Understanding Non-Response Patterns to Sexual Orientation and Gender Identity Questions: Insights from a Brazilian Survey

Mariel Gouvea Gruppi[1]

[1]Center for Development and Regional Planning (Cedeplar), Federal University of Minas Gerais, Belo Horizonte, Brazil.

Abstract

Collecting complete information has always been a concern in sample surveys, as well as refusal rates for items typically found in household surveys - i.e., income, education - have always been under the attention of experts. In parallel, few actions were done regarding the inclusion of questions about sexual orientation and gender identity in surveys. The implementation of these items has been slow, especially for developing countries such as Brazil, and little has yet been discussed in the literature about the quality of the data collected. Therefore, this work is dedicated to analyzing the non-response rate to the question of sexual orientation and gender identity in a sample survey carried out in one of the Brazilian regions. Hence, this work aims to characterize the pattern of non-response to questions about sexual orientation and gender identity based on sociodemographic variables, as well as to find the odds ratios for non-response considering covariates that may be important for predicting non-response. The results have shown the relevance of some of the covariates in the response refusal, such as race and marital status. Thus, it will be possible to contribute to the national and international literature regarding the implementation of questions that identify sexual and gender minorities, in addition to a better understanding for the implementation of measures that favor adherence and response to sexual orientation and gender identity.

Introduction

Collecting information on sexual orientation is a relatively recent development in several countries, particularly in Brazil. It was only in 2022 that data from the 2019 National Health Survey (NHS) became available, marking the first instance of a national survey inquiring about the sexual orientation of Brazilians as part of its diverse questionnaire. It represents the inaugural nationally representative study to collect self-declared data on sexual orientation from Brazilian people.

Regarding gender identity in Brazil, only the 2021 PDAD[1] has thus far collected information on transgender and non-binary individuals in its sample, alongside including the categorization of sexual orientation. The latter survey was a household survey conducted in one of Brazil's regions, the Federal District, home to the country's capital, Brasília.

The importance of identifying sexual and gender minorities[2] in household surveys is justified by various findings that indicate poorer mental health conditions for gay and lesbian individuals (Carpenter, Eppink, Gonzales, & McKay, 2021), disadvantages in the labor market for lesbian, gay and bisexual individuals (Drydakis, 2022; Suliano, Filho, & Irffi, 2021), as well as a higher risk of suicide and experiencing violence among sexual and gender minorities (LeVasseur, Kelvin, & Grosskopf, 2013; Pinto et al., 2020). Additionally, a Gallup[3] study on the United States suggests that 20% of Generation Z, those born between 1997 and 2003, and only 2.6% of Baby boomers, those born between 1946 and 1964, identify as a sexual or gender minority, which suggests that this is a population with a large growth in the coming years. However, there is much discussion about including a question that aims to collect sensitive information in a questionnaire, because it could affect the whole response rate. As it is the case of sexual orientation and gender identity questions, due to its private and stigmatized nature, not everyone would be willing to disclose their identity to someone else, which could affect the survey response rate, thereby impacting the quality of the data (Young & Rater, 2021).

---

[1] *Pesquisa Distrital por Amostra de Domicílios* is a representative survey for the Federal District region in Brazil.
[2] Sexual minorities or gender minorities are those who do not identify as heterosexual or cisgender.
[3] Available: < https://news.gallup.com/poll/389792/lgbt-identification-ticks-up.aspx >

Objectives and Focus

Considering the aforementioned, this study seeks to contribute to the literature on non-response to questions regarding sexual orientation and gender identity in sample surveys. The research is a novelty in the field, since was not found a study that has analyzed non-response in SOGI[4] questions in Brazilian surveys, particularly concerning both sexual orientation and gender identity within the same survey. Furthermore, it is important to highlight that PDAD represents only one of the regions of Brazil, however, it has a significant potential to contribute with empirical results for sexual and gender minorities in the country.

As objectives, we first intend to examine the differences in the non-response profile taking into account sociodemographic characteristics. Secondly, it aims to explore the set of characteristics in a binomial logistic model to generate Odds Ratios, thereby obtaining the likelihood ratio of refusal for the items related to sexual orientation and gender identity. In this way, this study is dedicated to filling gaps in national and international literature, providing a more in-depth understanding of the factors that allow predicting non-response to questions of sexual orientation and gender identity in sample surveys.

