

The Changing Meaning Of “No” In Canadian Sex Work

Lynn Kennedy^{1*}

¹Sex Work Population Project <https://populationproject.ca/>

*corresponding author

Email address: lynn.kennedy@populationproject.ca

Word count: 7465

Abstract

This study considers how Canadian contact sex workers communicate limits by examining how the word “no” was used by over one hundred thousand online advertisers over a 15 year period.

Source materials consisted of three collections of online advertising representing 214456 advertisers: 385729 ads collected between April 2007 and April 2009 (7939 advertisers), 2951642 ads collected between November 2014 and December 2016 (167539 advertisers), and 887698 ads collected between September 2021 and September 2022 (38978 advertisers).

Advertisers and demographic variables were identified from ad metadata. Common terms surrounding the word "no" were used to identify themes. The word “no” was used by 115127 advertisers. Five major themes were identified: *peace of mind* (54084 advertisers), *communication* (47130 advertisers), *client race* (32612 advertisers), *client behaviour* (23863 advertisers), and *service restrictions* (8545 advertisers). The likelihood of there being an association between an advertiser and a major theme was found to vary in response to several variables, including: time period, region, advertiser gender, and advertiser ethnicity.

Collectives, those advertising in multiple regions, and cis-female identified advertisers were more likely to use "no". Over time, the themes of *client behaviour*, *service restrictions*, and *client race* became more prominent.

Keywords: sex work, advertising, health and safety, information and communication technology

Introduction

This study explores how the word “no” is used in Canadian sex work advertising. The purpose of this research is to provide better information on what types of restrictions the majority

of sex workers actually use and to consider the demographic variables associated with those restrictions. It uses a combination of automatic document analysis and thematic analysis (Braun & Clarke, 2006) to identify who uses specific types of restrictions.

Most research documenting and health and safety strategies used by contact sex workers in industrialized democracies is based on questionnaire and interview data from relatively small or geographically constrained sample populations or small collections of websites or online ads (Argento et al., 2018; Bernier et al., 2021, 2021; Bungay & Guta, 2018; Cunningham & Kendall, 2010; Jiao et al., 2021; Kille et al., 2017; Kolar et al., 2014; Machat et al., 2022; Minichiello et al., 2013; Sanders et al., 2016, 2018; Strohmayer et al., 2019). There are many possible reasons for this but most likely the cost of research using questionnaires and interviews and possibly general lack of experience with text mining among criminologists and sociologists are reasons research has mainly focused on smaller scale exploratory studies.

Research at scale

There are exceptions to this rule. (Boecking et al., 2015, 2018) analysed two multi-million document datasets from one prominent North American classifieds site to identify temporal and geographic trends in advertising. (Kennedy, 2022) analysed 3.6 million ads from six prominent classified ad sites used by Canadian workers to estimate sex worker population dynamics. (Cunningham & Kendall, 2010) analyzed metadata from over 90000 worker profiles on one popular sex work review site in the US and followed this up with a survey of 685 workers to better understand the health and safety practices of workers over multiple years. Clearly, there are opportunities to create a more representative picture of the attitudes and communication practices of sex workers using publicly available online data. However, given the technical

challenges of research using document analysis on large document collections, few studies have used large samples of online data to better understand the health and safety practices of sex workers in industrialised democracies.

Limits and safety

As a safety strategy, advertising online significantly reduces risk (Argento et al., 2018; Machat et al., 2022; Sanders et al., 2018) and in industrialised democracies, web and telecommunications technologies have become the dominant way in which workers interact with clients. Workers describe investing considerable effort in getting to know clients before any in-person contact is initiated (Argento et al., 2018; Bungay & Guta, 2018; Jiao et al., 2021) and this can include more formal screening processes in some cases. Information sharing between workers is another important safety strategy (Strohmayr et al., 2019) and how workers share information has changed as the internet has become the dominant way that workers advertise (Argento et al., 2018; Bungay & Guta, 2018; Strohmayr et al., 2019). Communication in advertising plays an important part in this process as a critical first step in the worker-client relationship (Kille et al., 2017).

Methodological considerations

The motivation for this study emerged from the process of characterising systematic error in a sex worker population dynamics study (Kennedy, 2022). This required reading texts from thousands of ads where phrases starting with “no ...” would appear quite frequently. A review of older and newer datasets, included in this study, showed that this type of phraseology has been in use by sex work advertisers for quite a long time. An advantage of these common phrases is that

they tend to be short and many advertisers use similar phrases in a similar way making the analysis process easier. It is a relatively simple task to scan large numbers of documents for commonly used short phrases, tabulating statistics on how they are used. It is possible to then do thematic analysis (Braun & Clarke, 2006) on the most commonly used phrases, including context for clarity, to see whether sociologically useful information can be gathered from their patterns of use.

The advantages of this approach, compared to reviews of relatively small random samples of documents, are that the discovered themes are more likely to be relevant and complete. Furthermore, compared to convenience sampling often used in in-person qualitative sex work research, selection and social desirability biases are avoided as the source material is generated by advertisers in a naturalistic setting. Moreover, ad metadata can be used to provide context for who uses these phrases similar to the approach taken by (Cunningham & Kendall, 2010). In general this metadata can tell us when and where the advertiser is advertising, what gender and, sometimes, ethnicity they self-identify as. In addition, behavioural information such as how active they were on the advertising platform is readily available with a sufficiently complete dataset. Lastly, given that three collections spanning a 15 year period are available, it is possible to show the historic progression of how the use of the word “no” evolved, perhaps reflecting more general trends in the industry.

Research questions and objectives

The primary objective of this study is to catalogue how the word “no” was used in advertising. Is it always the case that this refers to a restriction? For uses which refer to restrictions, what types of prohibitions are imposed?

Secondary questions relate to differences in the advertisers who use “no” versus other advertisers: are there behavioural or demographic variables that differentiate these advertisers? Lastly, are there temporal, regional, gender or ethnic associations with the different ways this language is used?

Materials and Methods

This study provides evidence for common themes in sex work advertising surrounding how the word "no" is used. The likelihood of there being an association between an advertiser using “no” and a major theme is calculated for several variables, including: time period, region, advertiser gender, and advertiser ethnicity.

Extracting ad data

Source materials for the analysis were advertising texts from online classified ads. These ads were collected from six prominent online classified advertising sites described in (Kennedy, 2022) during three time periods: April 1, 2007 to March 31, 2009 inclusive, November 1, 2014 to December 31, 2016 inclusive, and September 15, 2021 and September 22, 2022 inclusive. The ad collections for 2014-2016 and 2021-2022 were nationally representative. The 2007-2009 collection is representative of one western Canadian province. Similar to (Cunningham & Kendall, 2010), this study uses the abstract entity of an *advertiser* (Kennedy, 2022) as a meaningful way to group ads. Advertisers were identified either from contact information in 2007-2009 and 2014-2016 or internal chat ids in 2021-2022 using methods described in (Kennedy, 2022).

Ad text was cleaned as described in (Kennedy, 2022) and ad texts were scanned for

relevant groups of terms. Pairs of words (bigrams), and triplets of words (trigrams) were extracted along with a count of the number of advertisers and ads using each. These were stored in a MariaDB database (MariaDB & Widenius, 2017) for further analysis. Supplemental materials S1 File contains a copy of the anonymized database.

Defining “no”

In this work, frequently used bigrams beginning with “no ” were combined with the most common words preceding and following them to generate a relatively small number of meta-documents that could be analysed for common themes. Context words were discovered based on common trigrams embedding each bigram. These files can be found in Supplemental materials S2 File. QualCoder (Curtain, 2023) was used to code underlying themes based on the generated files. For each identified code, advertiser frequencies for all related bigrams were tallied and the relative rank of each theme within each period was calculated. Figure 1 is an example of one of the QualCoder input files for the bigram “no low”. This bigram was included in the *client behavior* code described below as the embedded usage of the bigram refers to restrictions on negotiating prices. Note that in this case the words following the bigram are the most salient terms. Usually this was the case for most bigrams. In general the meanings of the bigrams were consistent based on context words.

