On Kinks and Candidates: Geo-Variation in Escort Ads and Electoral Outcomes

John P. Schoeneman Jr.*[1] and Christopher Zorn[2]

[1]School of Global Studies, Oklahoma State University
[2]Department of Political Science, Pennsylvania State University

Abstract

Existing work in the social sciences has documented a number of ties between psychological, physiological, and sociological components that make up political identities and individuals' corresponding sexual practices. However, this literature is far from conclusive, and has primarily relied on survey data, which are subject to well-known and serious sampling and response biases. We seek to overcome these biases by using machine learning algorithms to determine geo-variation in the characteristics of a dataset of 2.3 million online escort ads from Backpage and Craigslist, covering nearly 50,000 unique locations in the United States. We compare the resulting characteristics with the political outcomes of the locations in which these ads are listed to investigate the relationship between political geo-variation and such variation in manifest sexual practices. By using observed data on actual services provided by sex workers in an area, our study adds robustness over existing research that relies on survey data, and allows us to better investigate nuances of those findings.

Keywords— Political ideology, sexuality, survey bias, text analysis, machine learning, LLMs

*Paper prepared for presentation at the Annual Meeting of the Western Political Science Association, March 28-30, 2024, Vancouver, BC. Correspondence email: john.schoeneman@okstate.edu.

1 Introduction

Sex is – ideally – all around us. Beyond the biological imperative, sex is infused in every aspect of Western culture (e.g., Atwood, 2009), including (as Foucault famously observed) in systems of social power. Its ubiquity touches every social institution; Berlant and Warner, for example, note that "(H)eteronormative forms of intimacy are supported ... not only by overt referential discourse such as love plots and sentimentality but materially, in marriage and family law, in the architecture of the domestic, in the zoning of work and politics" (1998, p. 562).

It is perhaps a little ironic, then, that social norms around sex and sexual intimacy also often work to limit our understanding of the topic. Contemporary western views of sex emphasize its private nature while at the same time encouraging relatively narrow – and sometimes rigid – views of what constitute "normal" sexual desires, activities, and values. The effect of such double-edged norms is to render sexual behavior simultaneously entirely secret and fully subject to societal pressure. The result is that we often live our (private) sexual lives in ways that are governed to varying degrees by (public) non-sexual expectations and standards. Perhaps even more important from the perspective of social scientists, these pressures mean that beliefs and behaviors related to sex are often among the most challenging to investigate empirically.

Our goal in this paper is to consider the question of sex as it relates to politics, in an empirical sense. In short, we want to know to what extent sexual preferences and practices map onto political (and, particularly, ideological) beliefs in the U.S. Our aim here is primarily descriptive; we reserve (likely futile) efforts at establishing causality on this subject for future work. Instead, our primary focus is on questions of measurement and measurement validity. As we describe herein, doing so requires that we short-circuit the norms described above, ideally in a way that allows us to connect externally reliable, valid indicators of political preferences onto similarly grounded measures of sexual practices. We do this by relying on a unique database of sexual behavior, one that is knowingly public, largely anonymous, and driven (by market and other forces) to reflect local conditions. Some key results of this work cut against prevailing conclusions regarding the connection between ideology and sex, a fact we attribute at least partially to our ability to overcome the dynamics described above.

2 Theory And Background

2.1 Sex, Sexuality, and Political Ideology

As manifestations of the same fundamental values that drive much of human behavior, political / ideological beliefs have long been understood to correlate with people's beliefs, preferences, and actions in non-political realms. From humor (Wilson, 1990) to the cars they drive (Hetherington & Weiler, 2018) to their opinions about baseball's designated hitter rule (Gill & Zorn, 2007), individuals' ideological attitudes covary with – and, sometimes, shape – their personal proclivities in a host of areas having little or nothing to do directly with politics. In many if not most instances, these connections stem from more general associations among basic personality characteristics, political / ideological preferences, and individuals' general behavioral tendencies. To take but one example, it's generally well established that individuals' general tendency toward risk aversion – often motivated by fear of death and/or uncertainty, and manifested as a need for order and/or a lack of openness to new experiences – is associated with greater political conservatism (e.g., Jost et al., 2007).

It is therefore unsurprising to learn that among the human behaviors we often see linked to political and ideological attitudes, those related to sexuality feature prominently. Early studies of the link between political attitudes and sexual behavior tended to focus on political conservatism, and to draw parallels between conservatives' general tendencies towards risk aversion, submission to authority, and opposition to change and lower frequency and diversity of sexual activities. Cottle et al. (1970), for example, describe a pattern of relationships between conscious and unconscious sexual identification on the one hand and a series of political constructs relating to the Vietnam War, birth control, sex role morality, racial discrimination, and what they term "intolerance of expressivity" on the other. In another example, Joe et al. (1976) find that both men and women with more strongly conservative political attitudes had substantially fewer and less diverse sexual experiences than those exhibiting low conservatism, and that the former were also more likely to express that sexual activity was morally wrong (see also, e.g., Eysenck, 1976; Joe and Kostyla, 1975; Thomas, 1975). In a subsequent study of over 300 German university students, Brody et al. (1996) found that, while the magnitude of those associations was generally small, left-wing political beliefs were associated with greater levels of sexual disinhibition in both men and women.

More recent work has deepened our understanding of these links. In one important study, Hatemi et al.

(2017) revisit Eysenck's (1976) analysis of personal sexual preferences and political ideology. Their analysis of a web-based sample of U.S. adults finds a mixed pattern of marginal (controlling for demographics) associations between 15 different sexual behaviors and seven indicators of political ideology / preferences. Among the most robust findings was a robust positive association between ideological liberalism and a range of more "adventurous" sexual activity, including various forms of risky sexual behavior (e.g., having sex while drunk/high, sex with a stranger, infidelity, etc.), masturbation, and respondents' lifetime number of sexual partners. Conversely, self-reported conservatives were more likely to have begun having sex later in life, had fewer total partners, and less diversity in their sexual experiences.

