

**The Category is: Gender Diversity in Demographic Research
A beyond the binary perspective on Education, Life Course, and Dating**

Sabina Bercovich Szulmajster

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  <https://creativecommons.org/licenses/?lang=ca>

ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <https://creativecommons.org/licenses/?lang=es>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

The Category is: Gender Diversity in Demographic Research

*A beyond the binary perspective on
Education, Life Course, and Dating*

*Doctoral Thesis Submitted to Obtain the
Degree of Doctor in Demography*

by
Sabina Bercovich Szulmajster

Directors: Diederik Boertien and Fernanda Fortes De Lena
Tutor: Diederik Boertien

PhD in Demography, 2026
Universitat Autònoma de Barcelona
Centre d'Estudis Demogràfics (CED – CERCA)



Abstract

This thesis advances demographic research by adopting a diverse gender perspective, going beyond the binary. The main aim is to provide empirical evidence of the relation between gender diversity and life outcomes, and to question gender conceptualizations commonly used in demographic research. This work shows how both current and innovative data sources and methods can provide new perspectives on life outcomes for increasingly diverse populations.

The first paper demonstrates that measuring gender multidimensionally reveals educational inequalities that remain invisible when using binary classifications, showing how trans non-binary individuals experience distinct penalties in terms of educational attainment. The second paper presents how gender can be studied with a dynamic perspective by applying a life course perspective to examine gender identity milestones, focusing on the events of self-realisation and disclosure, and their interplay with birth cohort. The third paper uses behavioural data from a dating application to investigate how gender identification shapes digital dating experiences, revealing that transgender, non-binary, and gender-diverse users engage more actively in terms of liking, yet receive systematically lower returns in the form of matches.

These studies demonstrate that incorporating gender diversity is not merely about adding categories, but requires rethinking how demographers conceptualise and measure gender. The findings highlight persistent inequalities and differences for gender-diverse populations across educational, life course, and intimate relationship domains, with implications for understanding stratification, well-being, and the role of institutions, from schools to digital platforms, in reproducing or challenging gender-based exclusion.

Keywords: Gender, diversity, inequalities, measurement, life course.

Resumen

La presente tesis aporta al campo de la investigación demográfica incorporando una perspectiva de género amplia que va más allá de las concepciones binarias. El objetivo principal de la misma es proveer evidencia empírica sobre la relación que existe entre la diversidad de género y las distintas experiencias vitales, y cuestionar la forma en la que comúnmente se ha operacionalizado al género en la demografía. Este trabajo muestra cómo trabajar con una variedad de bases de datos, tanto de encuestas como provenientes de huellas digitales, y métodos innovadores puede aportar nuevas perspectivas y resultados sobre las experiencias de vida de poblaciones cada vez más diversas.

El primer artículo demuestra que conceptualizar el género de forma multidimensional revela desigualdades educativas que no serían visibles al utilizar clasificaciones binarias. Este trabajo evidencia que las personas trans no-binarias sufren las penalizaciones más grandes en términos de logros educativos. El segundo artículo presenta una mirada dinámica a la conceptualización del género al tomar una perspectiva del curso de vida. El objetivo es estudiar el desarrollo del género, con el foco puesto en los eventos de auto-percepción y comunicación del género. Estas trayectorias vitales se estudian en conjunto con la cohorte de nacimiento para estudiar la relación con distintos contextos de desarrollo. El tercer capítulo utiliza datos de comportamiento de una aplicación de citas para investigar cómo la identificación de género se relaciona con las experiencias de citas en entornos digitales. Este estudio muestra que las personas trans, no-binarias y de género diverso suelen dar mayor cantidad de *likes*, pero reciben sistemáticamente menores retornos en términos de *matches*.

Estos estudios demuestran en su conjunto que incorporar la diversidad de género a la investigación demográfica no pasa solamente por la adición de categorías, sino que implica repensar cómo conceptualizamos y medimos el género. Los resultados demuestran desigualdades persistentes y diferencias significativas en dominios tan distintos como son la

educación, las experiencias de citas digitales, y el desarrollo del género a lo largo de la vida. Estos resultados tienen implicaciones para la estratificación social, el bienestar, y especialmente el rol de las instituciones, desde las escuelas hasta las plataformas digitales, a la hora de reproducir, o desafiar, las normas de género de nuestras poblaciones.

Palabras clave: Diversidad, género, desigualdades, medición, ciclo de vida.

Resum

Aquesta tesi contribueix al camp de la recerca demogràfica incorporant una àmplia perspectiva de gènere que va més enllà de les concepcions binàries. El seu objectiu principal és aportar proves empíriques sobre la relació entre la diversitat de gènere i diferents experiències vitals, i qüestionar la manera com el gènere s'ha operativitzat habitualment a la demografia. Aquesta tesi mostra com treballar amb una varietat de bases de dades, tant d'enquestes com d'empremtes digitals, i amb mètodes innovadors pot proporcionar noves perspectives i coneixements sobre les experiències vitals de poblacions cada cop més diverses.

El primer article demostra que conceptualitzar el gènere de manera multidimensional revela desigualtats educatives que no serien visibles si es fessin servir classificacions binàries. Aquest treball mostra que les persones transgènere no binàries pateixen les penalitzacions més grans en termes d'assoliment educatiu. El segon article presenta una visió dinàmica de la conceptualització de gènere adoptant una perspectiva de curs de vida. L'objectiu és estudiar el desenvolupament de gènere, amb un enfocament en els esdeveniments de percepció de si mateix i la comunicació de gènere. Aquestes trajectòries vitals s'estudien conjuntament amb la cohort(sin la e) de naixement per examinar la relació amb diferents contextos de desenvolupament. El tercer capítol utilitza dades de comportament d'una aplicació de cites per investigar com la identificació de gènere es relaciona amb les experiències de cites en entorns digitals. Aquest estudi mostra que les persones transgènere, no binàries i de diversitat de gènere tendeixen a fer més likes, però constantment reben menys respostes en forma de matches.

En conjunt, aquests estudis mostren que incorporar la diversitat de gènere a la recerca demogràfica no és només una qüestió d'afegir categories, sinó que implica repensar com conceptualitzem i mesurem el gènere. Els resultats demostren desigualtats persistents i diferències significatives en àmbits tan diversos com l'educació, les experiències de cites

digitals i el desenvolupament de gènere al llarg del cicle vital. Aquests resultats tenen implicacions per a l'estratificació social, el benestar i, especialment, per al paper de les institucions, des de les escoles fins a les plataformes digitals, en la reproducció o el desafiament de les normes de gènere de les nostres poblacions.

Paraules clau: Diversitat, gènere, desigualtats, mesurament, cicle vital.

“No queremos encajar en un mundo que nos excluye; queremos transformarlo.”

-Marlene Wayar

*Dedicated to every soul who has ever felt **queer**,
at any point in their lives.*

Acknowledgments¹

I had the amazing opportunity to conduct my PhD thesis research at CED, where I started working in 2022 as a research assistant to the MINEQ Project. This amazing chance let me develop myself as a researcher, connected me with amazing colleagues, and helped me grow as a person. I am deeply grateful to my director, Diederik Boertien. I couldn't have ever imagined having such a wonderful mentor for this journey. Every meeting was insightful, la forma de acompanyar sempre desde el cuidado, desde lo humano, respetando sobre todas las cosas a mi persona y mi perspectiva. To my academic big sister, Fer, who turned into my supervisor para acompanyarme sobre el final de este camino. Gracias Fer, por todo lo construido juntas, por tu lucha, y por tu mirada. I also want to thank CED, for granting me the DEMOS Scholarship, which funded this PhD.

Probably one of the most fun parts of this journey was, and still is, the amazing MINEQ Family: Bob, Gioia, Anna, Mariel, and all those who came and went and left their mark. A la gent de CED que m'ha ajudat a arribar fins a aquest punt: Inés, Sergio, Teresa, Candi, Eulalia, Alex. To senior researchers who were always there for feedback and questions: Albert, Ariane, Kike, Juan, Maike, Jeroen, Sergi, Clara, and so many others. To my fellow pre-doctoral researchers for every birras de los jueves, la calçotada anual, conferences, and chismes de pasillo. It's been a real pleasure and honor to be a part of this institution for more than 4 years.

To the queer community for existing, for the revolution and for the inspiration. There is a big community LGBTQ+ researchers who challenge the status-quo and dream of a better world for everybody. I also thank all of them for valuable insights, comments, recommendations, and inspiration. To Oxford University, Francesco Rampazzo, The Leverhulme Centre for Demographic Science and the Sociology Department, for having me there for such an

¹ As my mind during this thesis, these *agradecimientos* happen to happen in a mix of Spanish and English, with a hint of Catalan.

Siendo fiel a mi forma de pensar durante esta tesis, los agradecimientos suceden en una mezcla de castellano e inglés, con una pizca de catalán.

incredible opportunity and experience. I felt I was living in a fantasy for a whole term, embodying my hermione and alice in wonderland dreams. To Alison, Solly, and The White Hart for welcoming us in such an amazing way.

Quiero agradecer y reconocer a toda la comunidad académica por seguir creyendo y luchando por la ciencia. A la Universidad de Buenos Aires, a la universidad pública, donde me enamoré de la demografía actuarial, y a mi mentor de grado, Pablo Caviezel. A Fundar, donde pude, por primera vez, usar mi expertise como data scientist para investigar temas de género. Gracias Daniel Yankelevich por ver y confiar en mi potencial. A Ame Botto por el empuje a ir a fondo. A la Universidad Pompeu Fabra y su excelente equipo de profesores por darme la oportunidad de formarme en demografía. A la Universitat Autònoma de Barcelona per ser un entorn tan acollidor per fer-hi el doctorat.

To my parents, Sandra y Diego, who always and still to this day invites me to think critically, to question, to be curious and to explore. Gracias por ser mi apoyo, mis mentores, y los que más festejan mis logros. To my brother, Ulises, for being my life-long companion on this adventure; Tina, my sister in law, a quien tengo el honor de llamar *mi amiga*; y al Ibecho, por ser el más capo del mundo. Al gordo y mis abuelas, por haberme hecho lo que soy. And my amazing, amazing extended family. La que no se elige pero elegiría cada vez, who are always curious about my research and experiences. To all of my cousins, con quienes amo crecer y ver las personas en las que nos vamos transformando.

To Jorge, mi principal aguante. My brilliant, sensitive, beautiful partner. Gracias por cuidarme tan bien, por hacerme reír tanto, por recordarme que no todo es la academia, por ayudarme a cuestionar hasta lo más profundo. Gracias por acompañarme a Oxfordshire y por divertirnos tanto flasheando LOTR while being there, gracias por tanto rock.

A mi banda querida de amigxs en Buenos Aires, who always remind me who I am and make me so so proud of where I come from: Caro, Lai, Mai, Maiu, Sebi, Chon, Javi, Flori, Gonza. Gracias por haberme acompañado en cada una de mis versiones y estar siempre tan cerca.

