian	Dummy, $=1$ for lesbian women
Lesbian mother	Interaction of <i>Mother</i> and <i>Lesbian</i>
(LT secondary education)	Dummy, $=1$ if individual did not complete high school
Secondary education	Dummy, $=1$ if individual's highest level of education is high school
Post-secondary education	Dummy, =1 if individual obtained post-secondary non-tertiary education
Tertiary education	Dummy, $=1$ if individual has a college degree or higher
Urban	Dummy, $=1$ if individual lives in a densely populated area
Age	Individual's age
Age^2	
Labor hours/week	No. of hours normally worked by the individual per week
Partner's hours	No. of hours normally worked by the individual's partner per week
Government	Dummy, $=1$ for public servants and heads of charities
Manager	Dummy, $=1$ for managerial jobs
Leisure	Dummy, $=1$ for hoteliers and other leisure-related jobs
STEM	Dummy, $=1$ for STEM and related fields
Medical	Dummy, $=1$ for the medical field
Education	Dummy, $=1$ for teachers and other educators
Skilled	Dummy, $=1$ for skilled work including lawyers, consultants, etc.
Technicians	Dummy, $=1$ for technicians and operators
Finance	Dummy, $=1$ for finance-related jobs
Services	Dummy, $=1$ for the service industry
Primary	Dummy, $=1$ for primary sector industries
Secondary	Dummy, $=1$ for secondary sector industries
Clerical	Dummy, $=1$ for office and administrative support occupations
Primary sector	Dummy, $=1$ for farming, fishing and forestry occupations
Construction	Dummy, $=1$ for construction workers
Secondary sector	Dummy, $=1$ for production occupations
Maternity leave	Weeks of full-pay maternity leave available in individual's country-year
Paternity leave	Weeks of full-pay paternity leave available in individual's country-year
Female LFP	Female labor force participation rate
Childcare enrolment	Average childcare enrolment rate $(\%)$ for the country and year

Income	Yearly earnings from salary and wages
$\log(\text{Income})$	Logarithm of yearly earnings from salary and wages
Motherhood	Dummy, $= 1$ if individual has a child
Lesbian	Dummy, $=1$ for lesbian women
Lesbian mother	Interaction of <i>Mother</i> and <i>Lesbian</i>
(LT secondary education)	Dummy, $=1$ if individual did not complete high school
Secondary education	Dummy, $=1$ if individual's highest level of education is high school
College education	Dummy, $=1$ if individual obtained an undergraduate degree
Post-graduate	Dummy, $=1$ if individual has obtained a master's degree or higher
(White)	Dummy, $=1$ if individual is White
Black	Dummy, $=1$ if individual is Black
Asian	Dummy, $=1$ if individual is Asian or Pacific Islander
Age	Individual's age
Age^2	
Labor hours/week	No. of hours normally worked by the individual per week
Housework hours/week	No. of hours individual normally spends on housework per week
AFF	Dummy, $=1$ for agriculture, forestry, and fishing
Construction	Dummy, $=1$ for construction industry
Manufacturing	Dummy, $=1$ for manufacturing jobs
Transport	Dummy, $=1$ for the transportation sector
Media	Dummy, $=1$ for entertainment and media
Utilities	Dummy, $=1$ for the utilities sector
Wholesale	Dummy, $=1$ for wholesale retail
Finance	Dummy, $=1$ for finance-related jobs
Real estate	Dummy, $=1$ for the real estate sector
Business	Dummy, $=1$ for business owners, managers, etc.
Security	Dummy, $=1$ for security and protective services
Repair	Dummy, $=1$ for technicians, maintenance and repair
Services	Dummy, $=1$ for the service industry
Leisure	Dummy, $=1$ for hoteliers and other leisure-related jobs
Medical	Dummy, $=1$ for the medical field
Law	Dummy, $=1$ for lawyers, paralegal assistants, etc.
Education	Dummy, $=1$ for teachers and other educators
Professional	Dummy, $=1$ for other white-collar jobs
Public sector	Dummy, $=1$ for jobs in the public sector

Table 1.A.2: Full list of PSID variables

1.B Primary vs. Secondary Workers

The tables in this section present the results of regressions splitting mothers into primary and secondary workers within the couple. I assign the categories of primary and secondary worker based on which partner reports spending more

	Random Effects		Fixed Effects		
	Coef.	Std. Err.	Coef.	Std. Err.	
Primary worker	0.048***	(0.01)	-0 274***	(0.01)	
straight mothers	0.040	(0.01)	0.214	(0.01)	
Secondary worker	-0.011*	(0.01)	-0.264***	(0.02)	
straight mothers					
Primary worker	-0.002	(0.02)	-0.105	(0.07)	
lesbian mothers		(0:02)		(0.01)	
Secondary worker	0.008	(0.03)	-0.057	(0.07)	
lesbian mothers	0.010***	(0.01)	0.00.1*	(0.01)	
Secondary education	0.213***	(0.01)	0.024*	(0.01)	
Post-secondary education	0.274***	(0.01)	0.037*	(0.02)	
Tertiary eduction	0.472***	(0.01)	0.066^{***}	(0.02)	
Urban	0.059^{***}	(0.00)	0.003	(0.01)	
Age	0.064***	(0.00)	0.171^{***}	(0.01)	
Age^2	-0.001***	(0.00)	-0.001***	(0.00)	
Labor hours/week	0.019***	(0.00)	0.009^{***}	(0.00)	
Marriage equality	0.012 +	(0.01)	0.018^{*}	(0.01)	
Maternity leave, weeks	-0.006***	(0.00)	-0.003***	(0.00)	
Paternity leave, weeks	0.007***	(0.00)	0.003^{***}	(0.00)	
Ratio of female to male LFP	0.005***	(0.00)	0.010^{***}	(0.00)	
Childcare enrolment	0.002***	(0.00)	0.001	(0.00)	
Constant	6.979***	(0.08)	3.790^{***}	(0.13)	
Occupation dummies	YES		NO		
Country FE	YES		NO		
Year FE	YES		NO		
No. of obs	280,537		280,537		
R^2	0.408		0	0.082	
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$	

 Table 1.B.1: Primary vs. secondary workers – EU-SILC data

	Random Effects		Fixed Effects	
	Coef.	Std. Err.	Coef.	Std. Err.
Primary worker straight mothers	-0.063*	(0.03)	-0.272***	(0.05)
Secondary worker straight mothers	-0.272***	(0.02)	-0.422***	(0.03)
Primary worker lesbian mothers	-0.139	(0.32)	0.173	(0.27)
Secondary worker lesbian mothers	-0.200	(0.25)	-0.216	(0.25)
Black	-0.031	(0.02)		
American-Indian	-0.024	(0.06)		
Asian	0.072	(0.07)		
High school	0.326^{***}	(0.03)	0.035	(0.07)
College	0.548^{***}	(0.03)	0.128 +	(0.08)
Post-grad	0.804^{***}	(0.04)	0.248^{**}	(0.09)
Age	0.098^{***}	(0.01)	0.135^{***}	(0.01)
Age ²	-0.001***	(0.00)	-0.001***	(0.00)
Labor hours/week	0.014^{***}	(0.00)	0.007^{***}	(0.00)
Housework hours/week	-0.013***	(0.00)	-0.014***	(0.00)
Marriage equality	-0.031	(0.03)	-0.073*	(0.03)
Constant	5.877***	(0.13)	6.577^{***}	(0.14)
Occupation dummies	YES		NO	
Country FE	YES		NO	
Year FE	YES		NO	
No. of obs	63,660		63,999	
R^2	0.425		0.080	
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$

Table 1.B.2: Primary vs. secondary workers - PSID data

hours on average per week working a job for a wage or salary. Individuals who are unemployed or not on the labor market are assigned a value of zero for this variable. I include only years prior to the birth of the couple's first child, and use the average for the entire period. As these results are qualitatively similar to those found in Sections 5.2 and 5.3, they are included here for reference only.

Chapter 2

Alcohol, Behavioral Norms and Sexual Violence on U.S. College Campuses

Julia Hoefer Martí¹

Paul Seabright¹

Abstract

This paper explores the role of social norms in influencing the incidence of sexual assault, and the contribution of alcohol to such events. We build a decision theoretic model where agents may use alcohol as a "disinhibitor" to undermine social norms discouraging consensual sexual encounters outside marriage. This makes non-consensual encounters more likely. Stronger norms against consensual sex might therefore increase the incidence of non-consensual sex.

We test the theory on data from US college campuses, using the presence of Planned Parenthood clinics in the county as an indicator of norms more accepting of consensual sex. Controlling for other factors, colleges in counties with fewer clinics have more incidents of rape and sexual assault in which alcohol is implicated. Colleges affiliated to the National Collegiate Athletic Association also have more such incidents, suggesting that sporting institutions also act as facilitators of a culture of sexual aggression. We provide suggestive evidence from attitudinal surveys and from campus religious affiliation that disapproval of consensual sex may indeed be involved. We explore rival explanations such as reporting and selection biases.

Keywords: social norms, violence, sexual assault, rape, alcohol

JEL codes: D74, D91, K42, Z13.

Funding from the French ANR (ANR-17-EURE-0010) is gratefully acknowledged. We are grateful for feedback from presentations at IAST in Toulouse, the ERINN Workshop on social norms, and seminars at the University of Frankfurt and University College, London. Slimane Dridi, Guido Friebel, Michael Kosfeld, Stephane Mechoulan and Michael Wilson all made very helpful suggestions. Ellen Stolzenberg and Jared Planas kindly arranged access to HERI data. All errors remain our own.

1 Introduction

What is the role of cultural factors, including social norms, in determining the incidence of sexual assault? The #MeToo movement has brought to the fore questions about the extent to which presuppositions about the acceptability or unacceptability of certain patterns of behavior shape the risks of sexual assault. It is not enough to consider legal rules and the resources invested in enforcement of the law: the expectations of actual and potential victims, perpetrators and enforcement authorities also have a large impact on the probability that individuals will face assaults in any given social and physical environment.

In this paper we examine the interaction between norms against consensual and non-consensual sex. We do so in the context of US college campuses, where there is evidence of a serious and widespread problem of sexual assault. An estimated 20-25% of college-age women are at risk of victimisation over the course of their degree (Fisher et al., 2000; Franklin, 2010). Female college students are three times more likely to experience sexual assault compared to the general population of women, and sexual violence is especially prevalent in colleges, compared to other crimes (Cantor et al., 2015).

A widely held view is that lax norms concerning consensual sex contribute to the weakening of norms against non-consensual sex. For instance, the *New York Times* columnist Ross Douthat, writing in 2014, blamed rape and sexual assault on college campuses on "a fun, even bacchanalian lifestyle... where teens and earlytwentysomethings are barely supervised and held to no standard higher than consent... a hard-drinking, sexually permissive culture" (Douthat, 2014). Similarly, retired Pope Benedict XVI has drawn fire for blaming sexual assault within the Catholic church on the "swinging sixties" (Associated Press, 2019)¹ However, the absence of strong norms against consensual sex is not the same thing as the absence of strong norms against sexual assault, and it is an open question whether the absence of the former has any impact on the presence and strength of the latter. Answering this question is the purpose of our paper.

We find evidence that the relation between norms governing consensual and nonconsensual sex is not as predicted by the "permissiveness" hypothesis, and may even be the opposite. Controlling for other relevant factors, stronger norms against

 $^{^1 {\}rm see}$ https://www.theguardian.com/world/2019/apr/11/ex-pope-benedict-xvi-blames-sexual-abuse-on-swinging-sixties

consensual sex are statistically associated with a *higher* incidence of sexual assault in the presence of alcohol. We find no robust association with incidents not involving alcohol, which would imply a likely increase in overall incidence, though our data are too noisy to support that overall conclusion. While we must be careful in inferring causality, it seems likely that such norms do in fact lead to an increased risk of assault in which alcohol is involved. Why?

The answer appears to lie in the role of alcohol as a "disinhibitor". It is well known that high alcohol consumption is associated with increased incidence of assault in a wide range of contexts, although the nature of the causal mechanism is far from settled. However, alcohol consumption does not strike individuals at random, like the weather. It is a choice, and one reason individuals choose to consume alcohol may be that they wish to weaken the hold that certain norms have over them, especially norms that might lessen their enjoyment of activities such as consensual sex. But weakening the hold of norms is risky - it may not weaken only the norms intended. A predictable consequence of alcohol consumption is to weaken norms against assault.

Our approach is as follows. First, we develop a decision-theoretic model of interactions between potential perpetrators and potential victims of sexual assault. These interactions are governed by an assessment of costs and benefits, and among these costs are the costs of violating various norms, including those governing consensual and non-consensual sex. However, these costs are not immutable; they can be subjectively manipulated by the use of various disinhibiting technologies - most obviously the consumption of alcohol. We assume that perpetrators would prefer to have consensual sex, and it is when their overtures are refused that they are tempted to resort to assault. In the presence of stronger norms against consensual sex there will be greater resort to alcohol as a disinhibitor, and this will lead to an increased incidence of assault.

We then test the model on a comprehensive dataset of sexual assaults and rapes on US college campuses from 1997 to 2019. An important feature of this dataset is that it records whether the victim believed that the perpetrator was under the influence of alcohol at the time of the assault. We find that, controlling for other relevant factors, colleges located in counties with a presence of Planned Parenthood clinics, which are likely to have weaker norms against consensual sex, have lower frequency of both rape and sexual assault in the presence of alcohol. There is suggestive but statistically weaker evidence that the opposite is true of campuses with a religious affiliation, which are likely to have stronger norms against consensual sex. As predicted by the model, these associations are strong for assaults where alcohol is recorded as a contributing factor, and absent or weakly negative for incidents without the involvement of alcohol. Restrictions on campus availability of alcohol do not appear effective in reducing incidents, though endogeneity may be a factor, as campuses with higher incidents of alcohol-fueled sexual assault may be more likely to ban alcohol in response.

Although our data document the use of alcohol as a disinhibitor, the lessons from the study are much wider. Many societies employ, deliberately or not, various institutional mechanisms for enabling individuals to disinhibit themselves from the influence of social norms the society itself imposes. The saying "what happens in Vegas stays in Vegas" indicates that the city of Las Vegas positions itself not just as a place where the *legal* restrictions on commercial sex and on gambling are looser than elsewhere in the US, but also where the moral norms surrounding such activities are looser as well. As a marketing pitch it invites people to choose strategically to visit the city in order to overcome the inhibiting effect of norms and not just the practical restrictions of the law. Even after they return home they can feel entitled to consider "what happened in Vegas" as less of a norm violation than it would have been had the same behavior occurred in another place.

In a different vein, the controversy surrounding the nomination of Brett Kavanaugh to the US Supreme Court in September 2018 cast a less than flattering light on the role of fraternities in many US universities, particularly as an environment in which young men could pursue sexual conquests with many fewer of the inhibiting norms they would encounter during interactions with women in other contexts. Similar things may be true of certain sporting associations on college campuses; we report evidence to that effect below. Many institutions in many countries have historically played such a disinhibiting role, including hazing rituals for newcomers, Hallowe'en and other festivals, and rituals permitting gestures of disrespect for figures of authority that would not be permitted at other times. Our analysis suggests that the effect of strengthening norms cannot be considered in isolation from technologies and institutions that permit people to circumvent those norms. This links to a more general literature on circumstances in which individuals strategically adjust their behavior to take account of the way in which norms will constrain their actions in the future (Ellingsen and Mohlin, 2019). The paper is structured as follows. In Section 2 we review the existing literature on the causes of sexual assault. In Section 3 we set out our model and derive the main empirical predictions. Section 4 describes the data and provides basic descriptive statistics; it also discusses questions about the reliability of the data, and considers to what extent reporting biases may generate spurious results. Section 5 tests our model predictions and considers a range of robustness tests, and Section 6 concludes.

2 Literature Review

2.1 The link between alcohol and violence

That a link between alcohol and sexual violence exists has been the subject of extensive documentation. Experimental evidence has shown men to display heightened sexual aggression after consuming alcohol (Davis, n.d.). This is supported by observational data showing that "party culture" and situations where alcohol is involved result in more cases of rape being reported to the police (Lindo et al., 2018). Further, substance use on the part of the offender is related to higher probability of a completed assault rather than an attempted one, and higher probability of injury to the victim, regardless of whether the victim drank or used drugs themselves (Brecklin and Ullman, 2010).

Several theories have been put forward to explain the exact mechanism by which alcohol results in heightened aggression. Apart from the fact that alcohol is a depressant, and therefore diminishes subjects' general sensitivity to external stimuli, there is evidence that alcohol differentially diminishes sensitivity to pains, including future pains (Gray and McNaughton, 2003, Chapter 4). This provides clinical support for a key feature of our model, which is that alcohol leads individuals to discount future psychological and social costs of their actions, more than it leads them to discount future benefits of their actions.

Expectations matter too. The alcohol expectancy theory suggests that the mere fact that individuals *expect* alcohol consumption to result in more aggression is enough for it to do so. Stappenbeck and Fromme, 2013 hypothesize that alcohol contributes to aggression by a) narrowing focus to most salient cues, b) lowering inhibitions, and c) interfering with the ability to regulate one's emotions. They find that alcohol is indeed significantly related to heightened intentions of physical and

verbal aggression, but are unable to untangle to what extent (if any) expectations play a role. However, a key result of the study is that, in the presence of alcohol, cognitive reappraisal² was effective at reducing intentions of verbal and physical aggression.

Exploring the mechanisms involved matters for understanding the clear statistical link between the availability of alcohol and the incidence of assaults. To echo an argument made in a different context, alcohol does not rape people - rapists do. But this observation does not imply that the ease with which alcohol is available has no impact on the choices made by rapists. In the case of firearms, there is growing awareness that an increase in the incidence of firearms in the population can make it more likely that a particular owner of a firearm will use it to kill someone. As O'Flaherty and Sethi, 2019 put it: "People sometimes kill simply to avoid being killed". Similarly, even if sexual assault results from choice rather than a purely mechanical process, an increase in the ease with which alcohol can be obtained will affect the choices that potential sexual aggressors make.

2.2 Other factors affecting the incidence of sexual violence on college campuses

Institutional policy and campus characteristics play a crucial role in the overall incidence of sexual assault. Among the most influential factors documented to date are the proportion of the student body living on campus, membership of the National Collegiate Athletic Association (NCAA), and campus alcohol policy (Stotzer and MacCartney, 2015). Empirical analyses have also provided some support for the idea that all-male peer groups such as fraternities serve as a disinhibitor, facilitating sexual assault in much the same way as alcohol, and often in conjunction with it (Franklin et al., 2012). Beyond this, colleges often do not provide adequate protection for victims or measures of redress, which are especially crucial given that, of men who self-report acts of rape, a majority report multiple such acts, with an average of around fourteen rapes per offender. Furthermore, "[t]he rate of campus peer sexual violence and the high non-reporting rate perpetuate a cycle whereby perpetrators commit sexual violence because they think they will not get caught or

 $^{^{2}}$ An emotional regulation technique that modifies an emotional response by reinterpreting the event that elicited it, e.g. by reinterpreting a disappointing exam result as an opportunity to challenge or better oneself.

because they actually have not been caught." (Cantalupo, 2011).

There is also evidence documenting attitudes towards casual sex in the U.S. in general, and among U.S. university students in particular. While the proportion of Americans who believe casual sex is "not wrong at all" has been steadily rising – from under 30% in the 80s to almost 50% and in the 2000s, and finally near 60% from 2010 (Twenge et al., 2015), a substantial portion remains which does not approve of "hookup culture". A significant double-standard among college students also exists, with men reporting more openness to casual sex than women, but also more judgmental attitudes toward women than toward men who have casual sex (England and Bearak, 2014). However, more permissive attitudes towards casual sex may not result in more casual sex overall: a study of the 1988-1996 and 2004-2012 waves of the General Social Survey (GSS) found that college students interviewed in the latter wave did not report having more sexual partners, despite this greater permissiveness (Monto and Carey, 2014).

2.3 Reporting Sexual Violence to Police

Although sexual violence is widespread, it is widely regarded as the most underreported crime of all, with estimates of reporting rates generally falling in the 25 - 35% range, even for recent years. Most rapes go unreported; in general, victims of sexual crimes do not report their victimisation via formal channels, and instead disclose to informal support networks, most usually female friends (Orchowski and Gidycz, 2012; Kilpatrick et al., 2007; Paul et al., 2013). Past research has found that rates of reporting to the police are substantially lower when alcohol or drugs were involved (Sabina and Ho, 2014). Conversely, there is a clear and positive association between injury to the victim and whether the incident is reported to the police (Du Mont et al., 2003).

The National Crime Victimization Survey (NCVS), a yearly survey conducted by the Bureau of Justice Statistics, sheds some light on victim underreporting; however, it has also been criticised for underestimating the true extent of victimisation, despite ostensibly being designed to do so. McCauley et al., 2009 show that using behaviorally specific questions (as opposed to simply asking individuals if they have been victims of a sexual assault) leads subjects to disclose higher levels of victimisation.

Most surveys of victims of sexual violence who did *not* report the crime to police tend

to find similar answers when asking why. One of the most commonly cited reasons is fear of reprisal by the perpetrator if the victim and perpetrator are acquainted; however, two other widespread reasons are that the victim "didn't think it was serious enough" or "handled the incident without police involvement" (e.g. Fisher et al., n.d.).

One possible theory put forward to explain this phenomenon is that victims "maintain a rape script that defines expectations for the roles, rules, and events that take place in a rape scenario", and any deviations from this 'script' reduce the likelihood not only of reporting the crime but even of recognizing it as a crime (Zinzow and Thompson, 2011). A study among female inmates who were victims of sexual assault found that, although belief in "rape myths"³ was low overall, women who reported more belief in these myths were less likely to have reported the crime to police (Heath et al., 2013). Denying or trivialising the incident is a psychological self-defense mechanism (Weiss, 2011; Zinzow and Thompson, 2011).

Expectations of the process also play an important part in the decision to report (James and Lee, 2014). Of reported cases, only around 50% result in positive identification of a suspect (Tasca et al., 2012) Frazier and Haney, 1996), and perhaps a third of these result in an arrest (Tasca et al., 2012). Sexual assault cases have lower odds of clearance by arrest, and higher odds of exceptional clearance, than other cases (Spohn and Tellis, 2011).

Here, too, alcohol plays a significant role. Experimental research has shown that the more intoxicated a (theoretical) victim is, the more likely police officers were to disbelieve the victim's version of events relative to the alleged perpetrator, and the less blame they allocated to the perpetrator (Schuller and Stewart, 2000). The relationship between victim and perpetrator also matters: though past research has found that acquaintance rape/sexual assault is more likely to lead to a positive identification of the suspect, controlling for a positive identification of the suspect, cases where the victim and offender were strangers were more likely to lead to an arrest (Bouffard, 2000; Tasca et al., 2012; Lafree, 1981).⁴

³Such as, for example, the belief that most rape is perpetrated by strangers to the victim.

⁴Past research has also attempted to investigate the sexual stratification hypothesis. This hypothesis regards sexual "access" as a commodity which is distributed according to a group's relative power; it argues that police and prosecutors will respond more severely to sexual assault cases with a white victim and black offender, and least seriously to those with a black victim and white offender. Although past research has provided some limited support for this hypothesis (LaFree, 1980; Spohn and Holleran, 2001; Stacey et al., 2016). If indeed the sexual stratification hypothesis was true in the past, its hold appears to have weakened considerably.

Other factors that influence likelihood of clearance via arrest are injury to the victim (crime seriousness), which is associated with higher likelihood of a successful identification and clearance by arrest (Tasca et al., 2012); evidentiary strength (including, for example, whether a rape kit test was conducted); and victim cooperation. Consistent with this hypothesis, cases where the victim resisted the offender have been shown to have higher rates of suspect identification (Horney and Spohn, 2006). Indeed there is reason to believe victims themselves take evidentiary strength into account when deciding whether to report (Fisher et al., 2003).

3 The Model

3.1 Basic set-up

An individual student on a college campus seeks to encounter possible sexual partners. In keeping with the statistical regularity that the majority of perpetrators of sexual assault are male and majority of their victims are female, we will refer to the student as "he" and the potential victim as "she", while acknowledging that assault perpetrators may be females and victims may be males, and that assault is an issue in a homosexual as well as in a heterosexual setting.

In a given time period, the student may be presented with an opportunity to make a sexual proposition to a potential partner. This occurs with some probability p, which may be influenced in various ways by both the student and by third parties, including college authorities. The potential partner may accept the sexual proposition or may refuse. If the partner refuses, the student may choose to accept the refusal, or to override the refusal by assaulting the partner. In this set-up we do not model explicitly the strategic determinants of the partner's decision whether or not to accept the proposition, but assume that this occurs with a certain probability q (which may be affected by a number of factors). Instead we focus on the decision problems of the student, since we are interested in the interaction between the student's decisions when sober and his decisions when drunk.

Various factors may influence how these encounters develop. We can think of these as divided into two main categories: temperament and circumstances. Tempera-

⁵It would be straightforward to make the partner's decision a function of the partner's objectives and constraints, but would not, we believe, bring additional insight.

ment consists of any psychological or other factors that might make two individuals behave differently in similar circumstances, and we summarize these factors in a single type variable $\theta \in \{C, T\}$. We do not claim that θ corresponds to a single psychologically identifiable condition - it is just a statistical aggregate of factors that contribute to different behavior in similar circumstances. C-type ('Calm') individuals will tend to be less likely to commit assaults in any given circumstances than T-type ('Turbulent') individuals.

Circumstances consist of a range of factors, of two broad kinds. There are persistent institutional characteristics (such as the system in place for the investigation and punishment of assault), which may be shaped by actions of the individuals in the model, but which, once chosen, remain unchanged. Then there are chance factors that may alter from one moment to the next.

The various possible outcomes of the student's interaction with others result in costs and benefits. Some of these are standard private costs and benefits. The student has a monetary endowment y out of which he may incur costs, and in particular a cost of alcohol consumption c, consuming the remainder. That cost is a random variable that is uniformly distributed on an interval [0, K], so that two students who have identical benefits from drinking might nevertheless make different decisions (this is to enable us to make predictions of the incidence of various outcomes as continuous functions of the parameters). If he is successful in initiating consensual sex he gains a benefit of F [6]

In addition to these standard costs and benefits, the student may incur certain costs of violating social norms. These norms fall into two main categories - those that impose purely psychological costs and those that impose physical or social costs. The student has a different attitude to these two types of norms: he accepts that the social costs are exogenous constraints, but purely psychological costs can be avoided by various means, notably by consuming alcohol.

Alcohol, in other words, is a "disinhibitor": once it has been consumed, it reduces the actual psychological costs of norm violation by γ , and also the perceived social cost of future norm violation by δ . However, in the sober state, the student knows

⁶This might vary by type, F_C being either greater or less than F_T , but we do not explore this complication here.

⁷We originally considered calling these these "guilt" and "shame", and our notation of g and S grew out of this, but the reality is more complex than this. The point about the distinction between psychological costs and social costs is that the sober student may resort to alcohol to reduce the former, but this will do nothing to reduce the latter.

that alcohol consumption will reduce only the future psychological costs of the norm violation, and will have no effect on real future social costs; the student also knows that once he is drunk he will forget that alcohol does not really reduce social cost. Alcohol also affects the probabilities of sexual encounters, namely by acting as a socially validated mechanism for bringing individuals into physical proximity in a way that increases the opportunities for sexual offers to be made. We represent this by a multiplier α on the probability that an encounter takes place.