Relying on nationally-representative samples in Denmark and the U.S., Petersen (2018) extends this work, exploring the associations between Simpson and Gangestad's (1991) sociosexual orientation inventory and measures of political ideology [including both self-reported ideology and psychological measures of right-wing authoritarianism (*RWA*) and social dominance orientation (*SDO*)]. The former measures the extent to which an individual pursues an "unrestricted sexual strategy," that is, one characterized by higher numbers of relatively short-term, uncommitted sexual encounters, with a focus on physical attractiveness and social status over commitment and relationship stability. Petersen hypothesizes that "people with an unrestricted sexual strategy are expected to be liberal as measured by RWA but conservative as measured by SDO." After controlling for other demographic and psychological factors shown to predict sociosexual orientation, his analyses find patterns consistent with these hypotheses in both the U.S. and Denmark, along with a relative absence of mediating effects for respondents' gender and social status.

While causal linkages are difficult if not impossible to establish, such behavioral associations are reflective of a larger pattern of dependencies between political and sexual attitudes. As far back as 1977, Hershey and Sullivan noted that "political conservatives on a wide variety of subjects support traditional sex roles to a greater extent than do political liberals" (Hershey & Sullivan, 1977, p. 55). Such attitudinal associations can be traced back to at least the nineteenth century (e.g., Freedman 1982), and reflects both the individual-level psychological dynamics described above as well as broader social institutions around the construction of gender and sexuality (see, e.g., Brown, 1988; Duerst-Lahti, 2008).

More contemporary empirical research has focused on the extent to which political preferences correlate

with negative attitudes toward members of sexual minorities.[1] Poteat and Mereish (2012) find that both RWA and SDO are strongly associated with support for restrictive policies against gays and lesbians, and that those connections were driven by prejudice in the former and a desire to perpetuate power hierarchies in the latter. Lytle et al. (2017) investigate the association between *essentialism* – the belief that sexual orientation is universal, discrete, and immutable – and anti-LGBT prejudice, finding that such beliefs mediate the pallative power of cross-orientation contact in mitigating such prejudice (see also Hoyt & Parry, 2018). In related work, Hoyt et al. (2019) describe how political conservatives rely upon "strategic essentialist" accounts of sexual orientation to justify their prejudice against LGBT individuals, and how conservatives are also more resistant than liberals to messaging designed to combat such essentialism.

2.2 Bias in Empirical Studies of Sexuality

A consistent concern with nearly all the studies referenced above is the potential for bias (and, thus, a lack of generalizability of their findings). For a host of reasons including social desirability, a desire for privacy, and incentives to misrepresent the truth, empirical work on sexual attitudes and behaviors that relies on self-reported data are ripe for bias. The sources of such bias fall into three main categories: *volunteer bias* (that is, bias due to the fact that those willing to take part in such studies are systematically different from those who are not), *nonresponse bias* (where study participants differentially refuse to answer survey and other self-report items), and *veracity effects* (where respondents misrepresent the truth of their attitudes and/or behaviors in their responses).

Volunteer bias (also called "sample bias") is among the longest-studied aspects of bias in sexuality research (e.g., Farkas et al., 1978; Kaats & Davis, 1971). Catania et al. (1986) characterize volunteer bias as "relatively less...than might be expected," and develop a disclosure measure designed to overcome such bias in face-to-face interview settings. Strassberg and Lowe (1995) are less optimistic; they note that only a third of students who initially participated in a general survey were also willing to take volunteer for a follow-up (questionnaire-based) study, and that differences between volunteers and non-volunteers were substantial, with the former exhibiting less sexual guilt and higher levels of sexual experience than the latter.

[1] There is also at least some evidence that ideological differences shape the way individuals perceive sexual orientation. Stern et al. (2013) show that conservatives are more likely than liberals to use "gender inversion" cues (e.g., feminine = gay) to differentiate between gay and straight individuals, and that conservatives are also more willing to agree with statements that conform to such stereotypes.

They conclude by noting that "those with less sexual experience and/or more sexual guilt would, if they could be persuaded to participate in sex research, produce results similar to those researchers have already obtained with volunteers. However, just how their results might differ is far from obvious" (Strassberg & Lowe, 1995, p. 379). Dunne et al.'s (1997) much broader survey of more than 9000 adult twins in Australia reaches similar conclusions:

> "In comparison to people who refused, consenters had less conservative sexual attitudes, more novelty-seeking and less harm-avoidant personalities, less regular church attendance, less conservative political preferences, were more likely to smoke tobacco and drink alcohol regularly, had an earlier onset of sexual behaviour, more adverse sexual experiences and poorer mental health." (Dunne et al., 1997, p. 851)

More recent studies of volunteer bias uncover similar patterns. Bouchard et al. (2019), for example, find that "(W)omen and men with more sex-positive attitudes tended to be more willing to participate in both unclothed and clothed studies. Sexual orientation was also associated with willingness to participate in sex research, particularly for women, and some personality traits from the five-factor model were associated with willingness to participate in sex research scenarios" (Bouchard et al., 2019, p. 35) (see also Roath et al., 2023).

We find similar patterns of bias in investigations of nonresponse and veracity effects. As in other sorts of surveys, those who are successfully recruited still frequently fall prey to item nonresponse (Evans et al., 2008; Wiederman, 1993). And even among those choosing to respond, misrepresentation among self-reported attitudes and behaviors surrounding sex frequently occur. Marelich et al. (2008), for example, develop a multivariate behavioral sexual deception scale, finding that those who rank highly on that scale tend to be more promiscuous (with higher numbers of partners and more frequent one-night stands); they also report that "(M)en were more likely to use blatant lies to have sex, while women were more likely to have sex to avoid confrontation" (Marelich et al., 2008, p. 27). More recent work shows that, among other influences, misrepresentation on surveys of sexual behavior is shaped by demographic factors (Liddon et al., 2021), interviewer effects (Koroknay-Palicz & Montalvao, 2020), and social desirability bias (King, 2022).

At the same time, it is important not to overstate the extent of the bias described. Existing studies, for example, suggest little difference in the degree or direction of bias in internet-based vs. in-person samples

(Ross et al., 2005), and recent work supports the idea that members of sexual minorities are at worst no more likely to be underrepresented than those in the majority (Lee et al., 2018). And as Dunne et al. concluded more than a quarter-century ago, "the modest effect sizes found here indicate that this bias may not seriously compromise the validity of population estimates. Furthermore, it is possible that this overestimation of sexual experience due to voluntarism counteracts, to some extent, the often observed tendency for respondents in sex surveys to minimize or underreport the frequency and diversity of sexual behaviour" (1997, p. 852).