A mi familia migrante en Barcelona, España y las Europas. A la *xarxa* por acompañarnos en esto del *crecer*. A Juli por ser mi primer hogar en esta ciudad de locos. A Mai, por la hermosa coincidencia de estar, al menos un rato, en el mismo lugar. A Octa, Marian, Flor y Cami, por ser el bastión de hinchada, lucha y salud mental en este equipo gigante que estuvo atrás de esta tesis. A Ursu, Ayu, Nik, Sofi, Mati, que aunque estemos cada quien donde le toca, siempre encontramos un dónde. Gracias por ser casa lejos de casa.

Thank you Oli, Lola, y Monchi por existir. And finally, to the unofficial co-author of this thesis, who is (of course) laying on my lap: Tito. Thank you for being the most loving, playful, little ball of fur.

This work would not have happened without even one of you.

Gracias, desde el fondo de mi ser, por eso.

Sabi

Table of contents

The Category is: Gender Diversity in Demographic Research

Abstract.....	1
Resumen.....	2
Resum.....	4
Acknowledgments.....	7
Table of contents.....	10

Chapter 1: Introduction..... 14

1. 1. Preface.....	15
1. 2. What do we know?.....	18
1. 2. 1. Gender diversity in demography and contemporary populations.....	18
1. 2. 2. Gender measurement and operationalization.....	20
1. 3. Motivations and research questions.....	24
1. 4. Gender diversity, life outcomes, and processes.....	25
1. 4. 1. Gender diversity and educational inequalities.....	26
1. 4. 2. Gender diversity and the life course.....	27
1. 4. 3. Gender diversity, intimacy and digital life.....	28
1. 5. Data and methodological choices.....	29
1. 6. Outline.....	33
1. 7. References.....	34

Chapter 2: Categorizing Gender beyond the Binary: Inequalities in Education from a Multidimensional Gender Perspective..... 40

2. 1. Introduction.....	41
2. 2. Theoretical Background.....	43
2. 2. 1. What do we mean when we say gender?.....	43
2. 2. 2. Existing measures of gender identity.....	45
2. 2. 3. Gender-based Inequalities.....	47
2. 2. 4. Mexican context.....	49
2. 3. Data and Methods.....	50
2. 3. 1. Data Source.....	50
2. 3. 2. Gender Identity Measures.....	51
2. 3. 3. Other Measures and Method.....	55
2. 4. Results.....	55
2. 4. 1. Descriptive Results.....	56
2. 4. 2. First categorization: Gender Trajectory.....	60
2. 4. 3. Second categorization: Binary Dimension.....	62
2. 4. 4. Third categorization: Combined Gender Identification.....	64
2. 5. Conclusions.....	66
2. 6. References.....	69
2. 7. Appendix (Online Supplementary Materials).....	73

Chapter 3: Gender Trajectory Milestones from a Life-Course Perspective: A Comparison across Cohorts.....	77
3. 1. Introduction.....	78
3. 2. Theoretical Background.....	80
3. 2. 1. Gender trajectories and milestones.....	80
3. 2. 2. Contextual factors.....	82
3. 3. Research Design.....	86
3. 3. 1. Data.....	86
3. 3. 2. Methods.....	88
3. 4. Results.....	91
3. 4. 1. Descriptives.....	91
3. 4. 2. Gender trajectory typologies.....	92
3. 4. 3. Cohort distribution across clusters.....	96
3. 4. 4. Additional analysis.....	98
3. 5. Conclusions and Discussion.....	99
3. 6. References.....	104
3. 7. Appendix.....	108
3. 7. 1. Appendix A. Additional Descriptive Characteristics.....	108
3. 7. 2. Appendix B. Country-level variation in cluster representation.....	109
3. 7. 3. Appendix C. Event History Analysis.....	113
 Chapter 4: Beyond the Swipe: Gender Diversity and Interactions in Dating Apps.....	 118
4. 1. Introduction.....	119
4. 2. Theoretical Background.....	121
4. 2. 1. Gender Diversity and Dating.....	121
4. 2. 2. Engagement with Dating Apps.....	122
4. 2. 3. Returns to Engagement in Dating Apps.....	125
4. 2. 4. United Kingdom context.....	127
4. 3. Research Design.....	128
4. 3. 1. Data.....	128
4. 3. 2. Methods.....	129
4. 4. Results.....	131
4. 4. 1. Descriptives.....	131
4. 4. 2. Likes given.....	133
4. 4. 3. Like-to-match ratio.....	136
4. 4. 4. Like-to-match ratio - Coarsened Exact Matching.....	137
4. 5. Conclusions and Discussion.....	139
4. 6. References.....	145
4. 7. Appendix.....	150
4. 7. 1. Appendix A. Ordinary Least Squares (OLS) regression results predicting number of likes given.....	150
4. 7. 2. Appendix B. Geographical variations.....	151

Chapter 5: Conclusions.....	156
5. 1. Overview of Findings.....	159
5. 2. Contributions.....	162
5. 2. 1. What role does gender diversity play in shaping educational and relational outcomes?.....	162
5. 2. 2. If gender diversity relates to people's outcomes in life, what are the relevant gender dimensions that we should consider?.....	164
5. 2. 3. Which data sources and methods help us understand how these lives unfold over time?.....	166
5. 3. Discussion and implications.....	169
5. 4. Final Remarks.....	173
5. 5. References.....	174

Chapter 1: Introduction

1. 1. Preface

Demography is a discipline concerned with understanding how population processes such as mortality, fertility, and migration intersect with individual life courses, including family, educational and labour trajectories, to shape the size and structure of human populations (Girard, 1968). Traditionally, this perspective has treated sex and age as fundamental characteristics for describing and analysing these populations. However, while sex and age have been long used as core demographic variables, the incorporation of gender as an analytical category remained largely absent from demographic research until recent decades (Watkins, 1993; Riley, 1997; Presser, 1997). Even as feminist demographers have worked to centre gender in population studies, this integration has often remained limited to examining differences between women and men, following a cis²-binary understanding of gender (Compton et al., 2018; Baumle & Nordmarken, 2022).

If demographic characteristics are used to understand variation in life experiences and population developments, we should be aware that these are shaped not just by sex but also by gender. If gender categories that differ from binary sex are related to life courses, we need to understand how to measure gender and which aspects of gender are related to relevant outcomes in life. Current increases in gender identification at the population level (Lagos 2022), advances, and backlashes, in visibility, public discourse, and legal frameworks (ILGA World, 2024) justify the increasing relevance of including gender variables in a broader perspective to study evolving populations.

The overarching goal of this thesis is to demonstrate how demographic research can meaningfully incorporate gender diversity beyond the man/woman and cis/trans³ binaries, and to examine if new insights emerge when we do so. In this thesis, I use the term gender diversity to refer to this broader set of self definitions, including both non-binary and binary

² Cis(gender) refers to individuals whose gender identity aligns with the sex assigned to them at birth.

³ 'Trans' is used here as an umbrella term for individuals whose gender identity differs from the sex assigned at birth, including transgender men and women, non-binary people and other gender-diverse identities (adapted from ILGA-Europe, 2020).

identified people whose trajectories fall outside cisnormative⁴ expectations. In many social science disciplines the relevance of gender diversity is already well established, particularly in qualitative and ethnographic research. However, gender diversity remains largely invisible in social and demographic scholarship, despite growing recognition in legal frameworks and in population-level data collection efforts (Baumle & Nordmarken, 2022; Compton et al., 2018; Lagos, 2022). This invisibility is not just a matter of omission, as it reflects deeper epistemological assumptions about how demographers conceptualise and measure populations, which life course processes we examine, and whose experiences count as demographically significant.

This thesis addresses these gaps by incorporating gender diversity into several core demographic topics. Rather than treating gender diversity as an exception, it takes it as a central part of population processes and asks what becomes visible when demographic research is designed with this in mind. To incorporate gender diversity into demographic research, several questions arise which are central to this thesis.

First of all, the consideration of gender diversity in a broader way could shed light on meaningful variation in people's life courses. In other words, how are outcomes across the life course shaped by experiences that depart from, or challenge, dominant gender norms? In this thesis, we answer this question by looking at how two essential spheres of life are structured by gender: education and dating. These outcomes are central to two of the empirical chapters of this thesis.

Secondly, if gender diversity relates to people's outcomes in life, what are the relevant gender dimensions that we should consider? Existing research has used a variety of measures that capture different aspects of gender, with an increasing number of surveys including separate questions for sex and gender, and various options for the latter (

⁴ Cisnormativity refers to the cultural assumption that all people are cisgender: that their gender identity aligns with the sex assigned to them at birth, and that this alignment is natural, normal, and preferable. This ideology positions cisgender identities as the default, rendering transgender and gender-diverse experiences as deviations requiring explanation or justification. (Schilt & Westbrook, 2009)

Westbrook & Saperstein, 2015). Therefore, gender is usually measured and analyzed as a single variable, and crosstabled with sex to identify gender-diverse individuals. However, gender experiences tend to be multidimensional, with multiple axes relating to life experiences and opportunities, such as identification, binary relation to gender, expression, legal and medical affirmation steps, among many others. In addition, sex and gender have often been treated as characteristics that are stable across people's life courses, but little attention has been paid to how gender identity can change over people's lives. Such measurement issues and gender trajectories are central to two empirical chapters of this thesis. One chapter will consider the multi-dimensional nature of gender, whereas another chapter will look at gender milestones: when people's realization and disclosure of gender identities occur in people's lives.

Finally, what are the data sources and methods that can be used to study gender diversity? Measures of gender identity have only recently become available in representative data sources, and such representative sources often have limited information on processes relevant for people whose gender falls outside of normative categories. An example is formed by the gender milestones, which are rarely included in representative surveys, and are studied in one of the empirical chapters of the thesis by leveraging a specialized survey at the European level. Another question is to what extent people disclose their gender identity to official surveys, and if we are able to capture gender diversity in current surveys. Based on this, this thesis explores the opportunities that new data sources, such as big data and digital traces, can provide to study these groups in society. The use of innovative data sources is central to one empirical chapter of the thesis, which uses behavioural data on dating practices among trans, non-binary, and gender diverse individuals.

I will start by giving a brief overview of the incorporation of gender diversity into demographic research, including a review of studies that have documented the prevalence of gender identities that go beyond cisgender and binary identities. After this section, the motivation and research questions for this thesis are presented. Subsequently I will briefly discuss what

existing research has found about the relevance of gender diversity for socioeconomic outcomes in life, dating, and an aspect of life that has set apart the life courses of people whose gender falls outside of normative categories: gender milestones throughout people's life courses. Thirdly, I will present the data and methodological choices made in the thesis in relation to existing approaches, and will end the introduction by giving a brief outline of the different empirical chapters of the thesis.

1. 2. What do we know?

1. 2. 1. Gender diversity in demography and contemporary populations

In the past twenty years, the increasing visibility and legalization of *queer* practices, such as same-sex marriage and gender identity recognition, is taking shape in many countries (ILGA world, 2023). However, the term *queer* is not limited to same-sex partnership or non-cis gender identities' recognition: it extends to any lived experience that challenges normative practices, such as polyamorous relationships and rainbow families.

