IDB WORKING PAPER SERIES Nº IDB-WP-01439

How Accurately are Household Surveys Measuring the Size and Inequalities for the LGBT Population in Bogota, Colombia?

Evidence from a List Experiment

Andres Ham
Angela Guarin
Juanita Ruiz

Inter-American Development Bank
Gender and Diversity Division
Department of Research and Chief Economist

January 2023

IDB
Inter-American
Development Bank

How Accurately are Household Surveys Measuring the Size and Inequalities for the LGBT Population in Bogota, Colombia?

Evidence from a List Experiment

Andres Ham
Angela Guarin
Juanita Ruiz

January 2023

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library
Ham, Andrés.
How accurately are household surveys measuring the size and inequalities for the LGBT population in Bogotá, Colombia?: evidence from a list experiment / Andrés Ham, Ángela Guarín, Juanita Ruiz.
p. cm. — (IDB Working Paper Series ; 1439)
Includes bibliographic references.
1. Sexual minorities-Colombia. 2. Household surveys-Colombia. 3. Sex discrimination-Colombia. I. Guarín, Ángela. II. Ruiz, Juanita. III. Inter-American Development Bank. Gender and Diversity Division. IV. Inter-American Development Bank. Department of Research and Chief Economist. V. Title. VI. Series.
IDB-WP-1439

JEL codes: C90, D10, J10, J21, J70.
Keywords: population, measurement, discrimination, household surveys, list experiment.

http://www.iadb.org

How accurately are household surveys measuring the size and inequalities for the LGBT population in Bogotá, Colombia?

Evidence from a list experiment[*]

Andrés Ham[†]
Ángela Guarín
Juanita Ruiz

School of Government
Universidad de los Andes

[*] This research was supported by the Gender and Diversity Knowledge Initiative (GDLab) of the Inter-American Development Bank (IDB). We are grateful for comments from Monserrat Bustelo, Veronica Frisancho, Carlos Scartascini, Karen Martínez, Samuel Berlinski, Emilio Gutiérrez, Adrián Rubli, and Christian Posso, as well as participants of the Second Seminar of Research Network Project: LGBTQ+ Persons in Latin America and the Caribbean held by the IDB in July 2022 and seminar participants at Universidad de los Andes and Universidad Javeriana. This project was reviewed and approved in advance by the Institutional Review Board for the protection of human subjects at Universidad de los Andes (IRB #1542-2022). Fieldwork for the list experiment was carried out by the *Centro Nacional de Consultoría*. All remaining errors and omissions are the authors' sole responsibility. The views expressed in this paper do not necessarily reflect the views of the Inter-American Development Bank or Universidad de los Andes.

[†] Corresponding author. Contact information for all authors: Ham, a.ham@uniandes.edu.co; Guarín, a.guarin@uniandes.edu.co; Ruiz, j.ruiz@uniandes.edu.co. Physical Address: Carrera 1 #19-27, Edificio Aulas, Piso 3; Bogotá, 111711, Colombia. Tel: +57 601 339-4949.

Abstract

This paper studies whether household surveys precisely identify the LGBT population and are suitable to measure labor market discrimination in Colombia. We first quantify the size of the LGBT population and estimate labor market inequalities from these data, highlighting potential pitfalls from using this approach. We then present findings from a list experiment in Bogotá, Colombia. Results show that household surveys underestimate the size of the LGBT population and may yield biased estimates of labor market inequalities. While survey estimates range between 1-4%, we find that LGBT people constitutes around 12-22% of the total population. We find heterogeneous reporting by sex, age groups, educational attainment, and marital status. Our findings suggest that while current measurement practices are a step forward for the LGBT population's statistical visibility, additional steps are required before household surveys may be used to consistently estimate discrimination and guide policy responses to protect this population.

Keywords: LGBT population, measurement, discrimination, household surveys, list experiment.
JEL Classification: C90, D10, J10, J21, J70.

1. Introduction

Identifying and reducing discrimination against vulnerable groups requires statistical visibility. Many statistics institutes have implemented inclusive practices to identify vulnerable populations in nationwide surveys to better estimate unfair inequality and discrimination, prioritizing individuals with disabilities, belonging to ethnic groups, and sexual self-identification (DANE, 2020b). However, most surveys directly ask respondents about these sensitive topics, which has been shown by previous research to underestimate the true size of these vulnerable groups (Coffman et al., 2017). Additionally, the effectiveness of statistical methods to estimate discrimination not only requires that groups are adequately identified but also that there is no other measurement error or bias that confounds these estimates. Therefore, data that precisely identifies who belongs to these groups, their attributes, and outcomes are essential to guide public policy that reduces discrimination in all domains of well-being to improve people's lives (OECD, 2019).

This paper studies whether household surveys are precisely identifying the LGBT population in Colombia and are therefore suitable to measure their wellbeing and experiences of discrimination in different areas, including the labor market. The National Administrative Statistics Institute (DANE, for its acronym in Spanish) has implemented an inclusive approach to better identify vulnerable groups in its nationally representative household surveys, which includes LGBT people (DANE, 2020b). We quantify the size of the LGBT population and estimate labor market inequalities from household survey data, highlighting the potential pitfalls from using this empirical approach. We then present findings from a list experiment with 2,025 respondents in Bogotá, Colombia to determine whether household surveys are accurately measuring the true size of the LGBT population. Half the sample is randomly assigned to answer questions identical to

current household surveys (*direct response group*), while the other half answers questions about sexual self-identification indirectly (*veiled response group*).

Household survey estimates indicate that about 1.57% of the Colombian population self-identifies as LGBT, with differences across locations. The LGBT population is above the national average in the capital city of Bogotá (2.31%) but below the mean in other cities (1.46%) and in rural areas (0.95%). Monthly trends in reporting vary significantly at the national level but are stable for the city of Bogotá, which is partly why we focus on this setting. Observing the attributes of LGBT and non-LGBT individuals in Bogotá shows that those who identify as LGBT are younger, live in better neighborhoods, have greater educational attainment, and report better labor market outcomes. Results from conditional regressions and Oaxaca-Blinder decompositions suggest that the LGBT population in Bogotá faces no disadvantages in the labor market. Given data limitations, we expect that the available variables in household surveys cannot fully explain labor market differences. However, having all the necessary variables is one of the two assumptions required to appropriately estimate labor market discrimination. The other requires that we can adequately identify the groups of interest, which we test using a list experiment.

Our experiment seeks to understand whether current practices used in household surveys, whereby individuals are directly asked about their sexual self-identification, elicit truthful reporting. We randomly assigned 2,025 individuals into a direct response group that provides a list of N statements and answers sexual self-identification questions as in household surveys and a veiled response group that is provided a list of $N+1$ items, where the additional "sensitive" item determines whether they belong to the LGBT. We closely follow Coffman et al. (2017), both in design and empirical strategy. Results indicate that household surveys underestimate the size of the LGBT population. While survey estimates range between 1-4%, we find that the LGBT people

constitutes around 12-22% of the city's population. These results are robust to design effects, as well as ceiling and floor effects. We explore heterogeneity in reporting by assigned sex at birth, age groups, educational level, and marital status, finding some suggestive differences between direct and veiled response groups by individual attributes.

This paper contributes to several strands of literature. First, we contribute to the body of work that measures the size of the LGBT population through indirect response methods. While similar research has focused on developed countries (Coffman et al., 2017), we are only aware of one other effort in Latin America for Mexico (Gutierrez & Rubli, 2022). We are among the first to compare official nationally representative statistics on the size of the LGBT population with alternative estimates using a list experiment in Latin America and specifically, in Colombia. Second, we test whether the available data in Colombia are suitable for measuring labor market discrimination and identify potential issues that may preclude accurate estimates. Last, we hope to highlight the inclusive steps that the Colombian National Statistical Institute has taken to promote greater visibility of the LGBT population, which may help guide other countries that are looking to apply an inclusive measurement strategy in their official statistics. Together, this evidence will highlight the limitations of current approaches and will provide lessons and guidance for the future.

The remainder of this paper is organized as follows. The next section summarizes advances in measuring the size of the LGBT population and the inequalities they face in Latin America and Colombia, also highlighting efforts by the latter's Statistics Institute to improve measurement for this population group. Section 3 quantifies the size of the LGBT population and estimates labor market inequalities from household survey data, highlighting potential pitfalls from using this approach to guide policy. Section 4 presents the results from our list experiment conducted in

Bogotá, Colombia, which was designed to better approximate the size of the LGBT population and be directly comparable to household surveys. Section 5 concludes.

2. Advances in measuring the LGBT population

There has been significant progress in LGBT rights and policies in Latin America and the Caribbean (LAC) since the 1990s, although this progress has been uneven across countries. Some of the progress in approved laws and policies include the decriminalization of homosexuality, non-discrimination laws, legalizing same-sex marriage, expanding health services, same-sex couple adoption, among others; which has positioned LGBT rights in the region as comparatively better when compared to the global context (Corrales, 2017). However, despite progress on paper, discrimination based on sexual orientation and gender identity is far from eradicated (Corrales, 2020). Part of the requirements to build more inclusive societies require adequately measuring the LGBT population in different contexts and consistently over time, ideally with official statistics.

Measuring the LGBT population is fraught with issues because most nationally representative surveys ask respondents whether they belong to groups directly or without providing privacy to answer these questions. In some cases, questions about gender identity and sexual orientation are asked to one household member, usually the head of the household (OECD, 2019; Urban et al., 2020). Previous work has shown that survey respondents do not always report truthful information when asked directly (Coffman et al., 2017), and a single reporting member may have incomplete information or bias, which would result in an inaccurate estimate of the size of the LGBT population. Additionally, similar measurement issues occur with regards to instances of discrimination or violence, as well as their determinants (Agüero & Frisancho, 2017).

Accurately measuring discrimination requires that two assumptions about the data hold: that groups are well-identified and lack of biases in the variables of interest (Elder et al., 2010).

Estimates on the size of the LGBT population vary significantly across countries and years. On average, in 14 OECD countries with available data in the past decade[3], 2.7% of adults self-identified as LGBT, ranging from 1.2% in Norway to 3.8% in the US. In Latin America (i.e., Barbados, Guyana, Trinidad and Tobago, Chile, México, Guatemala, and Brazil), a recent study suggests that the size of the LGBT population by subgroups ranges from 1.6% of survey respondents who identify as homosexual in Trinidad and Tobago (2013) to 3% of respondents in Barbados (2004). Similarly, 1.8% and 4% of survey respondents reported being bisexual in Trinidad y Tobago and Barbados, respectively (Urban et al., 2020). Despite some scattered progress in measurement, research and information gaps remain in the region, particularly related with the lack of representative surveys informed by methodologies that measure sexual orientation and gender identity more accurately and consistently between and within countries over time.

An increasing number of studies are documenting the characteristics, wellbeing, and experiences of discrimination by the LGBT population through different data sources and methodologies (e.g., Consejo Nacional para Prevenir la Discriminación, 2011; Dirección de Diversidad Sexual & Secretaría de Inclusión Social, 2012; Human Rights Watch, 2014; INEC Ecuador, 2013; Nieves & Mondragón, 2013; Træen et al., 2009). Regarding labor market inequalities, an analysis of 46 research papers using survey data and covering 11 OECD countries suggests that there are gaps in employment status, earnings, and access to managerial positions for LGBT workers when compared to non-LGBT persons (OECD, 2019). Disparities across

[3] Transgender individuals are not accurately included in these figures due to data gaps. The 14 countries include Australia, Canada, Chile, France, Germany, Iceland, Ireland, Italy, Mexico, New Zealand, Norway, Sweden, United Kingdom, and the U.S.

subgroups are also highlighted, with significant differences for lesbians, who experience an employment and wage premium in comparison to heterosexual women. Additionally, experimental evidence suggests that both lesbians and gay men experience discrimination and barriers to participate in the labor market (OECD, 2019). However, there remains uncertainty as to who is disclosing being LGBT in survey data, with usually those better off more willing to self-identify as part of this group, suggesting these estimates might constitute a lower bound for discrimination.