Taken together, work on bias in empirical work on sexuality paints a mixed picture. On the one hand, the mechanisms by which biases occur are intuitive and well-understood, and suggest that (in general) patterns of self-reported sexual preferences and activities might vary significantly from those of the general public. It is also straightforward to imagine instances where any bias would be correlated with other phenomena – including psychological traits – that would similarly drive commonly-used predictors of such preferences and activities, resulting in the canonical form of "selection bias" so commonplace in observational studies. On the other hand, empirical work on the subject is ambiguous both about the degree of bias present and the correlates of such bias.

2.3 Escort Advertisements

We highlight the literatures on political correlates of sexual attitudes / behaviors and on the bias present in those studies to motivate our current work. Because political beliefs are often strongly correlated with other factors that themselves are related to both sexual behavior and bias in the reporting of that behavior, we might expect that existing work examining those links would be particularly likely to suffer the consequences of such bias. Hatemi et al. (2017, p. 324) acknowledge as much, noting that "(O)ne concern that is omnipresent in studies of sexual mores is reporting bias, due to social desirability or motivated reasoning, for example. This possibility can never be fully vitiated." While Hatemi et al. go on to note that their own study takes steps to mitigate the potential impact of such biases, an omnipresent challenge is the inability to know with certainty the degree such bias and its effects on the findings in any given study.

The central challenge of the bias described above lies in empirical work's reliance on self-reporting (in various forms) as a means of generating data. The fact that subjects of studies take part voluntarily, that having done so they remain able to choose when and when not to respond, and that they are assumed to provide information that is truthful and valid all drive the mechanisms by which bias might occur. One

potential way to circumvent this bias, then, is to rely on sources other than self-reports when seeking to investigate sexual behavior (and, particularly, its links to political attitudes).

We do so here by examining public manifestations of private sexual dynamics, in the context of the marketplace for sexual services and sex work. More specifically, we focus on advertisements for *escort services*, understood broadly to mean individuals who exchange sexual access and / or companionship for money. Escort services are popularly viewed as a particular form of prostitution, albeit one that is frequently better compensated than other forms (Edlund et al., 2009). In the present context, however, a focus on escort ads offers a number of advantages in the examination of sexual behavior.

First among these is the general ubiquity of such services. While escort services have long available in almost every geographical location in the U.S., the spread of the internet in the late 20th century further extended their presence. In one early study, Parsons et al. (2004) note that male escorts reported higher social and economic class and reduced drug use among on-line customers, as well as higher pay, generally safer sexual practices, and fewer concerns about law enforcement when engaging with customers via the internet. By 2008, Castle and Lee were able to characterize the "typical" escort website, noting that most advertised female escorts for male customers, provided photos and descriptions of services offered, and were usually described from the viewpoint of the customer. Subsequent descriptive studies support the idea that the internet has remained the central locus of advertisements for escort services (Burghart, 2018; Pruitt & Krull, 2010).

A related advantage is that escort advertisements tend to be a reliable source for information about the availability of sexual services in an area. The aforementioned descriptive studies make clear that such advertisements broadly reflect the services on offer, a fact that extends outside the U.S. as well (e.g., Fedorenko, 2015; Meneses-Falcon et al., 2017). Detailed investigations of such advertisements also indicate that the information contained in them is largely responsive to local market forces; for example, DeAngelo et al.'s (2019) analysis of more than 30 million such ads finds that advertised prices largely reflect supply-side factors (e.g., with out-call services demanding higher prices than in-call), and that ad content (including prices) is largely responsive to local market conditions (see also Edlund et al., 2009). Taken together with the relative safety provided by the anonymity of the internet (e.g., Castle & Lee, 2008) and the increasing tendency for localities to normalize sex work of this type (e.g., Maticka-Tyndale et al., 2005), it's reasonable to think that the supply of such services will be largely responsive to local demand. The result is that the content

of advertisements for escort services offer the opportunity to reliably, validly measure and characterize different local sexual cultures, and to do so in a way that circumvents the biases associated with self-reported indicators of sexual attitudes and activities.

3 Data and Design

3.1 Data

We began by web-scraping 2,346,365 advertisements (or "listings") for "escort services" listed on the online services Backpage and Craigslist between January 12, 2013 and February 15th, 2018.[2] As we discuss below, such ads vary widely in their content; nearly all include textual descriptions, photographs, and contact information, while some also link to videos and other online content. Crucially for our purposes, ads are both roughly geolocated by their listing location and often contain additional location information (addresses, neighborhoods, etc.). Figure 1 shows examples of two such listings within the time frame of our data.

For each ad, the scraper records the hyperlink, the posting time, the posting body, and the location listed for the ad. We took a number of steps to prepare these raw data for analysis. First, many ads are duplicates of one another, either due to ads being posted multiple times in our time sample or because the same ad was scraped from multiple sites. To prevent our analyses from being overly influenced by such duplications, we remove duplicates in the same geographic location from our sample. In addition to identifying exact matches as duplicates, we also calculate a similarity percentage score using the Levenshtein distance divided by they number of characters in the longest post body of the two ads being compared. We define any percentage higher than 80% in the same location as a duplicate, and remove those as well.

Second, many of the ads are exceptionally short, and only contain a phone number, brief, generic language, or no text at all. We remove ads with fewer than 25 characters from our sample. We also remove those

[2]First created in 1995, Craigslist is an on-line classified advertisements site that currently hosts ads in more than 70 countries. It initially had a section for "erotic" services, which was changed to "adult" in 2009. Craigslist removed its "adult" section in 2010, but continued to accept advertisements for what were effectively escort services through its "casual encounters" section, under "personal." Craigslist removed its "personals" section entirely in March of 2018. Backpage was founded in 2004 as a competitor to Craigslist; it was a major beneficiary of the removal of adult services from Craigslist in 2010. In response to political pressure, Backpage closed its adult section in early 2017; Backpage was subsequently seized and shut down by the U.S. Department of Justice in April of 2018, following a a 93-count indictment of its former owners and executives on charges of facilitating prostitution, money laundering, and conspiracy (U.S. Department of Justice, 2018).