Figure 1: Example coding file for the bigram “no low” for the 2021-2022 collection. “Before” and “after” words are terms that preceded and followed the bigram ordered in descending order by advertiser frequency.

```
pp2021-0005.txt
collection: pp2021, rank: 5
advertisers: 2658, proportion: 0.0561
before:
only, gents, calls, available, please, friendly, _NNNN_, free, services, service, independent,
pictures, you, wasters, men, pics, time, real, outcall, fee, showered, classy, and, me, bb, must,
call, greek, gentlemen, rush, accepted, negotiations, play, more, negotiable, gentleman,
blocked, av

term: no low

after:
ballers, baller, balling, ball, ballersno, price, balls, service, bawlers, rates, prices, iballers, rate,
bait, ballerno, ballin, budget, baller's, ballers, ballrs, ballersreal, ballersif, ballersi, ballersclick,
ballersand, services, ballersupon, offers, bowlers, ballerdrama, ballerb
```

Three groups of 100 files containing the most common bigrams and common context words were generated as inputs, one for each collection period. These most frequent "no" bigrams were used as the basis for thematic analysis.

Statistical measures

Advertisers using “no” were compared with other advertisers on two behavioural measures: days from the first ad to the last ad, and number of ads produced. Significance was tested using Welch’s modified two sample t-test, *tsum.test* from the R BSDA package (Arnholt & Evans, 2021). R version 4.2.2 was used to perform all tests using a 95% confidence level.

The proportion of advertisers associated with each coded theme was calculated and segmented by the following variables: time period, region, self-identified gender, and

self-identified ethnicity. These proportions represent the probability that an advertiser fitting a specific demographic category was associated with a given code (e.g. $p(\text{communication} = \text{True} \mid \text{gender} = \text{male})$). For each code, the differences between advertiser proportions for each of the demographic variables and the proportions for all advertisers were tested for significance using the *R prop.test* function (R Core Team, 2021).

Advertiser frequencies for other terms were tabulated to better understand some of the emergent themes. To better understand how advertiser communication around safety has changed, the most common health related bigrams starting with the word “safe” (“safe play”, “safe service”, “safe services”, and “safe gfe”) were searched and advertiser frequencies were calculated. More recently, advertisers have started screening clients and asking for deposits before meeting them. To provide historic context for this phenomenon “screening” and “deposit required” were searched as well.

To better understand attitudes surrounding the emergent theme of *client race* described below, term frequencies of common racial epithets from (Wikipedia contributors, 2023) were gathered. In addition, the conditional probabilities of the terms “black” or “brown” with “thug” or “pimp” were calculated. To see how many advertisers accepted clients of any race, the terms “all races”, “all nationalities”, “all backgrounds”, and “all ethnicities” were also searched. Lastly, counts were generated of advertisers with *client race* restrictions who also used the words “bad experience” or “sorry” to better understand the attitudes of advertisers associated with this code.

Ethics statement

All source data used in this study consisted of publicly available data at the time it was

collected and was collected in accordance with the policies of the sites in effect at the time. The methods used are conformant with the ethical standards of the Canadian Sociology Association (section 4.10 II) and the American Sociology Association (section 10.5 c) (American Sociological Association, 2018; CSA-SCS Policy, Ethics, and Professional Concerns Subcommittee, 2018). As the replicability of the main results of this paper is important, a data set is provided as part of the supporting information. However, to protect the safety and privacy of advertisers and third parties, all identifying information has been removed including the names of the source websites.

Results

Advertising behaviour

Differences in advertising behaviour were found between advertisers who used the word “no” and others. Advertisers who used the word “no” advertised more frequently (mean 30 ads, SD 264 versus mean 8 ads, SD 67) and were found to advertise longer (mean 118 days, SD 172 versus mean 75 days, SD 131). Welch’s modified two sample t-test indicated that the differences between the “no” group and other advertisers were significant for both duration ($t = -65.731$, $df = 211139$, $p < 0.001$) and ads posted ($t = -27.317$, $df = 131888$, $p < 0.001$).

Collected data

Table 1 summarises the advertising data in the three collections. A total of 4225069 ads were used (2007-2009: 385729 ads, 2014-2016: 2951642 ads, 2021-2022: 887698 ads). Of these, 39% (N=1628698) were ads that contained the word “no”. More than half of all advertisers used

“no” in at least one ad (54%, N=115127) and on average these advertisers used 2.3 distinct “no” bigrams (SD 2.1). Figure 2 shows three word clouds illustrating the relative term frequencies of the words following "no" for each period.

Table 1: Ads and number of advertisers identified by time period.

period	region	sites	ads		advertisers	
			total	using “no”	total	using “no”
2007-2009	British Columbia	1	385729	160458 (42%)	7939	4146 (52%)
2014-2016	Canada	6	2951642	1051167 (36%)	167539	91106 (54%)
2021-2022	Canada	1	887698	417073 (47%)	38978	19875 (51%)
All		6	4225069	1628698 (39%)	214456	115127 (54%)

Figure 2: Top 100 terms following the word “no” by time period based on term frequency.

Larger words represent more frequently used terms.



Themes and codes

Table 2 shows the identified themes, the number of associated advertisers, and the most commonly associated bigrams for each theme. A total of 168 unique "no" bigrams were

represented in the top 100 bigrams from every period. A total of 14 codes were identified from the analysis of these bigrams and context words. The vast majority of advertisers using “no” statements are associated with both the top 100 bigrams for each period (92%, N=105838) and the top 5 codes (88%, N=100688).

Table 2: Coded themes relating to the usage of the word “no” based on the top 100 bigrams for each period. Advertiser counts represent all periods. Ambiguous bigrams are followed by the most common terms following them in parentheses.

theme codes	advertisers	common bigrams
<i>peace of mind</i>	54084 (25%)	“no games”, “no disappointments”
<i>communication</i>	47130 (22%)	“no blocked” (calls), “no emails”, “no text”
<i>client race</i>	32612 (15%)	“no black” (gents), “no african”, “no aa”
<i>client behaviour</i>	23863 (11%)	“no lowballers”, “no negotiations”
<i>service restrictions</i>	8545 (4%)	“no bareback”, “no anal”
<i>service location</i>	3532 (2%)	“no carcalls”, “no outcall”
<i>services offered</i>	2786 (1%)	“no limits”, “no restrictions”
<i>no pictures</i>	2713 (1%)	“no picture”, “no face”, “no free” (pictures)
<i>pimps or law enforcement</i>	2679 (1%)	“no law” (enforcement), “no thugs”, “no pimps”
<i>service time</i>	1426 (1%)	“no hh”, “no half” (hour)
<i>employment</i>	988 (<1%)	“no experience”
<i>client age</i>	917 (<1%)	“no young” (men)
<i>appearance</i>	294 (<1%)	“no tattoos”
<i>payment</i>	169 (<1%)	“no e” (transfer)

The top 5 themes were coded as *communication*, *peace of mind*, *client behaviour*, *client*

race, and *service restrictions*. Less common themes used by 2% or fewer advertisers were coded as *services offered*, *no pictures*, *pimps or law enforcement*, *service time*, *employment*, *client age*, *appearance*, and *payment*.

The *communication* code related to restrictions on how the advertiser wished to be contacted including requiring caller id for phone calls. *Peace of mind* refers to “no” statements that were intended to reassure prospective clients. *Client behaviour* refers to restrictions related to etiquette such as not negotiating prices, wasting the advertiser’s time, or using explicit language. *Client race* refers to advertisers restricting prospective clients based on racial background. *Service restrictions* refers to statements regarding disallowed types of services - the inverse of *services offered*.