The number of studies aimed at understanding the LGBTQ+ population is scarce but growing. Most of the research that studies these individuals focuses on the discrimination and inequalities suffered by this collective (Valfort, 2017; GATE, 2013; Badgett et al., 2021), and there is an increasing amount of demographic work on the characteristics and life courses of lesbian, gay and bisexual people, often through the study of sexual identity and same-sex couples in population-based data (Baumle & Compton, 2018). In this sense, the demography of sexual diversity has advanced further than the demography of gender diversity. We know much more about the composition and inequalities affecting LGB populations than about those whose gender identity does not align with the one assigned at birth. Systematic population studies of trans, non-binary, gender-diverse and other non-cisgender individuals are still relatively few, and most of them focus on the US context (Baumle & Nordmarken, 2022).

The need for a demographic re-evaluation is justified by a decades-long increase in the quantity and relative proportion of people whose gender identity differs from the sex they were assigned at birth. This trend, often described as a "transgender tipping point" (Lagos, 2022), has moved from being estimated through convenience samples to being quantified in nationally representative, population-level instruments. For instance, early population-level estimates in Europe already pointed to a notable presence of transgender and non-binary people, such as surveys in the Netherlands in 2014 reporting a higher prevalence of gender-diverse identification in younger cohorts (Kuyper & Wijsen, 2014). More recently, the 2021 Census in England and Wales recorded that 0.5% of the adult population identifies as trans, amounting to nearly 280,000 people, while Canada's 2021 Census found that 0.33% of those aged 15 and older (over 100,000 individuals) identify as transgender or non-binary (Statistics Canada, 2022). In Latin America, the 2022 Census in Argentina was the first in the region to include a gender identity question (INDEC, 2019), revealing that 1.4% of the population aged 14 and older identifies with a gender identity other than woman or man, representing over half a million people. A year later, Uruguay's 2023 Census reported that 0.6% of its population identifies as trans or non-binary.

These census-based estimates are complemented by administrative sources in Europe, where gender diversity prevalence is studied through the number of individuals who underwent medical and/or legal transitions. In Sweden, following administrative data, Kolk et al. (2025) reported that only 0.083% of Sweden's older than 18 population had a gender incongruence medical diagnosis, while only 0.034% had their legal gender rectified. These percentages are much lower than survey-based estimates, potentially related to desires and barriers to undergo medical and/or legal affirmation steps.

The sharp increase in non-cis identification among younger age groups further justifies the need to incorporate gender diversity into demographic analysis. While it is not yet possible to determine whether these differences reflect cohort effects, period effects, or a combination of both, national counts consistently show much higher rates of non-cis identification among

younger respondents. In Canada, for example, the proportion in Generation Z (0.79%) is up to seven times higher than in older cohorts. A similar pattern appears in the UK, where 1.0% of those aged 16 to 24 reported a non-cisgender identity, doubling the national average. In the United States, Lagos (2022) finds that people born after 1984 are substantially more likely to identify as transgender or gender non-conforming than those in earlier birth cohorts. Regardless of the underlying mechanisms, these age-graded differences reveal a major shift in how gender is understood and lived, and highlight the urgency of developing demographic tools capable of capturing this diversity.

In 2022, Baumle and Nordmarken published *“Demography of Transgender, Nonbinary and Gender Minority Populations”* (2022), compiling evidence on health, labour market outcomes, workplace discrimination and poverty among transgender, non-binary and gender-diverse people. Yet even in this extensive volume, several central demographic domains remain largely unstudied: migration and spatial distribution, family structure and childbearing, and disease and mortality. Recent work has noted that these gaps are not simply the result of research priorities but reflect broader limitations in demographic data infrastructures, which are still not well suited to identifying transgender, non-binary and gender-diverse people in population surveys (Fernández-Carro, 2025). Large-scale surveys often rely on binary sex or gender markers or include gender identity measures inconsistently across countries and time (Westbrook & Saperstein, 2015). Without sufficiently inclusive and stable categories, it becomes difficult to estimate population sizes, study inequalities across domains or analyse how gender identity intersects with key demographic processes.

1. 2. 2. Gender measurement and operationalization

Not only has the number of people who identify in ways that diverge from their sex assigned at birth increased, but the terms used to name gender experiences has also expanded. Survey and community-based studies document the growing use of terms such as

trans/transgender, non-binary, genderqueer, gender non-conforming, and gender fluid (Lagos, 2022; Richards et al., 2016; Westbrook & Schilt, 2014). These labels capture different relationships between gender identity and sex assigned at birth, as well as distinct ways of positioning oneself in relation to the gender binary.

Non-binary identification is especially relevant here, since it questions the assumption that all gender identities must be organised around the categories of man and woman and instead points to identities that combine, move between, or stand outside these categories. In this sense, there are some non-binary identifying individuals who also identify as trans, or who are included in the trans umbrella, but there are other non-binary individuals who do not (Darwin, 2020). This implies a multidimensional perspective to gender identification, as being cis, trans, or something else represents one part of the relation to gender, and identifying in a binary (man or woman) or non-binary (man and woman, none of them, or a fluidity) represents a different axis.

Despite these challenges in terms of definitions, many demographic studies have been including gender identification in their inquiries. To do so, a variety of data sources have been used, ranging from survey data to administrative, community-based convenience samples to census data, and, more recently, digital trace and big data. However, most large-scale surveys and demographic studies can only capture gender complexities imperfectly, since they rely on fixed binary response categories, with “select one only” options. Many demographic studies still collapse diverse identities into a small set of groups such as “cisgender” and “transgender”, or by grouping cis men, trans men, cis women, and trans women (Carpenter et al., 2020; Carpenter et al., 2024; Kolk et al., 2025; Wilkinson et al., 2021)

This is pretty much consequential of the ways gender has been measured and operationalized in recent surveys and censuses. Historically, surveys and censuses used only one question related to sex, and recently the two-step method got traction, considering

separate questions for sex and gender identification. In this measurement practice, there is a first question related to sex assigned at birth, and a second question related to gender. The range of answers to the gender question is quite diverse, with some surveys only including man, woman, something else; while others include a longer list of available categories, such as EU FRA LGBTQ+ Survey, where six different gender options are provided. A comparison between the answer to sex assigned at birth and gender responses is conducted, and those individuals whose gender differs from the reported sex are coded as trans, gender diverse, or in a similar way. By leveraging the new data generated by the two-step method, some of the studies presented above (Wilkinson et al., 2021; Carpenter et al., 2020) could study gender diversity in more detail.

The two-step measurement method has been criticized by activism and gender expansive individuals (Guyan, 2022), as it, for example, does not allow for multiselection, rendering gender as a unidimensional concept. As mentioned before, for example, there are many non-binary individuals who also identify as trans. Another problem is that this way of asking does not rely on the right to forgetfulness, inquiring about information that might be sensitive such as sex assigned at birth.

Some studies have relied on administrative data to identify gender diverse populations (Kolk et al., 2025). This means that individuals are identified through administrative records related to legal gender changes, gender dysphoria diagnoses, and medical procedures. Administrative data provides valuable information about the individuals once included in the sample, but the selection bias might be high when seeking to compare these samples to the gender diverse population as a whole.

Legal gender recognition and medical affirmation are not processes that most gender-diverse people undergo, whether due to lack of access, bureaucratic barriers, lengthy procedures or simply because they do not wish to do so. Evidence from the FRA LGBTI Survey II shows that only 14.85% of trans and non-binary respondents had changed

their legal gender. This proportion varies notably across groups: just 1.78% of non-binary people reported legal gender recognition, compared to 34.97% of trans women. Among those who had not changed their legal gender, nearly one third reported that they did not want to, while others cited not fulfilling legal requirements. These patterns indicate that administrative datasets primarily capture a selective subset of gender-diverse individuals, limiting their suitability for population-level comparisons or for studying gender diversity more broadly.

Based on these arguments, it is notable that a single data source can provide valuable information about gender diverse populations. However, all of these sources present different biases in terms of selection and identification of the population. Considering who is included, and excluded, from our samples is a key part of gender diversity research in demography. This limits the possibility of making comparisons across gender groups and constrains how previous research has represented the plurality of gendered lives.

In this sense, this thesis considers multiple data sources and measurement methods, along with a critical reflection on who is included in each of the studies. This thesis adopts a data strategy informed by queer theory, data feminism, and queer data scholarship, which collectively highlight that quantitative data sources are not neutral containers of information but products of political, institutional, and epistemic choices (Butler, 1990; Compton et al., 2018; D'Ignazio & Klein, 2020).

Queer theory provides the understanding of gender as a lived experience, constituted through norms, culture, repeated practices and self-construction. Butler (1990) criticizes the binary understanding of gender as male or woman, and argues that this gender conceptualization is highly related to power dynamics and hierarchies in our societies. Data feminism deepens further on how power operates through data infrastructures: who is counted, what gets asked, who is considered as the “reference” category. D'Ignazio & Klein (2020) call for a more plural, situated, and accountable data world. Queer data scholarship

relates these two critiques by relating how gender concepts, understood and lived with a gender expansive light, can hardly be translated to quantitative and technical methods. The way gender is retrieved, Guyan (2022) argues, changes the way the individuals opt to respond, choosing strategically when and what to disclose. Category boundaries, defaults, forced responses, and the handling of “other” or missingness are different arguments that provide complexity to gender measurement.

Studying gender diversity through demographic lenses then requires engaging critically with the conditions under which gender becomes measurable. Measurement is not a neutral practice. It determines who is counted, whose experiences become visible and which forms of gendered life can be analysed. They also remind us that data collection infrastructures privilege some experiences over others and that methodological decisions influence the conclusions researchers can draw. These frameworks guide the data and methodological decisions in this thesis, which will be explained after presenting the motivations and goals of this work.

1. 3. Motivations and research questions

Although surveys and censuses increasingly include questions on gender identity, demographic research still lacks solid ways of studying the lives of transgender, non-binary, and gender-diverse people. Much of what is currently known relies on simplistic categories, very small samples, or approaches that treat gender as a fixed attribute rather than something that develops and changes over time. As discussed in the previous sections, these limitations restrict our ability to examine inequalities, life-course patterns, and dynamics across diverse gender groups. They also mean that many of the processes that shape gendered lives remain analytically invisible in demographic work.

This thesis takes these gaps as its starting point. Rather than assuming that existing gender dimensions and demographic tools are sufficient, it approaches gender as multidimensional and processual, and uses different kinds of data to observe different aspects of gendered

experiences. The goal is not to propose a single framework, but to show how demographic research can adapt when categories, measures and methods are aligned with the complexity of the populations being studied. By combining population surveys, community-based data and behavioural digital traces, the thesis explores what becomes visible when gender diversity is considered a central part of demographic analysis.