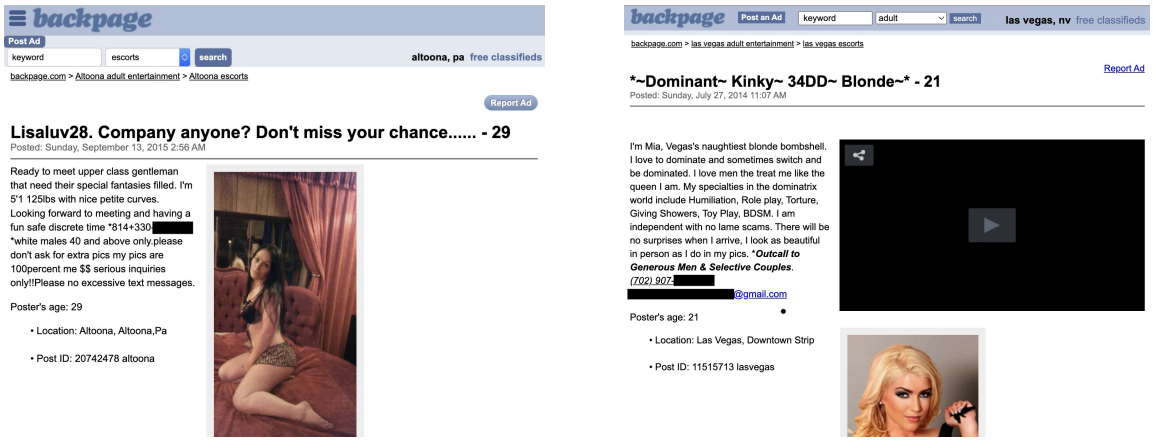

Figure 1: Example Ads From Backpage

over 1000 characters, as this is roughly when the tail of the skewed post lengths begin. Upon investigating posts of this length, we found that most were not escort ads, but a somewhat random collection of posts that had either nothing to with sexual proposition or non-escort site users that were looking for consensual, unpaid sexual encounters. Lastly, we drop observations that do not have location data.

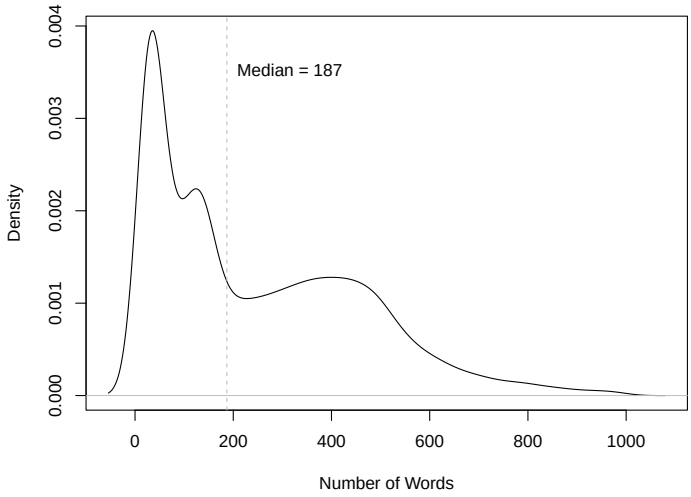

Figure 2: Density Plot of Ad Length

After dropping observations based on duplicates, length, and lack of location data we are left with a total of analyzable 552,424 posts. Figure 2 plots the density of the number of words per listing for our valid postings. For these remaining posts, we assign a county location using the latitude and longitude of the city the ad was listed in.[3] We then merge 2016 county-level presidential election data. In the remaining sample

[3]If multiple cities were listed for one ad, we used the mean coordinates. When multiple cities were listed, they were usually all nearby cities and this method should be a reflection of the intended audience.

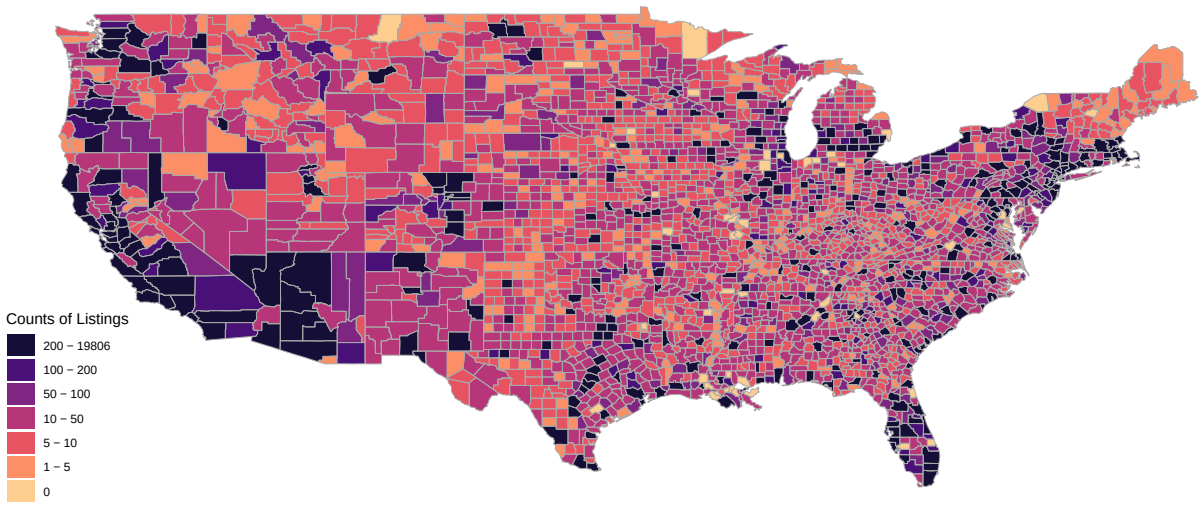

Figure 3: Geographic Distribution of Escort Ads

of ads, 3,031 counties in the continental U.S. are represented; we plot the number of such ads per county in Figure 3.

The pattern of ad frequencies in Figure 3 are of some interest, in that they show that, while such ads are ubiquitous across the U.S., there is also substantial geovariation in their prevalence. On the one hand, a cursory examination confirms that, because they reflect sum totals, they map closely to population numbers (cf. Munroe, 2012). On the other hand, notable exceptions to that pattern emerge; for example, Mountrail County, ND (2010 population 7,708), Colbert County, AL (pop. 54,428), and Hall County, NE (pop. 58,607) all have substantially higher numbers of listings than one would expect, given their size. Outside of population centers, regional differences become apparent, with rural New England, the Mountain West, and the Texas-Mexico border seeing fewer such ads than other areas.