In LAC many of the existing studies are still limited by the availability of data to identify sexual orientation, gender identity, and to conduct analysis across different dimensions and by different subgroups (Urban *et al*., 2020). Various initiatives have been launched to collect more data on this population to close these knowledge gaps. A first approximation has focused on identifying the size of the LGBT population, while the second is to document their experiences through a wide range of survey types, sample designs, questions, and concepts. Incorporating these groups into official surveys is the first step, which leads to an important discussion on how vulnerable groups are identified and what methods are well suited to estimate discrimination.

Colombia has recently joined the efforts to collect data on the LGBT population by incorporating questions about sexual orientation and gender identity in locally and nationally representative surveys. As a result of these surveys, recent estimates locate the prevalence of the LGBT population at the national level at 1.4% (DANE, 2022e). Below we provide details on the steps taken in Colombia to improve the visibility of the LGBT population in official surveys.

2.1. Current efforts to measure the LGBT population in Colombia

In 2020, the National Administrative Department of Statistics (*Departamento Administrativo Nacional de Estadística* or DANE, for its acronym in Spanish), officially launched the Guide for the Inclusion of a Differential and Intersectional Approach *("Guía para la Inclusión del Enfoque Diferencial e Interseccional")* where this institute recognizes the need to collect and disaggregate official statistics in the country by sexual orientation, gender identity, life cycle, disability status, ethnicity, and whether the person recognizes as a farmer, as well as the intersection of other relevant identities: displaced migrants, victims of the armed conflict, etc. (DANE, 2020b).

In addition to recognizing the importance of more disaggregated data to provide better measurements of Colombians' wellbeing, the guide provides concepts, tools, discussions, diagnoses, lessons, recommendations, and proposes next steps to motivate this inclusive approach in the statistical production of the country with the intent of eliminating the statistical invisibility of these groups and to better inform policy discussions that benefit all Colombians (DANE, 2020b).

The differential approach on sexual self-identification recommends the inclusion of a series of variables that allow the identification of the LGBT population. Although sex at birth is recognized as the minimum condition to conduct analysis related to gender gaps, the document provides additional guidelines to also include sexual orientation and gender identity questions as shown in Table A.1 in the Appendix (DANE, 2020b). In line with these guidelines, there have been different efforts at the local and national level to measure the LGBT population and gather information about their characteristics and experiences (See Table A.1 in the Appendix). These efforts have also complied with OECD recommendations to collect data on LGBT populations when at least two conditions are met: having a direct informant and privacy (OECD, 2019).

At the local level, the Multipurpose Household Survey of Bogotá *("Encuesta Multipropósito de Bogotá")* is carried out every three years and collects representative information for Bogotá and surrounding areas following a standard household survey format (DANE, 2022c). This was the first survey in Colombia to include questions on gender identity and sexual orientation. In 2014, this survey included a question about sexual orientation in addition to sex at birth. For its subsequent rounds, 2017 and 2021, it also included a question about gender identity (See Table A.1 in the Appendix).

National-level efforts to measure the LGBT population in Colombia include the Great Integrated Household Survey ("*Gran Encuesta Integrada de Hogares*" or GEIH, for its acronym in Spanish) and, more recently, the Voluntary Registry for the Visibility of Sexual and Gender Diversity in Colombia ("*Registro Voluntario para la Visibilidad de la Diversidad Sexual y de Género en Colombia*").[4]

The GEIH is collected monthly, following the standard format of labor force surveys in Latin America. In 2022, following a redesign of the survey after census data were collected in 2018, new questions were included to collect data on individual's sex, sexual orientation, and gender identity. Finally, the Voluntary Registry for the Visibility of Sexual and Gender Diversity in Colombia is an initiative launched by DANE in which, on a voluntary basis, Colombians are invited to fill out an online survey between June 26th and August 15th, 2022, to provide their name, identification, contact information, sexual orientation, gender identity, and location.[5] Once the registry is closed, the database will be used to send a targeted and more extensive survey to

[4] Other efforts at the national level are the inclusion of some questions in the Encuesta Nacional de Demografía y Salud 2015–ENDS ("The Colombia Demographic and Health Survey –DHS"), Encuesta Pulso Social ("Social Pulse Survey"), Encuesta de Uso de Sustancias Psicoactivas –ENCSP 2019 ("Survey on the Use of Psychoactive Substances"), and the Encuesta de Convivencia y Seguridad Ciudadana –ECSC 2020 ("Survey of Coexistence and Citizen Security".

[5] The Voluntary Registry for the Visibility of Sexual and Gender Diversity in Colombia may be found at https://www.dane.gov.co/index.php/estadisticas-por-tema/enfoque-diferencial-e-interseccional/enfoque-de-genero.

respondents and gather data on the experiences of the LGBTIQ+ population on education, work, health, subjective well-being, mental health, violence, discrimination, among others (DANE, 2022f). However, this voluntary registry is targeted specifically at the LGBTIQ+ people and would not be representative of the entire population of any specific city or the whole country.

Despite the recent inclusion of new questions in nationally representative household surveys, asking these questions is only part of the solution to promote the statistical visibility of vulnerable populations and to accurately estimate any inequalities or discrimination they may face. As there is still limited information on the consistency of the estimates for the size of the LGBT population and the inequalities they face, identifying how to better capture people's identities requires continuous work (DANE, 2020b). Some of the limitations already identified include the fact that people are not used to answer these questions in surveys, which leads to the need of using careful language, selecting the appropriate type of questions and their location within the survey, and providing the opportunity for a direct response in a private context (DANE, 2020b).

However, there has not been a thorough analysis of the patterns and trends in the size of the LGBT population with these new data. We quantify the size of the LGBT population and estimate labor market inequalities from these data in the next section to document the efforts by DANE to identify the LGBT people, highlighting potential issues from using this approach.

3. Estimating the size of the LGBT population and labor market differentials from Colombian household surveys

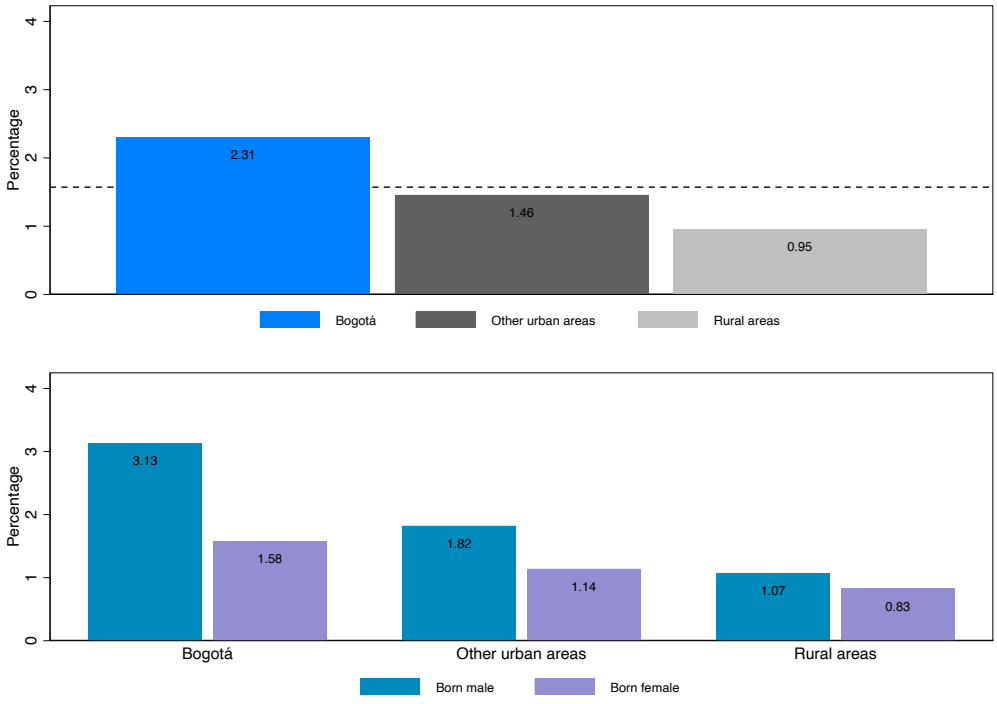
The National Administrative Statistics Department of Colombia (DANE, for its acronym in Spanish) implemented its new framework for the nationally representative household survey, GEIH, in 2021. As mentioned in the previous section, these methodological changes were also

meant to provide periodic statistics on vulnerable groups that had previously been omitted from official survey measurements, among them, the LGBT population (DANE, 2020b). In this section, we use pooled data from the GEIH spanning January 2021 to May 2022 to quantify the size of the LGBT population and compare their labor market outcomes with the non-LGBT population.

3.1 Size, composition, and attributes of the LGBT population

The top panel of Figure 1 shows the size of the LGBT population for the capital city of Bogotá, other urban cities, and rural areas. The dotted horizontal line presents the nationwide average. According to these estimates, about 1.57% of the Colombian population self-identifies as LGBT. The size of this population is above the national average in Bogotá, since 2.31% identify as LGBT. It is below average in other cities (1.46%) and in rural areas (0.95%).

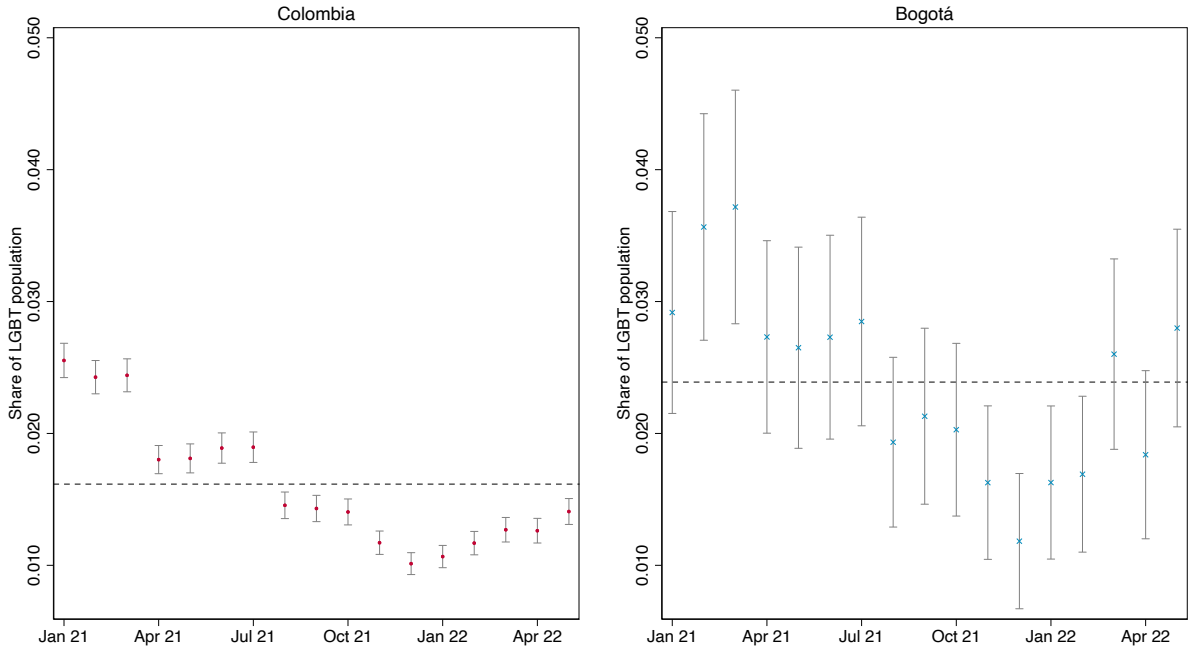
Figure 1. Size of LGBT population from household surveys



Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: All statistics are weighted using survey-provided expansion factors.

We also explore who belongs to the LGBT population based on their assigned sex at birth in the bottom panel of Figure 1 for the same locations. The evidence suggests that more individuals who are assigned as male at birth self-identify as LGBT compared to those assigned female. In Bogotá, 3.13% of males and 1.58% of females self-identify as LGBT. These differences fall when observing other urban centers (1.82 and 1.14 percent, respectively) and rural areas (1.07 and 0.83 percent, respectively). In Appendix Figure A.1, we also estimate the percentage of individuals who identify as lesbian, gay, bisexual, or trans; finding a greater percentage of self-identified gay and trans individuals compared to lesbian and bisexuals, across all three locations.