Data and Methods

For this study, I used data from the PDAD, which is a household sample survey for the Federal District region in Brazil. The only version of this survey available to date with information on sexual orientation and gender identity is from 2021. This survey is a household survey in which one person in the household is responsible for responding about all other residents. In this case, to identify individuals' sexual orientation, an open-ended question was used[5], allowing respondents to self-identify, and identify the others, as they understood the question. For gender identity, a two-step method was employed[6], first inquiring about the sex assigned at birth and then asking which gender the respondent identified with at the time of the survey. In both cases, the question could be freely answered, the respondent could indicate that they do not know, or they could refuse to answer the question for themselves or other household members.

For the purposes of this study, the variables of interest for non-response come from those individuals who received "Do not know" or "Refusal" as options for sexual orientation or gender identity. The control variables used include age, race, education, and marital status. The analyses were conducted using only observations with complete cases for the control variables. Ultimately, the final sample, with values expanded by the sampling weight, consisted of 16,476 (0.67%) refusals and 2,411,815 responses for sexual orientation, and 4,032 (0.16%) refusals and 2,424,259 responses for gender identity.

For analytical purposes, this study presents a description of the response and non-response profiles to questions about sexual orientation and gender identity, based on the selected covariates. Additionally, a binomial model is developed to find the Odds Ratios for refusal to answer questions about sexual orientation and gender identity, aiming to examine the effects of the covariates on predicting non-response.

Preliminary Findings

For patterns in the level of nonresponse for sexual orientation, the PDAD survey has a lower bar than other surveys. Kim and Fredriksen-Goldsen (2013) found a 1.86% prevalence of non-response to sexual orientation in a health survey for the United States. In another study, Fredkisen-Goldsen and Kim (2015) found a non-response rate of 1.93% for a sample survey in Washington (US). Using data from the NHIS (2013-2014), Lee et al. (2018) identified a 3.0% refusal rate for the sexual orientation question. Regarding gender identity, Zhang, Smith-Johnson, and Tumin (2023) found a 1.1% non-response rate for gender identity in the BRFSS survey (Washington, US).

Regarding the characteristics defining the non-response profile of the analyzed sample, Table 1 presents the distribution of non-response and response to the item by age group, race, education, and marital status. Regarding age,

[4] SOGI – sexual orientation and gender identity

[5] “What is the sexual orientation of (name of the resident)?”

[6] "What is the assigned sex at birth of (name of the resident)?" and then, "With which gender does (name of the resident) currently identify?"

the 30-49 age group accounts for 41% of refusals to respond to sexual orientation, while the same group comprises 54% of refusals to respond to gender identity. Regarding race, individuals identifying as Black or Brown (mixed-race) account for most refusals. The educational level of high school and higher education exhibits the highest concentration of refusals, and unmarried individuals represent the marital status with the highest refusal rates for both sexual orientation and gender identity.

Table 1. Prevalence Estimates of Characteristics by Responses to Sexual Orientation and Gender Identity Question in PDAD, 2021