What, if any, is the political valence of locales where escort listings are more or less prevalent? Figure 4 plots the natural logarithm of the number of such ads – normalized by population – in each county against

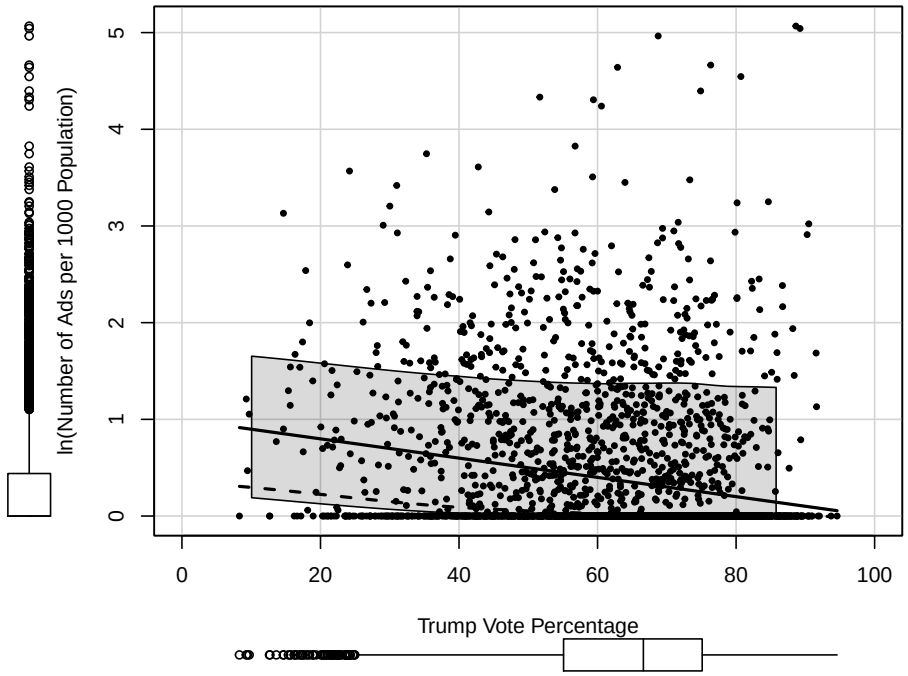

Figure 4: Number of Advertisements, by Trump Vote Percentage

the percentage of the vote received by former President Donald Trump in the 2016 Presidential general election.[4] We observe a very slight negative relationship between the two variables; the estimated $\hat{\beta}$ from the implicit linear regression in Figure 4 is -0.01 (with an estimated standard error of 0.0008), indicating that a one percentage-point increase in Trumps' vote share corresponds to an expected decrease in the number of such ads of around one percent. Note, however, that that model accounts for relatively little of the variation in ad prevalence (R^2=0.046), and much of the result is driven by sparsely-populated counties in which Trump received the lion's share of the two-party vote and in which Craigslist and Backpage did not exist. Limiting the analysis to counties where at least one ad appeared yields a $\hat{\beta}$ of 0.0005 (s.e. = 0.002, t = 0.3) and an R^2=0.00008. In short, in circumstances where they could occur, we find no evidence that escort ads appeared with either greater or lesser frequency in relatively conservative or liberal communities.

3.2 Research Design

Of greater importance than the prevalence of escort listings, however, is their content. To measure variation in escort ad characteristics, we employed two different text analysis approaches. In the first, we used the

[4]To deal with the presence of counties where no ads were listed, we add one to the value of every observation in the data; this means that on the plot, an observed value of zero corresponds to a county in which zero ads were run.

`LiblineaR` engine (2022) for a support vector machine (SVM) in the `tidymodels` package (2020) in `R` to assign feature weights for each word in the ads that occurs at least 50 times in the corpus, using Trump's vote share for the county the ad appeared in as the dependent variable. Because the language of escort ads is somewhat unique, we then trained GloVe word embeddings (Pennington et al., 2014) for the corpus, rather than using a pre-trained model. This method has been shown in the past to have success as a semi-supervised method for determining word similarity in unique corpi (Rice & Zorn, 2021). For our analysis, we then chose ten words for six categories and calculated the cosine similarity of the words with all the words that have SVM feature weights, and took the sum of the 1,000 best matched words similarity scores multiplied by the SVM feature weights to create a comparable score for how much a particular word correlates with Trump's vote share in 2016.

In the second approach, we use the Python package *BERTopic* (Grootendorst, 2022) to categorize the ads in a semi-supervised manner and then check the categories for geo-variation and correlation with county-level political outcomes. The *BERTopic* package uses bidirectional encoder transformation and class-based Term Frequency-Inverse Document Frequency (TF-IDF) to assign cluster membership for text topic modeling. Past work has shown this approach to be powerful for topic modeling short-text across versatile domains (e.g., Egger & Yu, 2022). We initially fitted the model fully unsupervised (using default settings) for all unique escort ads, to explore what topics the model was capable of identifying. While the model did find some topics of interest, many of the ads we categorized into nonsensical categories. After reading through the results and ads, we chose 14 topics that (a) had intuitive meaning, and (b) had at least 1000 occurrences in the corpus. For these 14 topics, we used *BERTopic*'s semi-supervised zero-shot modeling approach that embeds the 14 labels and assigns labels to the ads based on cosine similarity. For ads that were not 85% similar, they were assigned topics in an unsupervised manner. From these results, we checked the unsupervised assigned labels that have over 500 listings and merged those substantively similar with the 14 labels we chose. We then assign each topic a country level score based on the percentage of unique ads that are classified as the topic, and compare these scores to county-level measures of the percentage of the vote former President Trump won in the 2016 presidential election. That variable is a contemporaneous (with our ad data) measure of the extent to which a county is broadly conservative or liberal.

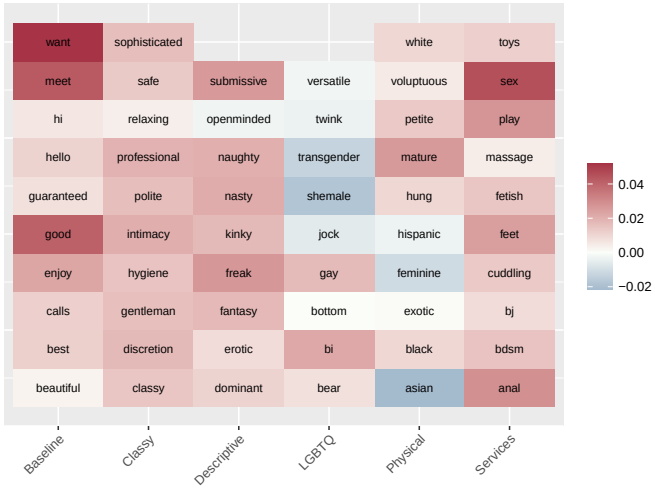

Figure 5: Escort Ad Word Scores for the 2016 Election Results.