Acts of consensual sex can generate both psychological costs and social costs, and so can acts of non-consensual sex. We write g and s for the psychological and social costs of consensual sex, and G_{θ} and S for the psychological and social costs of non-consensual sex, noting that the psychological costs may differ between types.

Finally, while most of the costs of norm violation are known to the student in advance, there is uncertainty about how much he will suffer psychological costs of engaging in non-consensual sex. These are not known to the student in advance of the decision to whether or not to assault a partner who has refused. Furthermore, this uncertainty differs between calm and turbulent types. Turbulent types have, for any given set of other parameters, a greater expected probability of having psychological costs low enough to make them choose to assault their partner. We represent this uncertainty by a uniform distribution of G_{θ} over the interval $[0, G_{\theta}^+]$, which is [0, H] for calm types and [0, L] for turbulent types, with H > L. This means that $E[G_C] = \frac{H}{2}$ and $E[G_T] = \frac{L}{2}$. We assume that L is large enough that, even for turbulent types, for normal values of the other parameters, the probability that the student assaults a partner who refuses is strictly less than one even when the student is drunk. Observed levels of S are such that, even for calm types, the probability that the student assaults a partner who refuses is strictly greater than zero even when the student is sober.

3.2 The student's choices

The timing of events and decisions

Actions take place in the following sequence. First, nature determines the type of the individual student, who is Calm with probability π_C and Turbulent with probability $1 - \pi_C$. Next the student decides whether or not to consume alcohol. We represent this as a binary choice of strategy $a \in \{0, 1\}$.

The student then searches for a partner, meeting one and making a sexual proposition to that partner with a probability p, which increases when he consumes alcohol by a factor $\alpha > 1$. The partner then decides whether or not to accept, with a probability q. If the partner refuses, the student decides whether or not to insist. We represent this last action as a binary choice of strategy $b \in \{0, 1\}$.

In modeling the student's payoff function we have to distinguish between the objective payoffs, which the student evaluates when sober, and the payoffs as they appear to the student after consuming alcohol. We begin with the objective payoffs, which we write as:

$$U = Y - a \cdot c + I(J(F - s - g(1 - a \cdot (1 - \gamma))) - (1 - J)b \cdot (F - S - G_{\theta}(1 - a \cdot (1 - \gamma)))) \quad (2.1)$$

where we can summarize the variables as:

- $Y \in \mathbb{R}^+$ is the student's endowment of financial resources.
- $a \in \{0, 1\}$ is the decision to consume alcohol or not.
- $c \in [0, K]$ is the cost of consuming alcohol.
- $I \in \{0, 1\}$ is an indicator variable indicating whether the student meets a partner and can make a sexual proposition to her.
- $J \in \{0, 1\}$ is an indicator variable indicating whether the partner accepts.
- F > 0 is the benefit from sexual intercourse.
- $\theta \in \{C, T\}$ is the student's type.
- $s \ge 0$ is the social cost of engaging in consensual sex.
- $g \ge 0$ is the psychological cost of engaging in consensual sex.
- $S \ge 0$ is the social cost of engaging in non-consensual sex.
- $G_{\theta} \in [0, G_{\theta}^+]$ is the psychological cost of engaging in non-consensual sex.
- $b \in \{0, 1\}$ is the decision to insist or not, if the partner refuses.
- γ is the discount factor on psychological costs due to alcohol.

When the student has consumed alcohol a subtle shift takes place, which is that the social costs as well as the psychological costs of norm violation appear to be discounted. Thus after alcohol consumption, when by definition a = 1, we write the utility function as V to distinguish it from the objective form U:

$$V = Y - c + I(J(F - \delta \cdot s - \gamma \cdot g) - (1 - J)b \cdot (F - \delta \cdot S - \gamma \cdot G_{\theta}))$$
(2.2)

where δ is the discount factor on social costs due to alcohol. We assume $\delta \geq \gamma$, in other words that psychological costs are discounted at least as much as social costs.

The order of events is depicted in Figure 1, with payoffs depicted using the $U(\cdot)$ function.

The decision to assault

Several conclusions follow immediately from this. First of all, alcohol makes the student more willing, once drunk, to engage in non-consensual sex than he would be without alcohol, to an extent that increases the smaller are γ and δ :

- If the student is sober, he will be willing if $F S G_{\theta} > 0$.
- If student is drunk, he will be willing if $F \delta \cdot S \gamma \cdot G_{\theta} > 0$.

We can write the conditional probability that the sober student, faced with a refusal, assaults his partner as the probability that $G_{\theta} < F - S$. For the calm type this is equal to $\frac{F-S}{H}$ while for the turbulent type it is equal to $\frac{F-S}{L}$. Our earlier assumptions on H and L ensure that $0 < \frac{F-S}{H} < \frac{F-S}{L} < 1$.



Figure 3.1: Timing of events and decisions in the model: objective payoffs

In the same way, the conditional probability that the drunk student, faced with a refusal, assaults his partner is the probability that $\gamma \cdot G_{\theta} < F - \delta \cdot S$. For the calm type this is equal to $\frac{F - \delta \cdot S}{\gamma \cdot H}$ while for the turbulent type it is equal to $\frac{F - \delta \cdot S}{\gamma \cdot L}$, with $0 < \frac{F - \delta \cdot S}{\gamma \cdot H} < \frac{F - \delta \cdot S}{\gamma \cdot L} < 1$.

Secondly, alcohol makes opportunities for non-consensual sex occur more often, to a greater extent if α is large (alcohol increases the student's probability of meeting a partner) and if q is small (the probability of his offer being accepted is low). However, when sober the student will be aware of these risks. We can write the probability R_{θ} that a student will engage in non-consensual sex, conditional on drinking, as

$$R_{\theta} = \alpha \cdot p(1-q) \cdot \left(\frac{F - \delta \cdot S}{\gamma \cdot G_{\theta}^{+}}\right)$$
(2.3)

while the probability Q_{θ} that a student will engage in non-consensual sex, conditional on remaining sober, is

$$Q_{\theta} = p(1-q) \cdot \left(\frac{F-S}{G_{\theta}^{+}}\right) \tag{2.4}$$

This means that drinking increases the overall probability of engaging in nonconsensual sex by:

$$R_{\theta} - Q_{\theta} = \left(\frac{p(1-q)}{G_{\theta}^+}\right) \left[\left(\frac{\alpha}{\gamma} - 1\right)F - \left(\frac{\alpha \cdot \delta}{\gamma} - 1\right)S\right]$$
(2.5)

This will always be positive, and is increasing in F, p and α and decreasing in S, q, γ and δ . It is also decreasing in G_{θ}^+ , meaning that turbulent types are not only more likely to engage in non-consensual sex, but their probability of doing increases by more when they consume alcohol.

The decision to consume alcohol

We can now calculate the student's incentive to consume alcohol, taking all these considerations into account.

His expected utility U_S of staying sober (evaluated when sober) is:

$$U_S = Y + pq(F - s - g) + Q_{\theta} \cdot (F - S) - E[Q_{\theta} \cdot G_{\theta}]$$

$$(2.6)$$

And his expected utility U_D of drinking (evaluated when sober) is:

$$U_D = Y - c + \alpha pq(F - s - \gamma g) + R_{\theta} (F - S) - \gamma E[R_{\theta} G_{\theta}].$$

$$(2.7)$$

Subtracting (2.6) from (2.7) and rearranging shows that the student will drink if:

$$(F-s)[p \cdot q(\alpha-1)] - g \cdot p \cdot q(\alpha\gamma-1) + (R_{\theta} - Q_{\theta})(F-S) - (\gamma \cdot R_{\theta} - Q_{\theta}) \cdot \frac{G_{\theta}^{+}}{2} > c \quad (2.8)$$

which yields:

$$c < (F-s)[p \cdot q(\alpha - 1)] - g \cdot p \cdot q(\alpha \gamma - 1) + (R_{\theta} - Q_{\theta})(F - S) - \frac{p(1-q)[(\alpha - 1)F - (\alpha \cdot \delta - 1)S]}{2}$$
(2.9)

The probability of drinking, which we can call D_{θ} , is therefore the expression on the RHS of inequality (2.9), divided by K. The first term in that expression represents the net benefit to the student from consensual sex (excluding psychological costs) multiplied by the extent to which drinking increases the probability of achieving consensual sex. The second term represents the net change in the psychological costs of consensual sex due to drinking, and will represent a reduction in these costs if γ is low enough. The third term represents the net benefit to the student from non-consensual sex (again excluding psychological costs), multiplied by the increased probability of its occurrence. The fourth term represents the expected psychological cost of non-consensual sex, multiplied by the increased probability of incurring it.

What factors affect the probability of drinking, and in what way do they do so? It's not possible to sign the derivatives of D_{θ} unambiguously with respect to α : a higher α makes both consensual and non-consensual encounters more likely to happen, so the student will be encouraged to drink by a higher α if they are more attracted by the former than afraid of the latter.

With respect to the other parameters the comparative statics are fairly straightforward. The student will be more likely to drink if he is the turbulent type, and, for either type, if:

- F is large, provided q is large enough or (if q is small) provided S is small enough - the student strongly desires sex, and either has a reasonable chance of obtaining it consensually or is prepared to risk obtaining it non-consensually.
- when G_{θ}^+ (that is, either *H* or *L*) is small fear of the psychological costs of non-consensual sex also discourages drinking.
- g is large, provided γ is small enough (so that $\alpha \gamma < 1$) the psychological costs of consensual sex are high and drinking substantially reduces these costs;
- S is small, provided δ is small enough (so that $\alpha \delta < 1$) fear of the social

costs of non-consensual sex discourages drinking, since drinking increases the risk of incurring these costs. When δ is larger than this threshold (so that $\alpha\delta > 1$), fear of the social costs of non-consensual sex does not discourage drinking, since the student knows that he is not likely to behave when drunk very differently from how he would behave when sober.

- s is small there are few social costs of consensual sex.
- *K* is small alcohol is cheap;
- γ is small alcohol is a major disinhibitor;
- p is large the student has a reasonable probability of meeting partners;
- q is large those encounters are likely enough to lead to consensual sex (provided γ is small enough, so that $\alpha \gamma < 1$) and S is large enough;

3.3 The incidence of assault

Finally, we can use these insights to calculate the determinants of the overall incidence of non-consensual sex, by both sober and drunk students, which we denote by A_S and A_D respectively. This is given by the overall proportions of the each of the two types of student, the probabilities that each type chooses to drink (D_{θ}) , and the probabilities for each type of assaulting the partner conditional on drinking or remaining sober. The incidence of assaults by drunk students is given by:

$$A_D = p(1-q) \left[\Pi_C \cdot D_C \cdot \alpha \frac{(F-\delta \cdot S)}{\gamma \cdot H} + (1-\Pi_C) D_T \cdot \alpha \frac{(F-\delta \cdot S)}{\gamma \cdot L} \right]$$
(2.10)

while assaults by sober students are given by:

$$A_S = p(1-q) \left[\Pi_C \cdot (1-D_C) \frac{(F-S)}{H} + (1-\Pi_C)(1-D_T) \frac{(F-S)}{L} \right]$$
(2.11)

The incidence of both types of assaults is unambiguously increasing in p, the probability of meeting other students. It is also unambiguously decreasing in Π_C , meaning that the fewer turbulent types there are in the population, the fewer assaults there are - and the incidence of drunken assaults is more sensitive to the proportion of turbulent types than is the incidence of sober assaults. The effect of q (the probability that the student's encounter leads to consensual sex) on the probability of drunken assaults is ambiguous. It reduces the probability of assault both when drunk and when sober, but also (under reasonable values of the other parameters) makes drinking more attractive and thereby increases the proportion of encounters with a higher probability of assault.

Comparing the two equations, we note that the cost of drinking, which reduces the probability of drinking, and thereby the incidence of assaults by drunk students, also increases the probability of sobriety, and to that extent the incidence of assaults by sober students. In other words, alcohol leads to some substitution between assaults while sober and assaults while drunk. However, the fact that $\alpha \frac{(F-\delta \cdot S)}{\gamma \cdot H} > \frac{(F-S)}{H}$ and that $\alpha \frac{(F-\delta \cdot S)}{\gamma \cdot L} > \frac{(F-S)}{L}$ means that this cost reduces the incidence of assaults by drunk students by more than it increases the probability of assaults by sober ones, and therefore also decreases the total incidence of assaults by all students combined. The same reasoning also implies that any factors which affect A_D and A_S only through affecting D_{θ} will also have an effect on the total incidence of assaults that is of the same sign, though smaller in absolute value, as their effect on the incidence of assaults while sober.

We can summarize the impact of different parameters on the incidence of assaults while drunk, assaults while sober, and total assaults as follows:

Proposition 1:

The incidence of assault by drunk students is **increasing** in:

- F, the benefit of sex to the student, provided q is large enough or (if q is small) provided S is small enough;
- p, the probability of meeting potential sexual partners;
- the effect of alcohol on the psychological cost of consensual sex low values of γ lead to more assaults.
- g the psychological cost of consensual sex provided $\gamma < \frac{1}{\alpha}$;

The incidence of assault by drunk students is **decreasing** in:

- Π_C , the proportion of calm students in the population;
- K/2 the average cost of alcohol;

- *s* the social cost of consensual sex;
- S, the social cost of non-consensual sex provided δ < ¹/_α, i.e. that alcohol is more effective as a disinhibitor than as a social lubricant;
- H and L, the psychological costs of non-consensual sex for the two types.

Proposition 2:

The incidence of assaults by sober students is **increasing** in:

- F, the benefit of sex to the student, provided q is large enough or (if q is small) provided S is small enough;
- *p*, the probability of meeting potential sexual partners;
- the cost of alcohol.

The incidence of assaults by sober students is **decreasing** in:

- Π_C , the proportion of calm students in the population;
- the social and psychological costs of consensual sex;
- the social and psychological costs of non-consensual sex.

Proposition 3:

The total incidence of sexual assault by all students is increasing in F and p and decreasing in s, S, H and L. It is affected by the factors K and g in the same direction as is the incidence of sexual assault by drunk students, but the absolute magnitude of the effect is smaller.





Of particular interest in relation to our data is the impact of g, the psychological cost of consensual sex, on drunken assaults. Figure 3.2 shows the way in which the incidence of drunken assault varies with g, with separate lines drawn for whether γ is greater than, equal to or smaller than $1/\alpha$, and according to whether p is low or high.

The observation that γ is likely to be small where alcohol is concerned motivates our choice of empirical specification. We will test the hypothesis that the incidence of assaults by drunk students is increasing, and the incidence of assaults by sober students is decreasing, in empirical measures of the psychological cost of consensual sex. We will also test the hypothesis that the total incidence of assaults is increasing in empirical measures of the psychological cost.

To do this we need to control for other potential confounding variables. We now consider our data sources and the potential confounding variables for which they help us control.

4 Data and Descriptive Statistics

4.1 Data sources and choice of variables

Our main dataset was constructed using data from a wide variety of sources. Crime data are sourced from the National Incident-Based Reporting System (NIBRS, Federal Bureau of Investigation, 1997-2019), a rich dataset containing a wealth of information on crimes reported to police. The NIBRS gathers data from various reporting agencies all over the U.S.; as our analysis focuses on campus sexual violence, we restrict our dataset to only those reporting agencies located on a college campus.

This leaves us with data from 370 unique reporting agencies across 34 states⁸. NI-BRS variables include the location and exact time of occurrence, as well as some limited perpetrator characteristics, including – crucial for this analysis – whether the suspected perpetrator was considered to be under the influence of alcohol (such information is not available about victim(s)).

The case-by-case dataset is then transformed into a datset recording the total number of incidents for each campus during a given day or month. Extreme underreporting of sex crimes means the NIBRS data do not present a complete picture of the problem of campus sexual violence; however, the purpose of this paper is not to comment on the overall scale of the problem, but to attempt to determine under which circumstances sexual violence is *relatively* more likely to occur.⁹

Data on university characteristics such as enrolment, athletic association membership and religious affiliation have been drawn from the National Center for Education Statistics' Integrated Post-secondary Education Data System (IPEDS, National Center for Education Statistics, 1997-2019); data on university alcohol policy was collected from university student handbooks. Data on Planned Parenthood clinic availability by county was collected from the Planned Parenthood website.

County-level voting data for the House, Senate, and Presidential elections, sourced

⁸Arkansas, Colorado, Connecticut, Delaware, Georgia, Iowa, Indiana, Kansas, Kentucky, Louisiana, Massachussetts, Michigan, Minnessota, Missouri, Mississippi, Montana, Nebraska, North Carolina, North Dakota, New Hampshire, Ohio, Oklahoma, Oregon, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Vermont, Washington, Wisconsin, and West Virginia.

⁹In the Appendix, section B, we describe a number of alternative data sources and explain why they were not appropriate for the purpose of this paper.

from the Atlas of U.S. Presidential Elections (Leip, 1996-2019), are used to control for each county's overall political environment; we construct dummies for campuses are located in entirely Democrat-voting or entirely Republican-voting counties.¹⁰ In addition, we collect data on state laws regarding alcohol and sexual assault. Table 4.1 provides definitions for the most important variables.

Variable	Definition		
Incidents of sexual assault	Outcome variable (total and with/without alcohol)		
Incidents of rape	Outcome variable (total and with/without alcohol)		
No. of Planned Parenthood clinics	Number of PP clinics in county		
NCAA membership	Dummy, =1 for members of the National Collegiate Athletic Association		
Religious affiliation	Dummy, $=1$ for universities with a religious affiliation		
Dry campus	Dummy, $=1$ if alcohol is banned in dorms or on campus		
Beer tax	Tax on commercial sale of beer (%age points)		
$\log(\text{enrolment})$	Log of 12-month enrolment		
Blue county	Dummy, =1 if county voted Democrat in last House, Senate and Presidential elections		
Red county	Dummy, =1 if county voted Republican in last House, Senate and Presidential elections		
Gender-inclusive housing	Dummy, $=1$ if university offers gender-neutral housing		
Freshman residency	Dummy, $=1$ if university requires freshmen to reside		
requirement	on campus		
Single-gender dorms offered	Dummy, $=1$ if university has single-gender dorms		
Single-gender dorms only	Dummy, $=1$ if university only has single-gender dorms		
Weekend	Dummy, $=1$ on Friday and Saturday		

Table 4.1: Main variables and definitions

Further data on university characteristics were collected directly from university web pages; this includes factors that might reflect or influence overall campus attitudes towards gender relations and casual sex, such as whether the university offers gender-inclusive housing^{III}, or factors that might affect opportunities for potential offenders to find victims, such as whether the university requires freshmen to live on campus or whether single-gender dorms are offered.

We construct a dummy variable indicating whether the university campus in question

¹⁰Note that, as whether a county is "blue" or "red" depends on voting behavior, campuses may be alternately located in a blue or a red county depending on the current election cycle.

 $^{^{11}}$ Gender-inclusive housing is generally aimed at LGBTQ+ students, and takes into account gender identification regardless of legal gender or assigned gender at birth.

is a "dry campus" or not. It is important to note that, (as data for previous years are largely unavailable), this variable is based on the university's alcohol policy at time of research in 2018-21, and should thus be taken as a proxy of the university's past policies, and overall attitude toward alcohol. Due to limitations in the availability and quality of data from earlier years, the final dataset spans from 1997 to 2019.

A series of descriptive statistics are reported in Table 4.2 below.

_

Variable	Mean	Std. Dev.
Average incidents of sexual assault per day	0.007	0.10
of which with alcohol	0.001	0.05
Average incidents of sexual assault per month	0.218	0.69
of which with alcohol	0.045	0.28
Average incidents of rape per day	0.004	0.07
of which with alcohol	0.001	0.04
Average incidents of rape per month	0.110	0.45
of which with alcohol	0.029	0.21
Dry campuses (%)	49.0	50.0
Counties with Planned Parenthood (%)		0.50
Number of Planned Parenthood clinics in county		1.08
National Collegiate Athletic Association members (%)	82.9	37.6
Religious campuses (%)	8.55	2.79
Located in blue county $(\%)$	37.9	49.7
Located in red county $(\%)$	34.5	37.3
12-month enrolment	$27,\!194$	24,777
% of student body that is female	57.7	10.5
% with freshman residency requirement	46.1	49.9
% offering single-gender dorms	34.6	47.6
% offering single-gender dorms		22.2
% with gender-inclusive housing		45.8

Table 4.2: Summary means

4.2 Using the model to derive an empirical specification

Our dependent and independent variables of interest

We use incidents of sexual assault per month as our main dependent variable of interest (regression results for rapes, a subset of sexual assaults, are reported in the Appendix). We subdivide these into incidents with and without alcohol, as well as using the sum of the two. Our model predicts a probability of assault per student, not a total number of assaults, so we clearly need to control for campus size. We use (the logarithm of) total enrolment for this purpose as a control variable, as our descriptive data clearly indicate a strongly non-proportional relationship between campus size and assaults; we discuss this further below.

For our independent variables of interest, we have three proxies for g, the strength of norms against consensual sex¹². The first main proxy we use is the number of Planned Parenthood clinics available in-county. By making both contraception and information on sexual health available to all who request it, these clinics offer a non-judgmental environment that enables both men and women to consider consensual sex outside marriage a reasonable lifestyle choice; we therefore expect this proxy to be negatively related to the strength of norms against consensual sex.

Consumption of alcohol is far from being the only way in which individuals disinhibit themselves with respect to involvement in consensual sex. One other likely candidate is involvement in high-level sporting activities, which are commonly considered to legitimate high-frequency sexual activity, especially among the most successful athletes.^[13]

However, it's not clear what exactly is the link between these two disinhibitors rather than simply serving as a substitute to alcohol, sports activities are likely to be a complement, providing an environment in which high-level consumption of alcohol is both tolerated and encouraged. Sporting events furthermore provide opportunities for students to intermingle, and college athletes are often expected to make sexual conquests. We therefore expect NCAA membership to have a positive effect on both types of sexual assault, but a disproportionate effect on incidents where alcohol is involved. Another way of expressing this is that NCAA membership lowers the psychological AND the social costs of consensual and non-consensual sex - that, is g, G, s and S. This seems to us importantly different from the effect of Planned Parenthood clinics, which lower the costs of consensual sex but certainly not those of non-consensual sex.

We test this by including college membership of the National Collegiate Athletic

 $^{^{12}}$ We exclude consideration of norms governing sex between married couples, which are not the focus of interest in this study, although there are also serious problems of sexual assault in such a setting.

¹³See the article 'Sex and the Olympic City" by Matthew Syed, in *The Times* (of London), August 22nd 2008.
Association as a control variable: if sports are a complement to alcohol consumption we should see a larger coefficient in the equation for incidents with alcohol, while if they are a substitute we should see a larger coefficient in the equation for incidents without alcohol. We describe our data on athletics below.

Our third proxy, which we expect to be positively related to the strength of such norms, is whether the campus has a formal religious affiliation. This is of course a highly imperfect measure - religious campuses vary as to whether and how much they seek to influence students' decisions about engaging in consensual sex outside marriage, but it seems fair to say that the majority of such campuses are likely to consider that pre-marital and extra-marital sex should be at least informally discouraged if not formally prohibited. Certainly so-called "hookup culture" is more actively and seriously discouraged on religious campuses than on most secular campuses.¹⁴

This is not to say that there are no norms discouraging hookup culture on secular campuses. Many students on secular campuses will attend religious services at which such norms are reinforced. But they typically do not form part of the explicit culture of the campus as a whole, and are therefore more easily avoided by students who do not wish to be bound by them. The key distinction therefore between religious and secular campuses is that on secular campuses such norms are voluntary whereas on religious campuses they are (usually) obligatory to some extent.

Confounding variables

Several parameters of our model are not readily empirically observable, but are also ones that we have no reason to think will vary systematically from one time to another or from one university campus to another. These include F, the benefit of sex to the student; γ , the effect of alcohol on the actual psychological cost of consensual and non-consensual sex; and δ , the effect of alcohol on the perceived social cost of consensual and non-consensual sex. Other variables in the model for which we do have reasonable proxies include:

• p - the probability of meeting potential partners. We expect this to be higher

 $^{^{14}}$ See "The voices of young people with different views of social justice are pushing the Mormon Church to modernize", *New York Times*, April 12, 2019. The article reveals that there may also be a negative correlation between g and S: among the claims recorded are that "After Brigham Young drew outrage for punishing sexual assault victims who were found to have violated the honor code, the school implemented a new "amnesty" policy in 2017" and "Another student said she overdosed on pills after she was punished for revealing she had been sexually assaulted."

in larger campuses, and those that have a freshman residency requirement, and higher at weekends than during the working week. We also expect it to be lower in campuses that have an asymmetric sex ratio, though this seems a less reliable measure to us, since an asymmetric sex ratio may also be associated with higher levels of sexual frustration among at least some individuals, and therefore with higher values of F.

Two other campus characteristics are also likely to affect p, but may also reflect other confounding influences. One is the presence of gender-inclusive housing on campus, which is likely to increase the ease with which opportunities arise for both consensual and non-consensual sex. But it is typically also associated with norms that accept the right of students to express both their gender identity and their sexuality – the latter of which consensual sex outside marriage. These two influences act in opposite directions, making it difficult to anticipate the direction of the net effect.

A second influence on p is the presence of separated dormitories, which in principle reduces the opportunities for both consensual and non-consensual sex, but may also be correlated with stricter norms against both. Once again this makes for an ambiguous net effect, which will have to be measured empirically.

- K/2 the average cost of obtaining alcohol. We do not have direct measures of this, but we do observe one variable that is likely to make a major difference to the time, energy and financial costs associated with obtaining alcohol. This is whether the campus operates a "dry" policy, forbidding the sale of alcohol at outlets located on campus. Other things equal, the cost of alcohol at a dry campus is likely to be significantly higher than at other campuses. However, this variable is also likely to be endogenous, because the decision to operate a dry policy may itself represent a response to certain behaviors, including the incidence of sexual assault. It may also be influenced by other unobserved factors that are correlated with the error term in our equation.
- S the social cost of non-consensual sex. This is the variable for which it is hardest to find reliable proxies. We do not have data that allow us to measure

these directly, so we use an indirect measure, which is the lagged incidence of sexual assaults (which is also likely to act as a signal to individuals about what is considered regular or normal in their community). To capture possible spillover effects of assaults in creating a culture of impunity, we use lagged assaults with alcohol as a regressor in our equation without alcohol, and vice versa.

Finally, because our proxies for g are only proxies, and not direct measures, we want to control for other confounding influences that might be correlated with them. The political alignment of the county in which the campus is located may have an independent influence on many factors that might affect the incidence of assaults other than through norms around consensual sex; to control for these we use voting indicators of the county in previous elections.

We now turn to some descriptive statistics revealed by our dataset.

4.3 Incidents by month and campus type

Figure 4.1 shows the pattern of incidents across the calendar year. Most reported incidents are concentrated toward the beginning of the academic term, in September and October, with two small peaks around February, and April, when many universities have their spring breaks. Figure 4.2 shows that incidents per month across the year are significantly lower in campuses that ban alcohol than in those that do not.