Figure 2. Size of LGBT population from household surveys over time



Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: All statistics are weighted using survey-provided expansion factors.

Given that the new methodology and survey questions to identify the LGBT population began in 2021, we analyze whether the size of the LGBT population has been consistent over time in Figure 2. We estimate the size of the LGBT population for Colombia and Bogotá, with their

respective 95% confidence intervals. The figure shows that the size of the LGBT population in Colombia has varied over time, beginning at around 2.5% and gradually reducing to around 1.5%. Trends for Bogotá are similar, but the differences are less pronounced, suggesting more stable reporting in the capital city. Given the stability in the size of the LGBT population in Bogotá and that our list experiment is representative for the capital city, we focus on studying the LGBT population in Bogotá for the remainder of this paper, which implies that the ensuing analyses apply only to Bogotá and not the entirety of Colombia.

What are the attributes for individuals who report being LGBT in Bogotá? Table 1 shows mean attributes and differences between LGBT and non-LGBT individuals for all survey respondents and by assigned gender at birth. We include demographic, socioeconomic, educational, and health characteristics routinely measured in the GEIH surveys.

When observing all individuals, we note that one in three individuals assigned female at birth self-identifies as LGBT. We also observe an age gradient, with younger people more likely to report belonging to the LGBT population, consistent with findings in other contexts such as Mexico (Gutierrez & Rubli, 2022). The LGBT population reports their marital status mainly as single, although several individuals are in civil union or married, the latter which is a legalized right since 2016. In Colombia, urban areas are often divided into economic strata, which classifies households according to their location to determine their utility bill categorization. Higher strata are considered more wealthy areas, and therefore pay higher utility bills to subsidize residents in lower income areas. As such, these categories provide a proxy of the income distribution. Our estimates show that while 8.6% of the non-LGBT population resides in the lowest strata, only 4% of the LGBT population lives in these areas. As economic strata increases, there is evidence of a larger LGBT population, suggesting that members of this group report living in better-off areas.

Table 1. Size of LGBT population from household surveys over time

	Full sample			Born male			Born female		
	Not LGBT	LGBT	Pr(i=ii)	Not LGBT	LGBT	Pr(i=ii)	Not LGBT	LGBT	Pr(i=ii)
Born female	53.5%	36.3%	0.000	-	-		-	-	
Age groups									
18-25	17.0%	23.8%	0.000	18.0%	18.9%	0.680	16.2%	32.3%	0.000
26-40	33.2%	41.4%	0.000	34.6%	40.2%	0.049	31.9%	43.6%	0.001
41-55	25.0%	20.5%	0.013	24.8%	26.6%	0.475	25.2%	9.9%	0.000
56 or older	24.8%	14.3%	0.000	22.7%	14.3%	0.000	26.6%	14.2%	0.000
Marital status									
Civil union	27.4%	28.1%	0.819	29.6%	28.5%	0.625	25.5%	27.2%	0.627
Married	24.9%	13.3%	0.000	26.6%	13.4%	0.000	23.5%	13.3%	0.000
Single	34.3%	50.9%	0.000	36.0%	51.6%	0.000	32.8%	49.6%	0.000
Separated or widowed	13.4%	7.7%	0.000	7.9%	6.4%	0.291	18.2%	9.8%	0.000
Economic strata									
1	8.6%	4.0%	0.000	9.0%	4.8%	0.000	8.2%	2.8%	0.000
2	39.5%	30.6%	0.000	39.9%	28.0%	0.000	39.2%	35.2%	0.210
3	34.5%	41.9%	0.001	33.8%	41.7%	0.006	35.1%	42.1%	0.040
4	10.9%	15.9%	0.003	10.8%	17.0%	0.004	11.1%	13.9%	0.278
5	3.9%	3.9%	0.932	3.9%	3.8%	0.984	3.9%	4.1%	0.858
6	2.5%	3.7%	0.193	2.6%	4.7%	0.131	2.5%	2.0%	0.578
Educational level									
Secondary incomplete or lower	22.4%	11.5%	0.000	21.8%	11.6%	0.000	22.9%	11.3%	0.000
Secondary complete	31.2%	23.1%	0.000	33.3%	22.7%	0.000	29.4%	23.6%	0.052
Technical and vocational	13.7%	16.0%	0.185	12.2%	14.9%	0.215	15.0%	18.0%	0.289
University or higher	32.7%	49.4%	0.000	32.7%	50.7%	0.000	32.7%	47.1%	0.000
Currently attending school	9.9%	15.6%	0.000	10.4%	15.0%	0.027	9.4%	16.6%	0.007
Access to healthcare	91.8%	91.8%	0.956	90.1%	91.5%	0.328	93.3%	92.3%	0.610
Health system									
Contributive	79.7%	83.7%	0.014	80.4%	83.8%	0.109	79.1%	83.6%	0.098
Special (Armed forces, etc.)	3.0%	2.6%	0.601	3.3%	3.4%	0.935	2.8%	1.2%	0.079
Subsidized (EPS-S)	17.3%	13.7%	0.014	16.3%	12.8%	0.058	18.1%	15.2%	0.264

Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: All statistics are weighted using survey-provided expansion factors.

Table 1 also looks at educational attainment and access to health for LGBT and non-LGBT individuals. Those who identify as LGBT are better educated, with only 34.6% reporting secondary schooling or lower compared to 46.4% for non-LGBT individuals. Almost half of self-identified LGBT individuals in Bogotá have at least some university education, suggesting greater

human capital accumulation for this group. LGBT individuals are also more likely to be currently attending school. Access to health care is not significantly different across groups, but there is evidence that LGBT individuals are more likely to be covered by the contributive health system, which is available only through the formal labor market. Given our findings that those who self-identify as LGBT tend to live in better-off areas of the city and accumulate greater human capital, greater access to contributive healthcare is consistent with better living conditions.

The above patterns are similar when looking at differences between LGBT and non-LGBT individuals conditional on their assigned sex at birth. We do find some differences depending on whether individuals are assigned male or female sex at birth. For instance, the age gradient is less pronounced for males compared to females. Similarly, there are larger differences in economic strata for those born male in lower income areas compared to individuals who are born female. Educational attainment shows similar patterns regardless of assigned sex at birth. Difference in healthcare access for the full sample, where access to contributive healthcare is greater, seems to be driven by individuals born as female since we find no significant differences for respondents who are born male.

3.2 Labor market outcomes for the LGBT population

The above results suggest that the LGBT population in Bogotá seems to be younger, lives in well-off neighborhoods, is better educated, and has greater access to contributive health through formal employment activities. We estimate a similar profile for labor market outcomes in Appendix Table A.2, which shows that LGBT individuals have slightly better labor market outcomes when analyzed unconditionally. On average, LGBT individuals have greater labor market participation,

higher employment rates, are more likely to be employed in a formal job and earn about 14% more per month compared to individuals who do not self-identify as LGBT.

However, to better approximate differences in labor market outcomes, we estimate linear regressions in Table 2 that control for several relevant attributes to compare conditional means instead of raw averages. We include age and its square; household size; as well as dummies for marital status, educational indicators, economic strata, and year-month fixed effects. The results from this conditional exercise show fewer differences compared to the unconditional analysis in Appendix Table A.2, denoting that while labor force participation and employment rates are significantly higher for the LGBT population in Bogotá, unemployment, formality rates, and earnings are not significantly different. When analyzing results by assigned sex at birth, we observe no differences for individuals born male in most outcomes, except for marginally significant lower earnings. LGBT individuals who are born female show greater labor force participation and formality rates, which are consistent with our profiles and unconditional means. Overall, these findings suggest that the LGBT population in Bogotá does not have systematically worse labor market outcomes than non-LGBT individuals when analyzing conditional differences.

Nevertheless, while these naïve estimates are suggestive, they are only conclusive if two assumptions hold. First, it assumes that the populations for which we estimate differences are precisely identified. Second, it assumes that the model is well-specified and therefore, unbiased. The first assumption will be tested in the next section by means of a survey list experiment. We close this section by investigating the validity of the second assumption, implementing commonly used decomposition methods to estimate what drives the differences we observe and how much of them are explained by the included controls in regressions from household survey data.

Table 2. Conditional differences in labor market outcomes

	Labor force participation	Employment rate	Unemployment rate	Informality	Log Earnings
Panel A: Full sample					
=1 if LGBT	0.042	0.036	0.002	0.013	-0.064
	(0.015)***	(0.019)*	(0.017)	(0.016)	(0.042)
Mean outcome	0.705	0.602	0.145	0.193	14.155
Adjusted R^2	0.271	0.206	0.053	0.080	0.456
Observations	30,216	30,216	20,806	30,216	15,540
Panel B: Born male					
=1 if LGBT	-0.003	0.007	-0.009	-0.016	-0.091
	(0.019)	(0.023)	(0.020)	(0.020)	(0.052)*
Mean outcome	0.799	0.693	0.133	0.230	14.211
Adjusted R^2	0.317	0.241	0.055	0.103	0.475
Observations	13,822	13,822	10,901	13,822	8,236
Panel C: Born female					
=1 if LGBT	0.064	0.037	0.015	0.048	-0.083
	(0.026)**	-0.032	-0.031	(0.028)*	-0.071
Mean outcome	0.621	0.522	0.159	0.159	14.089
Adjusted R^2	0.281	0.213	0.061	0.062	0.471
Observations	16,394	16,394	9,905	16,394	7,304

Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: This table presents conditional differences between LGBT and non-LGBT individuals in Bogotá, Colombia. Each column in the panel is a separate regression. Robust standard errors in parentheses. The regressions include controls for: age, age squared, household size; as well as dummies for marital status, educational attainment, and economic strata. All regressions are weighted using survey-provided expansion factors.
Significance levels: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$.

Appendix Tables A.3, A.4, and A.5 show the results of Oaxaca-Blinder decompositions comparing differences in labor market outcomes between LGBT and non-LGBT individuals in Bogotá. We conduct two decompositions following the suggestions in Jann (2008). The first breaks down the observed conditional differences in outcomes into endowment, price, and joint effects, while the second decomposes the difference into explained and unexplained components.

For the full sample, we find significant differences that favor the LGBT population in labor force participation, employment rates, and the logarithm of monthly earnings. In terms of what explains these differences, we find that endowment effects account for most of the difference, although price effects are also relevant for labor supply and employment. The importance of

endowment effects reinforces our results that LGBT individuals in Bogotá tend to have better socioeconomic conditions, which translate into slightly better observed labor market outcomes.

We also decompose the estimated differences into explained and unexplained components. Given that there are many variables that explain labor market outcomes that we cannot observe, there is likely omitted variable bias in our estimates, but we are unsure of its magnitude. The decompositions provide some insight as to their extent. Our estimates show that while we can explain up to 60% of the observed differences with the included explanatory variables, at least 40% of the total variation in earnings remains unexplained. Thus, this evidence suggests that while we can perform decompositions that try to mitigate existing biases, a large fraction of the differences in labor market outcomes between LGBT and non-LGBT individuals are due to unobservable factors. Implementing these decompositions therefore provides some approximation to inequality with estimations that are not free from limitations due to specification biases and other aspects pertaining to the suitability of decomposition methods (Elder et al., 2010).

Results by assigned sex at birth also indicate that LGBT individuals seem to be better off in terms of labor market outcomes. Decomposing these differences shows that endowment effects account for most of the observed differences in labor market outcomes for individuals born male and female. The analysis suggests that the included variables in our regressions better explain differences for individuals born male than female.