Other service related themes were *payment*, *service location*, and *service time* which described restrictions on methods of payment, venues, and service duration respectively. In addition to *client race*, *client age*, and *pimps or law enforcement* are restrictions on who can contact the advertiser and could be said to overlap with *client behaviour*. The *no pictures* and *appearance* themes relate to restrictions on presentation in the ad and presentation in person respectively. Finally, *employment* is unique to advertisers looking for prospective employees rather than clients.

Variables

The following sections describe how the variables of time period, region, advertiser gender, and advertiser ethnicity are associated with the 5 most prevalent codes. In all cases the *R prop.test* function showed that the overall proportions of subgroups of advertisers using “no” were significantly different from the proportions of all advertisers using “no” ($p < 0.001$). Except

where noted, the proportions of advertiser subgroups associated with individual codes were also significantly different from the proportion of all advertisers associated with that code ($p < 0.001$).

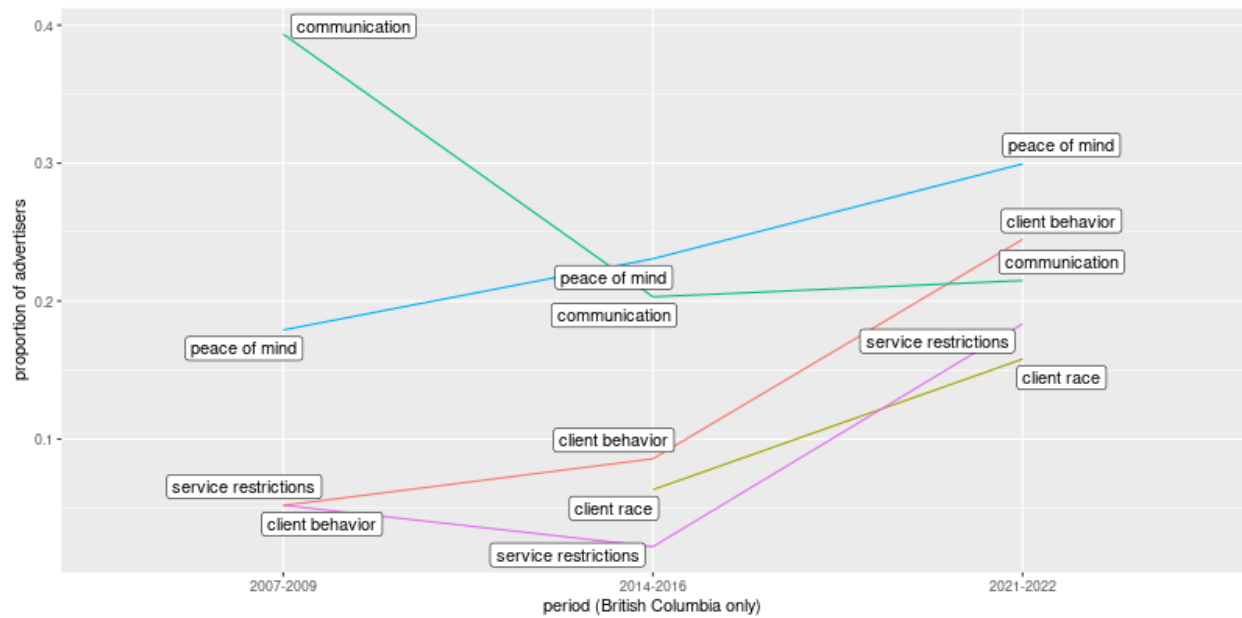
Time period

Data in 2007-2009 was only available for British Columbia (BC) therefore a comparison by time period was made for advertisers who had advertised in BC. Table 3 shows the proportions of advertisers using “no” by time period for these advertisers.

Table 3: Advertisers by time period for British Columbia. The p and CI values are the results of *prop.test* comparisons between the proportion of advertisers using “no” from BC versus the proportion of advertisers using “no” from all regions.

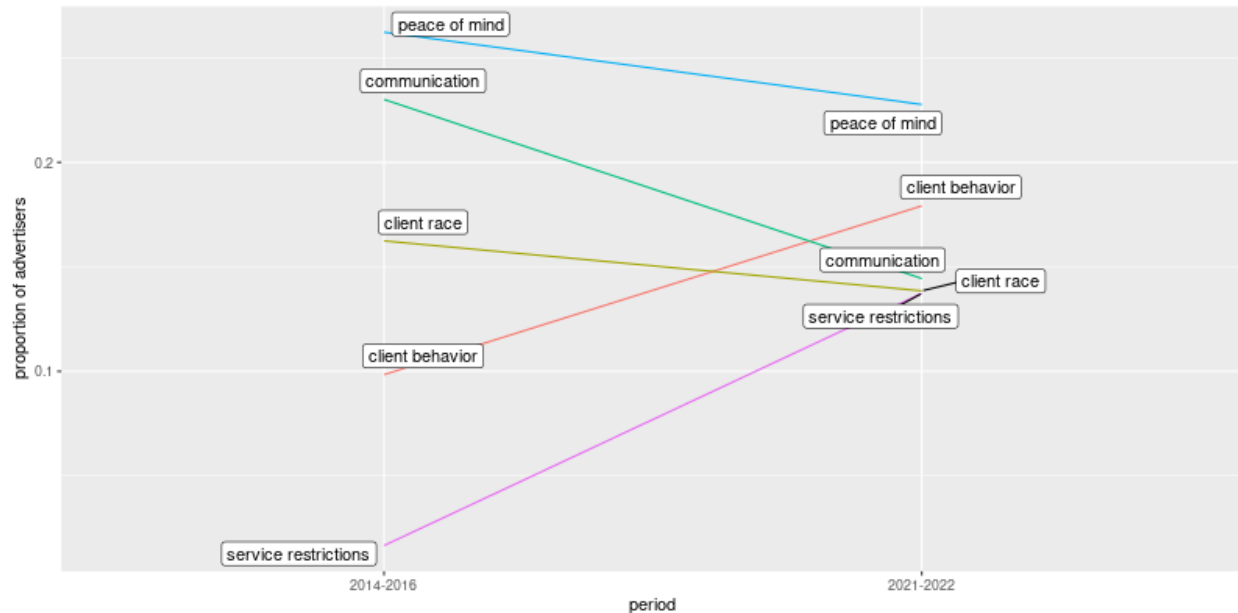
period	total	using “no”	p	CI
2007-2009	7939	4146 (52%)	–	–
2014-2016	24420	10759 (44%)	<0.001	[-0.11,-0.10]
2021-2022	7840	5175 (66%)	<0.001	[0.14,0.16]
all	40199	20080 (50%)	<0.001	[-0.04,-0.03]

The relative ranks of the top 5 codes changed significantly between periods for BC advertisers. Figure 3 shows the relative ranks of the top 5 codes by time period. Most themes became more prominent between 2007-2009 and 2021-2022. Three restriction-related themes had the largest increases during this period: *service restrictions* (13% increase), *client race* (16% increase), and *client behaviour* (19% increase).

Figure 3: Proportions of advertisers using the top 5 codes by time period in British Columbia.

How do advertisers from all regions in Canada compare? Figure 4 shows the relative ranks of the top 5 codes for periods where data was available from all regions. The relative rank of *client race* in 2014-2016 was higher for Canada (third) than it was for BC (fourth). However, *client race* became less prominent in 2021-2022 for Canada (12%, N=5397) compared to BC (16%, N=1239).

Figure 4: Proportions of advertisers using the top 5 codes by time period for periods including all regions.



Advertisers in the later collections were increasingly likely to advertise in multiple provinces. Between 2014-2016 and 2021-2022 the number of advertisers advertising in multiple provinces doubled both for all advertisers (7%, N=11929 in 2014-2016 versus 15%, N=5713 in 2021-2022) and for advertisers advertising in BC (14%, N=3309 in 2014-2016 versus 36%, N=2779 in 2021-2022).