The relationship between campus alcohol policy and incidents of sexual assault may be at least in part due to size effects: there is a clear decreasing relation between a university's size and the number of incidents of sexual assault per capita, as shown in Figure 4.3, and dry campuses are smaller by almost 7,000 students per year – an average of 23,000 students enrolled per year, compared to an average yearly enrolment of 31,000 students on campuses that do not ban alcohol.



Figure 4.1: Mean number of incidents by month

Figure 4.2: Mean incidents per month by campus alcohol policy





Figure 4.3: Incidents per capita by university size

Universities that ban alcohol in fact present a slightly higher rate of alcohol use in cases of rape than universities where alcohol is permitted on campus, and nearly identical rates of alcohol use for sexual assault more generally, as is shown in Table 4.3.

Table 4.3: Alcohol use by campus alcohol policy

Alcohol allowed on campus	Incidents of sexual assault involving alcohol	Percent of total	Incidents of rape involving alcohol	Percent of total
Yes	1,371	(20.5)	858	(25.6)
No	757	(20.8)	496	(26.4)

While at first glance it seems somewhat counter-intuitive to see such small differences in these rates between the two kinds of campus, it is important to note there is likely endogeneity at play in these statistics. Campuses that ban alcohol may have chosen to do so precisely because of problems with excessive alcohol consumption or alcoholfuelled sexual assault. It should therefore not be too surprising to see that campuses which ban alcohol do not have greatly reduced rates of alcohol-fuelled sexual assault relative to campuses that do not.

The importance of size effects makes it worth comparing campuses on a per capita

basis. Following our arguments about norms, Figures 4.4 and 4.5 show how assaults *per capita* across the year differ by two indicators of norms concerning consensual sex. Average assaults per capita per month are somewhat higher on campuses where there are no Planned Parenthood clinics in the county, and substantially higher on religiously affiliated campuses than on secular campuses.

Once again, though, size effects may be at work here. With the exception of Brigham Young University, which is affiliated with the Church of Latter-Day Saints and has an average 12-month enrolment of 26,764 students (median = 20,082), secular colleges are far larger than religious ones, with an average enrolment of 28,718 students (median = 22,244), compared to 4,401 for the average religious college (median = 3,760). For this reason we need to move beyond bivariate correlations and undertake multivariate regression, and the multivariate regression has to control for campus size.







Figure 4.5: Mean incidents per capita per month by campus religion

4.4 Regional characteristics

County-level voting data serve as a proxy for the general political environment at the time for each university's county. Rates of alcohol involvement are similar, but slightly higher in counties that voted Republican in each of the previous House, Senate and Presidential election cycle than in those that voted Democrat for all three races. This information is presented in Table 4.4.

Table 4.4: Alcohol use by red county/blue county

County election results	Incidents of sexual assault involving alcohol	Percent of total	Incidents of rape involving alcohol	Percent of total
Blue county	1,056	19.8	634	24.8
Red county	565	21.3	397	27.6
Purple county	507	21.5	323	26.2

Consistent with our model, counties with greater Planned Parenthood availability see lower rates of alcohol involvement, with each additional clinic associated with a 2 percentage point reduction in incidents that involved alcohol, as a proportion of the total. Finally, political climate is highly correlated with campus alcohol policy, as nearly two thirds of campuses in "red" counties ban alcohol, compared with less than a third in "blue" counties. Alcohol-free campuses are also more likely to be located in counties without a Planned Parenthood clinic – 64% of dry campuses do not have a Planned Parenthood clinic within the county, compared to 46% of campuses where alcohol is permitted. Finally, while alcohol prohibitions are common on U.S. campuses, universities with an official religious affiliation are somewhat more likely to prohibit alcohol on their premises: 59% of religious campuses do so, compared to 50% of secular ones.

 Table 4.5: Alcohol policy by county type

	Blue county	Red county	Purple county
Proportion banning alcohol	36.5%	63.0%	48.2%

4.5 Characteristics of perpetrators and victims

In 87% of all reported sexual assault and 92% of reported rape cases, the victim knew the attacker, a figure consistent with previous research (see e.g. Abbey, 2002). Alcohol was involved in 20% of all reported cases of sexual assault on college campuses, and is involved at similar rates whether the victim and perpetrator were acquainted or not. As might be expected for crimes occurring on college campuses, the mean age of all offenders is between 23 and 24 years old. The mean victim age is between 20 and 21 years old.

5 Regression Results

5.1 The effect of technical factors excluding social and psychological costs

Our main regressions take as their dependent variables the number of incidents of sexual assault and rape committed per campus per calendar month. As these variables take integer values and have many more zeroes than positive values, natural techniques to use are Poisson or negative binomial regressions. Our data display a degree of over-dispersion that makes the negative binomial the more appropriate of the two, but Poisson regressions not reported here show that this makes little difference to the qualitative results. The Appendix reports a number of other specifications that change the qualitative results only marginally if at all.

	Incidents	w. alcohol	Incidents w.o. alcohol Total inciden		cidents	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Dry campus	-0.269	(0.17)	-0.211	(0.13)	-0.228+	(0.13)
Beer tax	-0.274	(0.30)	0.079	(0.20)	0.003	(0.20)
Log of total enrolment	0.488^{***}	(0.09)	0.702^{***}	(0.06)	0.655^{***}	(0.06)
Freshman residency requirement	0.651***	(0.16)	0.433***	(0.13)	0.484***	(0.12)
Single-gender dorms offered	0.307^{*}	(0.15)	0.387**	(0.14)	0.374**	(0.13)
Single-gender dorms only	-0.282	(0.40)	0.122	(0.28)	0.054	(0.27)
Gender-inclusive housing	-9.088***	(0.92)	-9.415***	(0.53)	-8.803***	(0.55)
Month FE	Y	ΈS		YES	YI	ES
Year FE	Y	ΈS	-	YES	YI	ES
Regional Dummies	Y	\mathbf{ES}	-	YES	YI	ES
Constant	-8.311***	(0.87)	-8.843***	(0.54)	-7.737***	(0.49)
Ν	46	,294	4	6,294	46,2	294
$p > \chi^2 $	0.0	0000	0	.0000	0.00	000
α	6.807***	(0.95)	1.914***	(0.26)	1.851***	(0.22)
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001						

 Table 5.1: Regression excluding social and psychological costs

We begin by reporting a preliminary specification that uses only the "technical" variables affecting the opportunities for consensual and non-consensual sex, and excludes the variables representing the social and psychological costs. Table 5.1 illustrates. It reports results for assaults under three headings: assaults with alcohol, assaults without alcohol and total assaults. All specifications use year, month and region fixed effects.

As expected, a dry campus reduces both alcohol-related and total incidents, though the effect is significant at only the 10 per cent level and only in the equation for total incidents. Several control variables that increase the ease with which students of the opposite gender can meet each other have the expected positive sign: the presence of a freshman residency requirement, gender-inclusive housing. The presence of single-gender dorms also has a significant positive sign, which (as we discussed above) would be a surprise if this variable reduced purely reduced opportunities for sexual encounters. This suggests it may also capture an element of attitudes, as will be confirmed in the next specification.

5.2 Main findings: including social and psychological costs

Table 5.2 reports the results of our preferred specification including our three proxies for social and psychological costs of consensual sex, as well as the use of lagged assaults as a proxy for the social costs of non-consensual sex.

The results here are very striking. The number of Planned Parenthood clinics has a large and negative effect on the incidence of assaults with alcohol, one that is statistically significant at less than a tenth of one per cent. Consistently with our interpretation that alcohol serves as a disinhibitor in the presence of norms against consensual sex, and that the presence of Planned Parenthood clinics correlate with weaker norms against consensual sex, we see the negative effect only on incidents with alcohol. There is essentially no effect on incidents without alcohol. The effect on total incidents is in the same direction as the effect on incidents with alcohol but much smaller and statistically insignificant (our model predicts that the effect should be negative but smaller).

The dummy variable for campus membership of the National Collegiate Athletic Association has a very large and extremely significant coefficient. It is substantially larger (almost twice as high) in the equation for incidents with alcohol, indicating that the disinhibiting effect of this institution is complementary to alcohol consumption and not a substitute for it. However, it also has a large effect on incidents without alcohol.

The coefficient on the dummy variable for campuses with a religious affiliation has a positive sign and quite a large effect on incidents with alcohol (over twice that of the number of Planned Parenthood clinics). But this effect is not statistically significant, which may be unsurprising given that religious campuses make up under 10 per cent of our sample.

	Incidents	w. alcohol	Incidents	w.o. alcohol	Total in	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
No. of Planned Parenthood clinics	-0.223***	(0.07)	0.016	(0.04)	-0.022	(0.03)
National Collegiate Athletic Association membership	1.913***	(0.33)	1.038***	(0.18)	1.121***	(0.16)
Religious affiliation	0.523	(0.36)	0.196	(0.21)	0.237	(0.20)
Dry campus	-0.244+	(0.14)	-0.162	(0.11)	-0.168+	(0.10)
Beer tax	-0.165	(0.30)	0.171	(0.19)	0.076	(0.17)
Log of total enrolment	0.478^{***}	(0.08)	0.672^{***}	(0.06)	0.583^{***}	(0.05)
Freshman residency requirement	0.371**	(0.14)	0.271**	(0.10)	0.276**	(0.09)
Single-gender dorms offered	0.044	(0.14)	0.227 +	(0.13)	0.173	(0.11)
Single-gender dorms only	-0.392	(0.47)	0.000	(0.28)	-0.070	(0.27)
Gender-inclusive housing	0.370*	(0.15)	0.309*	(0.14)	0.296*	(0.12)
Lagged incidents of sexual assault w.o. alcohol	0.191***	(0.05)				
Lagged incidents of sexual assault w. alcohol			0.184***	(0.05)		
Lagged incidents of sexual assault					0.230***	(0.03)
Blue county	0.099	(0.12)	0.175^{*}	(0.08)	0.139 +	(0.07)
Red county	-0.198	(0.14)	-0.171+	(0.10)	-0.156+	(0.08)
Month FE	У	YES		YES	Y	ES
Year FE	У	YES		YES	Y	\mathbf{ES}
Regional Dummies	У	YES		YES	Y	ES
Constant	-8.311***	(0.87)	-8.843***	(0.54)	-7.737***	(0.49)
Ν	45	,924	4	5,924	45,	924
$p > \chi^2 $	0.	0000	0	0.0000	0.0	000
α	5.643***	(0.74)	1.729***	(0.22)	1.559***	(0.17)
+ p < 0.10, * p < 0.05, ** p	< 0.01, ***	p < 0.001				

Table 5.2: Main regressions - incidents of sexual assault, monthly data

Our proxy for the weakness of social norms against aggression, namely the lagged level of incidents, is strongly and significantly positive as predicted (whether we include this variable or not makes little difference to the other coefficients). Finally, the dummies for red and blue counties are weakly significant, indicating that campuses in blue counties have somewhat more, and campuses in red counties somewhat fewer incidents than would be predicted given their other characteristics in the regression, though this effect appears to be driven primarily by incidents without alcohol.

Including social and psychological costs in the regression makes very little difference to the coefficients on the "technical" parameters. The coefficient on single-gender dorms becomes smaller and loses its significance, which is consistent with the view that its importance in Table 5.1 reflects its correlation with attitudes. The coefficient on the dummy variable for a dry campus becomes weakly significant in the equation for incidents with alcohol, but essentially none of the other coefficients are affected in any meaningful way. This tells us, reassuringly, that the proxy variables for social and psychological costs are unlikely to be confounded by any of the technical factors we have been considering, with the mild exception of the presence of single-gender dorms.

5.3 Robustness checks

In a series of tables in the Appendix, we report a range of robustness checks of our main specification. The first check in Table 2.A.4 provides the same analysis for the incidence of rape; the results are qualitatively the same as for sexual assault. The second robustness check in Table 2.A.5 uses a dummy variable for the presence of Planned Parenthood clinics in the county, rather than the number of clinics, to verify that the results are not driven by outliers, and indeed the qualitative results hold up.

The third robustness check involves using daily rather than monthly data. A feature of Table 2.A.6 is the addition of the weekend variable, which is consistently positive and significant (and more than double the size of the planned parenthood coefficient for incidents with alcohol). It provides some perspective on the impact of the dummy variable for athletic association membership - this is associated with increases in incidents of assault of more than three times the difference in incidence on an average campus between weekdays and weekends. Otherwise there is no qualitative difference in these results compared to the use of monthly data.

A fourth robustness check, in Table 2.A.7 involves using incidents per capita per month instead of total incidents as the dependent variable. Here the fact that we

no longer have integer values means we can no longer use the negative binomial, so we use Ordinary Least Squares, but the presence of so many zeroes makes the estimation less reliable. The main difference with our preferred specification is that the coefficient on religious affiliation in the equation for incidents with alcohol becomes significantly positive at 5%. It is also much larger relative to the other coefficients of interest than in the main regressions - eight times as large in absolute terms as the coefficient on Planned Parenthood and three times as large as that on Athletic Association membership. It's hard to know how much weight to give to this specification compared to the others. However, if there is a tendency for religiously affiliated campuses to have higher incidence, as was suggested strikingly by Figure [4.5] it may be principally a by-product of the other characteristics of these institutions, and the small number of such campuses in our dataset makes it hard to generalize further.

Finally, because of the possible endogeneity of campus alcohol policy, we report in Table 2.A.3 in the Appendix an instrumental variable (IV) specification, where campus alcohol policy is instrumented using the first principal component of a series of measures of the strictness of states' laws on alcohol and a dummy for restrictions on the sale of alcohol at the county level. These instruments appear to satisfy the twin requirements of relevance and exogeneity: while campus alcohol policy is likely affected by the same overall attitudes towards alcohol that drive state and county laws and taxes, it is unlikely that the prevalence of sexual violence at any particular campus would affect legislature at the county or state level. However, the results are virtually unchanged from the specification without instrumentation, so we do not pursue this line of investigation further here.

A possible sources of bias: selection by individual temperament

One legitimate source of concern with our interpretation of our findings is that colleges that differ in their attitudes to consensual sex may attract different kinds of student. In some circumstances, this might reinforce our conclusions - if, for example colleges located in areas without Planned Parenthood clinics not only have higher values of g for their students through norms against consensual sex, but also attract students who have higher intrinsic levels of g than the rest of the population. However, this is not the only way in which selection might affect our results. Suppose, for example, that parents who recognize their children to have behavioral problems prefer to send them to colleges with more conservative norms governing relationships between students, in the hope that they will be better looked after in an environment where there is greater discipline. In that case the higher levels of assaults by drunk students in colleges with less Planned Parenthood presence might simply reflect the more troubled nature on average of the student intake.

There is some plausible evidence in favor of this hypothesis, as campuses without Planned Parenthood clinics do observe higher per-capita levels of crimes such as property damage or assault, as shown in Figures 5.1 and 5.2. So also do religious campuses compared to non-religious ones, as shown in Figures 5.3 and 5.4

Figure 5.1: Property damage and PP clinics



Figure 5.3: Property damage and reli- Figure 5.4: Assault and religious affiligious affiliation



Figure 5.2: Assault and PP clinics



ation



Fortunately our model allows us to test this hypothesis, and notably to distinguish it from the hypothesis that what distinguishes colleges with and without Planned Parenthood presence is principally the stronger norms against consensual sex. Table 5.3 shows the results. If colleges with more conservative norms regarding sex have more students with a turbulent temperament, the model predicts that this will increase assaults by both drunk AND sober students (albeit to a greater degree by drunk students). However, the effect of stronger norms against consensual sex is only to increase assaults by drunk students; the effect on assaults by sober students has the opposite sign, though it is not statistically significant.

The coefficients on incidents of property damage, burglary and assault are all statistically significant, though small. Property damage and burglary show slightly higher coefficients for incidents with alcohol, but the opposite is true for the most violent category, assault. Controlling for incidents of other types of crime also lowers the coefficient on Planned Parenthood very slightly for incidents with assault, but remains large and highly significant; similarly, the coefficients on National Collegiate Athletic Association membership are lower for both incidents with alcohol and without (as well as for total incidents). Overall the results do not provide support for the hypothesis that the gap in rates of alcohol-fuelled sexual assault between campuses with and without nearby Planned Parenthood clinics, or between members of the NCAA and non-members, is due to selection bias.

	Incidents	w. alcohol	Incidents	w.o. alcohol	Total ir	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
No. of Planned	0.001***	(0,06)	0.015	(0,02)	0.092	(0,02)
Parenthood clinics	-0.201	(0.00)	0.015	(0.03)	-0.023	(0.03)
National Collegiate Athletic	1 780***	(0.31)	0 860***	(0.15)	0 000***	(0.14)
Association membership	1.700	(0.01)	0.000	(0.10)	0.990	(0.14)
Religious affiliation	0.295	(0.34)	0.052	(0.18)	0.119	(0.18)
Dry campus	-0.242+	(0.13)	-0.130	(0.08)	-0.157^{*}	(0.08)
Beer tax	-0.242	(0.29)	0.106	(0.15)	0.029	(0.16)
Property damage	0.033***	(0.01)	0.020***	(0.00)	0.023***	(0.00)
(total incidents)	0.000	(0101)	0.020	(0.00)	0.020	(0.00)
Burglary	0.019^{*}	(0.01)	0.013**	(0.00)	0.015**	(0.01)
(total incidents)		()		()		()
Assault	0.028**	(0.01)	0.047***	(0.01)	0.044***	(0.01)
(total incidents)	0.000**		0 115***		0.001***	(0.05)
Log of total enrolment	0.223^{**}	(0.08)	0.445^{***}	(0.05)	0.394^{***}	(0.05)
Freshman residency	0.225 +	(0.12)	0.156^{*}	(0.08)	0.173^{*}	(0.07)
Circula and devidence a		. ,				. ,
Single-gender dorms	0.009	(0.12)	0.169 +	(0.09)	0.135	(0.09)
Single genden dermes						
Single-gender dorms	-0.244	(0.43)	0.104	(0.24)	0.038	(0.23)
Conder inclusive						
housing	0.396^{**}	(0.13)	0.256^{*}	(0.11)	0.282^{**}	(0.10)
Lagged incidents of						
sexual assault w.o. alcohol	0.134^{**}	(0.05)				
Lagged incidents of						
sexual assault w alcohol			0.196^{***}	(0.03)		
Lagged incidents of						<i>.</i>
sexual assault					0.186^{***}	(0.03)
Blue county	-0.001	(0.12)	0.056	(0.06)	0.038	(0.06)
Red county	-0.100	(0.13)	-0.075	(0.08)	-0.079	(0.07)
Month FE	Y	TES		YES	Y	ES
Year FE	Y	ΈS		YES	Y	ES
Regional Dummies	Y	ΈS		YES	Y	ES
Constant	-6.291***	(0.79)	-6.959***	(0.49)	-6.261***	(0.46)
N	45	.602	4	5.602	45,	602
$p > \chi^2 $	0.0	0000	0	0.0000	0.0	000
α	5.214***	(0.67)	1.528^{***}	(0.18)	1.438***	(0.15)
+ p < 0.10, * p < 0.05, ** p	< 0.01, *** ;	p < 0.001				

 Table 5.3: Regressions with other crimes as controls

	Incidents	w. alcohol	Incidents	w.o. alcohol	Total in	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Excess female students	-2.647^{*}	(1.06)	-0.253	(0.87)	-0.616	(0.75)
Excess male students	-3.056*	(1.33)	-1.088	(0.69)	-1.298*	(0.64)
No. of Planned	0.917***	(0, 06)	0.011	(0, 0.4)	0.027	(0.03)
Parenthood clinics	-0.217	(0.00)	0.011	(0.04)	-0.027	(0.03)
National Collegiate Athletic	1 000***	(0.22)	1 ()99***	(0.18)	1 111***	(0.16)
Association membership	1.090	(0.33)	1.055	(0.18)	1.111	(0.10)
Religious affiliation	0.460	(0.39)	0.198	(0.22)	0.237	(0.21)
Dry campus	-0.249+	(0.13)	-0.172	(0.11)	-0.171+	(0.09)
Beer tax	-0.195	(0.30)	0.173	(0.19)	0.073	(0.17)
Log of total enrolment	0.390^{***}	(0.09)	0.638^{***}	(0.06)	0.544^{***}	(0.05)
Freshman residency	0.959*	(0, 12)	0.250*	(0, 10)	0.941**	(0.08)
requirement	0.232	(0.13)	0.239	(0.10)	0.241	(0.08)
Single-gender dorms	0.044	(0.12)	0.200	(0, 12)	0 161	(0, 11)
offered	0.044	(0.13)	0.209	(0.13)	0.101	(0.11)
Single-gender dorms	0 422	(0.45)	0.044	(0.30)	0.036	(0.27)
only	-0.422	(0.43)	0.044	(0.30)	-0.030	(0.27)
Gender-inclusive	0 449**	(0, 14)	0.210*	(0.15)	0 291**	(0, 12)
housing	0.442	(0.14)	0.319	(0.13)	0.321	(0.12)
Lagged incidents of	0 171***	(0.05)				
sexual assault w.o. alcohol	0.171	(0.05)				
Lagged incidents of			0 199***	(0.05)		
sexual assault w. alcohol			0.182	(0.03)		
Lagged incidents of					0 226***	(0, 0.2)
sexual assault					0.220	(0.03)
Blue county	0.125	(0.11)	0.181^{*}	(0.08)	0.152^{*}	(0.07)
Red county	-0.126	(0.13)	-0.199*	(0.10)	-0.164*	(0.08)
Month FE	У	YES		YES	Y	ES
Year FE	У	YES		YES	Y	\mathbf{ES}
Regional Dummies	Y	YES		YES	Y	ES
Constant	-8.399***	(1.01)	-9.099***	(0.63)	-7.967***	(0.56)
N	41	,463	4	1,463	41,	463
$p > \chi^2 $	0.	0000	0	0.0000	0.0	000
α	5.087^{***}	(0.67)	1.700^{***}	(0.22)	1.492***	(0.16)
+ p < 0.10, * p < 0.05, ** p	< 0.01, ***	p < 0.001				

 Table 5.4:
 Regressions with excess female/male students

Gender ratios

A final test of robustness is to control for the gender ratios in the student body: these may affect p, the probability that students can meet, but may also be correlated with other factors including religious affiliation. There would be grounds for concern if the inclusion of these controls diminished significantly the impact of our variables of interest.

In fact, as can be seen in Table 5.4, including these makes essentially no difference to the impact of Planned Parenthood clinics. We control both for any excess of male students above parity, and for any excess of female students above parity (either of which would tend to reduce p). The coefficients on both measures are significantly negative, as expected. ¹⁵

Reporting bias: using the National Crime Victimization Survey Data

A major issue in the use of any statistics on sexual assault is the presence of reporting bias. A large part of the voluminous literature on this issue has attempted to establish accurate average rates of sexual assault, in the face of evidence of substantial under-reporting (see in particular Fisher et al., 2010). Our concern here is not so much with average levels of under-reporting as with *differences* in reporting rates which may bias our estimates of the contribution of our independent variables of interest.

To explore this question empirically, we collect data from the National Crime Victimization Survey (Bureau of Justice Statistics, 1992-2019), a nationally representative random survey. For comparability purposes, we use data only from the years that appear in our main dataset. Although this dataset has the important limitations we discussed above, it may still shed some light on whether differences in reporting rates might affect our empirical analysis, given that the survey asks respondents both whether they informed the police of a crime committed against them, and whether police were involved in general, as well as including questions on reasons for not reporting, if applicable.¹⁶

¹⁵However, it is important to bear in mind individuals may take into account information about the prevalence of sexual violence at a given institution when deciding which university to enrol in; this might <u>make</u> these variables to some degree endogenous.

¹⁶Table 2.A.1 in the Appendix summarizes definitions of the most important of our NCVS variables.

Table 5.5 reports descriptive statistics for sex crimes in the NCVS data.¹⁷ Table 5.6 sets out reporting rates for sexual assault and compares them with other crimes.

Outcome	Mean	Std. Dev.
Victim reported crime	0.178	0.38
Police involved	0.290	0.45
Alcohol involved	0.278	0.45
Offender is stranger	0.155	0.36
White offender	0.512	0.50
Black offender	0.177	0.38
White victim	0.800	0.40
Black victim	0.132	0.34
Weapon used	0.118	0.32
Victim has college degree	0.109	0.31
Urban	0.849	0.36
Northeast	0.098	0.30
Midwest	0.201	0.40
South	0.252	0.43
West	0.215	0.41

Table 5.5: Summary statistics for sexual assault (NCVS) - proportions of total cases

Table 5.6: Reporting rates by crime (NCVS)

	Self-r	reporting	Police	involvement
	Mean	Std. Dev.	Mean	Std. Dev.
Sexual assault	0.178	0.38	0.286	0.45
Assault and battery	0.258	0.44	0.503	0.50
Theft	0.262	0.44	0.354	0.48
Property damage	0.519	0.50	0.681	0.47

Unsurprisingly, sex crimes suffer from the lowest reporting rates, with only one in six NCVS respondents saying they reported their crime to police themselves, and police being involved in only 29% of cases.

An important issue concerns possible differences in reporting rates by race, of both perpetrator and victim. Though the majority of perpetrators of sexual assault are white, black perpetrators are over-represented in the NIBRS dataset, comprising

¹⁷Offender race is recorded only for multiple-offender crimes prior to 1Q2012.

over 20% of all perpetrators in reported incidents of rape versus 13% of the U.S. population overall. This over-representation is present in all sexual crimes taken individually (groping, sexual assault with an object, etc.) as well as together. Black individuals are also over-represented in violent crimes (43-44% for aggravated and simple assault). This proportion falls back in line with population averages, however, for non-violent crime such as property damage (13%).

It is possible that the races of both the perpetrator and the victim may play a role in whether the crime is reported, or whether a suspect is identified. Stacey et al., 2016] for example, find that arrests are more likely in cases of stranger rape when the perpetrator is black and the victim white. They also find that victims cooperate less with the police in black-on-black family assaults. Spohn and Holleran, 2001 find that prosecutors are more likely to file charges when the victim was white, but less likely when they had been engaging in risk-taking behavior. If victims form rational expectations of how police will deal with their case, this may affect propensity to report depending on the characteristics of the victim, the perpetrator, and the crime itself.

These differences may reinforce already existing differences in prevalence due to stereotypes in the minds of both perpetrators and victims. O'Flaherty and Sethi, 2019 suggest that common stereotypes of black males in the U.S. as violent may lead black males looking to commit a crime to self-select into crimes which involve face-to-face contact (such as robbery over burglary), knowing that the stereotype may work in their favour in ensuring cooperation from the victim.

Table 5.7 uses the NCVS data to estimate, using a probit estimation, the effect of various perpetrator and victim characteristics and characteristics of the incident on police involvement and victim self-reporting.