4 Results

The word scores from the SVM predicting Trump's vote share are normally distributed, with a mean of about 0.02 (as compared to a theoretical expectation of zero). The scores themselves were relatively small, as seen in Figure 5. Even the highest values of the word scores only approach around 0.05, while the lowest drop to around -0.02.

In addition, we also find no clear covariances with political outcomes. The two topic models did not yield substantively different results, with correlations between topics in counties being substantively small and with no particularly discernible pattern (Figure 6). The patterns of topics do tend to "cluster" in substantively comprehensible ways, with ads containing different linguistic content mapping relatively closely onto intuitive classifications. At the same time, we can discern no particular pattern of association between the topics themselves and vote shares for former President Trump.

For all the topics, the only one that had substantial coverage across the continental U.S. was "playful." Figure 7 shows a bivariate choropleth map depicting the "playful" ad coverage and Trump's vote share. Darker blue counties indicate higher incidences of "playful" ads, while darker red (appropriately) indicate higher proportions of votes cast for Trump. Counties which are either dark blue or dark red are those where the association is negative; those which are either light white or dark brown are those where the two covary positively. In our figure, greatest substantially meaningful geovariation comes from political outcomes, not ad classification. So, for example, many "red" counties (those where a high percentage of voters cat ballots

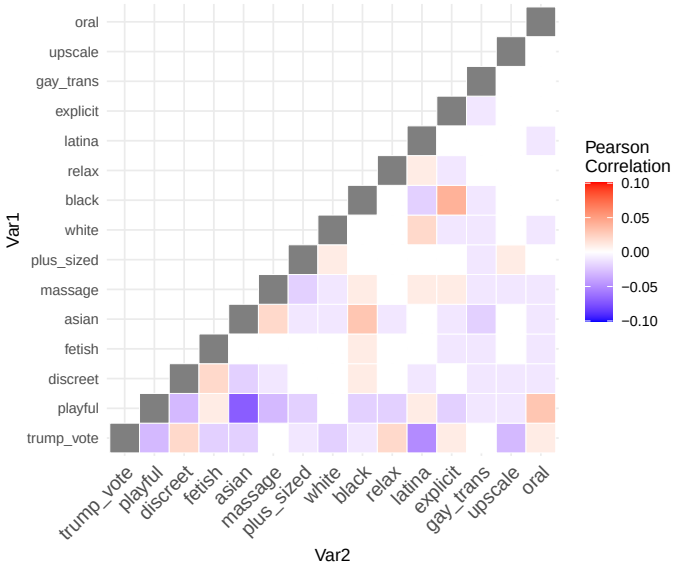

Figure 6: Cross Correlation: Topic Modeling and 2016 Election Results

for Trump) appear in red, many "blue" counties are rendered in light or medium grey, indicating that those counties did not have higher (or lower) than expected levels of "playful" ads. Exceptions to that pattern appear to be more-or-less randomly distributed across the map.

Collectively, these results indicate that the content of the escort ads comprising our data are relatively consistent across the U.S. More broadly, the lack of a clear association with revealed local political preferences suggest that sexual preferences do not correlate particularly strongly with political geovariation. While we recognize the ecological challenges associated with drawing individual-level conclusions from aggregate data, our findings here are largely inconsistent with a situation where there are substantial differences in sexual practices (or, at least, in demand for the market for sexual services) at the individual level.

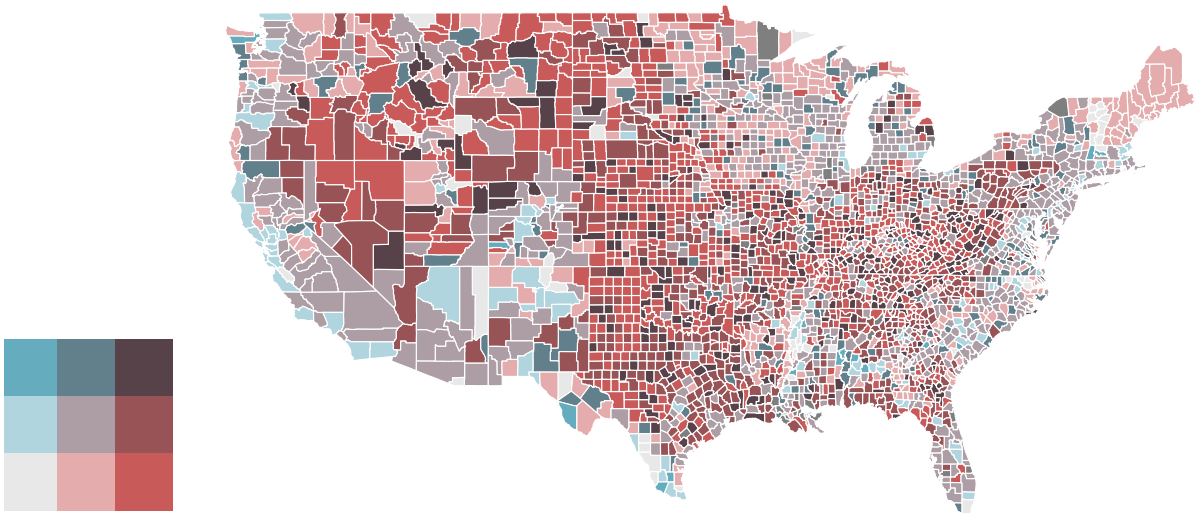

Figure 7: Bivariate choropleth of "playful" escort ads and 2016 election results. Blue is the percentage of ads that are classified as "playful," and red is Trump's county vote share. Bins were divided by equal value. See text for details.

5 Conclusion

Social mores around sex in the United States (and much of the world) generally dictate that sexual relations are private affairs between those involved. Furthermore, cultural and religious pressure can make individuals embarrassed to discuss sex and be less revealing of true behavior. Survey data is already afflicted by a number of challenges for more public, less governed topics and so survey data on sexual data behavior should be validated through other methods. In this paper, we work towards finding solutions to the challenges facing survey data by analyzing a dataset that is directly tied to actual, not claimed, sexual behavior to determine if local sexual behavior in the United States supports what past work has found, namely sexual variation across political ideologies, using survey data.