Safety-related messaging generally became more common between 2007 and 2022. In addition to safety related "no" terms such as those associated with the *service restrictions* code, the usage of the word "safe" itself provided further evidence for this trend. Over all periods 6% (N=16127) of advertisers used at least one of the tested "safe" bigrams ("safe service", "safe services", "safe gfe" and "safe play"). In 2007-2009 these bigrams were not common with only 1.3% (N=105) of advertisers using them. However, in 2014-2016 this increased to 7% of

advertisers (N=12166) and 10% (N=3856) in 2021-2022. Most advertisers using these "safe" bigrams also used "no" (84%, N=13522 all periods).

Safety messaging was not limited to health. The terms “deposit required” and “screening” were less commonly used but represent important risk reduction strategies. These became more prevalent between 2007 and 2022. In 2007-2009 only 2 advertisers (<1%) used the bigram “deposit required” however by 2021-2022 usage increased to 1725 advertisers (4%). Similarly, the term “screening” in 2007-2009 was used by 8 advertisers (<1%) but in 2021-2022 this grew to 522 (1.3%).

Region

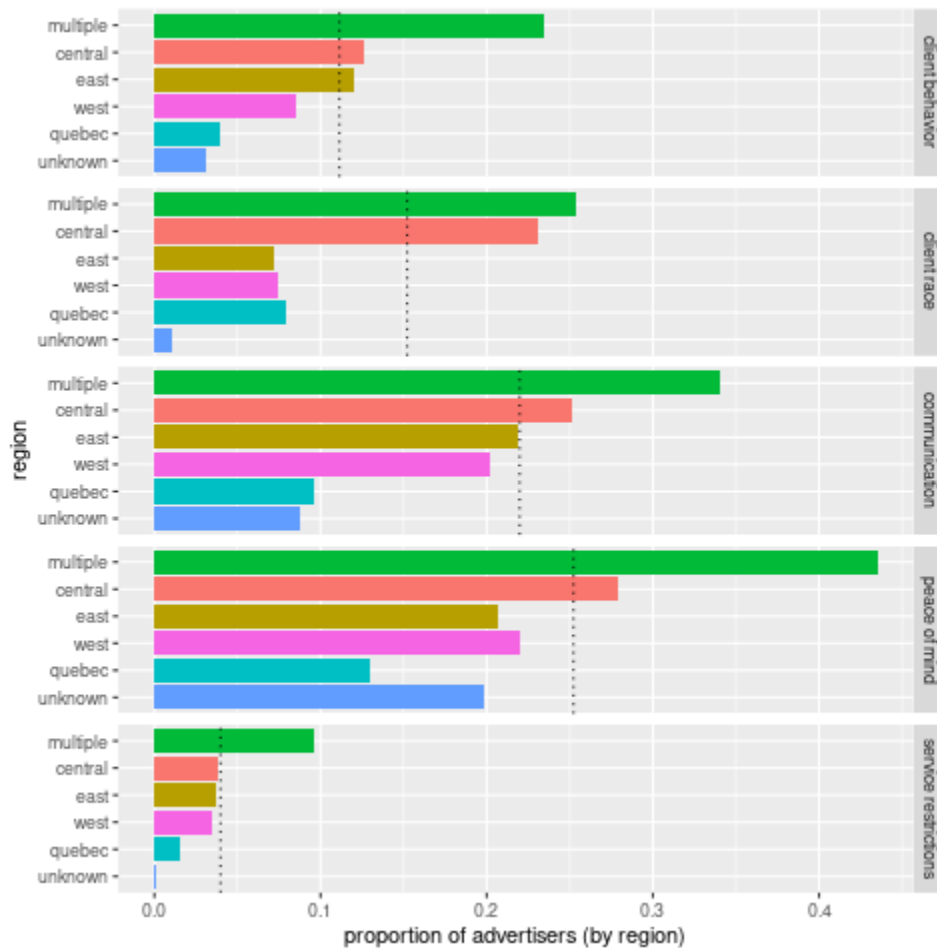
Table 4 shows the distribution of advertisers by region. Advertisers associated with multiple provinces, such as those who tour, used the word “no” more often than advertisers who only advertised in a single province (75%, N=17641 versus 51%, N=97201). Apart from those advertising in multiple provinces, the central region (61%, N=48491) and advertisers with no geographic information (56%, N=1734) were more likely than average to use “no”.

Table 4: Advertisers by region.

region	provincial codes	advertisers	using “no”
<i>multiple</i>	multi-province and the north (nw,nv,yk)	23676	17641 (75%)
<i>central</i>	on	79621	48491 (61%)
<i>unknown</i>	no geographic information	1734	970 (56%)
<i>west</i>	ab,bc,mb,sk	73369	36061 (49%)
<i>east</i>	nb,nf,nl,ns,pei	6927	3319 (48%)
<i>quebec</i>	qc	29131	8360 (29%)

Regional differences in code usage are illustrated in Figure 5. All proportions of advertisers were significantly different from the overall proportion for each code except for *service restrictions* in the *central* region ($p = 0.09$) and the *east* region where *service restrictions* ($p = 0.35$), *communication* ($p = 0.81$), and *client behaviour* ($p = 0.015$) were not significantly different from the average. Those advertising in *multiple* regions were significantly more likely to be associated with one of the *client behaviour*, *client race*, and *service restrictions* codes compared to other advertisers (41%, $N=9707$ versus 24%, $N=51567$, *prop.test* CI [0.25, 0.26], $p < 0.001$). Advertisers in the *central* region were also more likely to be associated with these codes (32%, $N=25092$, *prop.test* CI [0.16, 0.17], $p < 0.001$).

Figure 5: Proportions of advertisers by region associated with the top 5 codes. The dotted lines indicate the proportion of all advertisers associated with each coded theme.



For the advertisers associated with *multiple* regions, the most common provinces were Ontario (67%, N=11074) and Alberta (52%, N=8652) associated with 91% of advertisers.

Advertiser gender

Table 5 shows the number of advertisers associated with self-identified gender. Advertisers who advertised with multiple genders were the most likely to use “no” (62%, N=6289) followed by cis-female identified advertisers (57%, N=100865). Advertisers using “no” associated with multiple genders represented 1.8% of advertisers overall. These advertisers

were mostly associated with cis-female and cis-male identity (62%, N=2425) or cis-female and female-identified transgender identity (25%, N=984).

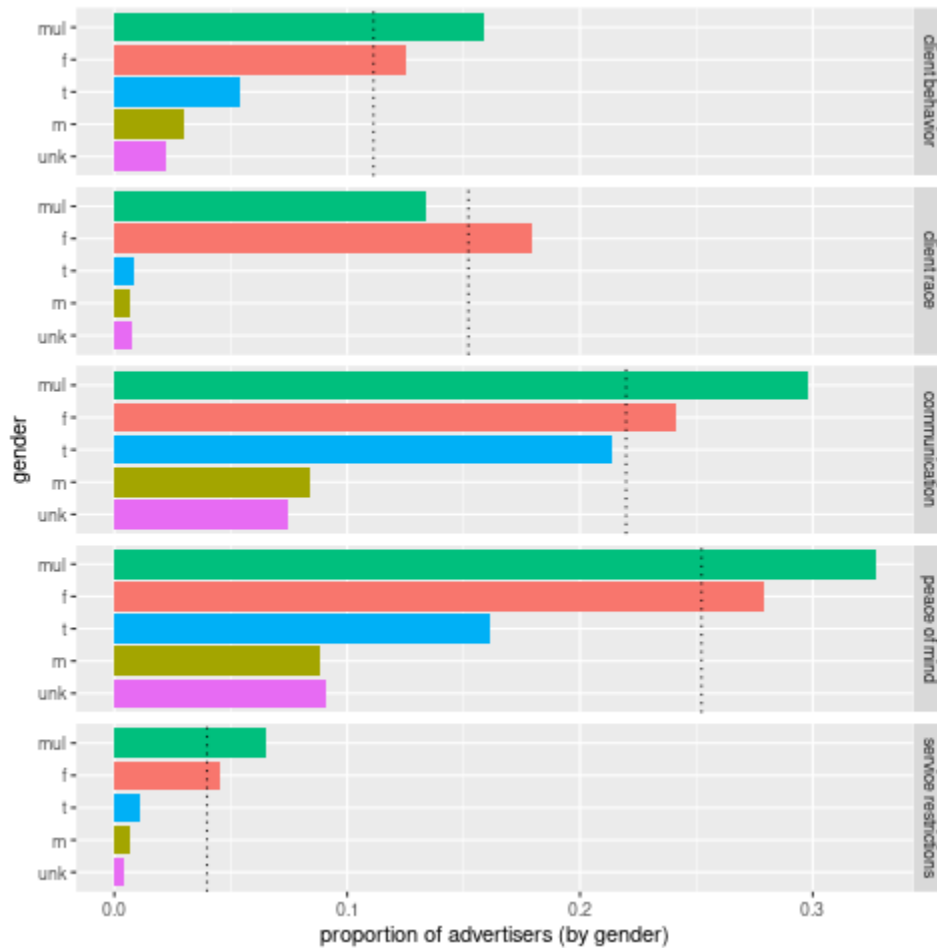
Table 5: Advertisers by gender. Some advertisers are associated with multiple genders.