	Police inv	volvment	Self-re	porting
	Coef.	Std. Err.	Coef.	Std. Err.
Alcohol involved	-0.065	(0.05)	-0.034	(0.06)
Offender is stranger	0.339^{***}	(0.07)	0.250^{**}	(0.08)
Offender is stranger*alcohol	-0.235+	(0.12)	-0.371^{**}	(0.14)
White victim	0.166^{**}	(0.06)	0.150^{*}	(0.07)
Black offender	0.234^{***}	(0.06)	0.107	(0.07)
Weapon used	0.467^{***}	(0.07)	0.326^{***}	(0.07)
Post-secondary education	-0.164***	(0.05)	-0.058	(0.05)
College	-0.298***	(0.08)	-0.232*	(0.09)
Urban area	-0.143*	(0.06)	-0.104	(0.07)
Northeast	-0.091	(0.09)	-0.150	(0.10)
South	0.010	(0.07)	-0.013	(0.07)
West	0.026	(0.07)	0.062	(0.08)
Year of incident dummies	YI	ES	Y	ES
Ν	3,6	54	3,6	354
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001				

Table 5.7: Probit marginal effects - cases of sexual assault

White victims are more likely to report sexual crimes against them to the police than those of other races. Black offenders are 23 percentage points more likely than white offenders to face police involvement, ceteris paribus, even though rates of reporting by victims do not significantly change. Stranger offenders are also more likely to be reported, though this tendency interacts in complex ways with the whether alcohol was involved (note that again, there is no information on victim's use of alcohol only the offender(s)').

In Figure 5.5 we compare offender race data from the NIBRS and NCVS to U.S. Census demographic data and student demographic data from the IPEDS. While black individuals are slightly underrepresented among U.S. college students, they are very much over-represented in our crime data. It is possible that this over-representation reflects higher rates of police involvement when a crime is committed by a person of color; and this may affect our results if the racial composition of the student body correlates with our variables of interest.



Figure 5.5: Black individuals' representation among offenders compared to general and student populations

Therefore, in Table 2.A.10 we control for the proportion of the student body that is black. The coefficient is insignificant at the 5% level for incidents of sexual assault both with and without alcohol, as well as for total incidents. It is, additionally, *negative* and weakly significant at the 10% level for incidents with alcohol. Including the variable makes essentially no difference to the sign or magnitude of the coefficients on any of our explanatory variables of interest.

6 HERI attitudinal data

The Higher Education Research Institute (HERI) at UCLA administers *The Fresh*man Survey (Higher Education Research Institute, 1997-2018), a yearly survey of incoming new university students which asks individuals about their activities in the past year, and attitudes towards various topics, as well as collecting data on students' overall political leanings, religion of preference, etc. Among the attitudinal data that have been collected in the past are views on the legality of abortion, and the morality of sex outside a committed relationship (in other words, hookups). We will use these data to investigate whether, and by how much, students attending university in areas without Planned Parenthood availability differ, as well as how much religious students differ from those with no religion. As universities are anonymized, it is impossible to assign responses to specific campuses; we therefore use students' home zip code instead. Responses by zip code serve as a measure of the environment in which individual campuses are embedded. To maximize comparability with our NIBRS data, we use HERI data only from 1997 onward only.

6.1 How do religious students differ from atheist ones?

Religious students, in general, are somewhat more likely than self-reported atheists to avoid alcohol or to drink it only occasionally. They also disagree far more strongly with the phrase "If two people really like each other, it is okay for them to have sex." They are also far more likely to choose "Disagree strongly" in response to the phrase "Abortion should be legal".

Figure 6.1: Beer consumption by religion



Figure 6.2: Wine consumption by religion



Figure 6.3: Attitudes towards abortion



Figure 6.4: Attitudes towards hookups



93

In terms of how they allocate their time, however, differences between religious and atheist students are less pronounced. They spend roughly equal hours on homework and partying.



Figure 6.5: Hrs/week on homework





6.2 How do students on campuses with a nearby Planned Parenthood clinic differ from those without?

As might be expected, students that attend universities in counties with at least one Planned Parenthood clinic look on abortion more favorably. What is rather more surprising is that, contrary to the idea that universities are centers of a "bacchanalian lifestyle", a majority of students disapprove of casual sex: nearly 62% reported either mild or strong disagreement overall. Here too, however, the presence of Planned Parenthood in the county correlates with a decreased rate of strong disapproval of casual sex, and small increases in the proportion of respondents reporting some level of favorability towards it. For both attitudinal variables, the presence or absence of Planned Parenthood appears to affect the incidence of strong approval or disapproval, while rates of mild approval or disapproval remain fairly similar.



Figure 6.7: Attitudes to abortion by campus

On the other hand, the availability of Planned Parenthood seems largely orthogonal to drinking behavior: a majority of respondents report no drinking at all (consistent with both the U.S. legal drinking age and the large proportion of campuses that ban alcohol on the premises). Less than ten percent of respondents report frequent consumption of wine or beer in the past year. Of course, it is consistent with our hypothesis that alcohol is a disinhibitor that students might resort to it only occasionally rather than frequently.

Figure 6.9: Beer consumption by campus

Figure 6.10: Wine consumption by campus



Table 6.1 therefore adds to our preferred specification the percentage of students on a campus that report strongly disapproving of casual sex. As this question was only part of the survey in 1997-2001 and 2004-5 – just 20% of our total dataset – we take the average percentage of respondents over the entire period. The coefficient is

Figure 6.8: Attitudes towards hookups by campus

	Incidents	w. alcohol	Incidents	w.o. alcohol	Total in	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
% Strongly disapproving	0 701	(0.00)	0.005	(0.00)		(0.07)
of casual sex	0.531 +	(0.32)	0.225	(0.29)	0.258	(0.25)
No. of Planned	0.910*	(0,00)	0.010	(0.05)	0.010	(0, 0, 4)
Parenthood clinics	-0.210	(0.09)	0.019	(0.05)	-0.019	(0.04)
National Collegiate Athletic	9 386***	(0.36)	1 071***	(0.20)	1 202***	(0.17)
Association membership	2.000	(0.50)	1.071	(0.20)	1.202	(0.11)
Religious affiliation	0.824^{*}	(0.35)	0.256	(0.26)	0.356	(0.24)
Dry campus	-0.286+	(0.16)	-0.111	(0.16)	-0.147	(0.13)
Beer tax	-0.301	(0.36)	0.223	(0.24)	0.082	(0.22)
Log of total enrolment	0.605^{***}	(0.09)	0.734^{***}	(0.07)	0.656^{***}	(0.06)
Freshman residency	0.509^{**}	(0.16)	0.418**	(0.14)	0.409***	(0.12)
requirement		(0120)	0.110	(*****)		(**==)
Single-gender dorms	-0.061	(0.16)	0.253 +	(0.15)	0.165	(0.13)
offered						
Single-gender dorms	-0.098	(0.39)	-0.014	(0.34)	-0.027	(0.31)
only Canden inclusive				. ,		
bousing	0.362^{*}	(0.15)	0.338^{*}	(0.16)	0.314^{*}	(0.13)
Lagrad incidents of						
sevual assault w.o. alcohol	0.175^{***}	(0.06)				
Lagged incidents of						
sexual assault w alcohol			0.197^{***}	(0.06)		
Lagged incidents of						
sexual assault					0.217^{***}	(0.03)
Blue county	0.151	(0.15)	0.177 +	(0.09)	0.148 +	(0.09)
Red county	-0.229	(0.16)	-0.292*	(0.13)	-0.259*	(0.10)
Month FE	У	TES Í		YES	Y	ES
Year FE	λ	ZES		YES	Y	ES
Regional dummies	λ	YES		YES	Y	ES
Constant	-9.541^{***}	(0.94)	-9.738***	(0.69)	-8.665***	(0.62)
N	35	5,979	3	5,979	35.	979
$p > \chi^2 $	0.	0000	0	0.0000	0.0	000
α	5.348***	(0.77)	1.650^{***}	(0.27)	1.519^{***}	(0.21)
+ p < 0.10, * p < 0.05, ** p	< 0.01, ***	p < 0.001				

Table 6.1: Regressions with HERI attitudinal data $% \mathcal{T}_{\mathrm{A}}$

positive as expected in the regression for incidents with alcohol, though significant only at the 10% level. It is insignificant in the other regressions - for incidents without alcohol and for total incidents. Given that attitudes evolve over time but our observations are averaged over the whole period, there is substantial measurement error in this variable and it is not surprising that only weak effects are observed.

Surprisingly, though, including the disapproval variable increases substantially (by nearly a half) the coefficient on religious affiliation. However, this appears to be due to the fact that limitations on the availability of HERI data restrict the sample size (to around 36,000 instead of 45,000 in our main specification), and it is this restriction that results in an increased coefficient on religious campuses. To see this, we report in Table 2.A.9 in the Appendix our main specification on the restricted sample - this shows that that increased coefficient is entirely the result of the sample size restriction and not at all to the fact of controlling for student attitudes.

7 Conclusions and policy implications

There has been growing awareness in recent years that sexual assault on college campuses, like sexual assault in many other contexts, is a major social problem that requires careful analysis and evidence-based policy prescriptions. The hypothesis that a culture of sexual permissiveness has contributed to the extent of this problem is one that has some initial plausibility, but such evidence as we have been able to collect provides no support for it. On the contrary, our findings suggest that, in the presence of a disinhibiting mechanism such as easy availability of alcohol, stricter norms against consensual sex are associated with somewhat more sexual assaults in which alcohol is implicated, and our behavioral model provides some grounds for thinking that such an association is causal. Nevertheless, we acknowledge that the evidence for our alternative hypothesis remains suggestive rather than definitive at this point - in particular we have not been able to find evidence that links student attitudes to consensual sex directly to the consumption of alcohol as a disinhibitor. This seems to reflect shortcomings in the available data that might enable such a link to be investigated - absence of evidence does not constitute evidence of absence - and it remains an important subject for future research.

It is important to make clear that, even if our hypothesis were more strongly sup-

ported by the data, this would not mean that colleges (religious or secular) would be wrong to implement stricter norms against consensual sex; that is a choice they might wish to make on a variety of other grounds. What it means is that, unless the implementation of these norms is accompanied by severe restrictions on the availability of alcohol (and perhaps of some analogous institutions such as fraternities), they may have damaging side-effects in an increased incidence of alcohol-fuelled cases of sexual assault.

Alongside restrictions on access to alcohol, various policy measures may be able to make more salient the possible consequences of alcohol for the risks of engaging in sexual assault. There may be valuable lessons to be drawn from the history of campaigns against drunken driving in many countries, which have radically changed the perceived social acceptability of consuming alcohol in any context where individuals may subsequently need to drive (see Potter, 2016). Reducing the social acceptability of consuming alcohol in contexts in which students may wish to engage in consensual sex would seem to be highly desirable - but probably easier to achieve if consensual sex is not itself considered a socially unacceptable activity.

Bibliography

- Abbey, A. (2002). Alcohol-related sexual assault: A common problem among college students. Journal of studies on alcohol. Supplement, 14, 118–28. https://doi. org/10.15288/jsas.2002.s14.118.
- Associated Press. (2019). Ex-pope Benedict XVI blames sexual abuse on swinging sixties. *The Guardian*. https://www.theguardian.com/world/2019/apr/11/ ex-pope-benedict-xvi-blames-sexual-abuse-on-swinging-sixties
- Bouffard, J. (2000). Predicting type of sexual assault case closure from victim, suspect, and case characteristics. Journal of Criminal Justice, 28, 527–542. https://doi.org/10.1016/S0047-2352(00)00068-4
- Brecklin, L. R., & Ullman, S. E. (2010). The Roles of Victim and Offender Substance Use in Sexual Assault Outcomes. Journal of Interpersonal Violence, 25(8), 1503–1522. https://doi.org/10.1177/0886260509354584
- Bureau of Justice Statistics. (1992-2019). National Crime Victimization Survey. https://bjs.ojp.gov/data-collection/ncvs
- Cantalupo, N. (2011). Burying Our Heads in the Sand: Lack of Knowledge, Knowledge Avoidance and the Persistent Problem of Campus Peer Sexual Violence. Loyola University Chicago Law Journal, (43), 205–266.
- Cantor, D., Fisher, B., Chibnall, S., Bruce, C., Townsend, R., Thomas, G., & Lee,
 H. (2015). Report on the AAU Campus Climate Survey on Sexual Assault and Sexual Misconduct. Association of American Universities (AAU).
- Davis, K. C. (n.d.). The influence of alcohol expectancies and intoxication on men's aggressive unprotected sexual intentions. *Experimental Clinical Psychophar*macology, 18(5), 418–428. https://doi.org/10.1037/a0020510
- Douthat, R. (2014). Rape and the College Brand. New York Times. https://www. nytimes.com/2014/05/11/opinion/sunday/douthat-rape-and-the-collegebrand.html

- Du Mont, J., Miller, K.-L., & Myhr, T. (2003). The Role of "Real Rape" and "Real Victim" Stereotypes in the Police Reporting Practices of Sexually Assaulted Women. Violence Against Women, 9, 466–486. https://doi.org/10.1177/1077801202250960
- Ellingsen, T., & Mohlin, E. (2019). Decency. Working Paper, Department of Economics, Lund UNiversity.
- England, P., & Bearak, J. (2014). The sexual double standard and gender differences in attitudes toward casual sex among U.S. university students. *Demographic Research*, 30, 1327–1338. https://doi.org/10.4054/DemRes.2014.30.46
- Federal Bureau of Investigation. (1997-2019). National Incident Based Reporting System. https://www.fbi.gov/services/cjis/ucr/nibrs
- Fisher, B. S., Cullen, F. T., & Turner, M. G. (2000). The Sexual Victimization of College Women.
- Fisher, B. S., Daigle, L. E., & Cullen, F. T. (n.d.). 'What Distinguishes Single from Recurrent Sexual Victims? The Role of Lifestyle-Routine Activities and First-Incident Characteristics. *Justice Quarterly*, 27(10), 102–129.
- Fisher, B. S., Daigle, L. E., & Cullen, F. T. (2010). Unsafe in the Ivory Tower: The Sexual Victimization of College Women. SAGE Publications.
- Fisher, B. S., Daigle, L. E., Cullen, F. T., & Turner, M. (2003). Reporting sexual victimization to the police and others: Results from a national-level study of college women. *Criminal Justice and Behavior*, 30(1), 6–38.
- Franklin, C. (2010). Physically forced, alcohol-induced, and verbally coerced sexual victimization: Assessing risk factors among university women. Journal of Criminal Justice, 38, 149–159. https://doi.org/10.1016/j.jcrimjus.2010.02. 004
- Franklin, C. A., Bouffard, L. A., & Pratt, T. C. (2012). Sexual Assault on the College Campus: Fraternity Affiliation, Male Peer Support, and Low Self-Control. Criminal Justice and Behavior, 39(11), 1457–1480. https://doi. org/10.1177/0093854812456527
- Frazier, P. A., & Haney, B. (1996). Sexual assault cases in the legal system: Police, prosecutor, and victim perspectives. Law and Human Behavior, 20(6), 607– 628. https://doi.org/10.1007/BF01499234
- Gray, J. A., & McNaughton, N. (2003). The neuropsychology of anxiety: An enquiry into the function of the septo-hippocampal system. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198522713.001.0001

- Heath, N., Lynch, S., Fritch, A., & Wong, M. (2013). Rape Myth Acceptance Impacts the Reporting of Rape to the Police: A Study of Incarcerated Women. Violence against women, 19, 1065–78. https://doi.org/10.1177/1077801213501841
- Higher Education Research Institute. (1997-2018). CIRP Freshman Survey. https://heri.ucla.edu/
- Horney, J., & Spohn, C. (2006). The influence of blame and believability factors on the processing of simple versus aggravated rape cases. *Criminology*, 34, 135–162. https://doi.org/10.1111/j.1745-9125.1996.tb01200.x
- James, V., & Lee, D. (2014). Through the looking glass. Journal of interpersonal violence, 30(14), 2447–2469. https://doi.org/10.1177/0886260514553116
- Kilpatrick, D. G., Resnick, H. S., Ruggiero, K. J., Conoscenti, L. M., & McCauley, J. (2007). Drug-facilitated, incapacitated, and forcible rape: A national study. National Criminal Justice Reference Service Charleston, SC.
- LaFree, G. D. (1980). The Effect of Sexual Stratification By Race on Official Reactions to Rape. American Sociological Review, 45(5), 842–854. http://www. jstor.org/stable/2094898
- Lafree, G. D. (1981). Official Reactions to Social Problems: Police Decisions in Sexual Assault Cases. Social Problems, 28(5), 582–594.
- Leip, D. (1996-2019). Dave Leip's Atlas of U.S. Presidential Elections. https://uselectionatlas.org/
- Lindo, J. M., Siminski, P., & Swensen, I. D. (2018). College Party Culture and Sexual Assault. American Economic Journal: Applied Economics, 10(1), 236–265.
- McCauley, J., Ruggiero, K. J., Resnick, H. S., Conoscenti, L. M., & Kilpatrick, D. G. (2009). Forcible, drug–facilitated, and incapacitated rape in relation to substance use problems: Results from a national sample of college women. *Addictive Behaviors*, 34(5), 458–462. https://doi.org/10.1016/j.addbeh. 2008.12.004
- Monto, M., & Carey, A. (2014). A New Standard of Sexual Behavior? Are Claims Associated With the "Hookup Culture" Supported by General Social Survey Data? Journal of sex research, 51(6), 1–11. https://doi.org/10.1080/ 00224499.2014.906031
- National Center for Education Statistics. (1997-2019). Integrated Postsecondary Education Data System (IPEDS). U.S. Department of Education. https://nces. ed.gov/ipeds/datacenter/DataFiles.aspx
- O'Flaherty, B., & Sethi, R. (2019). Shadows of doubt: Stereotypes, crime, and the pursuit of justice. https://doi.org/10.4159/9780674240162

- Orchowski, L., & Gidycz, C. (2012). To Whom Do College Women Confide Following Sexual Assault? A Prospective Study of Predictors of Sexual Assault Disclosure and Social Reactions. Violence against women, 18, 264–88. https: //doi.org/10.1177/1077801212442917
- Paul, L., Walsh, K., Mccauley, J., Ruggiero, K., Resnick, H., & Kilpatrick, D. (2013). College Women's Experiences With Rape Disclosure: A National Study. Violence against women, 19(4), 486–502. https://doi.org/10.1177/ 1077801213487746
- Potter, S. J. (2016). Reducing Sexual Assault on Campus: Lessons From the Movement to Prevent Drunk Driving. American journal of public health, 106(5), 822–9. https://doi.org/10.2105/AJPH.2016.303082
- Sabina, C., & Ho, L. (2014). Campus and College Victim Responses to Sexual Assault and Dating Violence: Disclosure, Service Utilization, and Service Provision. *Trauma, violence & abuse*, 15, 535–551. https://doi.org/10.1177/ 1524838014521322
- Schuller, R. A., & Stewart, A. (2000). Police Responses to Sexual Assault Complaints: The Role of Perpetrator/Complainant Intoxication. Law and Human Behavior, 24(5), 535–551. https://doi.org/10.1023/A:1005519028528
- Spohn, C., & Holleran, D. (2001). Prosecuting Sexual Assault: A Comparison of Charging Decisions in Sexual Assault Cases Involving Strangers, Acquaintances, and Intimate Partners. Justice Quarterly, 18, 651–688. https://doi. org/10.1080/07418820100095051
- Spohn, C. C., & Tellis, K. (2011). Justice Denied? The Exceptional Clearance of Rape Cases in Los Angeles. Albany law review, 74, 1379.
- Stacey, M., Martin, K., & Brick, B. (2016). Victim and Suspect Race and the Police Clearance of Sexual Assault. Race and Justice, 7, 226–255. https://doi.org/ 10.1177/2153368716643137
- Stappenbeck, C., & Fromme, K. (2013). The Effects of Alcohol, Emotion Regulation, and Emotional Arousal on the Dating Aggression Intentions of Men and Women. Psychology of addictive behaviors : Journal of the Society of Psychologists in Addictive Behaviors, 28(1), 10–19. https://doi.org/10.1037/ a0032204
- Stotzer, R., & MacCartney, D. (2015). The Role of Institutional Factors on On-Campus Reported Rape Prevalence. Journal of Interpersonal Violence, 31(16), 2687–2707. https://doi.org/10.1177/0886260515580367

- Tasca, M., Rodriguez, N., & Spohn, C. (2012). Police Decision Making in Sexual Assault Cases: Predictors of Suspect Identification and Arrest. Journal of interpersonal violence, 28, 1157–1177. https://doi.org/10.1177/ 0886260512468233
- Twenge, J., Sherman, R., & Wells, B. (2015). Changes in American Adults' Sexual Behavior and Attitudes, 1972–2012. Archives of sexual behavior, 44, 2273– 2285. https://doi.org/10.1007/s10508-015-0540-2
- Weiss, K. (2011). Neutralizing sexual victimization: A typology of victims' nonreporting accounts. *Theoretical Criminology*, 15, 445–467. https://doi.org/ 10.1177/1362480610391527
- Zinzow, H., & Thompson, M. (2011). Barriers to Reporting Sexual Victimization: Prevalence and Correlates among Undergraduate Women. Journal of Aggression, Maltreatment & Trauma, 20(7), 711–725. https://doi.org/10.1080/ 10926771.2011.613447

Appendix

2.A Supplementary tables

Table 2.A.1: NCVS variable definitions
--

Variable	Definition
Victim reported crime	Dummy, $=1$ if victim reported crime to police themselves
Police involved	Dummy, $=1$ if police were made aware of the incident by any means
Offender is stranger	Dummy, $=1$ if victim and perpetrator not acquainted
White offender	Offender is white (available for single offenders starting 2012Q1)
Black offender	Offender is black (available for single offenders starting 2012Q1)
White victim	Victim is white (available starting 2003Q1)
Black victim	Victim is black (available starting 2003Q1)
Weapon used	Dummy, $=1$ if we apon was used to threaten or harm victim
College	Dummy, $=1$ if victim has bachelor's degree or higher
Urban	Dummy, $=1$ if crime occurred in urban area
Northeast	Dummy, $=1$ if crime occurred in the Northeast
Midwest	Dummy, $=1$ if crime occurred in the Midwest
South	Dummy, $=1$ if crime occurred in the South
West	Dummy, $=1$ if crime occurred in the West

	Police inv	volvment	Self-reporting		
	Coef.	Std. Err.	Coef.	Std. En	
Alcohol involved	-0.010	(0.06)	-0.048	(0.06)	
Offender is stranger	0.393***	(0.09)	0.192^{*}	(0.09)	
Offender is stranger*alcohol	-0.278+	(0.14)	-0.341^{*}	(0.16)	
White victim	0.082	(0.07)	0.053	(0.07)	
Black offender	0.244^{***}	(0.07)	0.072	(0.07)	
Weapon used	0.458^{***}	(0.07)	0.343^{***}	(0.08)	
Post-secondary education	-0.177***	(0.05)	-0.068	(0.06)	
College	-0.222*	(0.09)	-0.147	(0.10)	
Urban area	-0.203**	(0.07)	-0.154*	(0.08)	
Northeast	-0.030	(0.10)	-0.138	(0.11)	
South	0.065	(0.07)	0.019	(0.08)	
West	0.061	(0.08)	0.089	(0.09)	
Year of incident	YES		YES		
N	3,018		3,018		

 Table 2.A.2: NCVS probit marginal effects - cases of rape

	Incidents w. alcohol		Incidents w.o. alcohol		Total incidents			
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.		
No. of Planned Parenthood clinics	-0.227***	(0.07)	0.015	(0.04)	-0.026	(0.04)		
National Collegiate Athletic Association membership	1.912***	(0.34)	0.998***	(0.17)	1.130***	(0.17)		
Religious affiliation	0.515	(0.36)	0.170	(0.21)	0.252	(0.21)		
Dry campus (instrumented)	-0.462	(0.77)	-0.067	(0.45)	-0.145	(0.45)		
Log of total enrolment	0.484^{***}	(0.08)	0.618^{***}	(0.05)	0.586^{***}	(0.05)		
Freshman residency requirement	0.377**	(0.14)	0.246**	(0.09)	0.277**	(0.09)		
Single-gender dorms offered	0.037	(0.14)	0.209 +	(0.11)	0.175	(0.11)		
Single-gender dorms only	-0.363	(0.48)	-0.011	(0.28)	-0.081	(0.28)		
Gender-inclusive housing	0.323 +	(0.19)	0.303*	(0.13)	0.308*	(0.13)		
Lagged incidents of sexual assault w.o. alcohol	0.184***	(0.05)						
Lagged incidents of sexual assault w. alcohol			0.256***	(0.04)				
Lagged incidents of sexual assault					0.248***	(0.04)		
	0.084	(0.14)	0.163 +	(0.08)	0.144 +	(0.08)		
Red county	-0.191	(0.16)	-0.150	(0.11)	-0.157	(0.11)		
Month FE	YES		YES		YES			
Year FE	YES		YES		YES			
Regional Dummies	YES		YES		YES			
Constant	-8.303***	(0.88)	-8.317***	(0.54)	-7.763***	(0.52)		
N	45,602		45,602		45,602			
$p > \chi^2 $	0.0000		0.0000		0.0000			
α	5.643***	(0.75)	1.651^{***}	(0.20)	1.573***	(0.17)		
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001								

Table 2.A.3: IV Negative binomial - incidents of sexual assault
	Incidents	w. alcohol	Incidents	w.o. alcohol	Total in	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
No. of Planned Parenthood clinics	-0.215**	(0.07)	-0.008	(0.05)	-0.051	(0.04)
National Collegiate Athletic Association membership	1.867***	(0.37)	1.490***	(0.23)	1.538***	(0.22)
Religious affiliation	0.171	(0.36)	0.188	(0.28)	0.158	(0.25)
Dry campus	-0.239+	(0.14)	-0.185	(0.11)	-0.190+	(0.10)
Beer tax	-0.046	(0.30)	0.207	(0.18)	0.131	(0.17)
Log of total enrolment	0.467^{***}	(0.09)	0.596^{***}	(0.07)	0.529^{***}	(0.07)
Freshman residency requirement	0.398**	(0.14)	0.354**	(0.11)	0.347***	(0.10)
Single-gender dorms offered	0.099	(0.14)	0.234 +	(0.12)	0.190 +	(0.11)
Single-gender dorms only	-0.380	(0.41)	-0.278	(0.36)	-0.286	(0.30)
Gender-inclusive housing	0.355*	(0.16)	0.341*	(0.16)	0.336*	(0.14)
Lagged incidents of rape w.o. alcohol	0.311***	(0.07)				
Lagged incidents of rape w. alcohol			0.269***	(0.07)		
Lagged incidents of rape					0.316***	(0.03)
Blue county	0.042	(0.13)	0.075	(0.09)	0.065	(0.08)
Red county	-0.090	(0.14)	-0.212+	(0.12)	-0.154	(0.10)
Month FE	У	ZES		YES	Y	ES
Year FE	γ	ΎES		YES	Y	\mathbf{ES}
Regional Dummies	У	ZES		YES	Y	\mathbf{ES}
Constant	-8.655***	(0.95)	-9.011***	(0.63)	-8.010***	(0.61)
N	45	5,602	4	15,602	45,	602
$p > \chi^2 $	0.	0000		_	0.0	000
α	6.714^{***}	(0.98)	2.387***	(0.34)	2.112***	(0.25)
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001						