	Sexual orientation		Gender Identity	
	Answered % [CI 95%]	Refused % [CI 95%]	Answered % [CI 95%]	Refused % [CI 95%]
Age (mean years)	42.1 [42-42.1]	40.5 [37.9-43.1]	42.1 [42-42.1]	39.8 [36.5 -43.1]
Age group				
18-29	0.26 [0.25 -0.26]	0.32 [0.27 -0.38]	0.26 [0.25 -0.25]	0.27 [0.18 -0.36]
30-49	0.44 [0.44 -0.44]	0.41 [0.35 -0.46]	0.44 [0.44 -0.44]	0.54 [0.43 -0.64]
50-64	0.20 [0.20 -0.2]	0.17 [0.12 -0.22]	0.20 [0.20 -0.20]	0.12 [0.05 -0.20]
65+	0.10 [0.10 -0.1]	0.10 [0.06 -0.15]	0.10 [0.10 -0.10]	0.07 [0.02 -0.12]
Race				
White	0.41 [0.40 -0.41]	0.36 [0.30 -0.42]	0.41 [0.40 -0.41]	0.40 [0.29 -0.51]
Black	0.12 [0.12 -0.13]	0.17 [0.12 -0.21]	0.12 [0.12 -0.13]	0.16 [0.08 -0.25]
Asian	0.01 [0.01 -0.02]	0.04 [0.02 -0.06]	0.01 [0.01 -0.02]	0.04 [0.00 -0.08]
Brown	0.45 [0.45 -0.45]	0.43 [0.36 -0.49]	0.45 [0.44 -0.46]	0.39 [0.27 -0.50]
Education				
< Elementary	0.15 [0.15 -0.16]	0.13 [0.08 -0.18]	0.15 [0.15 -0.16]	0.13 [0.04 -0.22]
Elementary	0.12 [0.12 -0.13]	0.11 [0.07 -0.15]	0.12 [0.12 -0.13]	0.09 [0.02 -0.17]
High School	0.40 [0.39 -0.43]	0.40 [0.34 -0.46]	0.40 [0.39 -0.40]	0.31 [0.21 -0.41]
Some college	0.32 [0.32 -0.33]	0.36 [0.3 -0.42]	0.32 [0.32 -0.33]	0.47 [0.35 -0.58]
Marital status				
Never married	0.32 [0.32 -0.33]	0.56 [0.49 -0.63]	0.33 [0.32 -0.33]	0.63 [0.52 - 0.75]
Cohabiting	0.06 [0.06 -0.07]	0.04 [0.01 -0.07]	0.06 [0.06 -0.07]	0.03 [-0.01 - 0.06]
Married	0.50 [0.49 -0.51]	0.29 [0.22 -0.35]	0.50 [0.50 -0.51]	0.22 [0.13 - 0.32]
Divorced	0.01 [0.01 -0.01]	0.01 [0.00 -0.02]	0.01 [0.01 -0.01]	0.01 [-0.01 -0.03]
Widowed	0.05 [0.05 -0.06]	0.04 [0.02 -0.06]	0.05 [0.05 -0.05]	0.06 [0.01 -0.10]
	0.04 [0.04 -0.05]	0.06 [0.02 -0.09]	0.04 [0.04 -0.05]	0.05 [-0.01 -0.10]

Note. % weighted values; CI = confidence interval

Regarding the Odds Ratios, Table 2 below presents the results along with the corresponding confidence intervals and p-values. Although not statistically significant, the analysis of age groups may suggest that the older the age group, the higher the likelihood of non-response regarding sexual orientation. However, for gender identity, the likelihood of a person refusing to answer the gender identity question decreases with age, as indicated by the ORs decreasing the value from the second to the last age group.

Concerning race, some cases were statistically significant, such as for Black individuals (OR = 1.68, p-value = 0.01) and Asians (OR = 3.69, p-value = 0.00), indicating that compared to white individuals, Black and Asian individuals have a higher likelihood of refusing to answer the sexual orientation question. For gender identity, only Asians were significant (OR = 6.12, p-value = 0.01), but to a much greater magnitude. This finding, suggest a need for a reassessment of the race covariate, given the intricate nature of racial issues in Brazil between Black/Brown and White people, and the small sample of Yellow/Asian people.

None of the education covariates were statistically significant, however, the ORs increases the higher the educational attainment, with individuals with some college education having a 1.51 likelihood to refuse the sexual orientation question compared to those with less than elementary education. This result is important since it contrasts with the findings of Kim and Fredriksen-Goldsen (2013) as they pointed to a stronger association between lower levels of education and non-response rate.

The results of the Odds Ratios for marital status indicated an association between refusal and marital status. Taking the group of never-married singles as the reference, all other classifications had an OR below 1.00 - for sexual orientation or gender identity -, meaning that all other groups were less likely to be non-respondents than never-

married individuals. The married status stands out as the lowest and most significant (OR = 0.28, p-value = 0.00). Only widowed individuals were not statistically significant for both models, and divorced individuals were not significant for the gender identity model.

Table 2. Odds Ratios of Response Refusal to Sexual Orientation and Gender Identity in PDAD, 2021