These motivations lead to the central questions of this thesis: First, how does gender diversity relate to life outcomes? How are educational and dating outcomes shaped by experiences that depart from, or challenge, dominant gender norms? Two different chapters of this thesis will be working on these questions. Second, if gender diversity relates to people's outcomes in life, what are the relevant gender dimensions that we should consider? How can demographic research better incorporate gender diversity to study them? Questioning current categorization methods and going beyond two-steps in one chapter, and focusing on gender as dynamic in another chapter, will seek to address these questions. Finally, which data sources and methodological approaches are most suitable for studying gender diversity in demographic research? A curated strategy that combines survey data, convenience samples, and digital trace data, with methods tailored to the research questions and data availability is present throughout the thesis. These main questions guide the empirical chapters and shape the data and methodological choices discussed in the following sections.

1. 4. Gender diversity, life outcomes, and processes

The first question of this thesis is to understand how gender diversity relates to outcomes in life, and, to do so, I will be focusing on three different aspects: education, dating, and gender development. Both relationships and education have been long proved to have an impact on well-being, life opportunities, and outcomes. Therefore, including a beyond the binary perspective to better understand these outcomes and processes is very much valuable. In this section I will be discussing previous research and specific motivations of each aspect.

1. 4. 1. Gender diversity and educational inequalities

Research on the socioeconomic (SES) outcomes of transgender, non-binary, and other gender-diverse populations has grown but remains unevenly distributed across disciplines and contexts. Existing studies consistently document disadvantages in key SES domains, including lower employment rates, higher poverty and food insecurity, and barriers in accessing stable labour market positions (Carpenter et al., 2022; Badgett et al., 2021; Badgett et al., 2024). Evidence from the United States shows that transgender adults are less likely to be employed and more likely to experience economic hardship than cisgender adults (Carpenter et al., 2020), while research in Europe highlights similar patterns of material deprivation and exclusion (FRA, 2020).

Findings on educational attainment, however, are more mixed. Some studies report substantial penalties for transgender and gender diverse people (Carpenter et al., 2020; Wilkinson et al., 2021), whereas others, including register-based analyses in Sweden, find little difference between transgender and cisgender adults (Kolk et al., 2025). A possible reason for this inconsistency lies in the way gender has been operationalised, and to who is included in the comparisons due to comparing surveys with administrative based studies. Many population surveys still collapse gender-diverse people into a single “transgender” category, treat non-binary identities as residual or marginal groups, or rely solely on binary gender categories. As several scholars have argued (Compton et al., 2018; Westbrook & Saperstein, 2015), these practices obscure meaningful differences within gender-diverse populations and restrict the types of inequalities that can be detected.

A recent study by Mittleman (2022) on educational inequalities based mainly on sexual identification presents a method to understand how individuals’ positioning to gender norms relates to gender, sexual identification, and educational outcomes. This multidimensional perspective introduces a “gender predictive” approach based on expected gender behaviour for men and women. This prediction positions the individual in a “masculinity/femininity” spectrum and compliments the result with stated gender and sexual identification. By

including a multidimensional perspective, advantages previously found for women and disadvantages for men are not as general as thought. The further away from gender norms an individual was, in terms of predicted gender and sexual identity, the more disadvantage, for women, and advantage, for men, was proved. This study provides very clear evidence of the role of gender norms on educational outcomes, specially for gender diverse individuals.

These ways of operationalizing gender have direct consequences for the study of educational inequality. Without gender measures that reflect the diversity of self-identifications and the relationship between gender identity and sex assigned at birth, it becomes difficult to determine where educational disadvantages emerge and for whom. In this thesis, I draw on this debate by using a multidimensional conceptualisation of gender that distinguishes between binary and non-binary identification and accounts for the relationship between gender identity and sex assigned at birth. This framework allows for a more nuanced understanding of educational inequalities and helps clarify why previous findings may diverge depending on the categories used.

1. 4. 2. Gender diversity and the life course

Recognising the diversity of gender identities also requires recognising that gender is not only a matter of labels, but a process that develops over time. For individuals whose gender does not align with the sex assigned at birth, this process typically involves a sequence of transitions shaped by when, how and in what context they unfold. A life course perspective is therefore particularly valuable, because it shifts attention from fixed categories to the timing and sequencing of key milestones such as self-realisation, disclosure and, when desired or possible, medical or legal affirmation (Elder et al., 2003; Wilkinson, 2018).

Existing research suggests that while the age at which people first realise that their gender differs from their assigned sex has remained relatively stable across cohorts (Wilkinson et al., 2018; Puckett et al., 2022; Scandurra et al., 2021), the pacing of external steps has changed substantially. Younger generations disclose their identity and access

gender-affirming pathways earlier than older cohorts, reflecting changes in social acceptance, visibility and access to information (Baams, 2025; Lagos, 2022; Nordmarken, 2023). These differences illustrate how gender development is embedded in historical context and shaped by broader social environments.

At the same time, they expose a limitation of demographic studies that rely on static classifications or treat gender as a single point-in-time attribute. Such approaches struggle to capture the temporal dimension of identity development and the heterogeneity of pathways individuals follow. Incorporating a life course lens makes it possible to examine gender diversity as a set of trajectories rather than discrete categories, and to analyse how people navigate internal and external transitions across different generations and social contexts.

1. 4. 3. Gender diversity, intimacy and digital life

Differences in family formation and intimate relationships represent another area where gender-diverse populations encounter distinctive challenges. Survey data consistently show that transgender, non-binary and gender-nonconforming people are substantially less likely to be partnered than cisgender adults (Badgett et al., 2024), a pattern linked to discrimination, stigma and concerns about safety in dating and relationship contexts (Blair & Hoskin, 2019; Webster et al., 2025). Online dating has become a crucial route to partnership formation for LGBTQ+ populations (Rosenfeld & Thomas, 2012; Sironi & Kashyap, 2022), yet research also documents widespread exclusion of transgender users, who are frequently rejected by potential partners and experience high levels of harassment and fetishisation. Only a small minority of cisgender individuals report openness to dating a transgender person (Blair et al., 2018), and this exclusion affects interactions both offline and online.

Despite the centrality of digital platforms in contemporary intimate life, quantitative evidence on the dating experiences of gender-diverse people remains scarce. Most studies rely on self-reported preferences or experiences, making it difficult to assess how interactions unfold in practice or how platform design shapes opportunities for connection. As a result, little is

known about the actual behavioural dynamics that structure online dating for transgender, non-binary and gender-diverse users. Addressing this gap requires data that can observe interactions directly, rather than relying solely on stated preferences or retrospective accounts.

1. 5. Data and methodological choices

Which dimensions of gender should we be incorporating to demographic research? What data sources can be used to do so? To answer the questions of how to better study gender diversity in demographic research, this thesis incorporates a variety of data sources, as well as research designs. Considering that data sources present both advantages and limitations, the thesis employs three complementary data sources (summarised in Table 1.1) that allow gender diversity to be studied from multiple angles and dimensions: a nationally representative survey (ENDISEG), a large community-based European survey (FRA LGBTQ+ Survey II, 2021) and digital trace data from a dating app (Zoe). Each source illuminates different dimensions of gendered experience. Population surveys allow the examination of structural inequalities with robust sample sizes; community-based datasets provide detailed information about lived experiences that are not present in general population surveys; and digital trace data make it possible to study relational behaviour as it unfolds in practice rather than through self-reports.

The first article focuses on gender categorisation and uses a nationally representative survey conducted in Mexico, with the aim of identifying sexual and gender diversity (ENDISEG), which includes multiple gender identity questions and a follow-up item for respondents whose identity differs from their sex assigned at birth. This structure enables a multidimensional operationalisation of gender, in line with critiques of how conventional measurement practices erase gender-diverse experiences (Compton et al., 2018; Westbrook & Saperstein, 2015). A work by Muñoz et al. (2024) used this data source to study

socioeconomic disparities by sexual orientation and gender identity, proving the relevance of the sample.

The second article challenges gender as a fixed attribute by applying a dynamic life course perspective to gender identity development. It uses the FRA LGBTQ+ Survey II, where specific questions related to gender milestones were included, thanks to the specific approach of the survey. This database was collected through common efforts between FRA and community organizations, leading to a large sample of individuals with trans, non-binary and gender diverse life experiences. This survey has been used to study sexual identity development (Caprinali & Vitali, 2025), justifying the value of this source to study gender development.

The third article turns to intimate life and relationship formation using behavioural digital trace data from Zoe. These data allow the study of online engagement and matching as they unfold in practice, rather than through self-reported preferences. This big data source has been previously used for demographic research on sexual identity (Rampazzo et al., 2025). Following scholarship that examines how platforms and algorithms shape visibility, recognition and exclusion (D'Ignazio & Klein, 2020; Noble, 2018; Seaver, 2019), the analysis considers not only differences in behaviour but also how platform design may structure the opportunities available to gender-diverse users.

The methodological approach follows the same logic of complementarity. To examine how measurement choices shape estimates of inequality, the thesis uses survey data that include detailed gender identity questions and allow for detailed categorisation strategies. To study gender identity development as a temporal process, it employs sequence analysis and inductive clustering methods that capture the timing and ordering of key milestones (Aisenbrey & Fasang, 2010; Elder et al., 2003). To analyse intimate life and digital interactions, it draws on behavioural records and engages with scholarship on how platforms and algorithms shape visibility, recognition and exclusion (Noble, 2018; Seaver, 2019).

Rather than organising the analysis around a single technique, the thesis aligns methods with the specific conceptual and empirical demands of each research question.

Taken together, these choices make it possible to study gender as multidimensional, dynamic and embedded in life trajectories. They also respond to the concerns raised by queer and feminist data perspectives by demonstrating how demographic tools, whether classification choices, sequence methods, or behavioural models, can be adapted to better capture the complexity of gender diverse lives, and some of the data sources that we could be using to do so.

Table 1.1 Overview of the datasets used in the thesis

Dataset	Total sample size	Analytical sample used in this thesis	Type of dataset	Geographical context	Key variables under study
ENDISEG – National Survey on Sexual and Gender Diversity (2021–2022)	~44,000 individuals aged 15+	341 non-cis respondents aged 18+, plus 40881 cis respondents	Nationally representative survey	Mexico	Gender identity, follow-up identity question when differing from SAAB, sex assigned at birth, education, sociodemographics
EU LGBTQ+ Survey II – FRA (2019)	~139,000 LGBTQ+ respondents	7,621 respondents with gender trajectory information	Large-scale specific survey (non-probability, EU-wide)	European Union + UK	Gender identity, age of self-realisation, age of disclosure, medical affirmation steps
Zoe – Digital trace data (2023–2025)	~4 million accounts worldwide	34,139 active users in the UK who gave ≥1 like	Digital trace data (behavioural data, dating app)	United Kingdom	Self-reported gender identity, likes given, matches received, profile characteristics

1. 6. Outline

This thesis is organised around three studies that together examine gender diversity from a demographic perspective. Each article addresses a different dimension of gender as a multidimensional, dynamic, and biographically embedded aspect of population life, and each relies on distinct data sources that allow for new analytical possibilities.