Table 2.A.4: Main regressions - incidents of rape, monthly data

	Incidents w. alcohol		Incidents	w.o. alcohol	Total ir	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Planned Parenthood in county	-0.350*	(0.17)	0.140	(0.13)	0.047	(0.12)
National Collegiate Athletic Association membership	1.839***	(0.33)	1.009***	(0.18)	1.079***	(0.16)
Religious affiliation	0.441	(0.36)	0.167	(0.21)	0.202	(0.20)
Dry campus	-0.257+	(0.15)	-0.150	(0.10)	-0.158+	(0.09)
Beer tax	-0.176	(0.30)	0.165	(0.19)	0.073	(0.17)
Log of total enrolment	0.474***	(0.09)	0.653***	(0.06)	0.568^{***}	(0.05)
Freshman residency requirement	0.426**	(0.14)	0.274**	(0.10)	0.291**	(0.09)
Single-gender dorms offered	0.095	(0.14)	0.233 +	(0.13)	0.188 +	(0.11)
Single-gender dorms only	-0.311	(0.47)	-0.022	(0.29)	-0.074	(0.27)
Gender-inclusive housing	0.397*	(0.16)	0.334*	(0.14)	0.321*	(0.13)
Lagged incidents of sexual assault w.o. alcohol	0.191***	(0.05)				
Lagged incidents of sexual assault w. alcohol			0.192***	(0.05)		
Lagged incidents of sexual assault					0.234***	(0.03)
Blue county	0.075	(0.13)	0.155^{*}	(0.08)	0.120 +	(0.07)
Red county	-0.209	(0.14)	-0.149	(0.09)	-0.142+	(0.08)
Month FE	Σ	YES		YES	Y	ES
Year FE	Σ	YES		YES	Y	ES
Regional Dummies	У	YES		YES	Y	ES
Constant	-8.277***	(0.89)	-8.753***	(0.56)	-7.673***	(0.51)
Ν	45	5,924	4	5,924	45,	924
$p > \chi^2 $	0.	0000	0	.0000	0.0	000
α	5.796***	(0.78)	1.742***	(0.22)	1.581***	(0.17)
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$						

 Table 2.A.5: Regressions using Planned Parenthood availability

	Incidents	w. alcohol	Incidents w.o. alcohol		Total ir	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
No. of Planned	-0.225***	(0.07)	0.006	(0.03)	-0.033	(0.03)
National Collegiate Athletic	1 020***	(0.99)	1 095***	(0.17)	1 161***	(0.16)
Association membership	1.950	(0.33)	1.025	(0.17)	1.101	(0.10)
Religious affiliation	0.514	(0.36)	0.181	(0.21)	0.264	(0.21)
Dry campus	-0.232+	(0.14)	-0.146	(0.10)	-0.167+	(0.10)
Beer tax	-0.165	(0.31)	0.156	(0.18)	0.088	(0.18)
Weekend	0.573^{***}	(0.06)	0.178^{***}	(0.04)	0.262^{***}	(0.04)
Halloween	0.993^{**}	(0.31)	0.506^{**}	(0.15)	0.635^{***}	(0.15)
Log of total enrolment	0.475^{***}	(0.08)	0.633^{***}	(0.05)	0.597^{***}	(0.05)
Freshman residency requirement	0.351**	(0.13)	0.243**	(0.09)	0.266**	(0.09)
Single-gender dorms offered	0.034	(0.14)	0.194 +	(0.12)	0.161	(0.11)
Single-gender dorms only	-0.390	(0.49)	0.000	(0.27)	-0.077	(0.28)
Gender-inclusive	0.363*	(0.15)	0.281*	(0.14)	0.294*	(0.13)
Lagged incidents of sexual assault w.o. alcohol	0.180***	(0.05)				
Lagged incidents of sexual assault w. alcohol			0.226***	(0.03)		
Lagged incidents of sexual assault					0.219***	(0.03)
Blue county	0.122	(0.12)	0.162^{*}	(0.07)	0.150^{*}	(0.07)
Red county	-0.190	(0.14)	-0.150+	(0.09)	-0.159+	(0.08)
Month FE	Y	ES	Ţ	YES	Y	ES
Year FE	Y	ES	,	YES	Y	ES
Regional Dummies	Y	ES	T.	YES	Y	ES
Constant	-11.882***	(0.87)	-11.893***	(0.51)	-11.339***	(0.50)
N	1,38	4,420	1,3	84,420	1,384	4,420
$p > \chi^2 $	0.0	0000	0.	.0000	0.0	000
α	100.6^{***}	(16.7)	21.40^{***}	(3.90)	19.86^{***}	(3.23)
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001						

 $\textbf{Table 2.A.6:} \ \text{Main regressions - incidents of sexual assault, daily data}$

	Incident	s w. alcohol	Incidents	w.o. alcohol	Total in	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
No. of Planned Parenthood clinics	-0.012**	(0.00)	0.001	(0.01)	-0.010	(0.01)
National Collegiate Athletic Association membership	0.027*	(0.01)	0.038	(0.03)	0.061 +	(0.04)
Religious affiliation	0.101^{*}	(0.05)	0.058	(0.07)	0.149	(0.09)
Dry campus	-0.018+	(0.01)	-0.021	(0.02)	-0.036	(0.03)
Beer tax	0.001	(0.03)	0.071	(0.07)	0.067	(0.08)
Log of total enrolment	-0.021**	(0.01)	-0.063***	(0.01)	-0.078***	(0.02)
Freshman residency requirement	0.023*	(0.01)	0.052*	(0.03)	0.070*	(0.03)
Single-gender dorms offered	-0.000	(0.01)	0.045 +	(0.03)	0.041	(0.03)
Single-gender dorms only	-0.038	(0.03)	-0.066	(0.07)	-0.097	(0.07)
Gender-inclusive housing	0.022	(0.02)	0.067 +	(0.04)	0.082 +	(0.04)
Lagged incidents of sexual assault w.o. alcohol p.c.	0.021	(0.01)				
Lagged incidents of sexual assault w. alcohol p.c.			0.057	(0.03)		
Lagged incidents of sexual assault p.c.					0.101**	(0.03)
Blue county	-0.002	(0.01)	0.027	(0.02)	0.023	(0.02)
Red county	-0.002	(0.01)	-0.027	(0.02)	-0.027	(0.03)
Month FE		YES		YES	YI	ES
Year FE		YES		YES	YI	ES
Regional Dummies		YES		YES	YI	\mathbf{ES}
Constant	0.191^{*}	(0.08)	0.576^{***}	(0.13)	0.715^{***}	(0.17)
N	4	5,602	4	5,602	45,	602
$p > F^2 $	0	.0000	0	.0000	0.0	000
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001						

 Table 2.A.7:
 Main regressions - incidents of sexual assault per capita

	Incidents Coef.	w. alcohol Std. Err.	Incidents Coef.	w.o. alcohol Std. Err.	Total in Coef.	ncidents Std. Err.
Proportion disapproving				~~~~		Sta. 111.
strongly of casual sex	0.826*	(0.36)	0.154	(0.37)	0.262	(0.31)
Proportion disapproving strongly of abortion	-0.833+	(0.45)	0.181	(0.58)	-0.011	(0.47)
No. of Planned Parenthood clinics	-0.211*	(0.09)	0.018	(0.05)	-0.019	(0.04)
Athletic association membership	2.362***	(0.36)	1.077***	(0.19)	1.202***	(0.17)
Religious affiliation	0.778^{*}	(0.35)	0.269	(0.26)	0.355	(0.24)
Dry campus	-0.302+	(0.16)	-0.108	(0.15)	-0.148	(0.13)
Beer tax	-0.293	(0.36)	0.220	(0.24)	0.082	(0.22)
Log of total enrolment	0.587^{***}	(0.09)	0.739^{***}	(0.07)	0.656^{***}	(0.06)
Freshman residency requirement	0.535***	(0.16)	0.414**	(0.14)	0.409***	(0.12)
Single-gender dorms offered	-0.029	(0.16)	0.248	(0.16)	0.165	(0.13)
Single-gender dorms only	-0.130	(0.39)	-0.004	(0.34)	-0.027	(0.31)
Gender-inclusive housing	0.310*	(0.15)	0.349*	(0.16)	0.313*	(0.13)
Lagged incidents of sexual assault w.o. alcohol	0.174***	(0.06)				
Lagged incidents of sexual assault w. alcohol			0.198***	(0.06)		
Lagged incidents of sexual assault					0.217***	(0.03)
Blue county	0.156	(0.15)	0.177 +	(0.09)	0.148 +	(0.09)
Red county	-0.165	(0.16)	-0.305*	(0.14)	-0.258^{*}	(0.12)
Month FE	Y	YES		YES	Y	ES
Year FE	Ŋ	YES		YES	Y	ES
Regional dummies	Ŋ	YES		YES	Y	ES
Constant	-8.654***	(1.05)	-9.616***	(0.83)	-8.384***	(0.72)
N	32	2,657	3	32,657	32,	657
$p > \chi^2 $	0.	0000	C	0.0000	0.0	000
α	5.302***	(0.82)	1.733***	(0.27)	1.594***	(0.21)
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001						

Table 2.A.8: Regressions with HERI attitudinal data (II)

	Incidents w. alcohol		Incidents w.o. alcohol		Total incidents	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
No. of Planned	0.01.04	(0.00)	0.01.0	(0.07)	0.000	
Parenthood clinics	-0.218*	(0.08)	0.016	(0.05)	-0.023	(0.04)
National Collegiate Athletic	0 /10***	(0, 2C)	1 000***	(0,01)	1 001***	(0, 10)
Association membership	2.419	(0.30)	1.089	(0.21)	1.221	(0.18)
Religious affiliation	0.863^{*}	(0.34)	0.277	(0.26)	0.379	(0.24)
Dry campus	-0.216	(0.16)	-0.085	(0.14)	-0.117	(0.12)
Beer tax	-0.332	(0.37)	0.210	(0.25)	0.066	(0.23)
Log of total enrolment	0.613^{***}	(0.09)	0.739^{***}	(0.07)	0.662^{***}	(0.06)
Freshman residency	0 /3/**	(0.15)	0 385**	(0.12)	0 379***	(0, 10)
requirement	0.494	(0.15)	0.000	(0.12)	0.012	(0.10)
Single-gender dorms	-0.060	(0.16)	$0.252 \pm$	(0.15)	0 163	(0.13)
offered	-0.000	(0.10)	0.202	(0.10)	0.105	(0.13)
Single-gender dorms	-0.070	(0.39)	-0.003	(0.34)	-0.014	(0.31)
only	0.010	(0.00)	0.000	(0.01)	0.011	(0.01)
Gender-inclusive	0.370^{*}	(0.16)	0.336^{*}	(0.16)	0.314*	(0.14)
housing	0.010	(0120)	0.000	(0120)	0.011	(0111)
Lagged incidents of	0.175***	(0.06)				
sexual assault w.o. alcohol		()				
Lagged incidents of			0.199^{***}	(0.05)		
sexual assault w. alcohol						
Lagged incidents of					0.219***	(0.03)
sexual assault	0.179	(0, 10)	0.157	(0,00)	0.1.40	
Blue county	0.173	(0.12)	0.157 + 0.105 + 0.105 + 0.105 + 0.105 + 0.105 + 0.00	(0.08)	0.140+	(0.07)
Red county Month EE	-0.179	(0.13)	-0.185+	(0.10) VEC	-0.163+	(0.09)
Month FE	ز د	ES ZEC		YES	Y.	ES ES
Year FE	۲ ۲	ES VEC		YES	Y. V	ES ES
Regional dummies	ز ۸ 207***	(0.06)	0 604***	(0, 60)	Y. 9 600***	ES
Constant	-9.397	(0.96)	-9.094	(0.69)	-8.009	(0.62)
N	35	5,979	3	5,979	35,	979
$p > \chi^2 $	0.	0000	C	0.0000	0.0	000
α	5.367***	(0.77)	1.656**	(0.27)	1.526**	(0.21)
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$						

Table 2.A.9:Main regressions - restricted dataset

	Incidents	w. alcohol	Incidents	w.o. alcohol	Total in	ncidents
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
Proportion of black students	-0.901+	(0.53)	0.366	(0.37)	0.172	(0.33)
No. of Planned	-0.212**	(0.06)	0.008	(0.04)	-0.029	(0.03)
Parenthood clinics National Collogiate Athletic						
Association membership	1.880^{***}	(0.33)	1.037^{***}	(0.19)	1.113^{***}	(0.17)
Religious affiliation	0.584	(0.36)	0.246	(0.21)	0.294	(0.21)
Dry campus	-0.169	(0.13)	-0.191+	(0.10)	-0.173*	(0.09)
Beer tax	-0.104	(0.30)	0.136	(0.20)	0.054	(0.18)
Log of total enrolment	0.452***	(0.08)	0.673^{***}	(0.06)	0.580***	(0.05)
Freshman residency requirement	0.255*	(0.13)	0.278**	(0.10)	0.257**	(0.09)
Single-gender dorms offered	0.087	(0.13)	0.187	(0.12)	0.152	(0.10)
Single-gender dorms only	-0.440	(0.49)	-0.019	(0.29)	-0.090	(0.27)
Gender-inclusive housing	0.413**	(0.15)	0.316*	(0.15)	0.313*	(0.13)
Lagged incidents of sexual assault w.o. alcohol	0.177***	(0.05)				
Lagged incidents of sexual assault w. alcohol			0.190***	(0.05)		
Lagged incidents of sexual assault					0.229***	(0.03)
Blue county	0.173	(0.12)	0.157 +	(0.08)	0.140 +	(0.07)
Red county	-0.179	(0.13)	-0.185 +	(0.10)	-0.163 +	(0.09)
Month FE	У	ZES		YES	Y	ES
Year FE	γ	ZES		YES	Y	ES
Regional dummies	λ	ZES		YES	Y	ES
Constant	-9.332***	(0.91)	-9.564***	(0.58)	-8.473***	(0.51)
N	41	,463	4	1,463	41,	463
$p > \chi^2 $	0.	0000	0	0.0000	0.0	000
α	5.154^{***}	(0.68)	1.701***	(0.22)	1.502***	(0.17)
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$						

 Table 2.A.10: Regressions controlling for student demographics

2.B Other sources of U.S. data considered

National Women's Study [1989-1991]

This survey, conducted by the National Institute of Drug Abuse (NIDA), was one of the earliest surveys to provide information on the prevalence of sexual violence at the national level, using a nationally representative sample of 4,008 adult women. However, its lack of behaviorally-oriented questions - as well as its age - make it unsuited for our analysis.

National Violence Against Women Survey (NVAWS) [1995-6]

This one-off telephone survey conducted by the National Institute of Justice (NIJ) and the Center for Disease Control and Prevention (CDC) interviewed adults, both male and female, from across the U.S. It aimed to collect data not only on sexual violence but also physical assault suffered either as a child at the hands of a caretaker, or as an adult at the hands of a partner.

As the NVAWS asked respondents for lifetime incidence of these crimes, it is unsuitable for our analysis.

National College Women Sexual Victimization Survey [1997]

This survey interviewed a national sample of women attending a two- or four-year college. Each college included in the survey was selected randomly, with the probability of inclusion being proportional to female enrolment. This survey contained ten behaviorally-specific questions designed to establish whether respondents had been victims of a sexual crime, whether or not they themselves realized it.

National Intimate Partner and Sexual Violence Survey (NISVS) [2010-present]

The National Intimate Partner and Sexual Violence Survey is the successor to the NVAW. It is an ongoing, nationally representative survey collects detailed information about intimate partner violence, sexual violence and stalking, including victim and perpetrator characteristics and details about the context in which the crime occurred. However, while the CDC conducts this survey annually, the raw datasets are currently unavailable.

Campus Climate Survey on Sexual Assault and Sexual Misconduct [2010-2015]

This survey, developed by the Association of American Universities (AAU), aims to improve understanding of both the experiences and attitudes of students with respect to sexual assault and sexual misconduct. The raw data are, unfortunately, unavailable for download.

Chapter 3

Religion and State

Julia Hoefer Martí¹

Abstract

States the world over support religions in a variety of ways, both monetarily and legislatively, but the effects of this sort of state intervention in the market for religion are not well understood. Research focusing on population-level effects risks occluding the true picture: state support is usually geared towards a specific religion, and effects may vary depending on whether the religious institution an individual belongs to is supported.

I run a multilevel mixed-effects regression using individual-level data from the World Values Survey, along with state religious legislation data and a selection of countrylevel control variables. Stronger government support of religion is associated with increased attendance at religious service in both the short and long term, but in contrast is associated with significant decreases in both private faith and levels of institutional confidence among the faithful, as well as decreasing propensity to consider oneself religious or belong to a religious denomination. These negative effects fall on members of preferred religions, and are mostly felt only over the longer term.

Keywords: politics, public policy, religion

JEL codes: Z12, Z18, N30, N40

1 Introduction

What are the effects of government intervention in the market for religious services? What is the impact on the motivation and commitment of its adherents, and on their internal beliefs? Does government support make a religious movement more attractive to current potential followers – or might it instead cause people to turn away? In this paper, I explore the effects of state support of religion on individual public religiosity, private faith, and the perceived legitimacy of the religious institutions in question. Using data from the World Values Survey, supplemented by religious legislation data and a range of economic development indicators, I find that government support of religions is associated with significant decreases in measures of individuals' private religiosity and their confidence in religious leadership – negative effects that may be obscured by the fact that public religiosity *in*creases with government support.

For most of human history, the relation between religion and the state has tended to fall somewhere on the scale between "highly intertwined" and "one and the same". Even today, many states the world over support religion (usually *a* specific religion) to a greater or lesser degree. The reasons for this may be ideological, but they may also be *quid-pro-quo*: governments and political candidates often attempt to leverage religious moral authority into "a degree of legitimacy that the politicians struggle to achieve through more conventionally political means" (Seabright, 2024). Consider, for instance, how U.S. presidential hopefuls on both the left and the right often court endorsements from religious leaders^[1], or the state's increasing attempts to lean on Judaism for legitimacy in Israel (Abulof, 2014). In exchange for lending their moral authority, religions are often provided a range of benefits such as favorable taxation, or legal enforcement of religious precepts.

Given how commonplace this sort of arrangement is, one might be forgiven for thinking the benefits of such a relationship to both parties must clearly outweigh the costs – but in fact things are not so clear. Although state support of religion provides clear and immediate advantages, the potential costs to the religion and to its adherents are harder to measure, and may only be felt over the longer term. This matters because religion is, among other things, a service people pay for – both monetarily in the form of donations, and in terms of behavioral adjustments that can be quite costly, like avoiding certain foods entirely or not working on certain

¹Goodstein, 2009

days. States, too, have a vested interest in knowing whether their support of a religion is likely to ultimately drive people away from that religion. This is not only because religion and religiosity affect a wide range of economic outcomes the state has an interest in, such productivity, entrepreneurship, savings behavior, or tax avoidance (Guiso et al., 2003, Boone et al., 2013, Campante and Yanagizawa-Drott, 2015, Henley, 2017, Cantoni et al., 2018) – but also simply because the favors offered to religions are often very costly to the state itself. It is worth asking whether economics can shed light on the effects of this sort of state intervention in the "religious economy". The answer, it appears, is yes – but only up to a point.

2 Literature Review

If religion is a service, then government support may be thought of as a subsidy. Often this subsidy literally financial, in the form of ecclesiastical taxes, subsidies, etc., though this need not always be the case. Like any other subsidy, then government support should decrease the price and increase demand. Of course, this is not the full story: quality is often endogenous, even for the most mundane of goods. Where quality is endogenous, subsidies often end up decreasing not only the price of a product or service, but its quality as well. In particular, monopoly power decreases the incentives to maximize operational efficiency, leading managers to expend lower efforts to ensure quality and push down costs (Hicks, 1935). This "quiet life hypothesis" has been found to apply to topics as diverse as utilities provision, healthcare, transport, and journalism (Gertler, 1989, Leroch and Wellbrock, 2011, Gómez-Lobo, 2014, McRae, 2015). If the goal of a state's support is to turn a given religion into a monopolist in the "market of faith", it is unsurprising that the quality of the services – and Services – it offers should decrease (Mussa and Rosen, 1978).

But religion is *not* a service like healthcare or utilities: not only is faith is a sui generis good, but religions themselves also serve at once as social clubs, marriage markets, mutual aid networks, and more. It is not only the case that government subsidy may result in reduced effort on the part of religious leaders, it is also possible that the price of accessing a religious denomination itself acts as a method of sorting high-quality members from low-quality ones (see e.g. Carvalho, 2016). The more difficult or costly membership is, the higher the quality of members is likely to be. Further, utilization of a service by one member in a club can decrease the quality of

service available to the rest (Olson, 1965; Sandler and Tschirhart, 1997). Both Hicks' theory of the quiet-life monopolist and the club goods model predict a decrease in quality as a result of government subsidy of religion, the former through the efforts of leadership and the latter through the direct effects of a lower price; both theories began to be applied to the topic of religion starting in the late 20th century.

Proponents of religious competition theory – a group populated mainly by economists (Iannaccone, 1991; Stark and Iannaccone, 1994; Iannaccone et al., 1998) – argue that faith is a market in which different religious institutions compete amongst themselves for followers. Rather than levels of religious belief being lower in countries with greater separation between state institutions and religious ones, this separation in essence guarantees the market framework by which religions may compete amongst themselves (Gruber and Hungerman, 2008), allowing religions to compete more vigorously. This increased competitive pressure increases the incentive for religious leadership to make themselves more attractive to current and potential adherents, strengthening their followers' faith and ultimately causing them to gain more "customers" overall. State support of religion, on the other hand, has an enervating effect, decreasing incentives for supported religions to compete and the ability of unsupported religions to do so (Finke and Stark, 1998; Finke and Stark, 2003; Chaney, 2008). Many of the arguments brought to bear by these studies are compelling, but not all of the evidence supports the religious competition theory: one key drawback, identified by Iannaccone (1991) himself, is that the argument hinges on state regulation of religion, but early attempts to test this hypothesis empirically were unable to measure state regulation directly, relying on religious pluralism as a proxy. Some research has challenged the (often implicit) idea that religious pluralism is associated with higher religious participation at all Land et al., 1991 Chaves and Gorski, 2001, though the balance of evidence overall seems to favor the negative association between religious participation and state regulation of religion (McCleary and Barro, 2006a; McCleary and Barro, 2006b).

As much as research into religion in the 1990s was guided by Hicks' (1965) monopoly theory of reduced effort, the idea that government support of religion may end up having negative effects on the "quality" of that religion was not new to the 20^{th} century - as was indeed widely acknowledged by academics of the time. Religious liberty, and the idea that citizens desire separation of religious institutions and the state for the good of both, formed a key tenet of the liberal philosophies of the Enlightenment. This is the view espoused by Adam Smith in *The Wealth of* *Nations*: that the costs to a religion of state support would, in the medium- to longterm, outweigh any short-run benefits. Preachers of the dominant religion would, in his view, become "apt gradually to lose the good and bad qualities that gave them authority and influence with the lower ranks of people" (Smith, [1776] 2002), losing the ability to effectively proselytize and draw a crowd. Smith directly attributes this loss of ability to the reduced zeal of the clergy. The resulting gap in the market would in turn create an opportunity for new underdog religious movements, much more motivated and responsive to their adherents, to take advantage of. (At the time of Smith's writing, the underdogs in question were the Methodists, who were, in his own words, "much more in vogue" than the "learned and elegant" Anglican priests).

However, there is an additional factor which is unique to religion and which, though it has has been the topic of much discussion, has thus far failed to receive much empirical study. The nature of the religion-state relationship may affect not only the incentives and competitive pressures faced by various religious institutions; it may also affect the moral legitimacy and authority accorded to them. This aspect is left almost implicit in Smith's writing, but was highlighted by Alexis de Tocqueville. de Tocqueville argued the same cause and (eventual) effect as Smith through a different channel, identifying the political aspect of religious liberty, not the economic one, as the main driver of increased religious participation. Separation of Church and state, he argued, engenders religious pluralism, which in turn allows all aspects of a political conflict to be associated with a religion, preventing religion itself from being associated with certain political views or classes such as the bourgoisie (an understandable view for a Frenchman not four decades removed from the Revolution to hold). When a religion allies itself with the state, "it sacrifices the future for the present, and by gaining a power to which it has no claim, it risks its legitimate authority" (de Tocqueville and Reeve, [1985] 1889). Unfortunately, without the proper data, untangling these similar but distinct theories is all but impossible; which of the two explanations is credited depends ultimately on the author's preference (see e.g. Martin, [1991], which addresses an issue found in Iannaccone, [1991] – that countries above 80% Catholic varied wildly in levels of religious participation - and argues the political aspect as an explanation).

Advances in the quality and coverage of data measuring religious attitudes and behaviors in the late 20th and early 21st centuries has finally made it possible to delve into this question directly. Recent research suggests that states that support religion may indeed lose *their* moral authority as a result (Fox and Breslawski, 2023); it is worth asking whether the same holds for religions supported by states. Evidence from the U.S. indicates that the sharp rise in the number of people with no denominational preference may be less a matter of declining faith amongst individuals, and more a backlash against politically-driven religion (Hout and Fischer, 2014). This religious legitimacy theory forms the core of the analysis undertaken in this paper.

It is worth mentioning one more school of thought on religion and state – one which dominated most of the 20th century – before turning to the predictions made each of these competing explanations. Secularization theory holds that economic development and increasing education lead people to desire separation of the state and religious institutions, at the same time as it decreases the strength and prevalence of religious beliefs in general. As early as 1917 Max Weber predicted the "disenchantment of the world", whereby religion would be rendered ever more irrelevant in the face of the implacable advance of science, rationality, and economic development (Weber, 1920, 1917); Parsons, 1974). A variant of this theory held that while religion would not actually disappear, it would become an entirely private affair (Luckmann, 1967). Secularization was further regarded as an absorbing state: according to proponents of the theory, once it was achieved, a resurgence of religion in the public sphere was all but impossible (Lechner, 1991). The logical implication stemming from this theory is that, in economically developed countries, citizens may dislike state support of religion and may turn away from both if the strength of this support increases – but in less-developed and less-educated countries, this relationship may be weaker or even inverted. It suggests, at the very least, that any analysis of this question should take care to control for indicators of economic development.

What, then, should we expect to see in the data, according to each of the theories previously laid out? Both the incentives-based view of Adam Smith and religious legitimacy theory predict disparate effects on religions that are favored (and their adherents) versus those that are not. Legitimacy theory posits that the moral authority of religious leadership, and the validity of their edicts, becomes increasingly tarnished the more that religion receives support from the state. Government support should therefore be associated with lower measures of private religiosity and trust in religious leadership among members of preferred religions, even if public religiosity remains the same – as indeed it may. In fact, it is perfectly possible these individuals might participate *more*, if state support for a given religion reduces the effective costs or increases the prestige of doing so – but the attendees will still be less likely to let religious leaders tell them what to do, and may eventually end up pulling away from religion entirely.

In contrast, Smith's view implicitly assumes that individuals, "act how they feel": he makes a clear prediction of decreased public religiosity, private faith, and confidence in religious leadership among members of religions that are favored by their respective states. Religious movements which are *not* favored should benefit from their lumbering competition, drawing increased crowds of motivated followers. In essence, this view ignores the possibility that government support may have contradictory effects on different measures of religiosity. Also implicit in this argument is the prediction that the effects of government support should not be different for men versus women – something that is very much not the case in actuality, as we will see in Section [5]

If the secularization theory of religion holds true, on the other hand, we should expect to see government support be associated with *higher* levels of both private and public religiosity overall, though this link would not be causal: government support of religion would be associated with greater religiosity only insofar as government support of religion tends to happen in less-developed countries. Lastly, religious competition theory predicts that 'monopoly' status for a given religion does not simply distort the incentives that leaders of *that specific religion* face: discriminatory state support of religion lowers religious competition across the board. State support of religion, therefore, should be associated with a generalized decrease in measures of religiosity among members of both favored and non-favored religions.