Transgender refers to female-identified transgender.

gender	advertisers	using “no”
multiple (<i>mul</i>)	6289	3895 (62%)
female (<i>f</i>)	175513	100865 (57%)
transgender (<i>t</i>)	3392	1511 (45%)
unknown (<i>unk</i>)	13359	4253 (32%)
male (<i>m</i>)	15905	4318 (27%)

Figure 6 shows the relative proportions of advertisers associated with the top 5 codes. All proportions of advertisers were significantly different from the overall proportion for each code except for female-identified transgender advertisers (*t*) associated with the *communication* code ($p = 0.38$). Exclusively cis-female identified advertisers were significantly more likely than other advertisers to be associated with one of the *client behaviour*, *client race*, or *service restrictions* codes (28%, N=48601 versus 8%, N=2989, *prop.test* CI [0.15, 0.16], $p < 0.001$). Exclusively cis-male advertisers were the least likely to be associated with *client behaviour*, *client race*, and *service restrictions* (4%, N=671).

Figure 6: Proportions of advertisers by gender using the top 5 codes. Gender codes: *f* cis-female, *m* cis-male, *t* female-identified transgender, *mul* more than one gender was associated with the advertiser and *unk* means that no gender was associated with the advertiser. The dotted lines indicate the proportion of all advertisers associated with that coded theme.



Advertiser ethnicity

The source site for the 2021-2022 dataset, had an "ethnicity" field which could be optionally filled with one of *Latino/Hispanic*, *Asian*, *Mixed*, *Caucasian/White*, *Black*, *Indo Canadian*, *Middle Eastern*, *Native*, or *Canadian Born Chinese*. These ethnic categories were grouped as shown in Table 6 as *Multiple* (advertisers associated with more than one ethnicity),

Mixed (advertisers self-identifying as multi-ethnic), *Hispanic*, *PoC* (People of Color), *White*, *Asian*, *Native*, and *Unknown*. Advertisers associated with more than one ethnic identity were the most likely to use “no” (75%, N=4766) followed by *Mixed* (60%, N=1912) and *Hispanic* (57%, N=1029). Advertisers who provided no ethnic identity information were much less likely to use “no” in ads (9%, N=609). Consistent data on self-identified ethnicity was only available for the 2021-2022 period.

Table 6: Advertisers by self-identified ethnicity in 2021-2022.

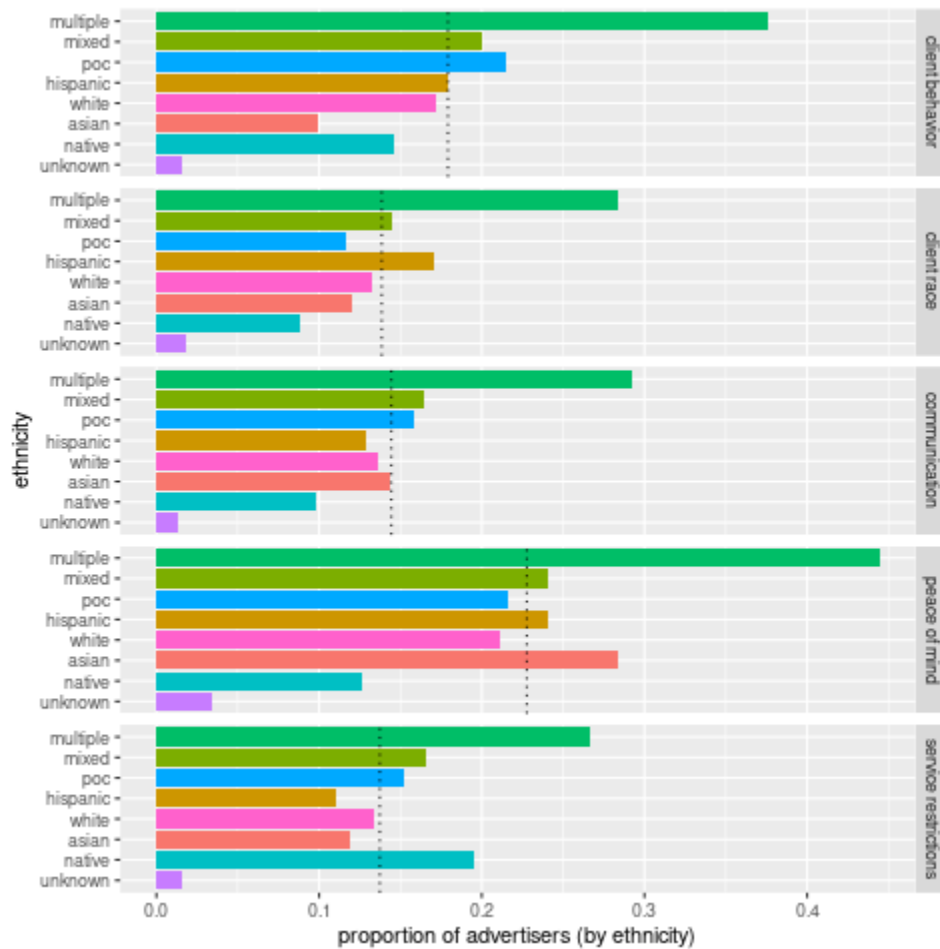
ethnicities	group	advertisers	using “no”
Advertiser uses multiple ethnicities	<i>Multiple</i>	6380	4766 (75%)
“Mixed” ancestry	<i>Mixed</i>	3165	1912 (60%)
Hispanic	<i>Hispanic</i>	1811	1029 (57%)
Black,Indo Canadian,Middle Eastern	<i>PoC</i>	2668	1509 (57%)
Non-Hispanic White	<i>White</i>	15551	8560 (55%)
Asian,Asian Canadian	<i>Asian</i>	2559	1384 (54%)
First Nations	<i>Native</i>	205	107 (52%)
No ethnicity provided	<i>Unknown</i>	6641	609 (9%)

In 2021-2022, 6380 advertisers (13%) were associated with more than one ethnic identifier. The Supplemental materials S4 File “ethnicity” tab shows how frequently these identifiers were used. The identifiers Caucasian/White (69%, N=3273) and Mixed (62%, N=2979) are associated with 95% (N=4463) of these advertisers.

Figure 7 shows the proportions of advertisers associated with each of the top 5 codes in 2021-2022. All proportions of advertisers were significantly different from the overall proportion

for each code. Advertisers using *Multiple* ethnic identifiers were significantly more likely to be associated with one of the *client behaviour*, *client race*, and *service restrictions* codes compared to other advertisers (93%, N=5905 versus 36%, N=11831, *prop.test* CI [0.55, 0.57], $p < 0.001$).