Here, we find that – at least with respect to the demand for commercial services related to sexual practices – there is little systematic variation in the volume and content of sexual activity between "blue" and "red" counties in the United States. More specifically, our results suggest the relationship between sexual preferences being catered to in online escort ads along a number of categories and political outcomes are too small to be substantively meaningful or considered different from noise. In concluding, we offer two implications of, and several important caveats to, these findings.

First: At the risk of overstating, we believe this finding indirectly supports the notion that survey-based

results are in fact confounded by social desirability or other biases. The lack of important differences across localities of widely-varying ideological composition is consistent with this result, though at this point we can only speculate about the mechanism(s) by which the bias occurs. For example, one possible explanation is that, while political conservatives face social pressure to answer that they are sexually more conservative than those on the left, behind closed doors sexual behavior is remarkably consistent across the ideological spectrum. Alternatively, social desirability pressures on those on the political left may lead them to overstate the diversity of their sexual preferences and practices.

Second, we believe our study suggests a way forward for dealing with the biases we describe. Conventional approaches address behavior at the individual level, but rely on self-selection and self-reporting. Here, we instead focus on public, market-based indicators, but can do so only at the aggregate level, raising questions about the ecological validity of our work were we to extent its conclusions to the level of the individual. Future studies might consider hybrid approaches that integrate data of both types, both to get a clearer picture of the extent and nature of potential bias, and to correct for the shortcomings of each.

Finally, while we have attempted to exhaustively measure the geovariation in escort ad characteristics for a relatively large dataset (survey datasets rarely exceed several thousand responses), we recognize that this study is just the beginning for investigating recorded sexual behavior. In particular, there a number of shortcomings that will need to addressed in future work. First, these findings should be compared to a larger sample of the population, as our sample only includes the population that uses escort services, and those for which we have data coverage during the period studied. This sample may not be representative of the general population. In contrast, other data sources may be more so; online pornography, for example, is used much more widely than escort services, and search terms and click rates for local areas could reveal preferences for a larger proportion of the population. Second, topic modeling to extract more detailed, specific services provided using LLMs should be explored.

References

Atwood, F. (2009). *Mainstreaming sex: The sexualisation of western culture*. Bloomsbury.

Berlant, L., & Warner, M. (1998). Sex in public. *Critical Inquiry*, *24*(2), 547–566. https://doi.org/10.1086/448884

Bouchard, K. N., Stewart, J. G., Boyer, S. C., Holden, R. R., & Pukall, C. F. (2019). Sexuality and personality correlates of willingness to participate in sex research. *The Canadian Journal of Human Sexuality*, *28*(1), 26–37.

Brody, S., Rau, H., Fuhrer, N., & Hillebrand, H. (1996). Traditional ideology as an inhibitor of sexual behavior. *The Journal of Psychology*, *130*(6), 615–626.

Brown, W. (1988). *Manhood and politics: A feminist reading in political theory*. Rowan; Littlefield.

Burghart, K. O. (2018). What's on sale? a discourse analysis of four distinctive online escort advertisement websites. *Sexuality & Culture*, *22*(1), 316–335.

Castle, T., & Lee, J. (2008). Ordering sex in cyberspace: A content analysis of escort websites. *International Journal of Cultural Studies*, *11*(1), 107–121. https://doi.org/10.1177/1367877907086395

Catania, J. A., McDermott, L., & Pollack, L. (1986). Questionnaire response bias and face-to-face interview sample bias in sexuality research. *The Journal of Sex Research*, *22*, 52–72.

Cottle, T. J., Edwards, C. N., & Pleck, J. (1970). The relationship of sex role identity and social and political attitudes1. *Journal of Personality*, *38*(3), 435–452.

DeAngelo, G., Shapiro, J. N., Borowitz, J., Cafarella, M., Re, C., & Shiffman, G. (2019). Pricing risk in prostitution: Evidence from online sex ads. *Journal of Risk and Uncertainty*, *59*, 281–305.

Duerst-Lahti, G. (2008). Gender ideology: Masculinism and feminalism. In G. Goertz & A. G. Mazur (Eds.), *Politics, gender, and concepts: Theory and methodology* (pp. 159–192). Cambridge University Press.

Dunne, M. P., Martin, N. G., Bailey, J. M., Heath, A. C., Bucholz, K. K., Madden, P. A., & Statham, D. J. (1997). Participation bias in a sexuality survey: psychological and behavioural characteristics of responders and non-responders. *International Journal of Epidemiology*, *26*(4), 844–854.

Edlund, L., Engelberg, J., & Parsons, C. A. (2009). The wages of sin. *Columbia University Department of Economics Discussion Paper*, *0809-16*, 1–27.

Egger, R., & Yu, J. (2022). A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in sociology*, *7*, 886498.

Evans, A. R., Elford, J., Bolding, G., & Wiggins, R. D. (2008). A comparison of item non-response in web and pen-and-paper surveys of sexual behaviour. *Methodological Innovations Online*, *2*, 6–17.

Eysenck, H. J. (1976). *Sex and personality*. Open Books.

Farkas, G., Sine, L., & Evans, L. (1978). Personality, sexuality, and demographic differences between volunteers and nonvolunteers for a laboratory study of male sexual behavior. *Archives of Sexual Behavior*, *7*, 513–520.

Fedorenko, O. (2015). Politics of sex appeal in advertising. *Feminist Media Studies*, *15*(3), 474–491. https://doi.org/10.1080/14680777.2014.930060

Freedman, E. B. (1982). Sexuality in nineteenth-century america: Behavior, ideology, and politics. *Reviews in American History*, *10*(4), 196–215. Retrieved February 9, 2024, from http://www.jstor.org/stable/2701827

Gill, J., & Zorn, C. (2007). The etiology of public support for the designated hitter rule. *Quarterly Journal of Political Science*, *2*(2), 189–203.

Grootendorst, M. (2022). Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Hatemi, P. K., Crabtree, C., & McDermott, R. (2017). The relationship between sexual preferences and political orientations: Do positions in the bedroom affect positions in the ballot box? *Personality and Individual Differences*, *105*, 318–325.

Helleputte, T., Gramme, P., & Paul, J. (2022). *Liblinear: Linear predictive models based on the liblinear c/c++ library* [R package version 2.10-22].