The first article examines how different ways of categorising gender shape estimates of educational inequality. Using the nationally representative ENDISEG 2021 survey from Mexico, which includes detailed gender identity questions and follow-up items for respondents whose identity differs from their sex assigned at birth, this chapter compares a range of classification strategies and assesses their impact on observed group differences. By drawing on critiques of conventional measurement practices (Compton et al., 2018; Westbrook & Saperstein, 2015), it shows that multidimensional approaches reveal inequalities that remain hidden in binary or cis/trans frameworks.

The second study adopts a life course perspective to analyse gender identity development. Using the FRA LGBTQ+ Survey II, which includes a large number of transgender, non-binary and gender-diverse respondents, it examines the timing and sequencing of key milestones such as self-realisation, disclosure and gender-affirming steps. Applying sequence analysis and inductive clustering methods (Aisenbrey & Fasang, 2010; Elder et al., 2003), this chapter identifies cohort differences in the pacing and structure of these trajectories and illustrates how gender can be studied demographically as a temporal process rather than a fixed characteristic.

The final article examines intimate life and relationship formation using behavioural digital trace data from the LGBTQ+ dating app Zoe in the United Kingdom. By analysing engagement and matching as they unfold in practice, the chapter provides quantitative evidence of relational inequalities affecting transgender, non-binary and gender-diverse users. Drawing on scholarship on algorithmic visibility and exclusion (Noble, 2018; Seaver,

2019), it also considers how platform design may shape the opportunities available to different gender groups within digital dating environments.

In this thesis, life outcomes such as education and relationships are studied with a broader than binary perspective, showing how life courses are linked to gender diversity. Chapter 2 reveals that trans non-binary individuals face the steepest educational penalties, a finding that only becomes visible through multidimensional measurement. Chapter 3 demonstrates that younger cohorts compress the timeline between realization and disclosure, reflecting changing social contexts. Chapter 4 documents how gender-diverse users engage more actively in dating apps yet receive systematically lower returns, even within LGBTQ+ spaces. The different gender dimensions broaden the demographic toolkit and offer a more grounded understanding of how social and structural forces shape the experiences of transgender, non-binary and gender-diverse people. Together, the three articles show how demographic research can better account for gender diversity by using gender measures that are multidimensional, by analysing identity development as a process unfolding over time and by drawing on data sources that capture aspects of life not visible in traditional surveys.

1. 7. References

Aisenbrey, S., & Fasang, A. E. (2010). New life for old ideas: The “second wave” of sequence analysis bringing the “course” back into the life course. *Sociological Methods & Research*, 38(3), 420–462.

Baams, L., Caprinali, A., & Fischer, M. (2025). Navigating time: A systematic review of the timing of sexual and gender identity milestones. In E. D. Rothblum (Ed.), *Sexually and gender diverse adolescents* (pp. 145–163). Amsterdam University Press.

Badgett, M., Carpenter, C., & Sansone, D. (2021). LGBTQ economics. *Journal of Economic Perspectives*, 35(2), 141–170. <https://doi.org/10.1257/jep.35.2.141>

Badgett, M. L., Carpenter, C. S., Lee, M. J., & Sansone, D. (2024). A review of the economics of sexual orientation and gender identity. *Journal of Economic Literature*, 62(3), 948–994.

Baumle, A. K., Compton, D. L., & Poston Jr, D. L. (2009). Same-sex partners: The social demography of sexual orientation. State University of New York Press.

Baumle, A. K., & Nordmarken, S. (Eds.). (2022). *Demography of transgender, nonbinary and gender minority populations*. Springer.

Blair, K. L., & Hoskin, R. A. (2019). Transgender exclusion from the world of dating: Patterns of acceptance and rejection of hypothetical trans dating partners as a function of sexual and gender identity. *Journal of Social and Personal Relationships*, 36(7), 2074–2095.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Caprinali, A., & Vitali, A. (2025). The demography of sexual identity development and disclosure among LGB people in Europe. *Demographic Research*, 52, 125–140.

Carpenter, C. S., Eppink, S. T., & Gonzales, G. (2020). Transgender status, gender identity, and socioeconomic outcomes in the United States. *ILR Review*, 73(3), 573–599. <https://doi.org/10.1177/0019793920902776>

Carpenter, C. S., Lee, M. J., & Nettuno, L. (2022). *Economic outcomes for transgender people and other gender minorities in the United States: First estimates from a nationally representative sample* (IZA Discussion Paper No. 15116). Institute of Labor Economics (IZA).

Carpenter, C. S., Goodman, L., & Lee, M. J. (2024). *Transgender earnings gaps in the United States: Evidence from administrative data* (NBER Working Paper No. w32691). National Bureau of Economic Research.

Compton, D. R., Meadow, T., & Schilt, K. (Eds.). (2018). *Other, please specify: Queer methods in sociology*. University of California Press.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Darwin, H. (2020). Challenging the cisgender/transgender binary: Nonbinary people and the transgender label. *Gender & Society*, 34(3), 357–380.
<https://doi.org/10.1177/0891243220912256>

Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 3–19). Springer.

European Union Agency for Fundamental Rights (FRA). (2020). *A long way to go for LGBTI equality*.

European Union Agency for Fundamental Rights (FRA). (2021). *The EU LGBTI II Survey, 2019* (GESIS Data Archive, ZA7604). <https://doi.org/10.4232/1.13733>

Fernandez-Carro, C. (2025). La diversidad sexo-genérica en demografía: una historia de inercias, vacíos y rupturas. *Papeles de identidad: Contar la investigación de frontera*, (2), 3.

GATE. (2013). *Identidad de género y derechos humanos*. GATE.

Girard, A. (1968). *Principles of Demography*.

Guyan, K. (2022). *Queer data*.

ILGA World. (2023). Legal gender recognition: Data and resources on laws, policies and practices. ILGA World. <https://database.ilga.org/legal-gender-recognition>

INDEC. (2019). *Nuevas realidades, nuevas demandas: Desafíos para la medición de la identidad de género en el Censo de Población*.

INEGI. (2021). *Encuesta Nacional sobre Diversidad Sexual y de Género (ENDISEG)*.
<https://www.inegi.org.mx/programas/endiseg/2021/>

Kolk, M., Tilley, J. L., von Essen, E., Moberg, Y., & Burn, I. (2025). The demography of Sweden's transgender population: A research note on patterns, changes, and sociodemographics. *Demography*, 62(2), 349–363.

Kuyper, L., & Wijsen, C. (2014). Gender identities and gender dysphoria in the Netherlands. *Archives of sexual behavior*, 43(2), 377-385.

Lagos, D. (2022). Has there been a transgender tipping point? Gender identification differences in U.S. cohorts born between 1935 and 2001. *American Journal of Sociology*, 128, 94–143. <https://doi.org/10.1086/719714>

Mittleman, J. (2022). Intersecting the academic gender gap: The education of lesbian, gay, and bisexual America. *American Sociological Review*, 87(2), 303–335.

Muñoz, E., Saavedra, M., & Sansone, D. (2024). *Socio-economic disparities by sexual orientation and gender identity in Mexico*.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Nordmarken, S. (2023). Coming into identity: How gender minorities experience identity formation. *Gender & Society*, 37(4), 584–613.

Presser, H. B. (1997). Demography, feminism, and the science–policy nexus. *Population and Development Review*, 23(2), 295–331.

Puckett, J. A., Tornello, S., Mustanski, B., & Newcomb, M. E. (2022). Gender variations, generational effects, and mental health of transgender people in relation to timing and status

of gender identity milestones. *Psychology of Sexual Orientation and Gender Diversity*, 9(2), 165.

Rampazzo, F., Winer, C., & Kovačič, M. (2025). Sexual orientation on Zoe: A global analysis of a lesbian dating app. *Demographic Research*, 53.

Richards, C., Bouman, W. P., Seal, L., Barker, M. J., Nieder, T. O., & T'Sjoen, G. (2016). Non-binary or genderqueer genders. *International Review of Psychiatry*, 28(1), 95–102.
<https://doi.org/10.3109/09540261.2015.1106446>

Riley, N. E. (1997). Gender, power, and population change. *Population Bulletin*, 52(1), 1–48.

Rosenfeld, M. J., & Thomas, R. J. (2012). Searching for a mate: The rise of the internet as a social intermediary. *American Sociological Review*, 77(4), 523–547.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big data & society*, 4(2), 2053951717738104.

Scandurra, C., Carbone, A., Baiocco, R., Mezzalana, S., Maldonato, N. M., & Bochicchio, V. (2021). Gender identity milestones, minority stress and mental health in three generational cohorts of Italian binary and nonbinary transgender people. *International Journal of Environmental Research and Public Health*, 18(17), 9057.

Sironi, M., & Kashyap, R. (2022). Internet access and partnership formation in the United States. *Population Studies*, 76(3), 427–445.

Valfort, M.-A. (2017). *LGBTI in OECD countries*. OECD.

Watkins, S. C. (1993). If all we knew about women was what we read in Demography, what would we know?. *Demography*, 30(4), 551-577.

Webster, A., Brooker, J., Misajon, R., & Wilson, C. (2025). Online dating experiences of transgender and gender-diverse people: A systematic review and thematic synthesis of qualitative evidence. *Psychology of Sexual Orientation and Gender Diversity*.

Westbrook, L., & Saperstein, A. (2015). New categories are not enough: Rethinking the measurement of sex and gender in social surveys. *Gender & Society*, 29, 534–560.

Wilkinson, L., Pearson, J., & Liu, H. (2018). Educational attainment of transgender adults: Does the timing of transgender identity milestones matter? *Social Science Research*, 74, 146–160.

Wilkinson, L., Shifrer, D., & Pearson, J. (2021). Educational outcomes of gender-diverse youth: A national population-based study. *Gender & Society*, 35(5), 806–837.
<https://doi.org/10.1177/08912432211038689>

Chapter 2: Categorizing Gender beyond the Binary: Inequalities in Education from a Multidimensional Gender Perspective⁵

Abstract

This study seeks to gain a better understanding of gender diversity and its implications for educational inequality research by exploring how different methods of gender categorization shape analytical outcomes. Using large-scale, nationally representative data from Mexico's National Survey on Sexual and Gender Diversity (ENDISEG), the study compares three gender categorization strategies: existing classifications of gender trajectories -cisgender, transgender, or something else-, a binary/non-binary framework, and a multidimensional model that considers both gender trajectory and binary identification simultaneously. These two axes reflect key dimensions of gender normativity: alignment with sex assigned at birth and conformity to binary gender categories. Linear models are employed to assess disparities in post-secondary education completion. Results show that trans individuals face significant educational disadvantages, while non-binary identification alone appears unrelated to educational attainment. However, a more nuanced picture emerges when accounting for gender trajectory: trans non-binary individuals face the steepest penalties. These findings underline the complexity of gender categorization and the analytical value of a multidimensional approach when measuring gender diversity and understanding inequalities across gender categories.

Keywords: Education, Gender, Transgender, Gender Identity, Non-binary, Gender fluid, Inequalities, LGBTQ+.

⁵ This chapter is published as:
Bercovich Szulmajster, S. (2025). Categorizing gender beyond the binary: inequalities in education from a multidimensional gender perspective. *Genus*, 81(1), 1-20.