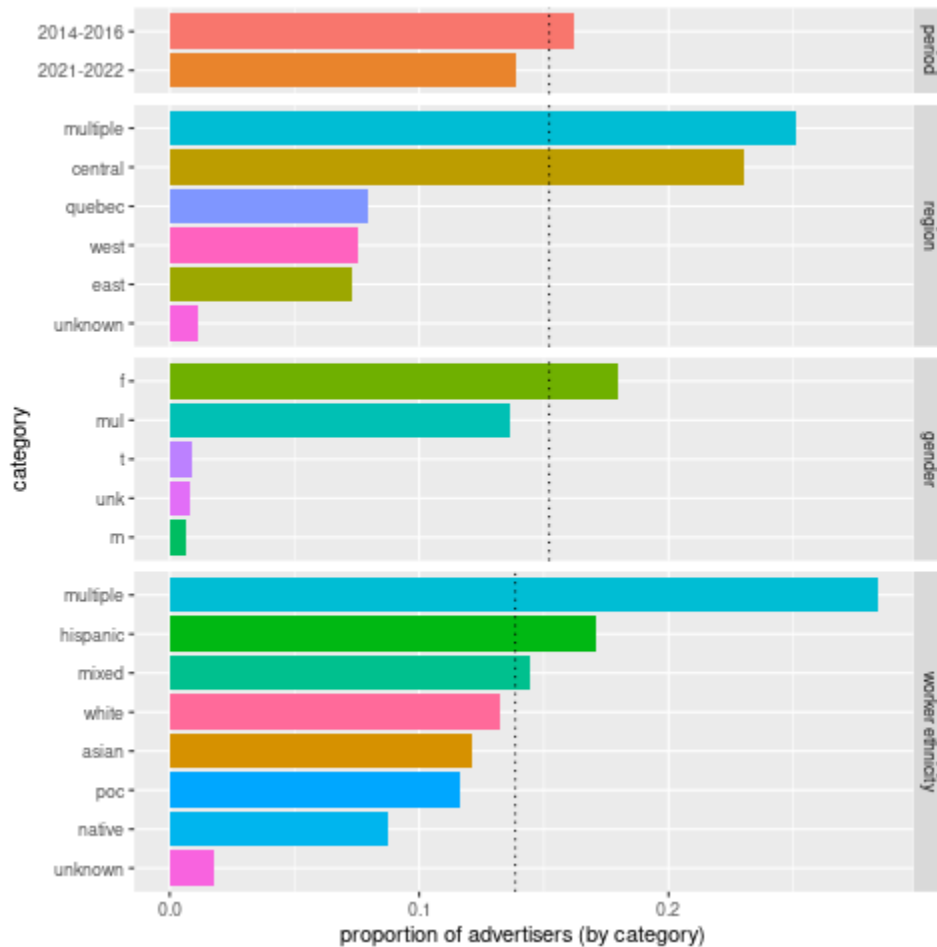
Figure 7: 2021-2022 advertisers associated with the top 5 codes by ethnicity. *Mixed* in this context is a single ethnic option but *Multiple* refers to advertisers who indicated more than one ethnicity in their ads. The dotted lines indicate the proportion of all 2021-2022 advertisers associated with that coded theme.



Client race

An emerging theme was that many advertisers restricted clients based on racial background. This is a disturbing finding therefore it warrants a more comprehensive review. In the later collections, 32610 advertisers (overall: 15%, 2014-2016: 16%, N=27215, 2021-2022: 12%, N=5397) were associated with the *client race* restriction. Figure 8 summarises the demographic characteristics of the advertisers associated with *client race*. Cis-female identified advertisers (18%, N=31540, *prop.test* CI [0.03, 0.03], $p < 0.001$), advertisers who advertised in multiple provinces (25%, N=5992, *prop.test* CI [0.10, 0.11], $p < 0.001$), or advertisers who only advertised in Ontario (23%, N=18353, *prop.test* CI [0.08, 0.08], $p < 0.001$) were significantly more likely to be associated with the *client race* code than all advertisers associated with this restriction (15%, N=32610). Those exclusively identifying as cis-men were the least likely to be associated with the *client race* code (<1%, N=101, *prop.test* CI [-0.15, -0.14], $p < 0.001$).

Figure 8: Proportions of advertisers by category who restrict clients based on *client race*. The dotted lines are the proportion of all advertisers with this type of restriction. Note that 2021-2022 was the only period with reliable data for advertiser ethnicity.



Ethnic self-identification affected the probability of being associated with the *client race* code. The majority of advertisers with ethnicity data (60%, N=23429) did not identify as *White*. *White* advertisers were slightly but significantly less likely to be associated with *client race* restrictions (*White* 13%, N=2062 versus not *White* 14%, N=3333, *prop.test* CI [-0.017, -0.003], $p = 0.007$). Advertisers associated with the *Multiple* ethnicities group were the most likely to be associated with client race restrictions (28%, N=1811) followed by the *Hispanic* (17%, N=309), *Mixed* (14%, N=457), *White* (13%, N=2062), *Asian* (12%, N=309), *PoC* (12%, N=310), *Native*

(9%, N=18) and *Unknown* (2%, N=119). The *PoC* subgroups were associated with *client race* with the following frequencies: Middle Eastern 14% (N=60), Black 11% (N=209), and Indo Canadian 10% (N=41).

This study's focus on restrictions may have biased the analysis. Are there advertisers who explicitly accept clients from all backgrounds? A search for the terms "all races", "all nationalities", "all backgrounds", and "all ethnicities" found 2139 advertisers (1% of all advertisers) indicating that this was the case for at least some advertisers. Furthermore, the number of advertisers using these terms increased by a factor of 7 from 2007-2009 to 2021-2022 (2007-2009: 0.3%, N=27, 2014-2016: 0.7%, N=1238, 2021-2022: 2.2%, N=874).

Few advertisers discuss why they used *client race* restrictions, but those who did spoke of bad experiences with previous racialized clients (e.g. "no black sorry multiple bad experiences"). However, these advertisers were a minority with only 63 using the term "bad experience" in ads with *client race* restrictions.

Is there evidence for stereotyping language in advertising data? All of the most commonly used bigrams relating to client race appear to refer to people of African ancestry ("no black", "no african", "no aa", "no blacks", "no blackgents", and "no blk"). To see if offensive language appeared in advertising, a list of related slurs, gathered from (Wikipedia contributors, 2023), was used as a basis for searching ad text. The Supplemental materials S3 File contains the regular expression (Perl 5 Porters, 2021) used to search for the terms. Occurrences were rare with only 2.4 documents in 1000 matching any of the terms. At least some of the time use of racial epithets appeared to be directed at other workers rather than prospective clients.

Discussion and conclusions

Over the 15 year period represented by the collections in this study, roughly similar proportions of advertisers and posts used the word “no”. However, how advertisers used “no” in the earliest collection is substantially different from how “no” was used in later collections. In the 2007-2009 collection, *communication* preferences were the most important consideration followed by *peace of mind*. However, in later collections the word “no” was more likely to refer to restrictions: *client behaviour*, *service restrictions*, or *client race*. Cis-female identified advertisers, advertisers who travel, and advertisers who were likely collectives (based on use of multiple genders, or use of multiple ethnic identities) tended to use the word "no" more often than others. It is notable that, over time, messages specifically relating to health such those associated with the *service restrictions* code and “safe” services became much more prominent.

Sex workers should always have the right to refuse service (Abel et al., 2007; International Planned Parenthood Federation, 2022) and even in a criminalised environment, as is the case in Canada (Protection of Communities and Exploited Persons Act, 2014), sex workers routinely exercise agency (Benoit & Millar, 2001; Bungay et al., 2011; Ham & Gerard, 2014; Jeffrey & MacDonald, 2006; Jiao et al., 2021). Nevertheless, it was disturbing to discover that at least 32610 advertisers used *client race* restrictions. Demographic research alone cannot answer why this was the case but it can provide clues for further investigation. This type of restriction appears to be mainly associated with cis-female advertisers; very few trans-female or cis-male advertisers are associated with *client race* restrictions.