Together, results from naïve regressions and decompositions from household surveys suggest that the LGBT population in Bogotá faces no disadvantages in the labor market. As shown throughout this section, the LGBT population seems to be in better overall conditions and the decompositions indicate that a sizeable fraction of the estimated differences cannot be explained by the variables we include in a conditional analysis. Given data limitations, we expect that the

available variables in household surveys are unable to fully explain labor market differences between any two groups. However, having all the necessary variables is one of the two main assumptions to appropriately estimate labor market discrimination. The other requires that we can adequately identify population groups. While this is straightforward with sex at birth, it is less forthright with respect to the LGBT people. These estimates use self-reported identification as LGBT, which is asked directly to respondents and has only been recently implemented by the National Statistics Department. While this is a giant leap in terms of LGBT visibility as described in Section 2, we cannot expect to retrieve unbiased estimates of labor market differences between LGBT and non-LGBT individuals if we are not precisely identifying these groups in the available data. To test if this is the case, we carried out a survey list experiment in Bogotá to formally test how direct reports from household surveys compare to indirect response techniques.

4. Measuring the size of the LGBT population using a list experiment

Our experiment seeks to understand whether current practices used in the GEIH survey, where individuals are directly asked about their gender and sexual self-identification, elicit truthful reporting. It also intends to provide alternatives for the National Administrative Department of Statistics (DANE) to measure this population more accurately by employing indirect response methods. Specifically, we investigate behaviors related to the individual's gender identity and sexual orientation, by asking participants how they self-identify and whether they are sexually or romantically attracted to individuals of their same sex, opposite sex, both sexes, and others.

4.1 Experimental Design

We use list experiment (LE) or item count technique (ICT), which is a methodology that intends to reduce potential social desirability biases by increasing respondents' perception of privacy. This method has been used in several studies to explore opinions and behaviors related to sensitive issues which tend to be more susceptible to underreporting (Agüero & Frisancho, 2017; Coffman et al., 2017; Fergusson et al., 2017). Previous research has shown that list experiments tend to provide increased estimates of prevalence only for stigmatized behaviors (Tsuchiya et al., 2007).

The population of interest in our study consists of adult men and women residing in Bogota. A total of 2,025 participants were selected from the National Consulting Center's (CNC, for its acronym in Spanish) over 2 million telephone records through probabilistic and stratified random sampling. Sex, age, and urban district quotas were included. The selected sample is representative of Bogotá, which allows us to compare our results with the those from the GEIH. Participants were contacted via WhatsApp, if they agreed to participate in our experiment and signed the informed consent (Appendix B.1)[6], they received a link that allowed them to take an online survey on their own cellphones. Participants were asked to answer demographic questions on their neighborhood, marital status, level of education, and labor market outcomes. Then, they were reminded of the anonymity and confidentiality of their answers right before starting the sensitive questions module.

We closely follow the design in Coffman et al., (2017), with an experiment using a between-subject approach, which means that participants were randomly assigned to one of two groups: "Direct Report" or "Veiled Report". Direct Report treatment was used as a control group and helped us to replicate the existing survey design used in GEIH, where participants are asked to directly answer questions regarding their gender identity and sexual orientation. Alternatively,

[6] This experiment was submitted to the Universidad de los Andes Institutional Review Board (IRB) and received approval before any data collection began. It is possible that the sample who consents may be different from the city's population since they agree to participate in a survey that measures "identity", which we test in Table 3 below.

the Veiled Report treatment was designed using the ICT and it sought to elicit truthful answers about sensitive behaviors without the participants having to disclose them to the researchers.

Individuals assigned to Direct Report saw a list of *N* neutral statements and were asked to indicate how many of the statements were true/valid for them without having to indicate which (See Appendix B.2 for the precise wording of the questions). Participants were also asked to respond directly to the sensitive question. It is important to mention that all direct questions were formulated in the same way they are asked in the GEIH household surveys, which allows us to estimate the survey population's rate of misrepresentation. Alternatively, individuals assigned to the Veiled Response treatment were given the same list of *N* neutral statements plus an additional one. The additional statement referred to a sensitive issue, e.g. "I have felt sexual or romantic attraction towards a person of my same sex". They were then asked to answer how many of the statements were true/valid for them. Thus, the difference between the mean number of statements reported by the Direct Report group versus the Veiled Report group approximates the estimated prevalence of the behavior that we are attempting to measure, i.e., the size of the LGBT population.

The experimental design meant to prevent the research team from perfectly inferring an individual's answer to the sensitive question under the veiled treatment, so ICT or LE would provide a veil that eliminated individual inference. Since this only occurs if the respondent did not report extreme value answers, the choice of the neutral items in each list had to be conducted carefully to eliminate participants' incentives to lie (avoid individual inference). According to the literature, one should introduce at least one non-sensitive behavior in which all participants have incurred, but also prevent respondents from having all the non-sensitive behaviors (Tsuchiya et al., 2007). Therefore, in our experiment, the four neutral items of each list were composed of one pair of items/behaviors that were negatively correlated and one frequent item or behavior. One

advantage of introducing negatively correlated items in the lists is that it reduces the estimator's variance (Glynn, 2013) and consequently increases the statistical power of the experiment.

All the lists were tested and validated during a pilot to rule out the existence of ceiling and floor effects.[7] Fieldwork was carried out between June 13th and July 15th, 2022; with a sample size of 2,025. It is important to note that, the neutral statements included in the lists were not entirely innocuous. Research has shown that using less innocuous baseline items may contribute to make respondents less suspicious of the researcher's intent and thus provide more accurate prevalence estimates (Chuang et al., 2021). That is why we took advantage of the pre-election context and tried to camouflage the behavior of interest among other neutral statements related to some of the main topics discussed during this pre-electoral period in Colombia. Moreover, previous research suggests that the order in which participants answer the sensitive question can affect their probability of revealing the sensitive answer (Coffman et al., 2017), hence, the order of the items within each list and the order in which the participants had to answer the lists were randomized.

4.2. Sample description and balance

The sample was chosen to be representative of the capital city of Bogotá, Colombia. Table 3 compares descriptive statistics from the GEIH surveys and our sample by direct and veiled response groups.

[7] Our experimental design was tested in a pilot study with 36 participants in May 2022, with full IRB approval.

Table 3. Descriptive statistics for list experiment sample

	GEIH	List experiment sample		
	Household survey	Direct response (C)	Veiled response (T)	Pr(C=T)
Labor market outcomes				
Labor force participation	70.4%	74.9%	73.8%	0.507
Employment rate	60.1%	64.5%	63.7%	0.728
Unemployment rate	14.7%	14.0%	13.6%	0.752
Born female	53.1%	55.4%	50.5%	0.054
Age groups				
18-25	17.2%	16.1%	17.7%	0.314
26-40	33.4%	34.5%	33.0%	0.496
41-55	24.9%	23.6%	25.5%	0.336
56 or older	24.5%	25.9%	23.8%	0.476
Marital status				
Civil union	27.4%	24.1%	21.7%	0.251
Married	24.6%	21.2%	22.5%	0.583
Single	34.7%	41.8%	44.3%	0.301
Separated or Widowed	13.3%	12.9%	11.5%	0.488
Educational level				
Secondary incomplete or lower	22.1%	12.8%	14.4%	0.378
Secondary complete	31.0%	16.4%	14.8%	0.358
Technical and vocational	13.8%	26.2%	22.7%	0.083
University or higher	33.1%	44.7%	48.1%	0.122
Observations	30,438	1,002	1,023	

Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022 and list experiment sample collected between June and July 2022.
Notes: All statistics are weighted using survey-provided expansion factors for household surveys and the list experiment sample.

Overall, the list experiment sample has largely similar attributes to the household survey sample in terms of labor market outcomes and age structure. There are some differences between the sample, since we observe a slightly greater fraction of list experiment respondents that are single and with higher educational attainment. The latter is the largest discrepancy between both samples, while almost half of the household survey sample have secondary schooling or lower, only around 30% of the list experiment sample has lower levels of schooling. Thus, our list experiment sample is mostly similar to the population of Bogotá, but slightly better educated. The last column of Table 3 compares means between direct response (control) and veiled response

(treatment) groups in the list experiment sample, rejecting any statistically significant differences in attributes with 95% confidence, lending support that the randomization was successful and that the selected sample has similar attributes to the capital city's population.

4.3. Empirical Approach

Following Coffman et al. (2017), we slightly adapted the traditional list experiment methodology. The control group (Direct Report) received a list with *N* neutral statements, they had to report the number of statements that hold true for them and answer the sensitive question of interest directly. Hence, for each question q and each individual *i* assigned to the Direct Report group we observe: i) C_{qi} , the number of *N* true/valid statements and ii) d_{qi} , a dummy variable that takes the value of one if the individual reported having engaged in the sensitive behavior (e.g., felt sexual or romantic attraction towards a person of my same sex). Adding both variables yields Y_{qi}^{D} , which is the number of true items reported by each participant under Direct Report treatment. Alternatively, participants in the Veiled Report group received the same *N+1* statements, as mentioned above, the additional statement was related to the sensitive behavior of interest, therefore for this group we observe Y_{qi}^{v} ; the number of *N+1* statements reported as valid/true.

Since participants were randomly assigned to each one of the treatments, under truthful report, $E\left[Y_{qi}^{V}\right] = E\left[Y_{qi}^{D}\right]$. However, in the presence of social costs associated with telling the truth, the expected value of the number of items may differ between treatment groups. Therefore, the change in reporting between groups corresponds to $\mu \equiv E\left[Y_{qi}^{V}\right] - E\left[Y_{qi}^{D}\right]$.[8] To obtain more accurate

[8] This parameter can also be understood as a measure of the social cost that arises when reporting the sensitive answer; a larger value would imply a higher cost (Coffman et al., 2017).

estimates of μ, we use regression analysis instead of comparing sample means. The model estimated is:

$$y_{qi} = \mu V_i + \beta X_i + u_i$$

where y_{iq} refers to either Y_{qi}^{V} or Y_{qi}^{D} , V_i is a treatment indicator that takes the value of one if the participant was assigned to Veiled Report or zero otherwise. μ is the parameter of interest, if it is greater than zero and statistically significant, it reveals current practices used to measure the LGBT population in Colombia are underestimating the size of this group. X_i is a vector of observed demographic characteristics that allows us to control for sex at birth, age groups, education level, marital status, and districts within Bogotá. We estimate regressions for each list both with and without controls and using robust standard errors to account for heteroskedasticity. Additionally, we explore heterogeneous effects by sex, age groups, educational attainment, and marital status.

4.4. Results

We present results using the above strategy to estimate the size of the LGBT population in Table 4. The first column presents the mean fraction of individuals who self-identify as LGBT in each list in the direct response group, an estimate of group size using household survey methodology. The second column presents estimates of μ, the difference between veiled and direct reporting. The last column sums the first two columns to provide a more accurate measurement of the size of different LGBT groups (direct + indirect), which we bootstrap 1,000 times to obtain standard errors and confidence intervals.

Table 4. The effect of veiled report treatment on LGBT population size

	Fraction Reporting Sensitive Answer (Direct Report)	Δ Reporting of Sensitive Answer (Veiled Report)	Estimated True Fraction for Sensitive Answer
Panel A. No controls			
Gender identity	0.010	0.117	0.126
	(0.002)	(0.054)**	(0.053)**
Same-sex attraction	0.041	0.198	0.239
	(0.004)	(0.057)***	(0.055)***
Attraction to both sexes	0.023	0.158	0.181
	(0.003)	(0.050)***	(0.049)***
Panel B. With controls			
Gender identity	0.010	0.106	0.116
	(0.002)	(0.054)**	(0.054)**
Same-sex attraction	0.041	0.180	0.221
	(0.004)	(0.058)***	(0.056)***
Attraction to both sexes	0.023	0.150	0.173
	(0.003)	(0.050)***	(0.050)***

Source: Authors' elaboration from list experiment sample collected between June and July 2022.
Notes: N=2,025, with 1,002 in Direct Ask condition. Column 1 is the sample mean. Column 2 is the coefficient μ on "Veiled Report" from a regression without controls (Panel A) and with controls (Panel B). Controls include sex at birth, dummies for age group, marital status, educational attainment, and zone fixed effects. Column 3 adds column 1 and 2. Standard errors in brackets: Column 2 presents heteroskedasticity-robust standard errors, Column 3 presents standard errors derived using the bootstrap with 1,000 repetitions.
Significance levels: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$.