Hershey, M. R., & Sullivan, J. L. (1977). Sex-role attitudes, identities, and political ideology. *Sex Roles*, *3*(1), 37–57.

Hetherington, M., & Weiler, J. (2018). *Prius or pickup?: How the answers to four simple questions explain america's great divide*. Mariner Books.

Hoyt, C. L., Morgenroth, T., & Burnette, J. L. (2019). Understanding sexual prejudice: The role of political ideology and strategic essentialism. *Journal of Applied Social Psychology*, *49*, 3–14.

Hoyt, C. L., & Parry, M. (2018). Sociocultural and individual manifestations of sexual stigma: The role of political ideology and prejudice in discrimination against sexual minorities. *Journal of Social and Political Psychology*, *6*(1), 92–128.

Joe, V. C., Brown, C. R., & Jones, R. (1976). Conservatism as a determinant of sexual experiences. *Journal of Personality Assessment*, *40*(5), 516–521.

Joe, V. C., & Kostyla, S. (1975). Social attitudes and sexual behaviors of college students. *Journal of Consulting and Clinical Psychology*, *43*, 430.

Jost, J. T., Napier, J. L., Thorisdottir, H., Gosling, S. D., Palfai, T. P., & Ostafin, B. (2007). Are needs to manage uncertainty and threat associated with political conservatism or ideological extremity? *Personality and Social Psychology Bulletin*, *33*(7), 989–1007. https://doi.org/10.1177/0146167207301028

Kaats, G., & Davis, K. (1971). Effects of volunteer bias in studies of sexual behavior and attitudes. *The Journal of Sex Research*, *7*, 26–34.

King, B. M. (2022). The influence of social desirability on sexual behavior surveys: A review. *Archives of Sexual Behavior*, *51*, 1495–1501.

Koroknay-Palicz, T., & Montalvao, J. (2020). Sex, lies, and surveys: The role of interviewer characteristics. *Economics Bulletin*, *40*, 3313–3324.

Kuhn, M., & Wickham, H. (2020). *Tidymodels: A collection of packages for modeling and machine learning using tidyverse principles.* https://www.tidymodels.org

Lee, S., Fredriksen-Goldsen, K. I., McClain, C., Kim, H.-J., & Suzer-Gurtekin, Z. T. (2018). Are sexual minorities less likely to participate in surveys? an examination of proxy nonresponse measures and associated biases with sexual orientation in a population-based health survey. *Field Methods*, *30*(3), 208–224.

Liddon, N., Pampati, S., Steiner, R. J., Hensel, D. J., Fu, T.-c., Beckmeyer, J., & Herbenick, D. (2021). Truth be told: Adolescents' disclosure of sexual activity to healthcare providers. *Journal of Adolescent Health*, *68*(3), 623–625.

Lytle, A., Dyar, C., Levy, S. R., & London, B. (2017). Essentialist beliefs: Understanding contact with and attitudes towards lesbian and gay individuals. *British Journal of Social Psychology*, *56*(1), 64–88.

Marelich, W. D., Lundquist, J., Painter, K., & Mechanic, M. B. (2008). Sexual deception as a social-exchange process: Development of a behavior-based sexual deception scale. *The Journal of Sex Research*, *45*, 27–35.

Maticka-Tyndale, E., Lewis, J., & Street, M. (2005). Making a place for escort work: A case study. *The Journal of Sex Research*, *42*(1), 46–53.

Meneses-Falcon, C., Uroz-Olivares, J., & Rua-Vieites, A. (2017). Flyers y anuncios de servicios sexuales en madrid. *Revista Latina de Comunicacion Social*, (72), 145–164.

Munroe, R. (2012). Heatmap. https://xkcd.com/1138/

Parsons, J. T., Koken, J. A., & Bimbi, D. S. (2004). The use of the internet by gay and bisexual male escorts: Sex workers as sex educators. *AIDS Care*, *8*, 1021–1035.

Pennington, J., Socher, R., & Manning, C. (2014). GloVe: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543.

Petersen, M. B. (2018). Reproductive interests and dimensions of political ideology. *Evolution and Human Behavior*, *39*(2), 203–211.

Poteat, V. P., & Mereish, E. H. (2012). Ideology, prejudice, and attitudes toward sexual minority social policies and organizations. *Political Psychology*, *33*(2), 211–224.

Pruitt, M. V., & Krull, A. C. (2010). Escort advertisements and male patronage of prostitutes. *Deviant Behavior*, *32*(1), 38–63.

Rice, D. R., & Zorn, C. (2021). Corpus-based dictionaries for sentiment analysis of specialized vocabularies. *Political Science Research and Methods*, *9*(1), 20–35.

Roath, O. K., Chen, X., & Kolacz, J. (2023). Predictors of participation for sexuality items in a u.s. population-based online survey. *Archives of Sexual Behavior*, *52*(4), 1743–1752.

Ross, M. W., Månsson, S.-A., Daneback, K., Cooper, A., & Tikkanen, R. (2005). Biases in internet sexual health samples: Comparison of an internet sexuality survey and a national sexual health survey in sweden. *Social Science & Medicine*, *61*(1), 245–252.

Simpson, J., & Gangestad, S. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology*, *60*(6), 870–883.

Stern, C., West, T. V., Jost, J. T., & Rule, N. O. (2013). The politics of gaydar: Ideological differences in the use of gendered cues in categorizing sexual orientation. *Journal of Personality and Social Psychology*, *104*(3), 520–541.

Strassberg, D. S., & Lowe, K. (1995). Volunteer bias in sexuality research. *Archives of Sexual Behavior*, *24*(4), 369–382.

Thomas, D. R. (1975). Conservatism and premarital sexual experience. *British Journal of Social & Clinical Psychology*, *14*, 195–196.

U.S. Department of Justice. (2018). Justice department leads effort to seize backpage.com, the internet's leading forum for prostitution ads, and obtains 93-count federal indictment [April 9, 2018]. *Office of Public Affairs*. https://www.justice.gov/opa/pr/justice-department-leads-effort-seize-backpagecom-internet-s-leading-forum-prostitution-ads

Wiederman, M. W. (1993). Demographic and sexual characteristics of nonresponders to sexual experience items in a national survey. *The Journal of Sex Research*, *30*(1), 27–35.

Wilson, G. D. (1990). Ideology and humor preferences. *International Political Science Review*, *11*, 461–472.