Outcome		Secular	economy	Religious	economy	Secularization
variable		Exogenous quality	Endogenous quality	Club goods model	Legitimacy theory	
Public religiosity	Preferred	+	-	-	+/-	+
0 1	Non-preferred	=	+	-	=	+
Private religiosity	Preferred	+	-	-	-	+
	Non-preferred	=	+	-	=	+
Legitimacy	Preferred	N/A	-	N/A	-	N/A
	Non-preferred	N/A	=	N/A	=	N/A
Extensive margin		+	+/-	-	-	+

 Table 2.1: Predicted effects of government support of religion

The rest of this paper is structured as follows. Section 3 below lays out the empirical approach to be used and outlines the variables used to measure public and private religiosity, and religious legitimacy. Section 4 covers a number of descriptive statistics, before the empirical results and their implications for the different theories mentioned above are discussed in Section 5 Section 6 concludes and discusses future avenues of research.

3 Data and Methodology

There are a number of ways in which public religiosity, private faith, and religious legitimacy and moral authority might be measured. This section outlines the outcome variables, variable of interest, and control variables used in the preferred estimation, before turning to the empirical approach.

The outcome variables used in this study are sourced from the Integrated Values Survey (IVS), which is a merging of the World Values Survey (Haerpfer et al., 2022) and European Values Survey (EVS, 2021), repeated cross-sectional surveys that began in the 1980s and run to the present day. This dataset contains information on individuals' values, attitudes, and behaviors.

Public religiosity is best conceptualized as the extent an individual's visible participation in religion. While the IVS does have some data on volunteering for religious organizations, charitable giving, etc., the highest quality data by far is information on how often individuals attend service on a 0-10 ordinal scale, with the lowest frequency being "never or almost never" and the highest being "once per week or more". Private religiosity is the perhaps the most challenging of the three outcome variables to evaluate, in part because the concept of faith is highly personal. The Integrated Values Survey contains several questions which might be used to evaluate an individual's private religiosity; in this analysis, I use data on how important individuals consider God to be in their lives on a 0-10 scale. The importance an individual accords to religion is a potential viable alternative measure, but it is more ambiguous: there are many ways in which a religion can be important to an individual that have very little to do with faith itself, as mentioned in Section [2]. This multipurpose nature makes the importance of religion more difficult to interpret than importance of God. Finally, religious legitimacy and moral authority is derived from individuals' trust in religious leadership and their guidance, making individuals' self-reported confidence in the Church/Temple/Mosque in their country (on a 0-4 ordinal scale) a valid measure.

Variable	Definition
Confidence: Church	0-4 ordinal (4 = complete confidence; $0 =$ none at all)
Attendance at service	0-10 ordinal (0 = never or almost never; $10 = $ once a week or more)
Importance of God	0-10 ordinal $(0 = \text{not at all important}; 10 = \text{extremely}$
1	1 important 1 dummy = 1 if individual considers themselves a
Religious person	religious person
Denomination member	dummy, $=1$ if individual belongs to a religious denomination

 Table 3.1: Outcome variables

While the legitimacy theory does not make clear predictions in this regard, it is nevertheless worthwhile to look at the effect state support of religion has at the extensive margin: does it increase the number of faithful in the population, or does it cause people to turn away from religion entirely? While the legitimacy theory and Smith's theory of incentives are ultimately agnostic as to the effects of state support of religion on global levels of religious adherence, the secularization and religious competition theories do make clear predictions as to population effects. In particular, secularization holds that state support of religion should be found in lessdeveloped countries, where more of the population tends to be religious. Religious competition, on the other hand, predicts that separation of state and religious institutions ultimately results in a greater portion of the population finding a religion they like. World Values Survey data contains two variables that might be used to measure this: the first is whether they consider themselves to be a religious person; the second is whether they report affiliation to any religious denomination. Together, these outcome variables help shed light on the effects on the whole population, and whether they represent a true case of "losing one's religion", or dissatisfaction with specific religious leadership.

Any investigation into the effects of government support of religion requires a measure of the strength of that support. In the past, research has tended to use binary variables indicating presence of a state religion, or any regulation of religion on the part of the state (see e.g. McCleary and Barro, 2006b). However, not all states that have official religions enforce them to the same degree; considering government regulation of religion as a binary state technically puts Sweden and Iran in the same category, despite their drastically differing treatments of religion in actuality. The outcome variable used in this analysis is sourced from the Religion and State Project data (Fox, 2020), which contains fine-grained information on state support for the preferred religion. This support is measured with 52 binary variables in five categories that capture different types of legislation designed to support the state's preferred religion. I take the first principal component of these 52 variables as my primary measure of the strength of state support. (Output tables for the first principal component of each of the five categories of support can be found in Appendix 3.B.) In addition, because I am primarily interested in effects over the longer term, I take as my primary regressor of interest the average of a country's religious support in the previous wave of the World Values Survey. While in an ideal world it would be preferable to use both current and lagged support as regressors, in practice, state support of religion tends to change very slowly, leading to concerns about collinearity between current and lagged values.

Variable	Definition
LX 1st PC	1st principal component of 52 binary variables measuring state support of the preferred religion
lag(LX)	Mean of $LX \ 1st \ PC$ for the previous IVS wave in which a country appeared
Enforcement	Enforcement of religious precepts including blasphemy laws, dietary restrictions, etc.
Funding	Government funding of religious schools, subsidies or grants to religious organizations, etc.
Restrictions on relationships	Laws governing relationships including those that restrict or ban homosexuality, abortion, etc.
Entanglement	Institutional entanglement including religious requirements for holding public office, etc.
Restrictions on women	Includes all laws not covered by the <i>Restrictions on relationships</i> category including modesty laws, etc.

 Table 3.2:
 Variables of interest

Furthermore, because legitimacy theory predicts disparate effects for members of preferred religions versus non-members, it makes sense to identify whether an individual is member of a favored religion or not. My third main source of data is the Government Religious Preference database (Brown, 2020), which contains detailed information on which specific religious denominations are supported in each country. From this I assign "preferred" or "non-preferred" status to each denomination by

country and year; this has the benefit of allowing me to identify (admittedly uncommon) cases where multiple religions are preferred by a government, eliminating the risk of accidentally classifying individuals as not being members of a preferred religion.

We turn now to the set of control variables, for which I draw on previous literature in the field. Certain individual attributes, such as age or gender, are known to have an effect on the strength of individuals' religious beliefs, whether the exact causal mechanism is understood or not (Schnabel, 2018; Kregting et al., 2019; Bryukhanov and Fedotenkov, 2023). The Integrated Values Survey data contains a wealth of individual-level covariates. Of these, I use Age, a *Female* dummy, and two dummies for education level: *Middle education* and *Upper education*. (Note that these standardized categories used by the WVS are designed to be applied worldwide, and therefore do not correlate cleanly with the high school and university levels of education most often used in analysis concerning the western world. "Middle education" as defined by the WVS includes both incomplete and complete secondary education, as well as some post-secondary *non*-university education. "Upper education" includes incomplete and complete undergraduate-level education, as well as graduate-level education.)

The secularization theory of religion argues that economic development and prosperity is the main driver of the presence and intensity of religious belief; I therefore include a number of country-level development indicators including the natural logarithm of real GDP, the logarithm of population, and the percentage of the labor force which is female. This last variable, while a useful indicator of a country's level of development, may be to an extent endogenous, given how often religious legislative agendas involve efforts to force women out of the public sphere and limit them to the domestic one. For this reason, a robustness test that swaps *Female labor force* (%) with a country's Gender Parity Index for enrolment at the primary school level. All country-level variables are sourced from the World Bank Development Indicators database (World Bank, 2022).

The religious competition theory, on the other hand, predicts that overall levels of religiosity should be higher the more robust competition between religions is; this makes the inclusion of a *Religious fractionalization* logical. This variable is expressed as the inverse Herfindahl index of adherence to different religious groups (among individuals who report belonging to any religion, and excluding atheists). I draw information on levels of adherence for each of the world's major religious denominations from the World Religion Project dataset (Maoz and Henderson, 2019). As these adherence figures are taken in five-year intervals, I impute missing IVS years with the most recent year for which adherence data exists (i.e., WRP data from 2000 is applied to all IVS observations from the years 2000 to 2004, etc.).

Interestingly, *both* the secularization and religious competition theories predict that levels of religiosity will be lower where state welfare capacity is higher. What supplyside proponents call competition with the state over services such as welfare (Cantoni et al., 2018; Masera, 2021), demand-side proponents call increased existential security? (Franck and Iannaccone, 2009, Franck and Iannaccone, 2014; Norris and Inglehart, 2015). However, data on state welfare expenditure is generally of poor quality, with many missing observations. I therefore use the Gini index, and an dummy for whether an individual's household is in the bottom income quintile for their country, instead. Together these serve as a proxy for the "demand side" of state welfare, and the overall extent (and/or effectiveness) of redistributive policies. It should be noted that Gini data suffers from missing observations, though to a much lesser extent than state welfare expenditure. I have imputed missing observations for the Gini index, following the method found in Fox and Breslawski, 2023 A discussion of the exact imputation procedure and its validity is included in Appendix **3.D**.

I use *Polity score* as a general measure of institutional quality, as well as a measure of much sway a country's citizens have over the state's behavior. This variable is drawn from the Polity5 Project, n.d. Finally, to account for possible unobserved differences between specific religions and their long-term effects, I include dummies for whether a country is historically an X Religion country (*Catholic country* is the reference category).

The final regression samples contain between 270,000 and 370,000 individual-level responses across 89 countries and six waves of the World Values Survey (from 1981 at the earliest to 2014 at the latest), depending on the survey question.

Because government support shows the largest variation between countries rather than over time, using fixed-effects at the country level risks omitting important information. However, using state-level data combined with individual-level data in an OLS regression generates errors that are "too small", as first shown in Moulton,

 $^{^{2}}$ As charity and welfare become less tied to religious identity.



Figure 3.1: State support percentiles for regression sample countries

1990, leading to over-rejection of the null for state-level variables. As noted in Bertrand et al., 2004, in order to correct the standard error for clustering one has to account for the presence of a common shock at the group level; Gelman, 2006 makes the argument that multilevel estimation – which which explicitly models clustering as a shared random effect for all individuals in a given country – is more effective than simply clustering standard errors.

The preferred estimation is a multilevel mixed-effects model, with the sample split into members of preferred religions vs. non-members. Because the full weight of government support of religion may not be felt immediately, it makes sense to also looked at possible long-term effects. However, the fact that state support of religion changes so slowly means that collinearity remains a concern. I therefore look at two estimations separately: the first using current government support of religion, and the second using its lagged value (i.e., the average level of state support for the previous IVS wave).

Also included in the Results section is a set of tables that interact government support with *Male* and *Female* dummies: as previously discussed, levels of religiosity tend to vary by gender, with women usually scoring higher than men on measures of internal beliefs and attitudes³, and religious-based legislation often im-

³Some countries prove an exception to this rule, most notably Israel (Schnabel et al., 2018).

poses discriminatory burdens and restrictions on women (see, for instance, Russia's new restrictions on abortion, designed to strengthen Putin's relationship with the Orthodox Church).⁴ It is therefore possible, and even likely, that the effects of state support on an individual's religiosity depend materially on whether that individual is male or female.

4 Descriptive Statistics

The vast majority of humanity, even in the relatively secular latter half of the 20th century, holds some level of belief in a god or gods: almost 80% of all individuals in the regression sample report membership to a religious denomination; *over* 80% report belief in God, and almost 30% of surveyed individuals attend religious service at least weekly. At the same time, however, there is a growing portion of the public that rejects religion: around one in every five individuals does not believe in any God, and similar proportions never or almost never attend church or pray.

IVS Survey Question	Response	No. responding	% of responses
Confidence: Churches	A great deal None at all	$117,127 \\ 42,021$	33.1 11.9
Attendance at religious services	Once a week or more	108,922	30.5
	Never/almost never	82,436	23.1
Importance of God	Extremely	160,308	45.3
	Not at all	$33,\!583$	9.49
Religious person	Yes	246,663	68.4
Religious person	No	$114,\!036$	31.6
Member of a religious denomination	Member	299,232	80.7
	Not a member	71,604	19.3

 Table 4.1: Religious attitudes and behaviors

These are figures which would have been beyond belief a century ago: while people have prophesied the decline of religion since at least the days of the ancient Greeks, the 19th and 20th centuries are perhaps the first time in human history these worries might have been well-founded.

 $^{^{4}} https://apnews.com/article/abortion-russia-women-rights-feminism-fc5eab75b5e3d028aeb1f70ec8a9a2b1$

Denomination	Number	% of responses
Roman Catholic	95,190	26.0
Protestant	$48,\!628$	13.3
Orthodox	49,065	13.4
Jewish	1,811	0.50
Muslim	67,496	18.5
Hindu	11,747	3.21
Buddhist	7,419	2.09
Other Christian	4,754	1.30
Other denomination	8,094	2.21
No denomination	$71,\!621$	19.6

 Table 4.2:
 Denomination membership (regression sample)

Unsurprisingly, a large proportion of the Integrated Values Survey data comes from the Western world – Western Europe alone accounts for 32.5% of observations in the survey, with Eastern Europe representing a further 14.6% on top of that; in comparison, the next largest regional category is Latin America and the Caribbean, representing 12.8% of the dataset. Sub-Saharan Africa represents just 5.6% of the IVS data, despite representing about 13% of the world's population. This disproportional representation is further exacerbated by the fact that other data, such as economic control variables, are more likely to be available for the exact countries that are already over-represented in the World Values Survey.

Countries themselves vary drastically in how much control the state exerts over religious institutes and practices, as well as the degree to which governments are willing to openly favor a particular religion. Less than 20% of country-year observations have any sort of official state religion, but over 30% have some form of restriction on minority religions. Over 40%, for instance, require political candidates to be members of a specific religion, or swear a religious oath in order to take office.

Tables 4.3 and 4.4 contain a selection of descriptive statistics from the Religion and State and Government Religions Preference datasets.

	No. of country-year observations	%	No. of IVS observations corresponding	%
Single state religion	57	18	76,839	16.7
Multiple state religions	5	1.6	4,761	1
Restrictions on public observance of religious holidays	97	30.6	147,572	32
Forced observation of religious laws of another group	39	12.3	61,527	13.3
Mandatory education in majority religion	78	24.6	117,999	25.6
Government control over clerical appointments	91	28.7	130,409	28.4
Blasphemy laws	91	28.7	134,968	29.3
Religious requirement/oath for taking office	132	41.6	197,237	42.7

Table 4.3: Religious governance descriptive statistics – RAS data

Table 4.4: Government religious preference – GRP data

	No. of country-year observations	%	No. of IVS observations corresponding	%
Christian preference	282	83.2	394,927	81
Catholic preference	194	57.2	278,701	57.1
Protestant preference	107	31.6	152,275	31.2
Jewish preference	1	0.3	1,199	0.3
Muslim preference	68	20	105,825	21.7
Hindu preference	11	3.2	19,691	4
Buddhist preference	15	4.4	20,658	4.2
Atheist preference	8	2.4	12,247	2.5
Multiple denominational preference	51	2.1	71,719	14.7

Note that the percentages in Table 4.4 do not add to 100 by design: in countries where there are multiple well-established religious denominations, governments will often prefer these over minority denominations even when only one of these is the "officially" preferred denomination. Of countries that favor multiple denominations, the most common combination is simultaneous preference for Christian and Muslim denominations, present in multiple countries in Eastern Europe, Sub-Saharan Africa,

the Middle East, and East and South East Asia. Country-years with a Buddhist preference have a relatively high probability of sharing government preference with at least one other religion (only 2 of 15 country-years show Buddhist preference alone).

	State religion		State s	support	
	Coef.	Std. Err.	Coef.	Std. Err.	
$\overline{\log(\text{GDP per capita})}$	-0.051	(0.04)	0.311	(0.04)	
Gini index	0.006	(0.04)	0.031	(0.04)	
$\log(\text{Population})$	-0.099	(0.04)	2.132	(0.04)	
Labor force, female $(\%)$	0.010	(0.01)	-0.033	(0.00)	
Religious fractionalization	-0.374	(0.11)	-0.931	(0.10)	
Protestant country	1.126^{***}	(0.11)	-0.801	(0.10)	
Orthodox country	-0.148 +	(0.12)	1.226	(0.08)	
Mixed Christian country	0.149	(0.12)	0.616	(0.09)	
New Christian country	-1.108^{***}	(0.10)	0.868 +	(0.06)	
Muslim country	1.234^{***}	(0.11)	5.000	(0.13)	
Buddhist country	0.086	(0.22)	-1.992	(0.10)	
Chinese rel. complex country	0.540	(0.23)	-10.42	(0.11)	
Constant	1.994	(0.65)	37.07	(0.37)	
N	295		295		
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

Table 4.5: In what kinds of countries do governments support religions?

5 Results

We now turn to the results of the empirical estimation. There are three sets of tables in this section. Table 5.1 presents the preferred estimation, which looks at the medium- and long-run effects of government support of religion, using a country's average level of this support in the previous wave of the Integrated Values Survey as the main regressor of interest. Table 5.2 splits this regressor in two: one for men, and one for women. Finally, Table 5.3 looks at the effects of government support on individuals' propensity to belong to a religious denomination and to consider themselves a religious person.

As mentioned previously in Section 3 due to concerns about collinearity I do not include the current value of government support of religion in the preferred functional form. Short run effects will be discussed in this section, however, and the relevant tables can be found in Appendix 3.A.

5.1 Preferred Estimation

The results of the preferred estimation are consistent with the predictions made by the legitimacy theory: the effects of government support of religion differ sharply between those whose religion is supported and those whose religion is not. Results also reinforce the necessity of looking beyond current government policies, as the picture changes significantly when looking at the effect that government support of religion in the previous wave in which a country appeared in the World Values Survey has on measures of religiosity today.

Government support *does* in fact give the favored religion a competitive advantage when it comes to getting people through the doors: public religiosity (as measured by attendance) increases with stronger government support among members of preferred religions, and if anything this effect grows stronger over time. The effect on members of non-preferred religions is a little less clear: while it is possible they might benefit from a "shot in the arm" effect when it comes to attendance at service, the positive coefficient on attendance is not statistically significant at the 5% level – and is smaller than the coefficient for members of preferred religions, clearly contradicting Adam Smith's predictions.

However, this increased attendance ultimately comes at a significant cost in terms of both individuals' private faith and their confidence in religious leadership – costs that may not make themselves immediately apparent. Government support of religion is associated with an immediate and significant decrease in the importance individuals attribute to God in their lives, for *both* members of preferred and non-preferred religions – but in the long run this effect persists only for individuals who belong to a preferred religion.

	Atte	endance	Importance of God		Confidence: Church		
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred	
lag(Gov't support)	0.345***	0.198+	-0.422***	-0.169	-0.043*	-0.064	
	(0.05)	(0.11)	(0.05)	(0.12)	(0.02)	(0.05)	
Female	0.121***	0.210***	0.503***	0.379***	0.124***	0.078***	
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)	
Age	0.015***	0.009***	0.017***	0.010***	0.005***	0.003***	
-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
Middle education	-0.002	-0.047	-0.125***	-0.178***	-0.131***	-0.087***	
	(0.01)	(0.03)	(0.01)	(0.04)	(0.00)	(0.01)	
Upper education	0.176***	-0.004	-0.132***	-0.249***	-0.171***	-0.126***	
	(0.01)	(0.04)	(0.01)	(0.04)	(0.01)	(0.02)	
First income quintile	-0.073***	-0.122**	0.087***	0.053	0.024***	0.020	
	(0.01)	(0.04)	(0.02)	(0.04)	(0.01)	(0.02)	
log (GDP per capita)	0.101**	-0.171**	0.180***	0.083	0.105***	0.020	
	(0.03)	(0.07)	(0.03)	(0.07)	(0.01)	(0.03)	
Gini (imputed)	0.018***	0.026***	0.007^{*}	0.021**	0.014***	0.018***	
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	
$\log(Population)$	-0.251^{***}	0.143^{*}	-0.324***	-0.002	0.236^{***}	0.036	
	(0.07)	(0.07)	(0.08)	(0.08)	(0.03)	(0.03)	
Labor force, female $(\%)$	-0.075***	-0.015	-0.088***	-0.065***	-0.029***	-0.032***	
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	
Religious fractionalization	-0.844***	0.342	-0.316^{***}	1.320^{***}	-0.357***	0.201 +	
	(0.09)	(0.26)	(0.10)	(0.28)	(0.04)	(0.12)	
Protestant country	-0.195	-0.292	-0.559	-0.533+	0.211	-0.033	
	(0.36)	(0.27)	(0.44)	(0.32)	(0.16)	(0.14)	
Orthodox country	-0.120	-0.448	0.472	0.957^{**}	0.661^{***}	0.544^{***}	
	(0.39)	(0.30)	(0.48)	(0.36)	(0.17)	(0.15)	
Mixed Christian country	-0.332	-0.486	-1.139	-0.911	0.183	0.057	
	(0.67)	(0.50)	(0.82)	(0.60)	(0.29)	(0.25)	
New Christian country	2.954^{***}	1.042^{*}	3.235^{***}	1.598^{**}	0.551^{*}	0.540^{*}	
	(0.58)	(0.45)	(0.71)	(0.54)	(0.25)	(0.23)	
Muslim country	-1.377^{***}	-0.436	2.290^{***}	1.473^{***}	0.078	0.118	
	(0.39)	(0.38)	(0.47)	(0.44)	(0.17)	(0.18)	
Buddhist country	0.319	-0.704	-2.137*	-0.581	-0.799*	0.383	
	(0.81)	(0.51)	(0.99)	(0.61)	(0.35)	(0.26)	
Polity score	-0.051^{***}	0.048^{**}	0.037^{***}	0.091^{***}	0.015^{***}	0.044^{***}	
	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.01)	
Constant	10.698^{***}	3.119^{*}	13.499^{***}	7.374***	-1.562^{**}	2.146^{**}	
	(1.20)	(1.39)	(1.31)	(1.60)	(0.49)	(0.66)	
N	186,582	27,767	184,519	27,129	183,801	26,732	
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001							

 ${\bf Table \ 5.1:} \ {\rm Main \ results \ - \ multi-level \ mixed-effects \ model, \ lagged \ support}$

State support of religion is also associated with a significant decrease in confidence in religious leadership amongst members of preferred religions, and this effect is only felt in the long-term.

Overall, government support does cause members of preferred religions to publicly participate by attending service, but considering the negative effects on levels of private faith and confidence in religious leadership, the crowd may not be quite as motivated and committed as they seem.

5.2 Interaction with Gender

The effects of government support of religion on measures of religiosity differ sharply by gender. Despite these differing effects, the overall picture remains broadly the same: government support of religion bolsters attendance at service, but comes at a cost of decreased measures of private religiosity and institutional legitimacy in the long run, even amongst men who adhere to a preferred religion – ostensibly the group most explicitly catered to by that government support.

The initial increase in attendance among members of preferred religions is driven almost entirely by men of the preferred religion, with a positive coefficient many times that of women of the preferred religion. It is likely some portion of this gap is a mechanical effect of restrictions on gender (see also Appendix 3.B). In the long run, the *decrease* in attendance among members of non-preferred religions is in fact driven by women. There is more evidence here of an energizing effect on attendance for non-preferred religions, specifically among men; the question remains open as to whether an overall positive effect may be present, but obscured by the mechanical effect of restricting women's movement.

There is an immediate negative effect of government support on the importance of God for members of preferred religions, and this effect is almost twice as strong for women as it is for men. In the long run, however, this initial negative impact disappears for members of non-preferred religions, but persists for both men and women who belong to a preferred religion, and – consistent with the loss of confidence in religious leadership – is larger for women than for men.

	Atte	endance	Importance of God		Confidence: Church	
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred
lag(Gov't support)*Male	0.578***	0.243*	-0.355***	-0.121	-0.030	-0.054
	(0.05)	(0.11)	(0.05)	(0.12)	(0.02)	(0.05)
lag(Gov't support)*Female	0.123**	0.139	-0.485***	-0.233+	-0.056**	-0.079
	(0.05)	(0.11)	(0.05)	(0.12)	(0.02)	(0.05)
Female	0.139***	0.192***	0.511***	0.361***	0.125***	0.073***
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)
Age	0.015***	0.009***	0.017***	0.010***	0.005***	0.003***
0	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle education	-0.015	-0.049	-0.129***	-0.180***	-0.132***	-0.087***
	(0.01)	(0.03)	(0.01)	(0.04)	(0.00)	(0.01)
Upper education	0.156***	-0.006	-0.138***	-0.252***	-0.172***	-0.126***
1 1	(0.01)	(0.04)	(0.01)	(0.04)	(0.01)	(0.02)
First income quintile	-0.080***	-0.124**	0.085***	0.051	0.024***	0.020
	(0.01)	(0.04)	(0.02)	(0.04)	(0.01)	(0.02)
log (GDP per capita)	0.106***	-0.170*	0.182***	0.085	0.105***	0.021
	(0.03)	(0.07)	(0.03)	(0.07)	(0.01)	(0.03)
Gini (imputed)	0.017***	0.027***	0.007*	0.022**	0.014***	0.018***
0 (F)	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
log(Population)	-0.261***	0.143*	-0.331***	-0.002	0.236***	0.036
	(0.07)	(0.07)	(0.08)	(0.08)	(0.03)	(0.03)
Labor force, female $(\%)$	-0.075***	-0.016	-0.088***	-0.065***	-0.029***	-0.032***
(, 0)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
Religious fractionalization	-0.856***	0.340	-0.318***	1.317***	-0.357***	0.200+
	(0.09)	(0.26)	(0.09)	(0.28)	(0.04)	(0.12)
Protestant country	-0.192	-0.288	-0.560	-0.528	0.211	-0.032
	(0.36)	(0.27)	(0.44)	(0.32)	(0.16)	(0.14)
Orthodox country	-0.132	-0.444	0.466	0.963**	0.660***	0.546***
y	(0.39)	(0.30)	(0.48)	(0.36)	(0.17)	(0.15)
Mixed Christian country	-0.341	-0.483	-1.144	-0.908	0.183	0.058
	(0.68)	(0.50)	(0.82)	(0.60)	(0.29)	(0.25)
New Christian country	2.969***	1.042*	3.243***	1.599**	0.552^{*}	0.541*
	(0.58)	(0.45)	(0.71)	(0.54)	(0.25)	(0.23)
Muslim country	-1.372***	-0.428	2.295***	1.484***	0.078	0.120
	(0.39)	(0.38)	(0.47)	(0.44)	(0.17)	(0.18)
Buddhist country	0.349	-0.701	-2.121*	-0.578	-0.796*	0.384
	(0.82)	(0.51)	(0.99)	(0.61)	(0.35)	(0.26)
Polity score	-0.052***	0.048**	0.036***	0.092***	0.015***	0.044***
	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.01)
Constant	10.871***	3.118*	13.600***	7.363***	-1.550**	2.142**
	(1.20)	(1.39)	(1.31)	(1.60)	(0.49)	(0.66)
N	100 500	07.707	104 510	07.100	109.001	0.0 720
IN .	186,582	27,707	184,519	27,129	183,801	20,732
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001						

 Table 5.2: Results by gender - multi-level mixed-effects model, lagged support

The fact that effects differ by gender is fairly unsurprising – gender differences in religiosity are a well-known and oft-studied phenomenon – but it is notable for being the clearest evidence in favor of the religious legitimacy theory over Adam Smith's incentives-based view. If the cause of religious leadership's loss of legitimacy is due to distorted incentives that limit their ability to proselytize, there would be no particular reason to expect such significant differences in how men and women react. On the other hand, it is perfectly reasonable to expect that direct negative effects of state support of religion would be strongest among those who most keenly feel the effects of that support, which – given that government legislation designed to support a religion often involves gender-based restrictions – means women.