2. 1. Introduction

Gender inequalities have been extensively explored in the fields of population studies and sociology. While much of this research has traditionally concentrated on the distinctions between men and women, recent studies have begun to shed light on disparities between cis and trans populations (Badgett et al., 2024; Carpenter et al., 2020). However, do the categories used in these studies capture the diversity of gender experiences exhaustively? How are gender categories defined, and who do they represent? What happens to individuals who defy both cisgender norms as well as the binary understanding of gender? Is there a bigger educational penalty for those who live their lives further away from these norms?

When conducting population-focused research, it is essential to adopt methods and categories that reflect gender-expansive realities (Compton et al., 2018) to ensure that studies accurately represent diverse gender populations. Non-normative gender experiences were not taken into consideration when developing surveys until recently, when two-step questions, with one related to sex assigned at birth and another asking about gender identity, began to be present in different surveys (Baumle and Nordmarken, 2022). This allowed for studies to go a step further and look at non-cis individuals: those whose gender differed from the sex registered at birth.

Yet again, the categories available in the gender identity questions are, in general, not enough to represent all individuals' stories (Suarez et. al., 2022). For instance, non-binary identities are often unavailable as answer options. Categorizing gender is not only important to capture gender diversity empirically, but also to better understand socioeconomic inequalities experienced by groups that defy norms. This raises the question: Do results regarding socioeconomic disparities change when we employ different gender categorization approaches?

The objective of this article is to explore the influence of varying gender categorization methods on educational attainment results. Building upon insights from trans and queer studies, I aim to transcend the traditional man–woman and cis–trans dichotomies, embracing a more inclusive approach that accommodates a broader spectrum of gender experiences. I seek to address two key questions: 1. Does educational attainment differ between genders? 2. How does the way we categorize gender affect research outcomes in this regard? Educational inequalities will be analyzed using the Mexican National Survey on Sexual and Gender Diversity (ENDISEG), which includes various questions that can be used to construct gender categories in different ways (INEGI, 2021).

This study examines three different methods of categorizing gender. It begins by comparing two approaches used in previous research on educational attainment: one that looks at cis, trans, and other gender diverse individuals, and another that compares binary and non-binary people (e.g., Carpenter et al. (2022) and Wilkinson et al. (2021)). Next, I document educational differences by considering multidimensional gender identification, combining gender trajectory -i.e. cis, trans, or something else- with binary and non-binary self-recognition, as proposed by Beischel et al. 's "3x3" framework (2023). Linear models are used to analyze disparities within these three methods of gender categorization in terms of educational attainment.

This study advances existing research in two ways. First, it analyzes educational attainment through the lens of queer and trans studies, exploring multiple approaches to categorizing gender identities. Owing to ENDISEG's innovative survey design, it offers a novel characterization of individuals, allowing for comparisons between trans, cis, and other gender diverse identities, as well as binary and non-binary identifications. This approach allows for a broader comprehension of the multidimensional nature of gender, as well as allowing for an in-depth comparison of categorization methods and their relation with educational attainment. This study is the first to compare educational attainment across different ways of categorizing gender.

Second, this article departs from the predominant focus on the United States in research on non-cisgender populations venturing into the Mexican context. In Mexico, the link between socioeconomic origin and life opportunities is notably pronounced and heavily influenced by traditional gender norms (Torche 2015). This presents an intriguing foundation for studying how access to origin resources may differ for individuals identifying in non-normative ways, potentially exacerbating disparities in life opportunities between cisgender and gender diverse populations. Among the few studies conducted in Mexico for this population, Muñoz et al. (2024) found that trans individuals who were assigned male at birth have similar levels of post-secondary education as cisgender individuals, while those who were registered as females at birth and identify as trans have lower rates of post-secondary education completion. In this study, all non-cis individuals are grouped, while I will be studying differentials by gender trajectory as well as binary identification, allowing for comparison across non-cis groups.

2. 2. Theoretical Background

The theoretical background begins with a presentation of how gender is commonly defined following broader-than-binary perspectives, including different ways of operationalizing gender, and later moves to address studies where gender inequalities in education are analyzed from such perspectives. This section ends with a description of the Mexican context in terms of educational inequalities and gender diversity.

2. 2. 1. What do we mean when we say gender?

Gender has been defined in various ways throughout history, with many scholars trying to put into words what this key part of our lived experience is. West and Zimmerman (1987) addressed gender as an emergent feature of social situations, with normative ideas about the attitudes and actions appropriate for each sex category. Butler (1990) understands gender as performative and fluid, in the sense of being a “copy for which there is no original.”

Fausto-Sterling (2019) discusses that the concept of what counts as “normal” is dynamic across time, with individuals acquiring their gender over intersubjective experiences.

These definitions leave the door open to understanding gender as a cornerstone of our identities and the way we relate with others, while also being highly situated and dynamic. Addressing gender in our studies must then conceptualize it in a broader way than the gender/sex individuals are assigned at the moment they are born. To frame the present study, inputs from queer and trans studies (Butler, 1990; Shuster, 2020; Stryker, 2023) will be used.

Identities that challenge cisnormative understandings of gender—such as transgender, non-binary, genderqueer, gender non-conforming, and gender fluid—have become increasingly visible in recent years (Flores et al., 2016; Lagos, 2022; Richards et al., 2016). While these identities are best understood through the self-definitions of the individuals who embody them, certain terms have gained ground as ways to conceptually group diverse gender experiences.

The term trans operates both as an umbrella category for those whose gender identity does not align with their sex assigned at birth, and as a specific gender identity, often combined with other terms such as man or woman (Darwin, 2020). In this sense, trans stands in contrast to cis, which refers to individuals whose gender identity corresponds with their sex assigned at birth. However, not all individuals identify as either cis or trans. Some reject both categories or use other terms, such as genderqueer or gender fluid, to name their gender experience. To account for this diversity, Beischel et al. (2023) introduce the concept of gender trajectory, which refers to the relationship between a person’s gender identity and their sex assigned at birth. This trajectory can be: Cis, when there is alignment; Trans, when there is a divergence; Or “something else”, when neither label captures the individual’s experience.

Non-binary identification represents a distinct dimension of gender. It challenges the assumption that all gender identities must align with either “man” or “woman,” and instead highlights the diversity of experiences that exist outside or beyond this binary framework. As Richards et al. (2016) note, *“some people have a gender which is neither male nor female and may identify as both male and female at one time, as different genders at different times, as no gender at all, or dispute the very idea of only two genders.”* This diversity, however, often comes into conflict with the binary logic embedded in institutional systems, where gender is typically registered, recognized, and regulated through male or female markers only. Even in countries with legal gender recognition, it is still common to only allow for choosing between male and female (ILGA, 2020). This institutional binarism limits access to rights and services—such as documentation, education, or healthcare—for those whose identities fall outside these normative categories (Westbrook & Schilt, 2014). Non-binary individuals are thus often rendered administratively invisible.

Despite this, most quantitative studies on gender inequalities—particularly in education—continue to treat gender as a singular axis of categorization, focusing either on the cis/trans distinction or on binary versus non-binary identities. This approach can obscure the lived experiences of individuals who, for example, identify as both trans and non-binary, or as neither cis or trans. By adopting a multidimensional framework that considers both gender trajectory (cis, trans, or something else) and binary identification (binary or non-binary), this study aims to capture a broader range of gender experiences and examine how they relate to educational inequalities.

2. 2. 2. Existing measures of gender identity

The question of how to capture gender in surveys is still developing and has a starting point in feminist scholars' critiques of quantitative methods (Westbrook and Saperstein, 2015). The first step that has been conducted in recent surveys is to include two separate questions regarding sex registered at birth and gender identification, currently known as the “two-step

approach.” In this survey design, gender is asked in various ways, with diverse response options: some only include men and women, while others include different ways of identification, such as trans or non-binary. Individuals whose sex assigned at birth differs from their gender identification are assumed to be non-cis, and presented as trans in many studies.

Some individuals' gender identity consists of multiple overlapping terms, and this identification can change during their life course (Ruberg & Ruelos, 2020). This illustrates gender as a multidimensional and dynamic concept, which means that multi-dimensional measurement methods might be better suited to study it. However, most of these surveys do not allow for selection of more than one option (Suarez et. al., 2022). For instance, non-binary individuals cannot also identify as trans, even though in practice these identities frequently overlap. This constraint results in a unidimensional classification, which fails to reflect the complexity of lived gender experiences. As Darwin (2020) argues, non-binary identification can hardly be sorted neatly into man/woman and cis/trans categories: non-binary and trans identification overlap, with some non-binary individuals identifying as trans, while others do not, representing a multidimensional aspect of individuals. In this context, creating mutually exclusive categories—such as cis, trans, and non-binary—may obscure important intra-group variation and produce analytically misleading results.

Responding to this need, Beischel et al. (2023) propose a multidimensional framework that incorporates two intersecting dimensions: gender trajectory—defined as cis, trans, or neither cis nor trans—and binary identification—coded as binary, non-binary, or neither binary nor non-binary. This innovative framework, known as the "3x3" model, classifies gender based on these two dimensions, allowing people to identify as both non-binary and trans for example, or even non-binary and cis. To the best of my knowledge, this novel categorization approach strategy has not yet been applied empirically in the study of educational inequalities.

In this paper, I adapt this framework to the structure of the ENDISEG survey in order to explore how intersecting dimensions of cisnormativity and binarism relate to educational outcomes. Before presenting the operationalization of this approach, the next sections provide an overview of the existing literature on gender-based educational inequalities and situate the study in the Mexican context.

2. 2. 3. Gender-based Inequalities

Individuals who depart from the gender binary have always existed, but have been granted legal recognition and rights quite recently (ILGA, 2020). In spite of these developments, gender-divergent individuals are still vulnerable in terms of discrimination, harassment, and violence (Valfort, 2017; GATE, 2013; Badgett et. al., 2021; Badgett et. al., 2024). Based on the volume edited by Baumle and Nordmarken “Demography of Transgender, Nonbinary and Gender Minority Populations” (2022), the available studies focusing on these populations are mainly related to health outcomes, labor market experiences and outcomes, discrimination in the workplace, and a few related to poverty. All these studies present a very unfavorable situation for gender experiences that diverge from cis-hetero norms.

Gender inequalities have been extensively explored in the field of educational attainment in population studies and sociology. While much of this research has traditionally concentrated on the distinctions between men and women, recent studies have begun to include a wider diversity of gender identities, such as trans, non-binary, and gender-nonconforming. Carpenter et al. (2020) found that transgender adults tend to achieve lower levels of education than their cisgender counterparts in the United States, with only 14% holding a college degree compared with 28% among cisgender individuals. In a later study using a nationally representative sample, Carpenter et al. (2022) found that non-cisgender individuals—those who informed a different gender than their sex assigned at birth—have lower employment rates and face higher levels of poverty and food insecurity compared to their cisgender counterparts.