Why has this not been found in previous research? Small sample sizes and non-random sampling strategies may be excluding relevant sources (Kennedy, 2022). However, the large body of research that uses interviews could also have been affected by social desirability bias where interviewees may not want to admit to behaviour that might be considered racist. It is also

possible that this is a more recent phenomenon that may not have been observed in older research.

Are these advertisers racist? Notably, the majority of advertisers in this study did not self-identify as White and advertisers who self-identified as *PoC* were associated with the *client race* code about as often as other advertisers. Outright derogatory terms were very infrequent.

Are racialized clients more sexist? Few advertisers spoke of why they employed *client race* restrictions. Those who did, mentioned bad experiences with previous racialized clients. This contrasts with recent opinion polls that show that gender equality is important to the majority of African Americans (Cox, 2023). Research on clients shows that in general clients are not more sexist than other people (Milrod & Monto, 2012; Monto & Milrod, 2014; Sanders, 2020). Having some sense of how race was associated with other types of sexual misconduct might have been helpful but there is a general lack of information on perpetrator race in sex-based harassment studies (Cotter & Savage, 2019; Kresnow et al., 2022; Saguy & Rees, 2021). The advertising evidence suggests that when advertisers make this type of restriction they do not do so lightly and client behaviour is the main motivation.

Black racialized clients appear to represent a minority in online spaces in North America (Milrod & Monto, 2012; Monto & Milrod, 2014; Sanders, 2020). More generally, racialized people who are identified as Black are often at a disadvantage in online dating spaces outside of sex work (Bany et al., 2014; Bedi, 2015; Feliciano et al., 2009; Robnett & Feliciano, 2011; Wade et al., 2022; Wade & Harper, 2021; Wade & Pear, 2022). How do other forms of discrimination, such as those experienced in dating, as well as the minority status of racialized clients online affect interactions with sex workers?

Advertiser attitudes changed with time. Not only did the overall number of advertisers

with *client race* restrictions drop significantly from 16% in 2014-2016 to 12% in 2021-2022 but the number of advertisers who explicitly accepted clients from any background increased substantially. The overall Canadian trend of advertisers becoming more tolerant of racialized clients in 2021-2022 contrasts with the results of a 2021 survey of 3698 Canadians on the topic of race (The Environics Institute for Survey Research, 2021). On multiple measures it was found that an increasing number of respondents believed discrimination was getting worse in 2021 compared to 2019.

Other questions were raised by the demographic analyses. It appears that the number of advertisers who travel is increasing and that these advertisers may have a heightened sense of risk. To what degree are touring workers more at risk than sedentary workers? A striking feature of the usage of the word “no” was the extent to which advertisers attempted to reassure clients (“no drama”, “no rush service”). Although this can be interpreted as a safety strategy (Bungay & Guta, 2018), this begs the question: how do prospective clients perceive risk in online spaces? Lastly, what health and safety strategies are the most effective in practice?

Limitations

Over time, advertising venues changed and advertising online has become more technically challenging especially for those advertisers paying with cryptocurrency. These structural changes could have affected who is represented in online classified ads.

The advertiser proportion measure used in this study identifies the relative prominence of each theme for the authors of the ads which may not be the attitudes of the represented workers. This research should be followed up by further mixed-methods qualitative research which grounds its samples in advertiser data to directly compare worker attitudes with how they are

portrayed in advertising.

Advertisers could represent one or more workers and in some cases, advertisers using different contact details or multiple chat names might be counted more than once. Advertisers identified by contact information, the case in 2007-2009 and 2014-2016, can change contacts over long periods, and small proportions, typically around 5%, are not sex workers (Kennedy, 2022). The fact that advertisers who use "no" advertised longer than those who did not mitigates this problem somewhat as these advertisers are less likely to change contact information and may be more likely to be relevant. Similarly, advertisers identified by online profiles, the case in 2021-2022, can have multiple profiles. The 2021-2022 dataset was extensively screened for non-sex work related advertisers and this appears to have also significantly reduced the detected number of duplicate profiles. However, even for this group, ~5% of profiles could also be duplicates.

While many advertisers in Quebec used “no”, the way similar types of statements are represented in French language ads is different from how these are represented in English resulting in undercounting of Francophone advertisers. More work is needed to understand the effect of linguistic identity and occupational health and safety communication for these advertisers. Similarly, even for Anglophone advertisers, “no” statements could be expressed in other ways. The proportions shown here should be considered a lower bound.

Conclusions

Some of the findings of this study are likely to be surprising. However, all of the findings in this study reflect the attitudes of thousands of advertisers identified in a naturalistic setting. Both (Cunningham & Kendall, 2010) and this study show that it is possible with a little creativity

to identify sociologically useful trends from extremely large samples of online data. We need these large scale perspectives to put other research in context.

Online sex work advertising is a balancing act between presenting an appealing image and setting limits. These limits are a reflection of sex worker agency and are an important source of information on how stigma manifests itself in the industry. Post #MeToo, it is clear that Canadian sex workers have become far less tolerant of bad behaviour online and far more assertive about their rights and sexual health even in the face of significant structural challenges.

Acknowledgements

Sadie Hudson from ANSWERS (<http://www.answersociety.org/>) and Robin C. A. Barrett provided editorial input on various versions of this manuscript.

Declaration of interest statement

The authors have declared that no competing interests exist.

Supplemental Materials

Supplemental materials including all source data can be found at <https://osf.io/hwzsn/>

S1 File. MariaDB database. <https://osf.io/2rjkm>

S2 File. QualCoder project with coded documents. <https://osf.io/k85pf>

S3 File. Racial epithet regular expression. <https://osf.io/5vbe2>

S4 File. Tables. <https://osf.io/cpsvm>

Bibliography

Abel, G., Fitzgerald, L., & Brunton, C. (2007). *The Impact of the Prostitution Reform Act on the Health and Safety Practices of Sex Workers* (p. 213). Department of Public Health and General Practice University of Otago, Christchurch.

<https://www.otago.ac.nz/christchurch/otago018607.pdf>

American Sociological Association. (2018). *ASA code of ethics*. American Sociological Association. https://www.asanet.org/sites/default/files/asa_code_of_ethics-june2018.pdf

Argento, E., Taylor, M., Jollimore, J., Taylor, C., Jennex, J., Krusi, A., & Shannon, K. (2018). The Loss of Boystown and Transition to Online Sex Work: Strategies and Barriers to Increase Safety Among Men Sex Workers and Clients of Men. *American Journal of Men's Health*, 12(6), 1994–2005. <https://doi.org/10.1177/1557988316655785>

Bany, J. A., Robnett, B., & Feliciano, C. (2014). Gendered Black Exclusion: The Persistence of Racial Stereotypes Among Daters. *Race and Social Problems*, 6(3), 201–213.

Bedi, S. (2015). Sexual Racism: Intimacy as a Matter of Justice. *The Journal of Politics*, 77(4), 998–1011. <https://doi.org/10.1086/682749>

Benoit, C., & Millar, A. (2001). *Dispelling Myths and Understanding Realities: Working Conditions, Health Status, and Exiting Experiences of Sex Workers* (p. 133).

Bernier, T., Shah, A., Ross, L. E., Logie, C. H., & Seto, E. (2021). The use of information and communication technologies by sex workers to manage occupational health and safety: Scoping review. *Journal of Medical Internet Research*, 23(6), e26085–e26085.

Boecking, B., Miller, K., Barnes, M., Boecking, B., & Kennedy, E. (2015). Leveraging publicly available data to discern patterns of human-trafficking activity. *Journal of Human Trafficking*, 1(1), 65–85.