In the short run, only men who are members of preferred religions express increased confidence in religious leadership when government support is higher; in the longer run, confidence in religious leadership is unchanged except among women of the preferred religion, where it decreases.

5.3 Extensive margin

The tables in this section present the results of three regressions: the first two columns in each table show individuals' willingness to identify as religious at all, for members of preferred and non-preferred denominations, respectively. The third column shows individuals' propensity to report belonging to any religious denomination, for the population as a whole. Together, these three regressions help clarify what the effects of government support of religion are at the extensive margin – i.e. whether it attracts new individuals to religion or pushes existing members away – as well as whether those changes reflect an actual decrease in individuals' religiosity or a mismatch between the individual and the country's actual religious institutions.

There is no effect on adherence to a denomination in the short run, or on propensity to consider oneself a religious person among individuals who adhere to a preferred religion. There *is* a small increase in likelihood of considering oneself religious among individuals who are members of non-preferred religions, though this coefficient is significant only at the 10% level. In the long run, on the other hand, government support decreases the likelihood of considering oneself religious among members

⁵As "member of a denomination" has to be equal to 1 in order to classify people as belonging to a preferred religion or not, it makes little conceptual sense to evaluate belonging to a denomination separately for preferred and non-preferred religions.

	Member of	Religious Person	Religious Person		
	Denomination	(Preferred)	(Non-preferred)		
lag(Gov't support)	-0.103***	-0.020*	-0.019		
	(0.01)	(0.01)	(0.02)		
Female	0.059***	0.073***	0.070***		
	(0.00)	(0.00)	(0.00)		
Age	0.002***	0.003***	0.002***		
5	(0.00)	(0.00)	(0.00)		
Middle education	-0.018***	-0.018***	-0.022***		
	(0.00)	(0.00)	(0.01)		
Upper education	-0.035***	-0.021***	-0.040***		
	(0.00)	(0.00)	(0.01)		
First income quintile	0.004 +	-0.003	-0.009		
1	(0.00)	(0.00)	(0.01)		
log (GDP per capita)	0.015***	0.029***	-0.016		
	(0.00)	(0.01)	(0.01)		
Gini (imputed)	-0.004***	0.004***	0.001		
<u>,</u>	(0.00)	(0.00)	(0.00)		
log(Population)	-0.083***	0.028**	0.001		
,	(0.01)	(0.01)	(0.01)		
Labor force, female $(\%)$	-0.006***	-0.008***	-0.004*		
	(0.00)	(0.00)	(0.00)		
Religious fractionalization	-0.336***	-0.100***	0.080 +		
	(0.01)	(0.02)	(0.04)		
Protestant country	-0.007	-0.038	-0.026		
	(0.07)	(0.05)	(0.04)		
Orthodox country	0.052	0.130^{*}	0.040		
	(0.08)	(0.05)	(0.04)		
Mixed Christian country	-0.168	-0.057	-0.032		
	(0.13)	(0.09)	(0.07)		
New Christian country	0.426^{***}	0.278^{***}	0.140^{*}		
	(0.11)	(0.08)	(0.06)		
Muslim country	0.434^{***}	0.026	-0.026		
	(0.07)	(0.05)	(0.06)		
Buddhist country	0.253 +	-0.516^{***}	-0.269***		
	(0.13)	(0.11)	(0.07)		
Polity score	0.013^{***}	0.009^{***}	0.003		
	(0.00)	(0.00)	(0.00)		
Constant	2.290***	0.079	0.903***		
	(0.19)	(0.18)	(0.21)		
N	315,022	188,411	27,639		
+ p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001					

of those supported religions, has no significant effect on members of non-preferred religions, and the population as a whole becomes less likely to report being a member of a denomination.

The results of these regressions tell us two things: firstly, that individuals abandon particular denominations more readily than they abandon faith – the negative coefficient on lagged support for *Denomination member* is far larger than the one for *Religious person*. Secondly, that state support of a given religion (or religions) does appear to have a negative effect at the extensive margin of faith and religious adherence, and is negative effect is likely driven by members of preferred religious movements no longer identifying with them, consistent with the predictions made religious legitimacy theory, and by Adam Smith, centuries ago.

6 Conclusion

What are the effects of government intervention in the market for religious services? This question is by no means a simple one: states support religions in a variety of ways, both monetarily and legislatively, and this support affects a wide array outcomes for both religious institutions and religious individuals. Previous research has tended to look at the effects on a country's entire population, but this risks occluding the true picture: state support is usually geared towards a specific religion, so it makes sense to ask whether the effects vary depending on whether the religious institution an individual belongs to is supported or not. Further, government support of a religion typically results in the implementation of religious behavioral restrictions, the burden of which falls disproportionately on women. Might the effects of state support on religious individuals therefore additionally vary by gender?

This paper has attempted to shed new light on an old question in these specific areas. Using individual-level data from the World Values Survey and European Values Survey, state religious legislation data from two separate datasets, and a number of state-level control variables informed by the literature, I run a multilevel mixed-effects regression that accounts for clustering at the country level by modeling clustering as a shared random effect for all individuals in a given country.

Stronger government support of religion *is* associated with increased attendance at religious service in both the short and long term, making it understandable why

so many religions might seek or accept a close relation with the state. However, stronger support is also associated with significant decreases in both private faith and levels of institutional confidence among the faithful. These negative effects fall on members of preferred religions, and are mostly only felt over the longer term. In practical terms, a one-standard deviation increase in the overall level of state support – the equivalent of moving from France to the United Kingdom, all else equal – increases the likelihood that members of preferred religions attend service at least weekly by five percentage points. However, it decreases the actual importance accorded to God by 0.4 points on the World Values Survey's 0-10 scale (with an even smaller decrease in confidence in religious leadership, on a 0-4 scale). More dramatically, it decreases the likelihood that an individual belongs to a religious denomination by 10 percentage points, a much larger effect than the 2 percentage point decrease in the propensity to consider oneself religious.

Type of support is important, and different types offer qualitative evidence in favor of different hypotheses that have been proposed to explain variations in levels of religious belief. Government support in the form of religious legislation is the driver of increased attendance at religious service, not funding. Although previous research has found that individuals tend to turn to religion as a form of ex-post insurance (Chen, 2010; Ager and Ciccone, 2018), increased attendance appears to be related to either the visibility and prestige associated with attending, or the reduced opportunity cost caused by religious legislation, rather than religious institutions' capacity to provide monetary insurance. What support in the form of funding *is* associated with is decreased confidence in religious leadership, entirely in keeping with the predictions of legitimacy theory: of the two categories of government support of religion, it is *monetary* support – the turning of state-religion relationships into business relationships – that has the clearest negative effect.

Finally, gender plays a clear mediating role: for every measure of religiosity examined in this paper, and for members of preferred religions and non-members both, the coefficients on the interaction of government support with gender are negative. These coefficients are significant for both preferred- and non-preferred religions, for every measure of religious belief with the sole exception of *Weekly attendance* for non-preferred religions. The main difference between the different outcome variables is whether the coefficient on the interaction of government support and gender merely mitigates a positive coefficient on general government support, reverses it, or exacerbates an already-negative coefficient. The positive and significant effect of government support on weekly attendance, for instance, is entirely balanced out by the negative coefficient on the interaction, the negative effects of government support on confidence in the Church are driven by women.

Overall, government support of a given religion may indeed increase attendance, but it decreases confidence in the Church, as well as most measures of private belief across a country's whole population. But there is an important caveat: government support increases attendance, but among non-members of the preferred religion! Similarly, it decreases belief among members of the preferred religion and *not* among nonpreferred individuals. Religions may have a variety of reasons to want government support, but if any should pursue it in an attempt to bolster their legitimacy or motivate their faithful: caveat emptor.

Bibliography

- Abulof, U. (2014). The roles of religion in national legitimation: Judaism and zionism's elusive quest for legitimacy. Journal for the Scientific Study of Religion, 53(3), 515–533.
- Ager, P., & Ciccone, A. (2018). Agricultural Risk and the Spread of Religious Communities. Journal of the European Economic Association, 16(4), 1021–1068.
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates?*. The Quarterly Journal of Economics, 119(1), 249–275. https://doi.org/10.1162/003355304772839588
- Boone, J. P., Khurana, I. K., & Raman, K. (2013). Religiosity and tax avoidance. The Journal of the American Taxation Association, 35(1), 53–84.
- Brown, D. (2020). Government Religious Preference 2.0 (GRP2.0). https://www. thearda.com/data-archive?fid=GRPCOMP
- Bryukhanov, M., & Fedotenkov, I. (2023). Gender differences in religiosity: A report on russian data. Journal of Gender Studies, 32(2), 107–123.
- Campante, F., & Yanagizawa-Drott, D. (2015). Does religion affect economic growth and happiness? evidence from ramadan. The Quarterly Journal of Economics, 130(2), 615–658.
- Cantoni, D., Dittmar, J., & Yuchtman, N. (2018). Religious competition and reallocation: The political economy of secularization in the protestant reformation. *The Quarterly Journal of Economics*, 133(4), 2037–2096.
- Carvalho, J.-P. (2016). Sacrifice and Sorting in Clubs. Forum for Social Economics.
- Chaney, E. (2008). Tolerance, religious competition and the rise and fall of muslim science.
- Chaves, M., & Gorski, P. S. (2001). Religious Pluralism and Religious Participation. Annual Review of Sociology, 27(1), 261–281.
- Chen, D. L. (2010). Club Goods and Group Identity: Evidence from Islamic Resurgence during the Indonesian Financial Crisis. *Journal of Political Economy*, 118(2), 300–354.
- de Tocqueville, A., & Reeve, H. ([1985] 1889). Democracy in America (Vol. 2). New York: Saunders; Otley.
- EVS. (2021). EVS Trend File 1981-2017 [ZA7503 Data file Version 3.0.0]. https: //doi.org/10.4232/1.14021
- Finke, R., & Stark, R. (1998). Religious choice and competition. American Sociological Review, 63(5), 761–766.
- Finke, R., & Stark, R. (2003). The dynamics of religious economies. Handbook of the Sociology of Religion, 96–109.
- Fox, J. (2020). The Religion and State Project, Minorities Module, Round 3. https: //doi.org/10.17605/OSF.IO/65WBU
- Fox, J., & Breslawski, J. (2023). State Support for Religion and Government Legitimacy in Christian-Majority Countries. American Political Science Review, 1–15.
- Franck, R., & Iannaccone, L. R. (2009). Why did religiosity decrease in the western world during the twentieth century. Bar-Ilan University Discussion Paper.
- Franck, R., & Iannaccone, L. R. (2014). Religious decline in the 20th century west: Testing alternative explanations. *Public Choice*, 159, 385–414.
- Gelman, A. (2006). Multilevel (Hierarchical) Modeling: What It Can and Cannot Do. *Technometrics*, 48(3), 432–435. https://doi.org/10.1198/004017005000000661
- Gertler, P. J. (1989). 'subsidies, quality, and the regulation of nursing homes. 'Journal of Public Economics, 38(1), 33–52. https://doi.org/https://doi.org/10.1016/ 0047-2727(89)90010-8
- Gómez-Lobo, A. (2014). Monopoly, Subsidies and the Mohring Effect: A Synthesis. Transport Reviews, 34(3), 297–315. https://doi.org/10.1080/01441647.2014. 902871
- Goodstein, L. (2009). Without a Pastor of His Own, Obama Turns to Five. Retrieved July 29, 2023, from https://www.nytimes.com/2009/03/15/us/politics/ 15pastor.html
- Gruber, J., & Hungerman, D. M. (2008). The church versus the mall: What happens when religion faces increased secular competition? The Quarterly journal of economics, 123(2), 831–862.
- Guiso, L., Sapienza, P., & Zingales, L. (2003). People's Opium? Religion and Economic Attitudes. Journal of monetary economics, 50(1), 225–282.

- Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., & et al. (eds.), B. P. (2022). World Values Survey Trend File (1981-2022) Cross-National Data-Set [Data File Version 2.0.0]. https://doi.org/10.14281/18241.23
- Henley, A. (2017). Does religion influence entrepreneurial behaviour? International Small Business Journal, 35(5), 597–617.
- Hicks, J. R. (1935). Annual Survey of Economic Theory: The Theory of Monopoly. Econometrica, 1–20.
- Hout, M., & Fischer, C. S. (2014). Explaining Why More Americans Have No Religious Preference: Political Backlash and Generational Succession, 1987-2012. *Sociological Science*, 1.
- Iannaccone, L. R. (1991). The consequences of religious market structure: Adam Smith and the economics of religion. *Rationality and society*, 3(2), 156–177.
- Iannaccone, L. R., Stark, R., & Finke, R. (1998). Rationality and the "Religious Mind". *Economic Inquiry*, 36(3), 373–389.
- Kregting, J., Scheepers, P., Vermeer, P., & Hermans, C. (2019). Why dutch women are still more religious than dutch men: Explaining the persistent religious gender gap in the netherlands using a multifactorial approach. *Review of Religious Research*, 61, 81–108.
- Land, K. C., Deane, G., & Blau, J. R. (1991). Religious pluralism and church membership: A spatial diffusion model. *American Sociological Review*, 237–249.
- Lechner, F. J. (1991). The Case Against Secularization: A Rebuttal. Social Forces, 69(4), 1103–1119. https://doi.org/10.1093/sf/69.4.1103
- Leroch, M. A., & Wellbrock, C. M. (2011). Saving newspapers with public grants: The effects of press subsidies on the provision of journalistic quality. Information Economics and Policy, 23(3), 281–286. https://doi.org/https: //doi.org/10.1016/j.infoecopol.2011.08.002
- Luckmann, T. (1967). The Invisible Religion (T. Kaden & B. Schnettler, Eds.). New York: MacMillan Company.
- Maoz, Z., & Henderson, E. A. (2019). World Religion Project Global Religion Dataset. https://www.thearda.com/data-archive?fid=WRPGLOBL
- Martin, D. (1991). The Secularization Issue: Prospect and retrospect. British Journal of Sociology, 465–474.
- Masera, F. (2021). State, religiosity and church participation. Journal of Economic Behavior & Organization, 186, 269–287.

- McCleary, R. M., & Barro, R. J. (2006a). Religion and economy. Journal of Economic perspectives, 20(2), 49–72.
- McCleary, R. M., & Barro, R. J. (2006b). Religion and political economy in an international panel. Journal for the Scientific study of religion, 45(2), 149– 175.
- McRae, S. (2015). Infrastructure Quality and the Subsidy Trap. American Economic Review, 105(1), 35–66. https://doi.org/10.1257/aer.20110572
- Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. The Review of Economics and Statistics, 72(2), 334–338. Retrieved October 1, 2023, from http://www.jstor.org/ stable/2109724
- Mussa, M., & Rosen, S. (1978). Monopoly and Product Quality. Journal of Economic Theory, 18(2), 301–317. https://doi.org/https://doi.org/10.1016/0022-0531(78)90085-6
- Norris, P., & Inglehart, R. (2015). Are High Levels of Existential Security Conducive to Secularization? A Response to Our Critics. The changing world religion map: Sacred places, identities, practices and politics, 3389–3408.
- Olson, M. (1965). The Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge: Harvard University Press.
- Parsons, T. (1974). Religion in Postindustrial America: The Problem of Secularization. Social Research, 41(2), 193–225. http://www.jstor.org/stable/40970177
- Polity5 Project. (n.d.). "political regime characteristics and transitions, 1800–2018.". https://www.systemicpeace.org/inscrdata.html
- Sandler, T., & Tschirhart, J. (1997). Club theory: Thirty years later. Public Choice, 93(3-4), 335–355.
- Schnabel, L. (2018). More religious, less dogmatic: Toward a general framework for gender differences in religion. Social science research, 75, 58–72.
- Schnabel, L., Hackett, C., & McClendon, D. (2018). Where men appear more religious than women: Turning a gender lens on religion in Israel. Journal for the Scientific Study of Religion, 57(1), 80–94.
- Seabright, P. (2024). The Divine Economy: How Religions Compete for Wealth, Power, and People. Princeton, NJ: Princeton University Press.
- Smith, A. ([1776] 2002). An Inquiry into the Nature and Causes of the Wealth of Nations. Retrieved from the Library of Congress. https://lccn.loc.gov/ 2002564559

Stark, R., & Iannaccone, L. R. (1994). A Supply-Side Reinterpretation of the "Secularization" of Europe. Journal for the Scientific Study of Religion, 33(3), 230–252. Retrieved August 1, 2023, from http://www.jstor.org/stable/1386688

Weber, M. (1920[1917]). The Sociology of Religion. Beacon Press.

World Bank. (2022). World Development Indicators. https://data.worldbank.org/

Appendix

3.A Current government support

This Appendix contains tables of results that are companions to Tables 5.1, 5.2 and 5.3 above. The tables in this section use the strength of current government support of religion as the primary regressor.

	Attendance		Importa	Importance of God		nce: Church
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred
Gov't support	0.302***	0.141 +	-0.193***	-0.504***	0.017	-0.173***
	(0.03)	(0.08)	(0.03)	(0.09)	(0.01)	(0.04)
Female	0.075***	0.183***	0.463***	0.348***	0.118***	0.081***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.01)
Age	0.015***	0.008***	0.015***	0.008***	0.005***	0.003***
0	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle education	-0.027*	-0.014	-0.142***	-0.188***	-0.130***	-0.083***
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)
Upper education	0.097***	-0.004	-0.154***	-0.271***	-0.175***	-0.138***
	(0.01)	(0.03)	(0.01)	(0.04)	(0.00)	(0.01)
First income quintile	-0.096***	-0.063+	0.076^{***}	0.074^{*}	0.023***	0.017
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)
$\log (GDP \text{ per capita})$	-0.065**	-0.371^{***}	0.597^{***}	0.015	0.056^{***}	0.114^{***}
	(0.02)	(0.05)	(0.02)	(0.06)	(0.01)	(0.02)
Gini (imputed)	-0.010***	0.017^{**}	0.017^{***}	0.028^{***}	0.017^{***}	0.028^{***}
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
$\log(\text{Population})$	-0.378***	0.042	-0.753***	-0.100	-0.026	0.039
	(0.06)	(0.06)	(0.06)	(0.07)	(0.02)	(0.03)
Labor force, female $(\%)$	-0.049***	-0.013	-0.034***	-0.063***	-0.017^{***}	-0.025***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Religious fractionalization	-0.315***	0.566^{**}	-0.200**	1.449^{***}	-0.202***	0.098
	(0.07)	(0.20)	(0.07)	(0.21)	(0.03)	(0.09)
Protestant country	-0.119	-0.172	-0.604	-0.356	0.170	0.156
	(0.39)	(0.24)	(0.59)	(0.31)	(0.11)	(0.14)
Orthodox country	-0.724	-0.932***	-0.075	0.415	0.289^{*}	0.504^{**}
	(0.44)	(0.28)	(0.66)	(0.36)	(0.13)	(0.16)
Mixed Christian country	-0.753	-0.499	-2.127+	-1.016	-0.061	0.101
	(0.79)	(0.48)	(1.19)	(0.63)	(0.23)	(0.28)
New Christian country	2.075^{***}	1.016^{**}	2.980^{***}	1.353^{**}	0.493^{**}	0.148
	(0.56)	(0.34)	(0.84)	(0.45)	(0.16)	(0.20)
Muslim country	-0.593	-0.450	2.999^{***}	1.658^{***}	0.332^{**}	0.630^{***}
	(0.38)	(0.30)	(0.56)	(0.37)	(0.11)	(0.16)
Buddhist country	0.653	-0.716	-0.493	-0.362	-0.444	0.235
	(0.96)	(0.49)	(1.44)	(0.64)	(0.28)	(0.28)
Polity score	-0.030***	0.002	-0.006*	0.067^{***}	0.008^{***}	0.009 +
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Constant	14.017^{***}	7.307***	14.640^{***}	9.604^{***}	2.716^{***}	0.762
	(0.98)	(1.17)	(1.07)	(1.39)	(0.33)	(0.59)
Ν	248,735	36,633	247,176	35,883	247,808	$35,\!805$
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$		

 Table 3.A.1: Main results - multi-level mixed-effects model, current support

	Attendance		Importa	ce of God	Confidence: Church		
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred	
Gov't support*Male	0.533***	0.147+	-0.134***	-0.471***	0.031*	-0.168***	
T T	(0.03)	(0.08)	(0.03)	(0.09)	(0.01)	(0.04)	
Gov't support*Female	0.073*	0.132	-0.256***	-0.547***	0.003	-0.179***	
	(0.03)	(0.08)	(0.03)	(0.09)	(0.01)	(0.04)	
Female	0.115***	0.182***	0.478***	0.343***	0.120***	0.080***	
	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.01)	
Age	0.015***	0.008***	0.015***	0.008***	0.005***	0.003***	
0*	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
Middle education	-0.039***	-0.014	-0.145***	-0.189***	-0.131***	-0.083***	
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)	
Upper education	0.078***	-0.004	-0.158***	-0.272***	-0.176***	-0.138***	
1 1	(0.01)	(0.03)	(0.01)	(0.04)	(0.00)	(0.01)	
First income quintile	-0.103***	-0.063+	0.074***	0.073*	0.023***	0.017	
Ŧ	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)	
log (GDP per capita)	-0.061**	-0.371***	0.598***	0.016	0.056***	0.114***	
	(0.02)	(0.05)	(0.02)	(0.06)	(0.01)	(0.02)	
Gini (imputed)	-0.011***	0.017**	0.017***	0.028***	0.017***	0.028***	
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	
log(Population)	-0.364***	0.042	-0.752***	-0.101	-0.025	0.039	
	(0.06)	(0.06)	(0.06)	(0.07)	(0.02)	(0.03)	
Labor force, female $(\%)$	-0.048***	-0.013	-0.033***	-0.064***	-0.017***	-0.025***	
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	
Religious fractionalization	-0.335***	0.566**	-0.205**	1.452***	-0.203***	0.098	
-	(0.07)	(0.20)	(0.07)	(0.21)	(0.03)	(0.09)	
Protestant country	-0.113	-0.171	-0.603	-0.352	0.170	0.157	
	(0.39)	(0.24)	(0.59)	(0.31)	(0.11)	(0.14)	
Orthodox country	-0.717	-0.931***	-0.076	0.420	0.289*	0.505**	
	(0.44)	(0.28)	(0.66)	(0.36)	(0.13)	(0.16)	
Mixed Christian country	-0.753	-0.499	-2.128+	-1.015	-0.061	0.101	
	(0.79)	(0.48)	(1.19)	(0.63)	(0.23)	(0.28)	
New Christian country	2.083^{***}	1.016^{**}	2.982^{***}	1.352^{**}	0.493^{**}	0.148	
	(0.56)	(0.34)	(0.84)	(0.45)	(0.16)	(0.20)	
Muslim country	-0.605	-0.448	3.005^{***}	1.671^{***}	0.332^{**}	0.631^{***}	
	(0.38)	(0.30)	(0.56)	(0.37)	(0.11)	(0.16)	
Buddhist country	0.639	-0.715	-0.489	-0.354	-0.445	0.236	
	(0.95)	(0.49)	(1.44)	(0.64)	(0.28)	(0.28)	
Polity score	-0.032***	0.002	-0.006*	0.068^{***}	0.008^{***}	0.009 +	
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	
Constant	13.796^{***}	7.307***	14.611^{***}	9.610^{***}	2.703^{***}	0.763	
	(0.97)	(1.17)	(1.07)	(1.39)	(0.33)	(0.59)	
Ν	248,735	36,633	247,176	35,883	247,808	35,805	
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$			

 Table 3.A.2: Results by gender - multi-level mixed-effects model, current support

	Member of Denomination	Religious Person (Preferred)	Religious Person (Non-preferred)
Gov't support	0.006	-0.005	0.020 +
	(0.00)	(0.01)	(0.01)
Female	0.053***	0.068***	0.063***
	(0.00)	(0.00)	(0.00)
Age	0.002^{***}	0.002^{***}	0.002^{***}
	(0.00)	(0.00)	(0.00)
Middle education	-0.014***	-0.017***	-0.013*
	(0.00)	(0.00)	(0.01)
Upper education	-0.032***	-0.022***	-0.033***
	(0.00)	(0.00)	(0.01)
First income quintile	0.004*	-0.003	-0.008
	(0.00)	(0.00)	(0.01)
log (GDP per capita)	0.013***	0.010**	-0.018*
	(0.00)	(0.00)	(0.01)
Gini (imputed)	-0.002***	0.005^{***}	0.002*
	(0.00)	(0.00)	(0.00)
log(Population)	-0.133***	0.000	-0.001
	(0.01)	(0.01)	(0.01)
Labor force, female $(\%)$	-0.004***	-0.001	0.002
	(0.00)	(0.00)	(0.00)
Religious fractionalization	-0.204	-0.113	(0.018)
Destate to at a constant	(0.01)	(0.01)	(0.03)
Protestant country	(0.015)	(0.001)	-0.010
Orthadam and the	(0.07)	(0.04)	(0.03)
Orthodox country	-0.053	(0.050)	(0.000)
Mirred Christian country	(0.08)	(0.04)	(0.05)
Mixed Christian country	-0.229	-0.087	-0.004
Now Christian country	(0.14) 0.262**	(0.08)	(0.00)
New Christian country	(0.10)	(0.06)	(0.043)
Muslim country	0.10)	0.00)	(0.04)
Wushin country	(0.07)	(0.04)	(0.04)
Buddhist country	(0.01) 0.226	-0.450***	-0.247***
Buddhist country	(0.14)	(0.10)	(0.06)
Polity score	0.005***	0.008***	-0.001
rondy score	(0.00)	(0.00)	(0.00)
Constant	3.069***	0.389**	0.750***
	(0.14)	(0.13)	(0.16)
N	370,836	248,886	36,363
+ p < 0.10, * p < 0.05,	** p < 0.01,	*** p <0.001	

3.B Funding vs. regulatory support

In the long run, funding decreases confidence among both preferred and non-preferred, consistent with legitimacy theory: of all types of government support, financial support carries the greatest risk of turning the relationship between state and religious into a business one. Funding also carries a negative coefficient on public religiosity in the long run for members of preferred religions – though this coefficient is only significant at the 10% level, and should therefore be interpreted only as not *in* consistent with Smith's predictions, in which it is funding itself that causes preachers to lose their ability to effectively proselytize. Regulatory support has a sustained positive effect on public religiosity, and a sustained negative effect on private religiosity, highlighting the gap between externally perceivable behaviors and private belief.