In recent years, studies based on administrative data have begun to emerge (Kolk et al., 2025; Carpenter et al., 2024; Thomsen et al., 2024). These studies focus on individuals who have legally changed their gender markers and/or undergone medical transition. While valuable, they are subject to a selection bias: the populations they capture tend to reflect a more privileged subset of the transgender community—likely reflecting existing barriers to access legal and medical transition.

Despite this limitation, these studies provide a clear picture of socioeconomic inequalities, consistently showing that transgender individuals experience substantial earning penalties in contexts such as the United States (Carpenter et al., 2024). However, when it comes to educational attainment, both Carpenter et al. (2024) and Kolk et al. (2025) find little to no difference between transgender and cisgender individuals. As noted earlier, this likely reflects the selectivity of the samples—i.e., those who complete legal or medical transitions may already have higher levels of education, allowing them to navigate institutional systems more easily.

Wilkinson, Shifrer, and Pearson (2021) utilized a broader conceptualization of gender, and analyzed the educational outcomes of binary transgender, non-binary, and gender-unsure youth compared with those of cisgender youth, while also exploring the differences within the non-cisgender groups. They discovered differences in terms of educational outcomes, with binary transgender and gender-unsure youths facing disadvantages compared to their cisgender and non-binary peers.

However, none of these studies have developed a multidimensional gender perspective to address these inequalities. For example, trans non-binary experiences are hardly taken into consideration, only allowing for trans or non-binary identification separately. Adopting a multidimensional perspective is essential to understanding how different combinations of gender trajectory and binary identification interact with structural conditions—particularly within institutions like education, which remain deeply shaped by cisnormative and binary

logics. Individuals with non-cisgender trajectories may face barriers related to recognition, stigma, and legitimacy within educational spaces, while those who identify as non-binary may find themselves invisible or unintelligible within systems that continue to classify students, records, and facilities strictly as male or female. Exploring these dimensions jointly allows us to capture how educational inequalities are not only based on who departs from gender norms, but also on how they do so.

2. 2. 4. Mexican context

In the Mexican context, genders beyond the male/female binary have existed since precolonial times. Some of these identities are still present in some indigenous communities, such as *muxes* in the Zapotec community (Ramirez and Munar, 2022). However, legal gender identity recognition varies across the territory, with some districts having developed special procedures to grant this right: Mexico City, Sinaloa, Nayarit, Michoacán, and Coahuila (Ortiz, 2019). In states where name and gender marker modification is possible, the main document that is rectified is the birth certificate, and the procedure is expected to be mainly administrative.

Despite the legal situation, access to name and gender rectification is still not guaranteed. In many cases, judicial processes and multiple proofs are requested for individuals who seek these procedures. In addition, access to choosing identities outside of the male and female binary is still very limited. Non-binary documentation has been provided in very few situations: The first non-binary identity document was issued judicially in Guanajuato, while Veracruz granted the first administrative recognition in 2022, followed by federal acknowledgment in 2023 with an 'X' marker on voter IDs and the country's first non-binary passport (ILGA Database, 2025).

Acceptance and recognition of non-cis identities also vary throughout society, with higher acceptance found among young people, women, and those with higher education, and higher levels of discrimination perpetuated by individuals of median age and low education

(Valencia et. al. 2023; Gutierrez and Rubli 2024). These results suggest that non-cis individuals can access better environments through education, an argument that has already been made for non-heterosexual individuals (Mittleman, 2022; Boertien et. al., 2024).

As Torche presented in her 2015 study on intergenerational mobility in Mexico, a strong relationship with social origin is still present in this context. There is a wide gender difference between men and women, which becomes even larger and more beneficial for married men. This suggests a patriarchal structure, where men who adhere to norms benefit from socioeconomic transfers, whereas women and single men do so at a lower level. Muñoz et al. (2024) found that trans+ individuals assigned female at birth were the only group with lower levels of post-secondary education completion, whereas trans+ individuals assigned male at birth presented similar levels to the cisgender population. Given this context, it is relevant to study how non-normative individuals relate to this situation in detail, and whether considering their identities in a multidimensional way provides a clearer picture.

2. 3. Data and Methods

2. 3. 1. Data Source

The number of representative surveys and censuses inquiring about gender identity is increasing, following the legal recognition of non-cis identities in many countries. Two-step questions, with one related to sex registered at birth and another inquiring about gender identification, have been recently included in the 2020s round of censuses by countries such as Argentina, Canada, and Scotland, among others.

In Mexico, a representative survey was conducted between 2021 and 2022 to identify diversity in terms of gender and sexual identities within their population (INEGI, 2021). This survey was conducted in person using a digital questionnaire and included the use of headphones to answer sensitive topics. The use of headphones ensured privacy and sensitivity, which have been shown to be important points to consider when asking questions on gender and sexuality (Miller and Wilson, 2022). One individual per household, aged 15 or

older, was randomly chosen to complete the entire individual interview. Non-responses at the household and individual levels were considered in the expansion factors to ensure representativity.

2. 3. 2. Gender Identity Measures

What is novel about ENDISEG—and what enables this study’s exploration of multiple gender categorizations—is its distinctive approach to collecting information on gender identity. The questionnaire first asks about sex assigned at birth, with the question: “What is your sex assigned at birth?” offering the options Male, Female, and a follow-up incise on intersex characteristics.

Then, gender identity is assessed through two separate questions. The first one is framed as: “*Do you consider yourself...*” with five single-choice options: *Male*, *Female*, *Both of them*, *None of them*, and *Something else*. If respondents provide a gender identity that differs from their sex assigned at birth, they are presented with an additional follow-up question: “*Given that your gender is different from your sex registered at birth...*” This question includes two mutually exclusive response options:

1. “*Are you transgender or transexual*”, or
2. “*Are you queer, gender-fluid or other*”.

These two options are treated as mutually exclusive within the survey instrument, meaning respondents are required to select only one of them. While these categories are not mutually exclusive in many theoretical frameworks or community understandings—where individuals might identify as both trans and gender fluid, for example—the survey design imposes a mutually exclusive logic likely aimed at simplifying analysis⁶. Variables related to gender identity were thus created based on the responses to these questions.

⁶ This limitation is important to acknowledge, as it reflects not necessarily the lived identities of respondents, but rather how survey structure constraints self-expression. It is also worth noting that these two broad categories are not widely established as mutually exclusive in Mexican LGBTQ+ discourse, making the ENDISEG’s wording potentially ambiguous for some participants. However, we assume here that respondents choose the answer that captures the most salient part of their identity.

Following our main research question, three different categorization methods are studied to test whether the outcome is stable throughout these categories. It is important to note that, as is common in studies analyzing diverse gender identities, the number of individuals in some categories is small. However, the sample's national representativeness and the survey's design ensure that the study remains valuable despite this limitation. The significance of the models' outcomes will be taken into special consideration given this situation.

The first categorization method relates to recent literature to study non-cis populations (Carpenter et. al., 2020, 2024; Lagos, 2022). This categories are based on gender trajectory, distinguishing individuals who identify as cisgender (n=40.881), transgender (n=120), or as neither cis nor trans—a meaningful category that includes those identifying as gender queer, gender fluid, or in some other way⁷ (n=221). To create these groups, the second question related to gender identity is used, where non-cis individuals had to choose between their identification as *transgender or transsexual* versus identifying as *queer, gender-fluid, or something else*.

The second categorization focuses on the binary dimension of gender identification to explore how identification with, or departure from, the binary system relates to educational outcomes (Wilkinson et. al, 2021). This method will follow the question where individuals were asked how they consider themselves, either *male, female, both, none, or something else*. Males and females were coded as binary identifying, while those who chose both, none, or something else as their identities were coded as non-binary. For the second method of categorization, the comparison is conducted between cis binaries, non-cis binaries (n=158), and non-cis non-binaries (n=183).

Lastly, I develop a multidimensional classification that combines gender trajectory and binary identification, adapting the “3x3” framework originally proposed by Beischel et al. (2022) to

⁷ For simplification reasons, this third gender trajectory will be presented as *diverse* when needed.

the structure of the ENDISEG survey. This framework makes it possible to examine how intersecting dimensions of cisnormativity and binarism shape educational inequalities. In particular, it allows an analysis of whether non-binary individuals who also identify as trans face different disadvantages than those who identify outside both cis and trans labels. This multidimensional approach contributes to recent efforts in queer demography to move beyond current gender categorizations, and to critically engage with how survey design and categorization practices both reflect and shape the realities they aim to measure (Lagos, 2022; Westbrook & Schilt, 2014; Guyan 2022).

The “3×3” framework developed by Beischel et al. (2023) combines a gender trajectory (trans, cis, or something else) and a binary (binary, non-binary, or neither binary or non-binary) dimension. Adapted from this framework, a third method of categorization is conducted by combining both gender-related questions. 5 different groups are then created: cis binary, trans binary (n=58), trans non-binary (n=62), diverse binary (n=100), and diverse non-binary (n=121). Even though this categorization does not replicate exactly the categories used in the 3x3 framework, the approach is comparable by incorporating the two most important features of the framework: consider both dimensions of gender trajectory and gender identity, as well as allowing for third options that are not well captured by more established categories (e.g. those who do not identify as neither cis or trans).

Table 2.1 shows a cross tabulation of the frequencies with which answers to the two gender questions were given. As can be noted, many trans individuals do not identify as males or females, nor do all individuals who do not identify as cis or trans identify in a non-binary way. This justifies the need to explore different ways of grouping, as well as to critically examine current ways of categorization, as the latest literature does not explore the different gender identities present in these lived experiences.

Table 2.1. Number of Individuals by Combined Gender Identification

		Response to: "Do you consider yourself..."					
		Female	Male	Both	None	Other	Total
Response to: "Given that your gender is different from your sex registered at birth, are you..."	Cisgender	22453	18428	-	-	-	40881
	Transgender or transexual	23	35	37	8	17	120
	Diverse (Queer, gender-fluid or something else)	55	45	66	20	35	221
	Total	22531	18508	103	28	52	44189

From the collected sample, 441 individuals reported a gender identity different from their sex assigned at birth, representing approximately 1% of the total sample ($n = 44,189$). Since this study focuses on educational outcomes related to post-secondary education completion, I excluded respondents under the age of 18. The final analytical sample consequently includes 341 non-cis individuals.

Based on the information on Table 1, a detailed description of the different categorization methods can be seen. In the first method, groups are created *horizontally*: a comparison between cisgender (40881), transgender ($n=120$), and diverse ($n=221$) is conducted. For the second method, categories are coded from the *vertical* groups: female and male are considered binary, while the rest of the options are considered as non-binary. Cis and non-cis identification is then considered to create the 3 resulting groups: cis binary ($n=40881$), non-cis binary ($n=158$), and non-cis non-binary ($n=183$). For the last categorization method, a combination of both *vertical* and *horizontal* responses is taken into consideration, resulting in 5 groups: cis binary ($n=40881$), diverse binary ($n=100$), trans binary ($n=58$), diverse non-binary ($n=121$), and trans non-binary ($n=62$).