We find that current practices employed in household surveys are significantly underestimating the fraction of LGBT individuals in Bogotá. Panel A presents results without controls and Panel B including controls for sex at birth, age, education level, marital status, and city districts. Findings for List 1, which measures gender identity, indicate that about 1% of individuals would self-identify with a different sex that the one assigned at birth under current measurement practices. The second column suggests that asking individuals about their gender identity, while protecting their privacy, increases the size of this group by 10.6 percentage points, implying that the percentage of individuals who do not identify with their assigned sex at birth is 11.6%. For List 2, which measures same-sex attraction, 4.1% of direct respondents report non-

heterosexual preferences, while the veiled response group has a share 18 percentage points higher, for a total percentage of 22.1%. List 3, which measures attraction to both sexes, suggests that the bisexual group is 2.3% when asked to self-identify. This share grows by 15 percentage points for a population size of approximately 17.3%. Estimated differences are statistically significant at 95% or higher and suggest sizeable underreporting in the LGBT population. Direct response would quantify the LGBT population between 1-4%, while our estimates suggest that this group represents between 12-24% of individuals in the city of Bogotá.

The household survey results in Section 3 show that there is some heterogeneity in sexual self-identification and labor market outcomes according to assigned sex at birth. We explore whether reporting differs by this attribute in Table 5, where Panel A shows results for the list experiment findings for individuals born male and female, respectively. We find that most of the effect observed in Table 4 is driven by individuals who are born female. For individuals born male, we find no significant effects between direct and veiled response groups for gender identity and same-sex attraction, but a slightly significant difference for attraction to both sexes. The results for individuals who are born female are all statistically significant and suggest that the LGBT population for these individuals ranges between 20-32 percent, but direct questioning would erroneously underestimate the size of this population to be between 1.1-3.3 percent.

Table 5. The effect of veiled report treatment on LGBT population size by sex at birth

	Fraction Reporting Sensitive Answer (Direct Report)	Δ Reporting of Sensitive Answer (Veiled Report)	Estimated True Fraction for Sensitive Answer
Panel A. Individuals born male			
Gender identity	0.008	0.009	0.019
	(0.003)	(0.070)	(0.066)
Same-sex attraction	0.050	0.051	0.092
	(0.007)	(0.070)	(0.070)
Attraction to both sexes	0.022	0.116	0.139
	(0.005)	(0.065)*	(0.066)**
Panel B. Individuals born female			
Gender identity	0.011	0.200	0.210
	(0.003)	(0.081)**	(0.081)***
Same-sex attraction	0.033	0.280	0.321
	(0.006)	(0.089)***	(0.085)***
Attraction to both sexes	0.024	0.175	0.198
	(0.005)	(0.075)**	(0.073)***

Source: Authors' elaboration from list experiment sample collected between June and July 2022.
Notes: N=1,011 born male with 483 in Direct Ask condition; and N=1,014 born female with 519 in Direct Ask condition. Column 1 is the sample mean. Column 2 is the coefficient μ on "Veiled Report" from a regression with controls (Panel B). Controls include gender at birth, dummies for age group, marital status, educational attainment, and zone fixed effects. Column 3 adds column 1 and 2. Standard errors in brackets: Column 2 presents heteroskedasticity-robust standard errors, Column 3 presents standard errors derived using the bootstrap with 1,000 repetitions.
Significance levels: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$.

To evaluate whether out results are robust and not a reflection of the sample we surveyed, we carried out two robustness exercises. First, participants may have difficulties understanding or complying with the instructions of the ICT (Höglinger and Jann, 2018), or there could also be design effects (i.e., Responses to control items change when the sensitive statement is added). To evaluate whether any of these issues affected our experiment, we conducted a Double List Experiment (DLE) test. This test consists in including the sensitive behavior of interest in two lists. In this case, we introduced the statement "I have felt sexual or romantic attraction towards a person of my same-sex" in Lists 2 and 4. The Veiled Report group received List 2 with the sensitive statement, while the Direct Report group received the same list without the sensitive item.

Alternatively, List 4 was given to the Direct Report group with the sensitive statement and to the Veiled Report group without the sensitive item. Therefore, two measures of the prevalence of the sensitive behavior can be estimated and compared to determine whether they are statistically different. Figure A.2 in the appendix reveals that there are no significant differences between the expected value of the prevalence of the sensitive behavior between the lists. Hence, there does not seem to be evidence of compliance issues or design effects, and we can have higher confidence in the validity of our estimates (Chuang et al., 2020).

We also assess if our ICT effectively provided a veil that mitigated individual inference, and thus helped to protect respondents' privacy. Figure A.3 in the appendix shows the distribution of the number of valid items for each question. Since there is no evidence of a high concentration of answers in extreme values (zero or five) in any of the questions, it seems that our ICT did help to protect participants' privacy and therefore elicit more truthful answers than direct questioning. The ceiling and floor effects for each list are reported in Table A.7.

Additionally, we also explore differences between direct and veiled response groups by age, educational level, and marital status to provide suggestive evidence of heterogeneous effects. We summarize the results for age in Figure A.4, educational level in Figure A.5, and marital status in Figure A.6 in the Appendix. Overall, we find an age gradient in reporting, with no significant differences between younger individuals but greater differences for individuals aged over 40 years old. For simplicity, we divide the sample into low and high educational categories, the former being individuals who report secondary complete or less and the latter those with at least some higher education. We find significant differences between direct and veiled report for those in the high skill category, which is most of our sample as shown in Table 3. We also find statistically relevant differences in List 3 (attraction to both sexes) for low-skilled individuals. In terms of

marital status, we divide the sample into individuals who are single and married or in civil union, finding some differences in Lists 2 (same-sex attraction) and List 3 (attraction to both sexes). These results are robust when we observe the placebo List 4 asked to the direct response group.

5. Conclusion

This paper studies whether household surveys precisely identify the LGBT population and are suitable to measure labor market discrimination in Colombia. We first quantify the size of the LGBT population and estimate labor market inequalities from these data, highlighting potential pitfalls from using this approach. We then present findings from a list experiment in Bogotá, Colombia. Results show that household surveys underestimate the size of the LGBT population and may yield biased estimates of labor market inequalities. While survey estimates of the LGBT people range between 1-4%, this fraction may be around 12-22% of the total population.

Using existing data, we may be likely to conclude that the LGBT population in Bogotá does not have systematically worse labor market outcomes than non-LGBT individuals when implementing conditional analysis using regressions or more complex decomposition methods. This is due to both imprecise identification of the groups and potential biases from omitted variables and measurement errors. Obtaining better statistical measurements on sensitive issues requires time and methodological developments. The Colombian National Administrative Statistics Institute has taken valuable first steps to include and identify the LGBT people in its household surveys. However, other forms of questioning may be required to better bridge the gap we find between direct and veiled reporting of sexual self-identification. While list experiments have their limitations, mainly that we cannot identify whether the individual is or is not part of the LGBT population and their limited representativity, they do provide a metric for how much work

remains.

In Colombia the rights of LGBT individuals have improved considerably, starting with the decriminalization of homosexuality in 1981. Progress in the legal protections for this population has placed the country as one of the most advanced in terms of inclusivity (Choi et al., 2020). LGBT rights in Colombia include same-sex marriage, retirement pensions, social security, property rights, legal protection from discrimination, the right for same-sex couples to adopt children, and the right to change their name and gender identification (Choi et al., 2020). Despite advances in legal protections, there remain barriers in practice to guarantee such rights, leading to experiences of discrimination and potential disadvantage in some areas for the LGBT population (Ripoll, 2009). But identifying and reducing discrimination against vulnerable groups also requires statistical visibility. Our findings suggest that while current measurement practices are a step forward for the LGBT people's visibility, additional steps are required before household surveys may consistently track this group, measure discrimination, and guide policy responses.

Identifying LGBT individuals and documenting the inequalities and discrimination they face in their daily lives requires further research. Other methods that inquire about sexual self-identification should be tested to gauge the relative effectiveness of different methodologies. This involves not only quantitative work, but also qualitative studies that help us to understand why truthful reporting carries more stigma with respect to some attributes in comparison to others. Additionally, differences between direct and veiled reports may vary outside of Bogotá, especially in other urban cities and rural areas, which were not included in our list experiment sample. Better measurement not only applies to group belonging, but also to the determinants of outcomes, including the labor market, but also in terms of education, health, and other facets of well-being. We hope that these findings motivate other researchers to continue the conversation on how to

identify vulnerable individuals and better understand their situation to improve their lives.

References

Agüero, J., & Frisancho, V. (2017). Misreporting in sensitive health behaviors and its impact on treatment effects: An application to intimate partner violence. *IDB Working Paper Series*.

Choi, S. K., Divsalar, S., Flórez-Donado, J., Kittle, K., Lin, A., Meyer, I. H., & Torres-Salazar, P. (2020). *Stress, health, and well-being of LGBT people in Colombia: Results from a National Survey*. 110.

Chuang, E., Dupas, P., Huillery, E., & Seban, J. (2021). Sex, lies, and measurement: Consistency tests for indirect response survey methods. *Journal of Development Economics*, *148*. https://doi.org/10.1016/j.jdeveco.2020.102582

Coffman, K. B., Coffman, L. C., & Marzilli Ericson, K. M. (2017). The Size of the LGBT population and the magnitude of antigay sentiment are substantially underestimated. *Management Science*, *63*(10), 3168–3186. https://doi.org/10.1287/mnsc.2016.2503

Consejo Nacional para Prevenir la Discriminación. (2011). *Encuesta nacional sobre discriminación en México: Resultados generales*. https://www.conapred.org.mx/userfiles/files/Enadis-2010-RG-Accss-002.pdf

Corrales, J. (2017). Understanding the uneven spread of LGBT rights in Latin America and the Caribbean, 1999-2013. *Journal of Research in Gender Studies*, *7*(1), 52–82. https://doi.org/10.22381/JRGS7120172

Corrales, J. (2020). The Expansion of LGBT rights in Latin America and the backlash. In *Handbook of global LGBT and sexual diversity politics* (pp. 184–200). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190673741.013.14

DANE. (2020a). *Encuesta Nacional de Consumo de Sustancias Psicoactivas (ENCSPA): Resultados 2019*. https://www.dane.gov.co/files/investigaciones/boletines/encspa/presentacion-encspa-2019.pdf

DANE. (2020b). *Guía para la inclusión del enfoque diferencial e interseccional en la producción estadística del Sistema Estadístico Nacional*. https://www.dane.gov.co/files/investigaciones/genero/guia-inclusion-enfoque-difencias-intersecciones-produccion-estadistica-SEN.pdf

DANE. (2022a). *Boletín técnico Gran Encuesta Integrada de Hogares (GEIH) Abril 2021—Marzo 2022*. https://www.dane.gov.co/files/investigaciones/boletines/ech/ech_poblacion_lgbt/boletin_GEIH_poblacion-lgbt_abr21_mar22.pdf

DANE. (2022b). *Encuesta de Convivencia y Seguridad Ciudadana (ECSC 2020)*.

DANE. (2022c). *Encuesta Multipropósito Bogotá, Cundinamarca, 2021*.

DANE. (2022d). *Encuesta Pulso Social: Población LGBT y bienestar subjetivo (Periodo enero-mayo 2022)*. https://www.dane.gov.co/files/investigaciones/boletines/pulso-social/presentacion_EPS_poblacion_lgbt-ene2022-may2022.pdf

DANE. (2022e). *Mercado laboral de la población LGBT Junio 2021-Mayo 2022* (Boletín Técnico Gran Encuesta Integrada de Hogares (GEIH)).

DANE. (2022f). *Registro Voluntario para la Visibilidad de la Diversidad Sexual y de Género en Colombia*. http://www.dane.gov.co/files/investigaciones/genero/diversidad_sexual_y_de_genero/preguntas-frecuentes-del-registro.pdf

Dirección de Diversidad Sexual & Secretaría de Inclusión Social. (2012). *Consulta nacional sobre realidades LGBTI en El Salvador*.