- Boecking, B., Miller, K., Kennedy, E., & Dubrawski, A. (2018). Quantifying the relationship between large public events and escort advertising behavior. *Journal of Human Trafficking*, 5(3), 220–237. <https://doi.org/10.1080/23322705.2018.1458488>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Bungay, V., & Guta, A. (2018). Strategies and challenges in preventing violence against canadian indoor sex workers. *American Journal of Public Health (1971)*, 108(3), 393–398.
- Bungay, V., Halpin, M., Atchison, C., & Johnston, C. (2011). Structure and agency: Reflections from an exploratory study of Vancouver indoor sex workers. *Culture, Health & Sexuality*, 13(1), 15–29.
- Cotter, A., & Savage, L. (2019). *Gender-based violence and unwanted sexual behaviour in Canada, 2018: Initial findings from the Survey of Safety in Public and Private Spaces* (No. 85-002-X; Juristat). Statistics Canada.
<https://www150.statcan.gc.ca/n1/pub/85-002-x/2019001/article/00017-eng.htm>
- Cox, K. (2023). *Black Americans Firmly Support Gender Equality but Are Split on Transgender and Nonbinary Issues*. Pew Research Center.
<https://www.pewresearch.org/race-ethnicity/2023/02/16/black-americans-firmly-support-gender-equality-but-are-split-on-transgender-and-nonbinary-issues/>
- CSA-SCS Policy, Ethics, and Professional Concerns Subcommittee. (2018). *Statement of Professional Ethics*. Canadian Sociological Association-Societe canadienne de sociologie.
https://www.csa-scs.ca/files/www/STATEMENT_OF_PROFESSIONAL_ETHICS.pdf
- Cunningham, S., & Kendall, T. D. (2010). Risk behaviours among internet-facilitated sex

- workers: Evidence from two new datasets. *Sexually Transmitted Infections*, 86(Suppl 3), iii100–iii105. <https://doi.org/10.1136/sti.2010.044875>
- Curtain, C. (2023). *QualCoder* (3.2) [Python]. <https://github.com/ccbogel/QualCoder>
- Feliciano, C., Robnett, B., & Komaie, G. (2009). Gendered racial exclusion among white internet daters. *Social Science Research*, 38(1), 39–54.
- Protection of Communities and Exploited Persons Act, (2014). https://laws.justice.gc.ca/eng/AnnualStatutes/2014_25/FullText.html
- Ham, J., & Gerard, A. (2014). Strategic in/visibility: Does agency make sex workers invisible? *Criminology & Criminal Justice*, 14(3), 298–313.
- International Planned Parenthood Federation. (2022). *IPPF Policy on Sex Work*. International Planned Parenthood Federation. <https://www.ippf.org/resource/ippf-policy-sex-work>
- Jeffrey, L. A., & MacDonald, G. (2006). “It’s the Money, Honey”: The Economy of Sex Work in the Maritimes*. *The Canadian Review of Sociology and Anthropology*, 43(3), Article 3.
- Jiao, S., Bungay, V., & Jenkins, E. (2021). Information and communication technologies in commercial sex work: A double-edged sword for occupational health and safety. *Social Sciences (Basel)*, 10(1), 1–13.
- Kennedy, L. (2022). The silent majority: The typical Canadian sex worker may not be who we think. *PloS One*, 17(11), e0277550–e0277550.
- Kille, J., Bungay, V., Oliffe, J., & Atchison, C. (2017). A Content Analysis of Health and Safety Communications Among Internet-Based Sex Work Advertisements: Important Information for Public Health. *Journal of Medical Internet Research*, 19(4), e111. <https://doi.org/10.2196/jmir.6746>
- Kolar, K., Atchison, C., & Bungay, V. (2014). Sexual safety practices of massage parlor-based

sex workers and their clients. *AIDS Care*, 26(9), 1100–1104.

<https://doi.org/10.1080/09540121.2014.894611>

Kresnow, M., Smith, S. G., Basile, K., & Chen, J. (2022). *The National Intimate Partner and Sexual Violence Survey: 2016/2017 Methodology Report* (p. 28). Centers for Disease Control and Prevention National Center for Injury Prevention and Control.

<https://www.cdc.gov/violenceprevention/pdf/nisvs/nisvsMethodologyReport.pdf>

Machat, S., Lyons, T., Braschel, M., Shannon, K., & Goldenberg, S. (2022). Internet solicitation linked to enhanced occupational health and safety outcomes among sex workers in Metro Vancouver, Canada 2010–2019. *Occupational and Environmental Medicine (London, England)*, 79(6), 373–379.

MariaDB, & Widenius, M. (2017). *MariaDB* (10.3.31).

<https://mariadb.com/kb/en/about-mariadb-software/>

Milrod, C., & Monto, M. A. (2012). The Hobbyist and the Girlfriend Experience: Behaviors and Preferences of Male Customers of Internet Sexual Service Providers. *Deviant Behavior*, 33(10), 792–810.

Minichiello, V., Scott, J., & Callander, D. (2013). New Pleasures and Old Dangers: Reinventing Male Sex Work. *The Journal of Sex Research*, 50(3–4), 263–275.

<https://doi.org/10.1080/00224499.2012.760189>

Monto, M. A., & Milrod, C. (2014). Ordinary or Peculiar Men? Comparing the Customers of Prostitutes With a Nationally Representative Sample of Men. *International Journal of Offender Therapy and Comparative Criminology*, 58(7), Article 7.

<https://doi.org/10.1177/0306624X13480487>

Perl 5 Porters. (2021). *perltre—Perl regular expressions*. <https://perldoc.perl.org/perltre>

- Robnett, B., & Feliciano, C. (2011). Patterns of Racial-Ethnic Exclusion by Internet Daters. *Social Forces*, 89(3), 807–828.
- Saguy, A. C., & Rees, M. E. (2021). Gender, Power, and Harassment: Sociology in the #MeToo Era. *Annual Review of Sociology*, 47(1), 417–435.
<https://doi.org/10.1146/annurev-soc-090320-031147>
- Sanders, T. (2020). *Paying for sex in a digital age: US and UK perspectives / Teela Sanders, Barbara G. Brents and Chris Wakefield*. Routledge.
- Sanders, T., Connelly, L., & King, L. J. (2016). On Our Own Terms: The Working Conditions of Internet-Based Sex Workers in the UK. *Sociological Research Online*, 21(4), 133–146.
- Sanders, T., Scoular, J., Campbell, R., Pitcher, J., & Cunningham, S. (2018). Characteristics and Working Practices of Online Sex Workers. In *Internet Sex Work: Beyond the Gaze* (pp. 55–85). Springer International Publishing. https://doi.org/10.1007/978-3-319-65630-4_3
- Strohmayer, A., Clamen, J., & Laing, M. (2019). Technologies for Social Justice: Lessons from Sex Workers on the Front Lines. *Conference on Human Factors in Computing Systems - Proceedings*, 1–14.
- The Environics Institute for Survey Research. (2021). *Race relations in Canada 2021*. Canadian Race Relations Foundation.
https://www.crrf-fcrr.ca/images/Environics_Study_2021/Race_Relations_in_Canada_2021_Survey_-_FINAL_REPORT_ENG.pdf
- Wade, R. M., Bouris, A. M., Neilands, T. B., & Harper, G. W. (2022). Racialized Sexual Discrimination Who Seek Intimate Partners Online. *Sexuality Research & Social Policy*, 19(3), 1341.
- Wade, R. M., & Harper, G. W. (2021). Racialized Sexual Discrimination (RSD) in Online Sexual

Networking: Moving from Discourse to Measurement. *The Journal of Sex Research*, 58(6), 795–807. <https://doi.org/10.1080/00224499.2020.1808945>

Wade, R. M., & Pear, M. M. (2022). A Good App Is Hard to Find: Examining Differences in Racialized Sexual Discrimination across Online Intimate Partner-Seeking Venues. *International Journal of Environmental Research and Public Health*, 19(14), 8727.

Wikipedia contributors. (2023). *List of ethnic slurs and epithets by ethnicity*—Wikipedia, The Free Encyclopedia.

https://en.wikipedia.org/w/index.php?title=List_of_ethnic_slurs_and_epithets_by_ethnicity&oldid=1136627460