	Sexual orientation			Gender Identity		
	O.R.	CI 95%	p-value	O.R.	CI 95%	p-value
Age group						
18-29 (ref)	1.00			1.00		
30-49	1.01	[0.72 - 1.41]	0.97	1.78	[0.98 - 3.25]	0.06
50-64	1.17	[0.76 - 1.81]	0.47	1.22	[0.55 - 2.72]	0.63
65+	1.10	[0.61 - 1.98]	0.76	1.03	[0.37 - 2.9]	0.95
Color/race						
White (ref)	1.00			1.00		
Black	1.68	[1.12 - 2.52]	0.01**	1.82	[0.88 - 3.76]	0.11
Asian	3.69	[1.71 - 7.98]	0.00***	6.12	[1.67 - 22.41]	0.01**
Brown	1.21	[0.85 - 1.72]	0.30	1.10	[0.63 - 1.9]	0.74
Education						
< Elementary (ref)	1.00			1.00		
Elementary	1.07	[0.58 - 1.98]	0.84	0.90	[0.30 - 2.75]	0.85
High School	1.15	[0.71 - 1.86]	0.57	0.74	[0.32 - 1.71]	0.48
Some college	1.51	[0.92 - 2.48]	0.10	1.75	[0.74 - 4.13]	0.20
Marital status						
Single(ref)	1.00			1.00		
Cohabiting	0.40	[0.19 - 0.83]	0.01**	0.19	[0.05 - 0.67]	0.01**
Married	0.28	[0.20 - 0.40]	0.00***	0.22	[0.12 - 0.39]	0.00***
Divorced	0.38	[0.20 - 0.73]	0.01**	0.52	[0.22 - 1.21]	0.13
Widowed	0.67	[0.34 - 1.29]	0.23	0.70	[0.18 - 2.73]	0.60

Note. *** p<0.001, ** p<0.01, * p<0.05; Ref. = reference group; CI = confidence interval; regression estimated considering weighted sample

Discussion and Future Steps

Firstly, the analytical strategy proposed by this work to analyze non-response to the sexual orientation and gender identity item presented fundamental results that were successful with the initially established objectives. The differences in non-response prevalence by age, race, education, and marital status help in understanding the profile of individuals more likely to refuse to answer questions that are crucial for identifying intersectional vulnerabilities. This could contribute to public policies aimed at reaching more people in data collection through sample surveys, mainly the sexual and gender minorities.

The findings in this work present some agreement and disagreement with what is already published in the literature. However, the implementation of SOGI questions has been slowly emerging in surveys, there is still a lot of work to be done in the future to have a better understanding of the importance of sexual orientation and gender identity questions in sample surveys. To this end, this work proposes to complement the results found here with a comparison with the refusal of other survey items, such as income, in addition to a comparison with the refusal of the sexual orientation item in the 2019 PNS Furthermore, we will move on to a more complete analysis of the family composition and non-response rates, to assess the impact of family composition on non-response. Ultimately, this research aims to contribute to the literature on non-response to questions about sexual orientation and gender identity in Latin American countries.

References

Carpenter, C. S., Eppink, S. T., Gonzales, G., & McKay, T. (2021). Effects of access to legal same-sex marriage on marriage and health. Journal of Policy Analysis and Management, 40(2), 376-411.

Drydakis, N. (2022). Sexual orientation and earnings: a meta-analysis 2012–2020. Journal of Population Economics, 35(2), 409-440.

Fredriksen-Goldsen, K. I., & Kim, H. J. (2015). Count me in: Response to sexual orientation measures among older adults. Research on Aging, 37(5), 464-480.

Kim, H. J., & Fredriksen-Goldsen, K. I. (2013). Nonresponse to a question on self-identified sexual orientation in a public health survey and its relationship to race and ethnicity. American journal of public health, 103(1), 67-69.

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.

LeVasseur, M. T., Kelvin, E. A., & Grosskopf, N. A. (2013). Intersecting identities and the association between bullying and suicide attempt among New York city youths: results from the 2009 New York city youth risk behavior survey. American journal of public health, 103(6), 1082-1089.

Pinto, I. V., Andrade, S. S. D. A., Rodrigues, L. L., Santos, M. A. S., Marinho, M. M. A., Benício, L. A., ... & Canavese, D. (2020). Perfil das notificações de violências em lésbicas, gays, bissexuais, travestis e transexuais registradas no Sistema de Informação de Agravos de Notificação, Brasil, 2015 a 2017. Revista Brasileira de Epidemiologia, 23(suppl 1), e200006-SUPL.

Suliano, D., Jesus Filho, J. D., & Irffi, G. (2021). Sexual orientation and wage differentials using anthropometric and health measures. Estudos Econômicos (São Paulo), 51, 111-142.

Young, L. J., & Rater, B. R. (2021). The Farm Producer Survey: Unit and Item Nonresponse (No. 1496-2022-1759).

Zhang, Z., Smith-Johnson, M., & Tumin, D. (2023). Contextual Influences on Nonresponse to Health Survey Questions About Sexual Orientation and Gender Identity. LGBT health.