Elder, T. E., Goddeeris, J. H., & Haider, S. J. (2010). Unexplained gaps and Oaxaca–Blinder decompositions. *Labour Economics*, *17*, 284–290. https://doi.org/10.1016/j.labeco.2009.11.002

Fergusson, L., Molina, C., & Riaño, J. F. (2017). I sell my vote, and so what? A new database and evidence from Colombia. *Documentos CEDE*. https://doi.org/10.2139/ssrn.2935385

Glynn, A. N. (2013). What can we learn with statistical truth serum? *Public Opinion Quarterly*, *77*, 159–172. https://doi.org/10.1093/poq/nfs070

Gutierrez, E., & Rubli, A. (2022). *Challenges for measuring LGBT+ population and homophobia in Mexico*.

Human Rights Watch. (2014). *Not safe at home: Violence and discrimination against LGBT people in Jamaica*.

INEC Ecuador. (2013). *Estudio de caso sobre condiciones de vida, inclusión social y cumplimiento de derechos humanos de la población lGBTI en el Ecuador*.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, *8*(4), 453–479.

Nieves, A., & Mondragón, A. (2013). *Diversidad sexual: Derechos, discriminación y violencia. Encuesta Marcha del Orgullo*.

OECD. (2019). *Society at a glance 2019: OECD social indicators*. OECD. https://doi.org/10.1787/soc_glance-2019-en

Ripoll, J. L. (2009). El amor en los tiempos de cólera: Derechos LGBT en Colombia. *SUR -Revista Internacional de Derechos Humanos*, *6*(11), 79–97.

Træen, B., Martinussen, M., Vitters, J., & Saini, S. (2009). Sexual orientation and quality of life among university students from Cuba, Norway, India, and South Africa. *Journal of Homosexuality*, *56*(5), 655–669. https://doi.org/10.1080/00918360903005311

Tsuchiya, T., Hirai, Y., & Ono, S. (2007). A study of the properties of the item count technique. *Public Opinion Quarterly*, *71*(2), 253–272. https://doi.org/10.1093/poq/nfm012

Urban, A.-M., Flor Ágreda, M. J., Ramos Moreno, A., & Ortiz, D. (2020). *Hacia un mejor entendimiento de la discriminación por orientación sexual e identidad de género* (IDB-TN-01885; Nota Técnica). Banco Interamericano de Desarrollo.

Appendix A: Supplementary Tables and Figures

Table A.1.Proposed and implemented efforts to measure the LGBT population in Colombia

Local/National	Proposed/Implemented	Questions sexual self-identification
National	**Proposed** Guide for the Inclusion of a Differential and Intersectional Approach *Guía para la Inclusión del Enfoque Diferencial e Interseccional* [a]	**Sex:** * Option a: Sex (man/male or woman/female) * Option b: Sex assigned at birth (male, female, indeterminate) **Sexual orientation:** * Option a: Sexual orientation is the attraction a person can have for the opposite sex (heterosexual), the same sex (homosexual) or towards both sexes (bisexual). Which of the following alternatives best defines your sexual orientation? Heterosexual, gay/lesbian, bisexual, other -which one? * Option b: Do you feel sexual or romantic attraction towards? Men, women, both sexes, other -which one? **Gender identity:** * Option a: Gender identity refers to how the person feels or identifies towards its gender (male or female) and this can align or not with their biological sex. In terms of how you feel about your gender, how do you identify? Male, female, transgender (trans male and trans female, other -which one? * Option b: Do you recognize as? Man, woman, trans man, trans woman, other. * Option c: Do you feel identified with the sex that was assigned to you at birth? Yes or no. * Option c.1: Do you feel identified with the sex that was assigned to you at birth? Yes or no (ask the next question). Do you recognize as? Trans man, trans woman, other –which?
	Implemented The Colombia Demographic and Health Survey (DHS) *Encuesta Nacional de Demografía y Salud (ENDS 2015)*[b]	**Sexual orientation:** Sexual orientation is the attraction a person can have for the opposite sex (heterosexual), the same sex (homosexual) or towards both sexes (bisexual). Which of the following alternatives best defines your sexual orientation? Heterosexual, gay/lesbian, bisexual, other -which one? **Gender identity:** Gender identity refers to how the person feels or identifies towards its gender (male or female) and this can align or not with their biological sex. In terms of how you feel about your gender, how do you identify? Male, female, transgender (trans male and trans female, other -which one?
	Implemented Great Integrated Household Survey *Gran Encuesta Integrada de Hogares (GEIH 2022)*[c]	**Sexual orientation:** Do you feel sexual or romantic attraction towards? Men, women, both sexes, other -which one? **Gender identity:** Do you recognize as? Man, woman, trans man, trans woman, other -which one?

	Implemented Social Pulse Survey *Encuesta Pulso Social (EPS 2022)*[d]	The EPS is a subsample of the GEIH; therefore, the LGBT population is identified using the same questions. **Sexual orientation:** Do you feel sexual or romantic attraction towards? Men, women, both sexes, other -which one? **Gender identity:** Do you recognize as? Man, woman, trans man, trans woman, other -which one?
	Implemented Survey of Coexistence and Citizen Security *Encuesta de Convivencia y Seguridad Ciudadana (ECSC 2020)*[e]	**Sexual orientation:** Do you feel sexual or romantic attraction towards? Men, women, both sexes, other -which one? **Gender identity:** Do you recognize as? Man, woman, trans man, trans woman, other
	Implemented Survey on the Use of Psychoactive Substances *Encuesta de Uso de Sustancias Psicoactivas (ENCSPA 2019)*[f]	**Sexual orientation:** Sexual orientation is the attraction a person can have for the opposite sex (heterosexual), the same sex (homosexual) or towards both sexes (bisexual). Which of the following alternatives best defines your sexual orientation? Heterosexual, gay/lesbian, bisexual, other -which one? **Gender identity:** Gender identity refers to how the person feels or identifies towards its gender (male or female) and this can align or not with their biological sex. In terms of how you feel about your gender, how do you identify? Male, female, transgender (trans male and trans female, other -which one?
	Implemented National Survey of Information Technologies *Encuesta Nacional de Tecnologías de la Información (ENTIC 2020-2021)*[g]	**Sexual orientation:** Pending to confirm, waiting for publication of questionnaires **Gender identity:** Do you feel identified with the sex that was assigned to you at birth? Yes or no.
Local	**Implemented** Multipurpose Household Survey of Bogotá *Encuesta Multipropósito de Bogotá (EM 2014, 2017 & 2021)*[h]	**Sexual orientation (2014 onwards)** What is your sexual orientation? Heterosexual (attraction to opposite sex), homosexual (same sex attraction), bisexual (attraction to both sexes) **Gender identity (2017, 2021 onwards):** Do you feel identified with the sex that was assigned to you at birth? Yes or no (ask the next question). Do you recognize as? Trans man, trans woman, other –which?

Source: [a](DANE, 2020b) [b](DANE, 2020b) [c](DANE, 2022a) [d](DANE, 2022d) [e](DANE, 2022b) [f](DANE, 2020a) [g](DANE, 2020b) [h](DANE, 2022c)

Table A.2. Labor market outcomes for non-LGBT and LGBT population in Bogotá

	Full sample			Born male			Born female		
	Not LGBT	LGBT	Pr(i=ii)	Not LGBT	LGBT	Pr(i=ii)	Not LGBT	LGBT	Pr(i=ii)
Labor force participation	70.2%	80.5%	0.000	79.9%	82.4%	0.226	61.8%	77.2%	0.000
Employment rate	59.9%	68.6%	0.000	69.1%	71.9%	0.211	51.9%	62.7%	0.001
Unemployment rate	14.7%	14.9%	0.932	13.5%	12.7%	0.634	16.1%	18.8%	0.401
Hours worked per week	47.8	46.9	0.263	49.5	46.2	0.000	45.9	48.4	0.013
Labor relationship									
Salaried	67.4%	68.5%	0.706	67.1%	67.7%	0.886	67.8%	70.1%	0.610
Self-employed	29.8%	29.6%	0.975	32.4%	31.2%	0.748	26.8%	26.3%	0.969
Unpaid worker	0.6%	0.3%	0.247	0.4%	0.3%	0.653	0.8%	0.3%	0.210
Domestic worker	2.2%	1.6%	0.394	0.1%	0.8%	0.140	4.7%	3.3%	0.420
Has a contract	74.5%	77.6%	0.133	71.9%	76.0%	0.126	77.4%	80.7%	0.321
Contributes to pension (formal)	67.2%	73.3%	0.006	65.7%	75.6%	0.000	68.9%	68.7%	0.987
Monthly labor income	2,175,650	2,485,106	0.080	2,294,594	2,755,609	0.055	2,039,366	1,929,597	0.541
Not happy at current job	30.6%	34.6%	0.514	31.2%	36.8%	0.418	29.9%	30.2%	0.976
Willing to change current job	18.4%	20.6%	0.451	17.8%	20.5%	0.422	19.2%	20.7%	0.701
Has problems in current job	7.5%	11.1%	0.421	7.1%	12.5%	0.337	8.0%	8.5%	0.809

Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: All statistics are weighted using survey-provided expansion factors.

Table A.3. Oaxaca-Blinder decompositions for labor market outcomes in Bogotá (Full sample)

	Labor force participation	Employment rate	Unemployment rate	Informality	Log Earnings
Not LGBT	0.702	0.600	0.145	0.192	14.155
	(0.003)***	(0.003)***	(0.003)***	(0.003)***	(0.008)***
LGBT	0.807	0.694	0.141	0.180	14.299
	(0.017)***	(0.020)***	(0.017)***	(0.017)***	(0.053)***
Difference	-0.105	-0.093	0.004	0.013	-0.144
	(0.017)***	(0.020)***	(0.018)	(0.017)	(0.054)***
Decomposition 1					
Endowments	-0.057	-0.056	0.023	0.041	-0.262
	(0.016)***	(0.018)***	(0.015)	(0.014)***	(0.051)***
Coefficients	-0.042	-0.036	-0.001	-0.013	0.063
	(0.015)***	(0.019)*	(0.017)	(0.016)	(0.042)
Interaction	-0.006	-0.001	-0.017	-0.016	0.055
	(0.013)	(0.016)	(0.014)	(0.014)	(0.038)
Decomposition 2					
Explained	-0.063	-0.057	0.006	0.026	-0.208
	(0.008)***	(0.008)***	(0.004)	(0.004)***	(0.031)***
Unexplained	-0.042	-0.036	-0.002	-0.013	0.064
	(0.015)***	(0.019)*	(0.017)	(0.016)	(0.042)
Observations	30,216	30,216	20,806	30,216	15,540

Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: This table presents Oaxaca-Blinder decomposition results for differences between LGBT and non-LGBT individuals in Bogotá, Colombia. Each column in the panel is a separate regression. Robust standard errors in parentheses. The regressions include controls for: age, age squared, household size; as well as dummies for marital status, educational attainment, and economic strata. All regressions are weighted using survey-provided expansion factors.
Significance levels: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$.

Table A.4. Oaxaca-Blinder decompositions for labor market outcomes in Bogotá (male at birth)

	Labor force participation	Employment rate	Unemployment rate	Informality	Log Earnings
Not LGBT	0.798	0.692	0.133	0.232	14.208
	(0.004)***	(0.005)***	(0.004)***	(0.004)***	(0.011)***
LGBT	0.826	0.731	0.115	0.168	14.403
	(0.021)***	(0.025)***	(0.021)***	(0.020)***	(0.066)***
Difference	-0.028	-0.038	0.017	0.064	-0.195
	(0.021)	(0.026)	(0.021)	(0.021)***	(0.067)***
Decomposition 1					
Endowments	-0.061	-0.061	0.016	0.045	-0.437
	(0.022)***	(0.024)***	(0.019)	(0.019)**	(0.070)***
Coefficients	0.002	-0.007	0.009	0.016	0.088
	(0.019)	(0.024)	(0.021)	(0.020)	(0.051)*
Interaction	0.031	0.030	-0.007	0.002	0.154
	(0.020)	(0.022)	(0.019)	(0.018)	(0.056)***
Decomposition 2					
Explained	-0.030	-0.031	0.008	0.048	-0.286
	(0.011)***	(0.011)***	(0.006)	(0.007)***	(0.040)***
Unexplained	0.003	-0.007	0.009	0.016	0.091
	(0.019)	(0.023)	(0.020)	(0.020)	(0.050)*
Observations	13,822	13,822	10,901	13,822	8,236

Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: This table presents Oaxaca-Blinder decomposition results for differences between LGBT and non-LGBT individuals in Bogotá, Colombia. Each column in the panel is a separate regression. Robust standard errors in parentheses. The regressions include controls for: age, age squared, household size; as well as dummies for marital status, educational attainment, and economic strata. All regressions are weighted using survey-provided expansion factors.
Significance levels: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$.