	Atte	Attendance Importanc		nce of God	ce of God Confidence	
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred
Funding	0.025	-0.043	0.137***	0.312***	0.020*	-0.054+
	(0.02)	(0.06)	(0.03)	(0.07)	(0.01)	(0.03)
Regulation	0.295***	0.163 +	-0.219***	-0.638***	0.011	-0.154***
<u> </u>	(0.03)	(0.09)	(0.03)	(0.09)	(0.01)	(0.04)
Female	0.075***	0.183***	0.463***	0.348***	0.118***	0.081***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.00)	(0.01)
Age	0.015***	0.008***	0.015***	0.008***	0.005***	0.003***
-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle education	-0.027*	-0.013	-0.144***	-0.191***	-0.130***	-0.083***
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)
Upper education	0.096***	-0.004	-0.155***	-0.273***	-0.176***	-0.138***
	(0.01)	(0.03)	(0.01)	(0.04)	(0.00)	(0.01)
First income quintile	-0.096***	-0.063+	0.077***	0.073*	0.024***	0.018
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)
log (GDP per capita)	-0.076**	-0.359***	0.539***	-0.089	0.047***	0.137***
	(0.02)	(0.05)	(0.03)	(0.06)	(0.01)	(0.03)
Gini (imputed)	-0.010***	0.016**	0.017***	0.033***	0.017***	0.028***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
log(Population)	-0.363***	0.042	-0.669***	-0.091	-0.017	0.036
	(0.06)	(0.06)	(0.06)	(0.07)	(0.02)	(0.03)
Labor force, female $(\%)$	-0.049***	-0.013	-0.035***	-0.060***	-0.016***	-0.026***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Religious fractionalization	-0.320***	0.561^{**}	-0.216**	1.446***	-0.205***	0.105
-	(0.07)	(0.20)	(0.07)	(0.21)	(0.03)	(0.09)
Protestant country	-0.112	-0.181	-0.562	-0.265	0.175	0.145
	(0.39)	(0.24)	(0.56)	(0.31)	(0.11)	(0.15)
Orthodox country	-0.719	-0.929***	-0.037	0.398	0.290^{*}	0.519^{**}
	(0.44)	(0.28)	(0.63)	(0.35)	(0.12)	(0.17)
Mixed Christian country	-0.743	-0.494	-2.074+	-1.006	-0.055	0.099
	(0.79)	(0.48)	(1.14)	(0.62)	(0.22)	(0.29)
New Christian country	2.064^{***}	1.018^{**}	2.935^{***}	1.331^{**}	0.485^{**}	0.159
	(0.56)	(0.34)	(0.80)	(0.44)	(0.16)	(0.21)
Muslim country	-0.619	-0.468	2.827^{***}	1.708^{***}	0.314^{**}	0.633^{***}
	(0.38)	(0.30)	(0.53)	(0.36)	(0.11)	(0.17)
Buddhist country	0.624	-0.739	-0.653	-0.228	-0.462 +	0.228
	(0.95)	(0.49)	(1.37)	(0.63)	(0.27)	(0.30)
Polity score	-0.030***	0.001	-0.007*	0.074^{***}	0.008^{***}	0.008 +
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Constant	13.872^{***}	7.267^{***}	13.886^{***}	10.013^{***}	2.646^{***}	0.689
	(0.98)	(1.16)	(1.06)	(1.37)	(0.33)	(0.61)
Ν	248,735	36,633	247,176	35,883	247,808	35,805
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$		

Table 3.B.1:Funding vs. regulatory support - current values

	Attendance Importance		ce of God Confid		nce: Church	
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred
lag(Funding)	-0.043 +	-0.076	0.077^{**}	0.093	-0.034***	-0.085**
	(0.02)	(0.06)	(0.03)	(0.07)	(0.01)	(0.03)
lag(Regulation)	0.369***	0.247*	-0.456***	-0.225+	-0.024	-0.021
	(0.05)	(0.11)	(0.05)	(0.12)	(0.02)	(0.05)
Female	0.121***	0.210***	0.503***	0.379***	0.124^{***}	0.078***
	(0.01)	(0.03)	(0.01)	(0.03)	(0.00)	(0.01)
Age	0.015***	0.009***	0.017***	0.010***	0.005^{***}	0.003***
-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle education	-0.001	-0.047	-0.127***	-0.177***	-0.130***	-0.087***
	(0.01)	(0.03)	(0.01)	(0.04)	(0.00)	(0.01)
Upper education	0.178^{***}	-0.002	-0.134***	-0.251***	-0.170***	-0.124***
	(0.01)	(0.04)	(0.01)	(0.04)	(0.01)	(0.02)
First income quintile	-0.073***	-0.123**	0.087^{***}	0.053	0.024^{***}	0.021
	(0.01)	(0.04)	(0.02)	(0.04)	(0.01)	(0.02)
log (GDP per capita)	0.120^{***}	-0.160*	0.153^{***}	0.066	0.118^{***}	0.045
	(0.03)	(0.07)	(0.03)	(0.08)	(0.01)	(0.03)
Gini (imputed)	0.018^{***}	0.026^{***}	0.007^{*}	0.021**	0.015^{***}	0.018***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
$\log(\text{Population})$	-0.280***	0.140^{*}	-0.276***	0.002	0.212^{***}	0.034
	(0.08)	(0.07)	(0.08)	(0.08)	(0.03)	(0.04)
Labor force, female $(\%)$	-0.075***	-0.015	-0.088***	-0.065***	-0.029***	-0.034***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)
Religious fractionalization	-0.823***	0.358	-0.348***	1.292^{***}	-0.339***	0.236^{*}
	(0.09)	(0.26)	(0.10)	(0.28)	(0.04)	(0.12)
Protestant country	-0.217	-0.309	-0.527	-0.509	0.194	-0.056
	(0.36)	(0.27)	(0.43)	(0.32)	(0.15)	(0.15)
Orthodox country	-0.135	-0.454	0.501	0.961^{**}	0.648^{***}	0.561^{***}
	(0.40)	(0.29)	(0.47)	(0.36)	(0.17)	(0.16)
Mixed Christian country	-0.356	-0.477	-1.104	-0.914	0.163	0.056
	(0.68)	(0.49)	(0.80)	(0.60)	(0.29)	(0.27)
New Christian country	2.988^{***}	1.027^{*}	3.190^{***}	1.614^{**}	0.577^{*}	0.552^{*}
	(0.59)	(0.45)	(0.69)	(0.54)	(0.25)	(0.24)
Muslim country	-1.340^{***}	-0.433	2.209^{***}	1.463^{***}	0.109	0.130
	(0.40)	(0.37)	(0.46)	(0.44)	(0.17)	(0.19)
Buddhist country	0.373	-0.736	-2.228*	-0.546	-0.755*	0.373
	(0.82)	(0.50)	(0.97)	(0.61)	(0.35)	(0.28)
Polity score	-0.051^{***}	0.048^{**}	0.037^{***}	0.091^{***}	0.015^{***}	0.045^{***}
	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.01)
Constant	11.018^{***}	3.038^{*}	12.976^{***}	7.505^{***}	-1.296^{**}	1.978^{**}
	(1.22)	(1.38)	(1.31)	(1.61)	(0.49)	(0.69)
N	186,582	27,767	184,519	27,129	183,801	26,732
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$		

 Table 3.B.2:
 Funding vs. regulatory support - lagged values

3.C Robustness test: Primary enrolment Gender Parity Index

This section presents results of the main regression using both current and lagged government support, but exchanging *Female labor force* (%) for an alternative measure of gender equality: the Gender Parity Index (GPI) for enrolment in primary school. This is a more conservative measure than female labor force, in the sense that it is only meaningfully expected to vary in the least economically developed countries. Indeed, most countries score north of 0.95 on this measure; the only countries that scored an average of less than 0.9 for the period in which they participated in the World Values Survey are India, Iraq, Nigeria, and Pakistan. The data are also of lower quality, with a greater number of missing observations, which results in Bangladesh, Iraq, Japan, and Moldova dropping out of the regression dataset entirely. It does, however, address concerns of potential endogeneity with female labor force: female economic competition can provide a powerful incentive for men to pass legislation restricting women, as indeed religious laws tend to do (for instance, see the fate of female workers in Afghanistan for a particularly clear recent example of women being forced out of the public sphere). However, even relatively strict religious legislation tends to allow for primary education for girls.

Overall, results are robust to the use of this alternate measure in the sense that government support of religion causes individuals' private religiosity to decrease even as their public participation increases. However, much of the difference in effects between members of preferred versus non-preferred religions decreases, and the negative coefficient on *Confidence in Church* for members of preferred religions in the long run becomes insignificant. The effects of government support of religion here appear to be felt in similar ways by members of both preferred and non-preferred religions, the difference being one of degree.

	Confidence: Church		Atte	ndance	Importa	ance of God
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred
Gov't support*Male	0.007	-0.048***	0.165***	0.067*	-0.026*	-0.112***
	(0.00)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)
Gov't support*Female	-0.001	-0.052***	0.079***	0.061*	-0.061***	-0.140***
	(0.00)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)
Female	0.122^{***}	0.081***	0.170^{***}	0.176^{***}	0.492^{***}	0.340***
	(0.00)	(0.01)	(0.01)	(0.02)	(0.01)	(0.03)
Age	0.005^{***}	0.003^{***}	0.014^{***}	0.009^{***}	0.016^{***}	0.009^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle education	-0.125***	-0.084***	-0.061***	-0.010	-0.153^{***}	-0.213***
	(0.00)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)
Upper education	-0.167***	-0.129***	0.072^{***}	-0.011	-0.141***	-0.283***
	(0.01)	(0.02)	(0.01)	(0.04)	(0.01)	(0.04)
First income quintile	0.015^{**}	0.020	-0.099***	-0.038	0.086^{***}	0.084^{*}
	(0.01)	(0.02)	(0.01)	(0.04)	(0.01)	(0.04)
$\log (GDP \text{ per capita})$	-0.014	0.060*	-0.066**	-0.379***	0.489^{***}	-0.155*
	(0.01)	(0.03)	(0.02)	(0.06)	(0.02)	(0.07)
Gini (imputed)	0.017^{***}	0.024^{***}	0.002	0.020**	0.023***	0.023***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
$\log(\text{Population})$	-0.692***	0.032	-0.764***	0.039	-1.279***	-0.161*
	(0.03)	(0.03)	(0.08)	(0.06)	(0.08)	(0.08)
Primary enrollment GPI	1.856^{***}	0.762^{***}	-0.720***	0.235	1.533^{***}	2.660^{***}
	(0.08)	(0.20)	(0.21)	(0.47)	(0.20)	(0.50)
Religious fractionalization	-0.103**	0.118	-0.706***	-0.176	-0.613***	1.204^{***}
	(0.03)	(0.09)	(0.08)	(0.21)	(0.08)	(0.23)
Protestant country	0.046	0.069	-0.188	0.019	-0.564	-0.506
	(0.33)	(0.13)	(0.48)	(0.25)	(0.76)	(0.33)
Orthodox country	-0.426	0.332^{*}	-1.175^{*}	-0.947**	-0.700	-0.204
	(0.38)	(0.15)	(0.56)	(0.29)	(0.88)	(0.39)
Mixed Christian country	-0.451	-0.013	-0.926	-0.158	-2.152	-1.072
	(0.66)	(0.26)	(0.96)	(0.49)	(1.53)	(0.66)
New Christian country	0.775 +	0.137	2.113^{**}	1.097^{**}	3.156^{**}	1.273**
	(0.46)	(0.18)	(0.68)	(0.35)	(1.08)	(0.47)
Muslim country	1.068^{**}	0.838***	0.019	-0.744*	3.536^{***}	1.656^{***}
	(0.33)	(0.16)	(0.48)	(0.32)	(0.76)	(0.40)
Buddhist country	1.054	0.466	1.226	-1.199*	0.820	-0.739
	(1.10)	(0.31)	(1.61)	(0.60)	(2.56)	(0.81)
Polity score	0.016***	0.020***	-0.034***	-0.019	-0.011***	0.055***
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)
Constant	11.754^{***}	-0.348	18.859^{***}	6.989***	21.303^{***}	7.291***
	(0.55)	(0.58)	(1.36)	(1.20)	(1.41)	(1.49)
N	216,166	31,726	21,4528	32,252	214,947	31,777
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$		

 $\label{eq:Table 3.C.1: Primary enrolment GPI - multi-level mixed-effects model, current support$

	Confiden	ce: Church	Atte	ndance	Importa	ance of God
	Preferred	Non-preferred	Preferred	Non-preferred	Preferred	Non-preferred
lag(Gov't support)*Male	-0.003	0.002	0.239***	0.093*	-0.069***	-0.069+
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)
lag(Gov't support)*Female	-0.011+	-0.008	0.133***	0.054	-0.111***	-0.117**
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)
Female	0.126***	0.078***	0.180***	0.194***	0.524^{***}	0.343***
	(0.00)	(0.01)	(0.01)	(0.03)	(0.01)	(0.03)
Age	0.005***	0.003***	0.015***	0.009***	0.018***	0.010***
-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Middle education	-0.128***	-0.084***	-0.029*	-0.027	-0.136***	-0.212***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)	(0.04)
Upper education	-0.161***	-0.114***	0.148***	0.003	-0.111***	-0.264***
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)
First income quintile	0.011+	0.024	-0.100***	-0.084+	0.093***	0.058
1	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)	(0.05)
log (GDP per capita)	-0.032*	-0.062+	0.139***	-0.206**	0.027	-0.147+
0 1 1 /	(0.01)	(0.03)	(0.03)	(0.07)	(0.04)	(0.09)
Gini (imputed)	0.016***	0.022***	0.022***	0.032***	0.008*	0.020*
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)
log(Population)	-0.074**	0.027	-0.182*	0.154*	-1.133***	0.023
	(0.03)	(0.03)	(0.08)	(0.06)	(0.12)	(0.08)
Primary enrollment GPI	2.758***	1.471***	-1.659***	0.202	2.691***	4.236***
	(0.10)	(0.25)	(0.26)	(0.60)	(0.29)	(0.66)
Religious fractionalization	-0.342***	0.035	-1.468***	-0.221	-0.923***	1.172***
0	(0.04)	(0.12)	(0.10)	(0.28)	(0.11)	(0.31)
Protestant country	0.132	-0.038	-0.222	-0.148	-0.819	-0.566+
	(0.12)	(0.12)	(0.37)	(0.26)	(0.75)	(0.33)
Orthodox country	0.143	0.393**	-0.304	-0.492	-0.737	0.471
y	(0.14)	(0.13)	(0.43)	(0.30)	(0.87)	(0.38)
Mixed Christian country	0.025	0.068	-0.357	-0.247	-1.576	-0.900
	(0.23)	(0.21)	(0.69)	(0.48)	(1.41)	(0.61)
New Christian country	0.472^*	0.315+	2.507^{***}	1.025*	3.495**	1.099*
iten emilieran eeanery	(0.20)	(0.19)	(0.60)	(0.43)	(1.21)	(0.53)
Muslim country	0.443**	0.224	-1.207**	-0.598	3.234***	2.076***
indonin country	(0.14)	(0.17)	(0.42)	(0.40)	(0.85)	(0.47)
Buddhist country	(0.11) 0.172	0.562*	-0.107	-0.943	-1 929	-0.931
Duddhibt country	(0.38)	(0.26)	(1.15)	(0.59)	(2.35)	(0.74)
Polity score	-0.006***	(0.20) 0.044***	-0.069***	0.048**	(2.00)	(0.14) 0.104***
I only booto	(0,00)	(0.01)	(0,00)	(0.02)	(0.01)	(0.02)
Constant	1 046*	0.175	8 011***	$2.386 \pm$	22 500***	2 171
Constant	(0.46)	(0.57)	(1.35)	(1.31)	(1.89)	(1.57)
N	165,079	24,023	164,736	24,782	162,880	24,162
	+ p < 0.10,	* $p < 0.05$,	** $p < 0.01$,	*** $p < 0.001$		

 Table 3.C.2:
 Primary enrolment GPI - multi-level mixed-effects model, lagged support

3.D Gini imputation

This appendix contains an analysis of the quality of World Bank data on the Gini index, and a discussion of the validity of the imputation conducted for this paper.

Including data on economic inequality is highly useful to any analysis of the drivers of public and private religiosity, as discussed in Section [3] but these data are often of poor quality. While the Gini index is one of the most widely available measures of overall country-level economic inequality, availability is patchy at best. Attempting to assign unimputed Gini index data by year causes a number of countries to drop from the regression dataset altogether. Even for countries that are generally thorough about recording economic indicators, assigning the unimputed Gini index can result in entire waves of the Integrated Values Survey vanishing, if the responses were collected in a year in which the Gini index is missing. This is, for example, the case of Czechia, which collected its responses for the 1990 wave of the IVS in 1991, but began recording its Gini index in 1992.

It is no great surprise that the pattern of missing observations in Gini index data is non-random: countries with lower levels of economic development are more likely to have missing data for any given year, and these are generally also countries with higher values of government support of religion. Using Gini data without accounting for this non-random missingness therefore risks biasing results – but imputing missing Gini data also carries potential problems with it, if the imputation method happens to produce values that are substantially different from the actual historical values for the period in which the country appears in the Integrated Values Survey.

The method I use to impute missing observations is identical to that used in Fox and Breslawski, 2023: missing observations are filled in using the last non-missing one; any remaining missing observations after this process are imputed using subsequent non-missing values. The STATA code used to impute missing values of the Gini index is as follows:

```
* Down
sort iso_alpha3 year
replace gini = gini[_n-1] if iso_alpha3==iso_alpha3[_n-1] & gini==. ///
& gini[_n-1]!=.
```

```
* Up
gsort -iso_alpha3 -year
replace gini = gini[_n-1] if iso_alpha3==iso_alpha3[_n-1] & gini==. ///
& gini[_n-1]!=.
```

In R the equivalent code is:

```
gini <- gini %>%
dplyr::group_by(ISO3) %>%
fill(gini, .direction="downup")
```

Table 3.D.1 below compares the averages of the raw and imputed Gini index for countries in the first main regression sample (i.e. for confidence in the Church/Temple/Mosque).

Country name	Gini	Imputed Gini	Country name	Gini	Imputed Gini
Albania	30.85	30.85	Korea (Republic of)	32	31.775
Algeria		27.6	Kyrgyzstan	27.8	27.8
Argentina	45.4	45.4	Latvia	37.2	36.8
Armenia	29.3	29.3	Lithuania	35.7	36.35
Australia		33.9	Luxembourg	32.6	32.6
Austria	30.05	30.05	Malaysia		43.9
Azerbaijan		26.6	Mexico	51.25	51.25
Bangladesh		33.4	Moldova	35.3	35.3
Belarus	28.733	28.733	Morocco		40.7
Belgium	28.6	27.7	Netherlands	28.7	28.7
Brazil	55.8	55.8	Nigeria	35.5	44.133
Bulgaria	36.567	36.22	North Macedonia		42.8
Canada	33.75	33.75	Norway	27.05	26.7
Chile	51.667	50.25	Pakistan	28.7	29.2
China	39.7	40.3	Peru	48.667	48.667
Colombia	53.25	53.25	Philippines	46.5	47.1
Croatia	30.4	31.867	Poland	33	34.667
Cyprus	32.15	32.15	Portugal	36.6	37.7
Czechia	25.6	25.7	Romania	36.45	38.34
Denmark	26.95	25.633	Russian Federation	39.38	39.217
Egypt		28.3	Rwanda		47.2
Estonia	32.2	33.867	Slovakia	24.6	25.433
Finland	27.425	26.64	Slovenia	24.35	24.5
France	31.733	31.733	South Africa		61.325
Georgia	38.1	38.1	Spain	34.95	35.2
Germany	30.433	30.433	Sweden	27.525	26.75
Ghana	42.4	42.4	Switzerland	33.6	33.6
Greece	34	34	Thailand	37.8	37.8
Hungary	27.25	28.575	Trinidad and Tobago		40.3
India		33.375	Turkey	39.133	40
Indonesia	35.3	35.3	Ukraine	27	29.05
Iran		44.8	United Kingdom	35.8	35.8
Iraq		29.5	United States of America	40.68	40.68
Ireland	30.9	33.25	Uruguay	44.05	44.05
Italy	33.8	34.767	Viet Nam	35.8	35.8
Japan	32.1	33.9	Zimbabwe		43.2
Jordan	•	33.8			

Table 3.D.1: Gini imputation comparison

Imputing the Gini index data adds sixteen countries and between 100,000 and 160,000 observations to the regression sample (depending on the outcome variable being studied, as not all survey questions were asked in all countries or all waves of the survey). Of the fifty-seven countries that appear in the first main regression sample for which there is at least some raw Gini data, only for seven does the mean

imputed Gini differ from the unimputed one by more than 5%: Estonia, Ireland, Japan, Nigeria, Poland, Romania, and Ukraine. These represent just over 12% of the countries in the regression, and over 11% of the total observations in the regression sample. It should be noted that all of these countries bar two are European, and most are highly developed; it therefore does not appear to be the case that imputed Gini values are more noisy or error-prone for less-developed countries. However, it is still possible that those countries that enter the regression sample once Gini is imputed are doing so with values substantially outside the norm. For this reason, Table 3.D.2 compares the average raw Gini index for the entire period that a country participated in the World Values Survey with the imputed Gini, for those countries where assigning Gini data directly proved impossible.

	Year range	Average Gini for the period	Imputed Gini	Absolute difference	% deviation from WB mean
Algeria	2002-2014	27.6	27.6	0	0
Australia	1981 - 2018	33.517	33.9	0.383	1.143
Azerbaijan	1997-2018	28.36	26.6	1.76	6.206
Bangladesh	1996-2018	32.775	33.4	0.628	1.916
Egypt	2001-2018	30.783	28.3	2.483	8.066
India	1990-2012	34.3	33.375	0.925	2.697
Iran	2000-2020	40.99	44.8	3.81	9.295
Iraq	2004-2018	29.05	29.5	0.45	1.549
Jordan	2001-2018	34.3	33.8	0.5	1.458
Malaysia	2006-2018	42.967	43.9	0.933	2.171
Morocco	2001-2021	40.1	40.7	0.6	1.496
North Macedonia	1998-2019	36.609	42.8	6.191	16.91
Rwanda	2007-2012	47.2	47.2	0	0
South Africa	1982-2013	61.66	61.325	0.335	0.543
Trinidad and Tobago	2006-2010	40.3	40.3	0	0
Zimbabwe	2001-2020	45.933	43.2	2.733	5.950

Table 3.D.2: Imputed Gini vs. World Bank average

Of the sixteen countries that enter the regression sample when Gini is imputed, five have average imputed Gini values that differ by more than 5% from the average for the entire timespan the country participated in the World Values Survey: these are Azerbaijan, Egypt, Iran, North Macedonia, and Zimbabwe. (Further, for Trinidad and Tobago, Gini index data is available only from 1988 and 1992, over a decade before the country entered the World Values Survey. Therefore, while the deviation is zero, it is impossible to evaluate how accurate the figure actually is. A similar case is Algeria which, while not out of sample, only recorded one value of the Gini index for the entire period, in 2011.) Together these represent over 31% of the countries in Table 3.D.2, but just 3% of the overall regression dataset, mostly owing to the fact these countries tended to enter the World Values Survey later on.

Table 3.D.3 does a similar calculus for the outcome variable *Confidence in Church*: the first column reports the average value of this variable⁶ for the first main regression sample when the Gini is not imputed; the second column reports the average of the outcome variable in the sample where the Gini *is* imputed. The average value of the outcome variable in the sample with the imputed Gini differs from the value in the unimputed Gini sample by more than 5% for just three countries: China, Croatia, and Turkey. None of these are countries for which the imputed Gini itself differed significantly from the mean of the unimputed Gini.

⁶The outcome variable *Confidence in Church* is an inversion of the original variable, E069, in the IVS data, so that higher values correspond to greater confidence.

	Raw Gini	Imputed Gini	Absolute	%
	sample	sample	difference	deviation
Albania	2.658	2.658	0	0
Algeria		3.085		
Argentina	2.601	2.601	0	0
Armenia	3.225	3.225	0	0
Australia		2.249		
Austria	2.279	2.279	0	0
Azerbaijan		2.279		
Bangladesh		3.884		
Belarus	2.933	2.933	0	0
Belgium	2.157	2.241	0.084	3.897
Brazil	2.921	2.921	0	0
Bulgaria	2.405	2.400	0.005	0.190
Canada	2.743	2.743	0	0
Chile	3.100	2.985	0.115	3.708
China	1.972	2.082	0.111	5.604
Colombia	3.250	3.250	0	0
Croatia	2.191	2.481	0.289	13.199
Cyprus	2.929	2.929	0	0
Czechia	1.791	1.870	0.079	4.411
Denmark	2.691	2.679	0.012	0.438
Egypt		3.180		
Estonia	2.573	2.523	0.500	1.939
Finland	2.555	2.561	0.005	0.214
France	2.272	2.272	0	0
Georgia	3.572	3.572	0	0
Germany	2.127	2.127	0	0
Ghana	3.530	3.530	0	0
Greece	2.700	2.700	0	0
Hungary	2.341	2.345	0.004	0.155
India		3.456		
Indonesia	3.628	3.628	0	0
Iran	.	3.158		

 Table 3.D.3:
 Confidence in government: Mean in raw vs. imputed sample

Iraq		3.135		
Ireland	2.653	2.700	0.047	1.768
Italy	2.867	2.868	0.001	0.045
Japan	1.638	1.625	0.013	0.815
Jordan		3.378		
Korea (Republic of)	2.529	2.489	0.040	1.579
Kyrgyzstan	2.485	2.485	0	0
Latvia	2.726	2.757	0.031	1.138
Lithuania	2.939	2.882	0.056	1.911
Luxembourg	2.255	2.255	0	0
Malaysia		3.386		
Mexico	3.068	3.068	0	0
Moldova	3.117	3.117	0	0
Morocco		3.551		
Netherlands	2.064	2.064	0	0
Nigeria	3.628	3.614	0.015	0.403
North Macedonia		2.760		
Norway	2.405	2.457	0.052	2.172
Pakistan	3.582	3.686	0.105	2.928
Peru	2.916	2.916	0	0
Philippines	3.672	3.610	0.063	1.705
Poland	2.777	2.822	0.045	1.626
Portugal	2.995	3.042	0.047	1.563
Romania	3.249	3.296	0.048	1.473
Russian Federation	2.775	2.777	0.002	0.070
Rwanda		2.710		
Slovakia	2.655	2.716	0.061	2.301
Slovenia	2.160	2.185	0.025	1.137
South Africa		3.295		
Spain	2.215	2.232	0.017	0.779
Sweden	2.434	2.432	0.002	0.092
Switzerland	2.262	2.262	0	0
Thailand	3.058	3.058	0	0
Trinidad and Tobago		2.953		
Turkey	3.090	2.924	0.165	5.356
Ukraine	3.008	2.970	0.037	1.241

United Kingdom	2.334	2.334	0	0
United States of America	2.799	2.799	0	0
Uruguay	2.341	2.341	0	0
Viet Nam	2.689	2.689	0	0
Zimbabwe		3.421		