Table A.5. Oaxaca-Blinder decompositions for labor market outcomes in Bogotá (female at birth)

	Labor force participation	Employment rate	Unemployment rate	Informality	Log Earnings
Not LGBT	0.618	0.520	0.158	0.158	14.094
	(0.004)***	(0.005)***	(0.004)***	(0.003)***	(0.012)***
LGBT	0.776	0.633	0.185	0.199	14.100
	(0.030)***	(0.035)***	(0.033)***	(0.028)***	(0.084)***
Difference	-0.158	-0.112	-0.027	-0.041	-0.006
	(0.030)***	(0.035)***	(0.033)	(0.028)	(0.085)
Decomposition 1					
Endowments	-0.051	-0.066	0.006	0.009	0.081
	(0.028)*	(0.036)*	(0.033)	(0.026)	(0.080)
Coefficients	-0.064	-0.037	-0.015	-0.048	0.086
	(0.027)**	(0.033)	(0.032)	(0.027)*	(0.072)
Interaction	-0.044	-0.009	-0.019	-0.001	-0.173
	(0.025)*	(0.033)	(0.032)	(0.025)	(0.067)***
Decomposition 2					
Explained	-0.094	-0.075	-0.012	0.007	-0.089
	(0.013)***	(0.012)***	(0.007)	(0.006)	(0.052)*
Unexplained	-0.064	-0.037	-0.015	-0.048	0.083
	(0.026)**	(0.031)	(0.030)	(0.026)*	(0.067)
Observations	16,394	16,394	9,905	16,394	7,304

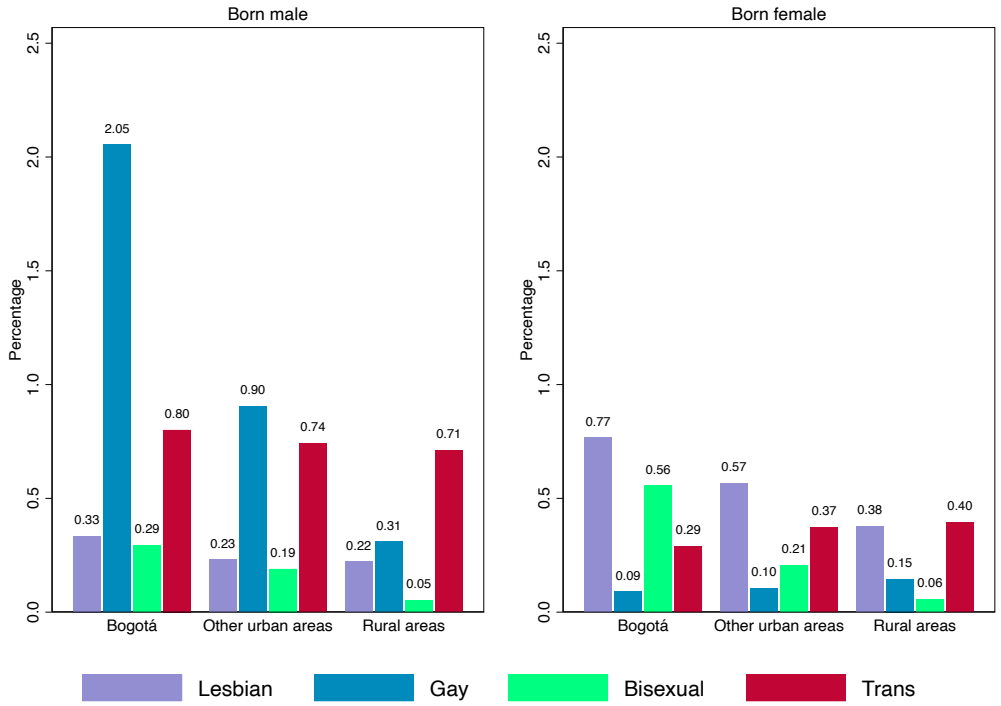
Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: This table presents Oaxaca-Blinder decomposition results for differences between LGBT and non-LGBT individuals in Bogotá, Colombia. Each column in the panel is a separate regression. Robust standard errors in parentheses. The regressions include controls for: age, age squared, household size; as well as dummies for marital status, educational attainment, and economic strata. All regressions are weighted using survey-provided expansion factors.
Significance levels: $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.10$.

Table A.6. Prevalence of Ceiling and floor effects

	% of Individuals Reporting 0 Statements are True		% of Individuals Reporting all Statements are True	
	Direct	Veiled	Direct	Veiled
L1	1.40	2.74	10.08	5.67
L2	1.60	3.23	10.08	5.08
L3	2.30	2.05	10.08	3.42
L4	1.10	1.17	3.79	8.80
Observations	1002	1023	1002	1023

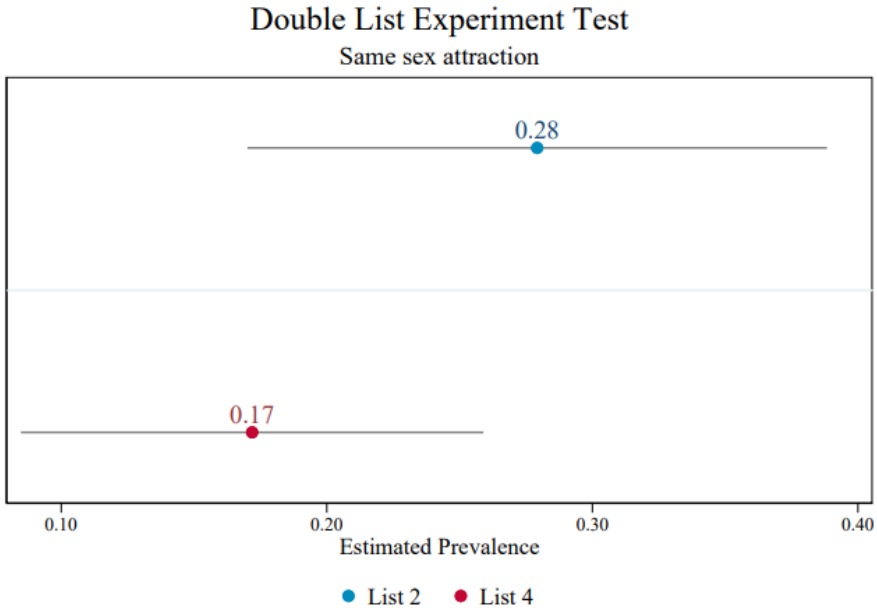
Source: Authors' elaboration from list experiment sample collected between June and July 2022.

Figure A.1. Composition of LGBT population by sex at birth



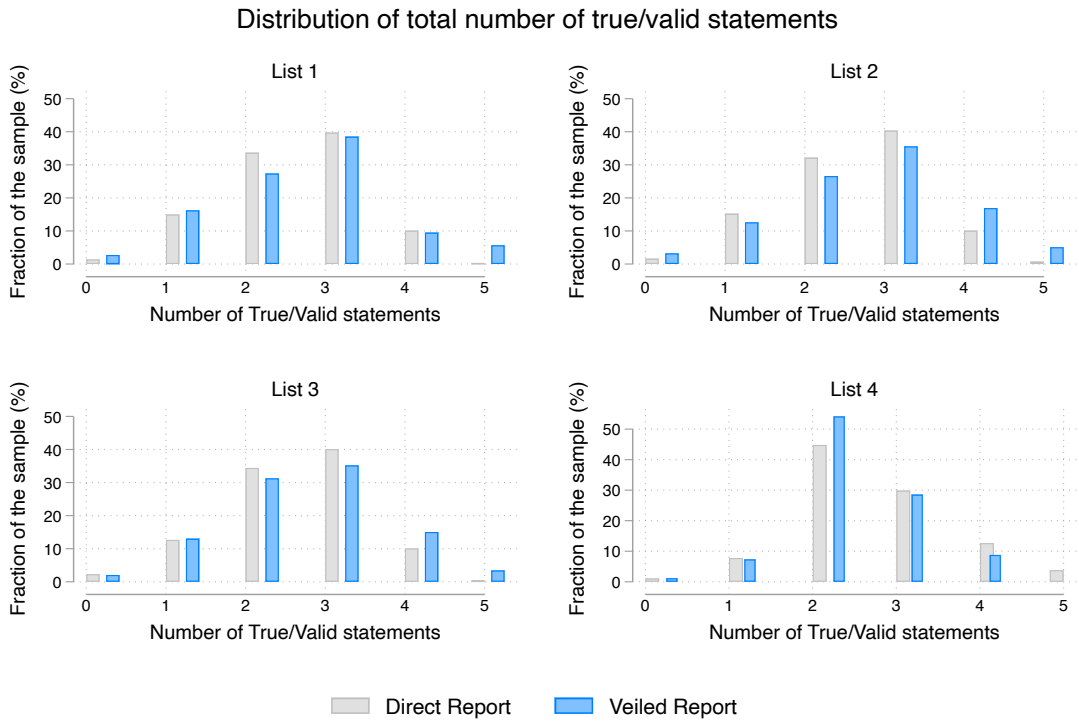
Source: Authors' elaboration from Colombian household surveys (GEIH) from January 2021 to May 2022.
Notes: All statistics are weighted using survey-provided expansion factors.

Figure A.2. Robustness test for list experiment



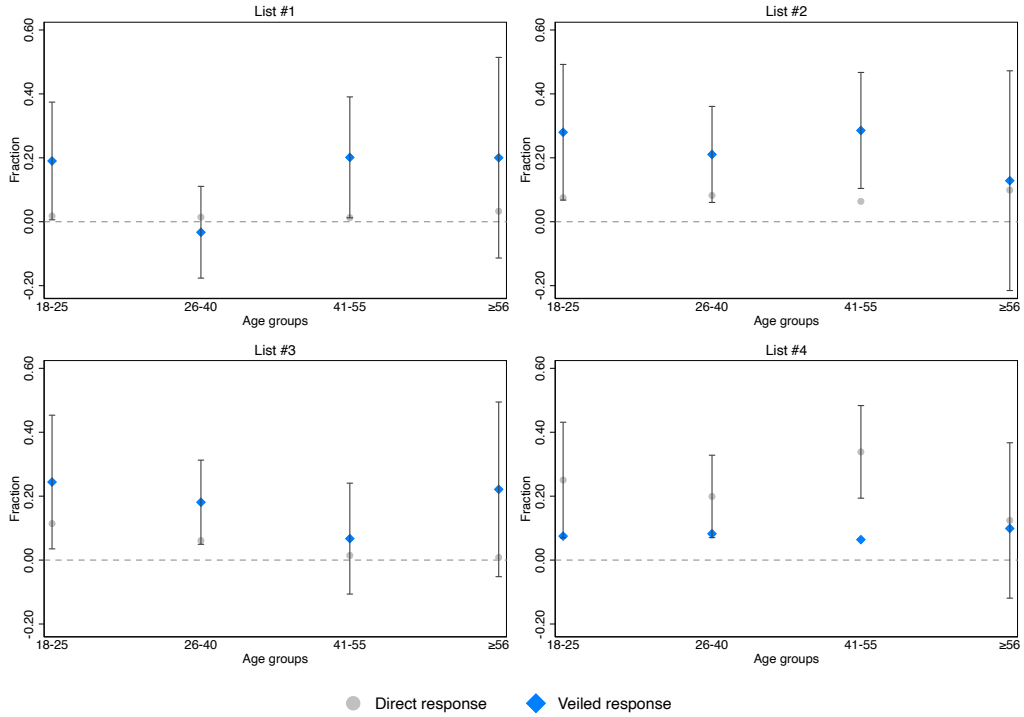
Source: Authors' elaboration from list experiment sample collected between June and July 2022.
Notes: The graphs show the estimated prevalence of the behavior "same sex attraction" according to each list with its respective 95% confidence interval.

Figure A.3. Distribution of number of true/valid items for each list



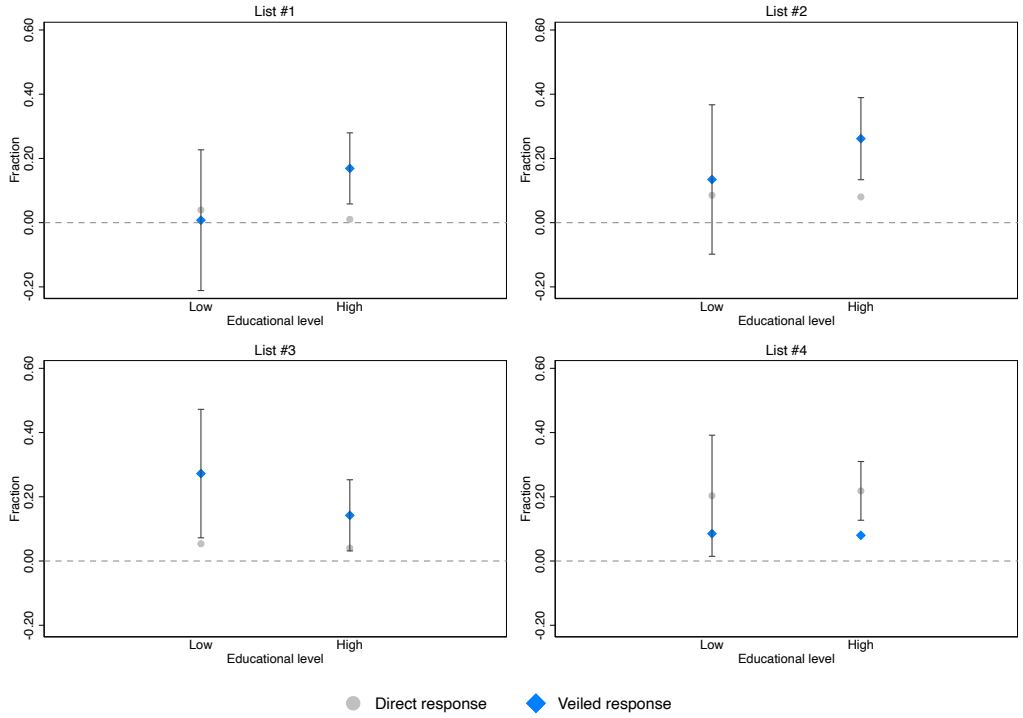
Source: Authors' elaboration from list experiment sample collected between June and July 2022.

Figure A.4. Size of LGBT population by age group



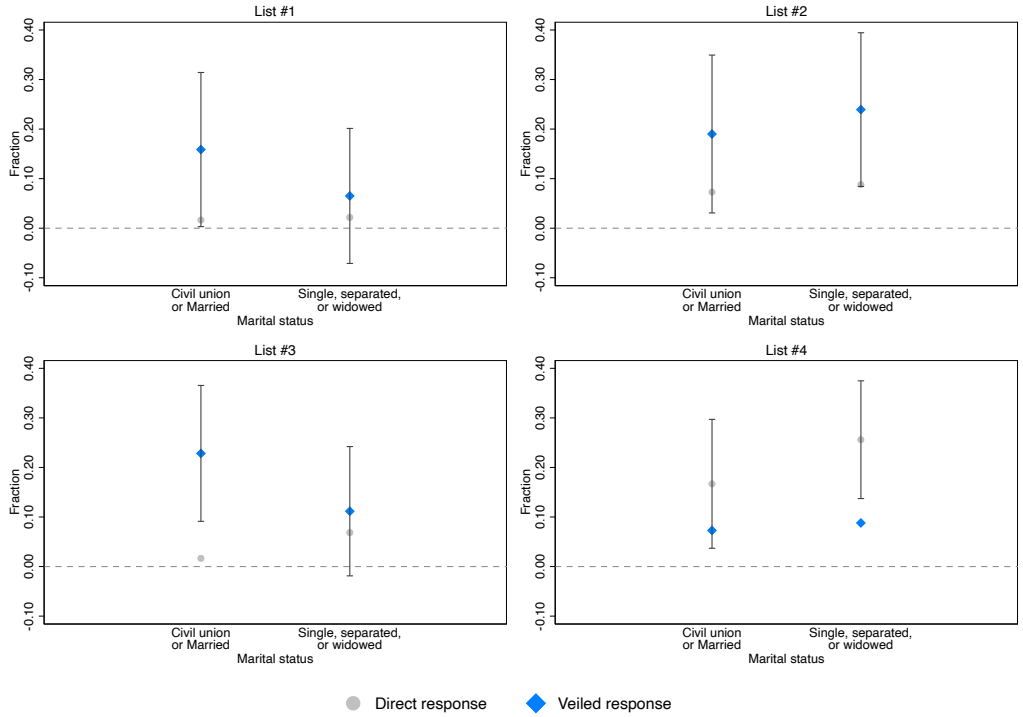
Source: Authors' elaboration from list experiment sample collected between June and July 2022.
Notes: The graphs show the mean size of the LGBT population for the direct response group (in grey) and the veiled response group (in blue) with its respective 95% confidence interval.

Figure A.5. Size of LGBT population by educational level



Source: Authors' elaboration from list experiment sample collected between June and July 2022.
Notes: The graphs show the mean size of the LGBT population for the direct response group (in grey) and the veiled response group (in blue) with its respective 95% confidence interval.

Figure A.6. Size of LGBT population by marital status



Source: Authors' elaboration from list experiment sample collected between June and July 2022.
Notes: The graphs show the mean size of the LGBT population for the direct response group (in grey) and the veiled response group (in blue) with its respective 95% confidence interval.

Appendix B: Sample Instruments (in original Spanish)

B.1. Informed Consent:

Mensaje inicial WhatsApp:
¡Buen día [nombre]! El Centro Nacional de Consultoría agradece su disposición a participar en el **estudio de percepciones en temas de identidad en la ciudad de Bogotá.** Este estudio se compone de un experimento de lista para identificar si las encuestas a hogares miden adecuadamente las características sociodemográficas y la identidad de las personas en diferentes dimensiones. Por favor responda en las próximas 48 horas la encuesta en el siguiente enlace [enlace]. Hacer esto le tomará alrededor de 15 minutos y, así, habrá terminado la primera parte del estudio. Sólo si usted completa las dos partes del estudio, automáticamente recibirá un bono de Helados Popsy a su celular. [Sujeto a cambios según el desempeño de la prueba piloto: bonos del éxito o recargas a celular]

Agradecemos guardar este contacto o responder cualquier palabra a esta conversación para que pueda abrir el enlace directamente desde el mensaje. En caso de que no abra el enlace, puede copiarlo y pegarlo en el navegador (Google).
Mensaje automático de respuesta (WhatsApp):
☝ Este es un mensaje automático.
¿Tiene dudas? Para cualquier inquietud o soporte que requiera, puede comunicarse al Tel.: 3394888 en Bogotá o al correo xxxxx@cnccol.com

Consentimiento informado:
Saludo al primer contacto. Buenos días/tardes, mi nombre es (nombre y apellido), pertenezco al Centro Nacional de Consultoría, una empresa privada dedicada a la investigación de mercados, social y de opinión pública. Actualmente nos encontramos desarrollando un **estudio de percepciones en temas de identidad en la ciudad de Bogotá.** Este estudio se compone de un experimento de lista para identificar si las encuestas a hogares miden adecuadamente las características sociodemográficas y la identidad de las personas en diferentes dimensiones y para hacerlo requerimos contactar a la siguiente persona _____.

Saludo al entrevistado. Buenos días/tardes mi nombre es (nombre y apellido), del Centro Nacional de Consultoría, una empresa privada dedicada a la investigación de mercados, social y de opinión pública. **Actualmente realizamos un estudio de percepciones en temas de identidad en la ciudad de Bogotá.** Este estudio se compone de un experimento de lista para identificar si las encuestas a hogares miden adecuadamente las características sociodemográficas y la identidad de las personas en diferentes dimensiones. Su participación es voluntaria y **su identidad, así como las respuestas suministradas serán de carácter confidencial, por lo que no serán compartidas con ningún individuo u organización fuera del equipo de investigadores. Los datos se utilizarán únicamente para fines estadísticos de investigación**. El beneficio de su participación es hacer valer en el estudio la opinión de personas que piensan como usted.

El Centro Nacional de Consultoría pide su autorización para recolectar y compartir únicamente sus datos personales con fines de contacto para la aplicación de encuestas. Esta solicitud se hace de conformidad con la Ley 1581 de 2012. Nuestra Política de Tratamiento de Datos Personales puede encontrarla en la página www.centronacionaldeconsultoria.com y cualquier consulta puede dirigirla al correo electrónico cnc@cnccol.com o comunicarse con el Tel: 3394888 de Bogotá.

B.2. List Experiment (translated from original Spanish)

2.1 Direct response Questionnaire: Direct questions on gender identity and sexual orientation

Gender Identity
You self-identify as:

1. Male
2. Female
3. Male Trans
4. Female Trans
5. Other

Sexual Orientation
You are sexually or romantically attracted to:

1. Men
2. Women
3. Both Sexes
4. Other

2.2 Direct response Questionnaire: List Experiment Questions

Gender identity

1. If I could decide, I would vote to legalize abortion
2. I spent time as a child playing outdoors
3. The use of contraception should be illegal
4. We all have the right to decide about our body
   How many of the above statements apply/are true for you?

Sexual Orientation (Same sex)

1. I voted YES in the peace plebiscite
2. All drugs or psychoactive substances should be illegal
3. I learned how to ride a bicycle when I was a child
4. If I could decide, I would vote to legalize marihuana
   How many of the above statements apply/are true for you?

Sexual Orientation (Both sexes)

1. I have participated in a protest or social demonstration to support a cause
2. If I could decide, I would vote to eliminate the right to freedom of speech
3. I graduated from high school or at least I completed elementary school
4. I voted in the recent presidential elections on May 29th
   How many of the above statements apply/are true for you?

2.3 Direct response Questionnaire: Compliance test

1. I spent more than eight hours on my phone
2. Young people should receive sex education at school
3. Young people's sexual education is the sole responsibility of their parents
4. Mental, psychological, and emotional health is important at all stages of life

5. I have felt sexual or romantic attraction towards a person of my same sex

2.4 Veiled Report Questionnaire: List Experiment Questions

Gender Identity

1. If I could decide, I would vote to legalize abortion
2. I spent time as a child playing outdoors
3. The use of contraception should be illegal
4. We all have the right to decide about our body
5. I feel identified by a different sex that the one that was assigned to me at birth

Sexual Orientation (Same sex)

1. I voted YES in the peace plebiscite
2. All drugs or psychoactive substances should be illegal
3. I learned how to ride a bicycle when I was a child
4. If I could decide, I would vote to legalize marihuana
5. I have felt sexual or romantic attraction towards a person of my same sex

Sexual Orientation (Both sexes)

1. I have participated in a protest or social demonstration to support a cause
2. If I could decide, I would vote to eliminate the right to freedom of speech
3. I graduated from high school or at least I complement elementary school
4. I voted in the recent presidential elections on May 29th
5. I have felt sexual or romantic attraction towards people from both sexes

2.5 Veiled Report Questionnaire: Compliance test

1. I spent more than eight hours on my phone
2. Young people should receive sex education at school
3. Young people's sexual education is the sole responsibility of their parents
4. Mental, psychological, and emotional health is important at all stages of life