2. 3. 3. Other Measures and Method

As non-cis individuals are highly concentrated among ages where many people go to university (Carpenter et. al, 2024), as well as having restricted access to post-secondary education in the Mexican context, having completed education above the secondary level will be used as the dependent variable. To study the differences in terms of educational attainment, a variable related to post-secondary level education completion was created and analyzed, including individuals with a technical career or program, preparatory school, bachelor's degree or undergraduate program, specialization, and a master's degree or doctorate. This information is based on the question asking *“Until which year and grade did you pass in school?”*, with ten different educational levels, going from none to master/doctorate.

Age and sex assigned at birth were added as control variables to the models in a step-by-step manner. Younger individuals generally achieve higher levels of education than older generations, and in the Mexican context, women have historically exhibited lower educational attainment. This suggests that sex assigned at birth may influence educational outcomes even among those who do not identify as women (Hausman et al., 2012). While other potentially relevant factors—such as parental education or broader socioeconomic indicators—were not available in the dataset, we consider descriptive differences between groups to be an important first step towards understanding gender differences in education.

Once the distinct ways of categorization are coded, linear probability models are used to gain a better understanding of the differentials found among the groups, and across the different ways of categorization in terms of post-secondary completion (Mood, 2010).

2. 4. Results

This section presents the findings of the study, beginning with a descriptive comparative analysis between cis and non-cis populations, as well as across the multidimensional gender categorization. It then turns to the results of the OLS models, structured around the three

different gender categorization methods: First, a comparison between cisgender, transgender, and individuals who identify with another gender trajectory; Second, an analysis based on binary identification, distinguishing cis binary, non-cis binary, and non-cis non-binary individuals; And finally, the outcomes derived from applying the multidimensional framework, which combines gender trajectory with binary/non-binary identification.

2. 4. 1. Descriptive Results

Table 2.2. Descriptive Statistics Comparing Cisgender and Non-Cisgender Groups.

	Cis	Non-cis
	Mean/Share	Mean/Share
Obs.(n)	40.881	341
Mean Age	45.21	33.97
<u>Gender id (Share)</u>		
Male	0.45	0.23
Female	0.55	0.23
Both	0.00	0.30
None	0.00	0.08
Other	0.00	0.15
<u>Sex Assigned at Birth (Share)</u>		
Male	0.45	0.50
Female	0.55	0.50
<u>Education (Share)</u>		
None	0.05	0.02
Primary	0.23	0.13
Secondary - High School	0.50	0.62
University or more	0.22	0.23
<u>Occupation (Share)</u>		
Employed	0.62	0.66
Unemployed	0.02	0.04
Student	0.02	0.06
Unpaid domestic worker	0.25	0.20
Other non-employed	0.09	0.04
<u>Partnership Status (Share)</u>		
Single	0.21	0.48
Free union	0.20	0.24

Married	0.39	0.21
Separated	0.08	0.04
Divorced	0.04	0.02
Widow	0.09	0.01

When descriptively comparing non-cis individuals with their cis counterparts, several key distinctions emerge (Table 2.2). Non-cis individuals have a younger mean age than cis individuals, and their education level is more concentrated at higher levels, whereas the education pattern is more dispersed for the cis population. Despite similar levels of employment, non-cis individuals experience higher rates of unemployment and are more likely to have student status, potentially linked to their younger age structure. Additionally, nearly half of the non-cis individuals are single, more than doubling the proportion of single cis people. However, it is crucial to note that grouping non-cis individuals into a single category may obscure important disparities among gender identities, highlighting the need for a more granular examination of gender groups.

Table 2.3. Number of individuals per combined identification, non-cis population.

		Coding from response to: "Do you consider yourself..."		
		Binary	Non-Binary	Total
Response to: "Given that your gender is different from your sex registered at birth, are you..."	Trans	58	62	120
	Queer, gender fluid or something else	100	121	221
	Total	158	183	341

By looking at Table 2.3, presenting the number of individuals per combined gender location, an noteworthy image appears: Binary/non-binary identification does not relate clearly to trans identification. Out of the total sample of non-cis individuals, 35% identified as trans, while 65% identified as gender queer, gender fluid, or with some other gender. Of those in this last group, 45% identify in a binary way, while 55% identify as non-binary. Trans

individuals have a similar pattern, with 48% having a binary identity, and 52% a non-binary one. This picture justifies the need to address gender identity in a multidimensional way, as these identities present a complex form of lived experiences regarding their gender.

Table 2.4. Descriptive Statistics Comparing multidimensional Gender Groups

	Cis Binary	Diverse Binary	Diverse Non-Binary	Trans Binary	Trans Non-Binary
	Mean/Share	Mean/Share	Mean/Share	Mean/Share	Mean/Share
Mean Age	45.21	33.85	31.90	34.02	38.16
<u>Gender identification</u> (Share)					
Male	0.45	0.45	-	0.60	-
Female	0.55	0.55	-	0.40	-
Both	-	-	0.55	-	0.60
None	-	-	0.17	-	0.13
Other	-	-	0.29	-	0.27
<u>Sex Assigned at Birth</u> (Share)					
Male	0.45	0.55	0.45	0.40	0.61
Female	0.55	0.45	0.55	0.60	0.39
<u>Expression - Do you openly present according to your gender?</u> (Share)					
Yes	-	0.30	0.51	0.40	0.61
No	-	0.60	0.48	0.60	0.34
Unspecified	-	0.10	0.01	0.00	0.05
<u>Education</u> (Share)					
None	0.05	0.01	0.01	0.03	0.03
Primary	0.23	0.12	0.07	0.14	0.27
Secondary - High School	0.50	0.60	0.62	0.64	0.63
University or more	0.22	0.27	0.30	0.19	0.06
<u>Occupation</u> (Share)					
Employed	0.62	0.67	0.66	0.64	0.68
Unemployed	0.02	0.06	0.02	0.05	0.03
Student	0.02	0.05	0.11	0.03	0.00
Unpaid domestic worker	0.25	0.17	0.18	0.24	0.26
Other non-employed	0.09	0.05	0.03	0.04	0.03

<u>Partnership Status</u> (Share)					
Single	0.21	0.34	0.62	0.34	0.55
Free union	0.20	0.32	0.19	0.26	0.19
Married	0.39	0.27	0.13	0.26	0.21
Separated	0.08	0.03	0.03	0.10	0.02
Divorced	0.04	0.01	0.02	0.03	0.00
Widow	0.09	0.03	0.00	0.00	0.03

To get an initial picture of the characteristics of each multidimensional group, table 2.4 provides the main descriptive statistics by gender location, following a combined perspective. The table reveals several distinctions across these groups. The mean age varies significantly, with cis binary individuals being the most aged group at 45 years and diverse non-binary individuals the youngest at 32 years. Gender identity distribution shows that cis binary and gender diverse binary groups are evenly split between male and female identification, while non-binary individuals exhibit more diverse identities, including people identifying as both man and woman, none of these, and other ways of identification. Sex assigned at birth (SAAB) proportions also differ, with a higher percentage of people assigned male at birth among gender diverse binaries and trans non-binaries. The visibility of gender expression differs across groups: while a majority of trans non-binary individuals report expressing their gender openly, most individuals in the other gender-expansive categories do not present their gender identity publicly.

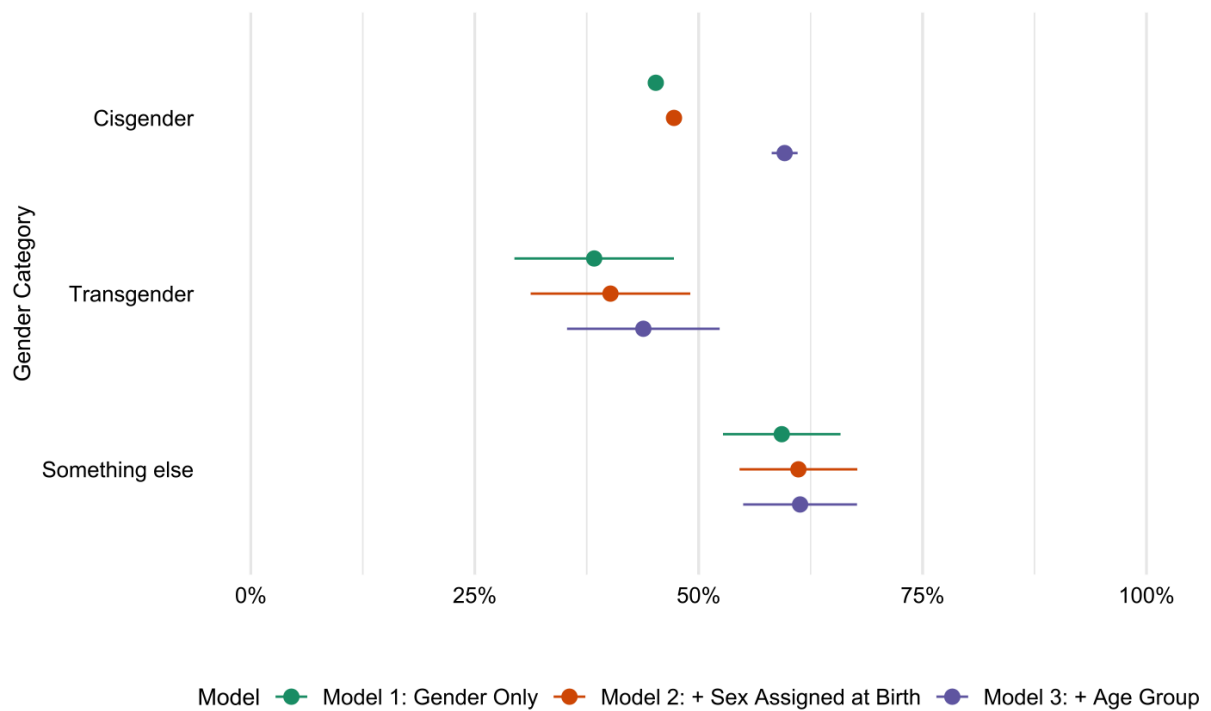
Employment rates are similar across all groups, but those who identify as queer, gender fluid, or something else have higher student status and lower unpaid domestic worker proportions. Partnership status reveals that non-binary individuals, regardless of their gender trajectory, are more likely to be single, while cis individuals have higher marriage rates. These descriptives present the complex and varied experiences within non-cis populations, emphasizing the importance of gender categorization for understanding the differences across groups.

2. 4. 2. First categorization: Gender Trajectory

Using ordinary least squares linear regressions (OLS) to study educational differentials, three models are conducted: a null model considering only the effect of gender on educational attainment, a second one including sex registered at birth, and a third one considering age. Results are visualized in Figure 2.1, where the predicted shares and their respective confidence intervals related to post-secondary education completion by gender group are presented, and in more detail in Table 2.5, where the models' results are presented with their significance levels and standard errors.

In the first model, a significant positive differential is observed for those who identified as queer, gender fluid, or something else compared to the cis-population, while a non-significative penalty is observed for trans-identifying individuals. However, when considering age, the results vary significantly: the predicted probability of completing post-secondary education for cis individuals becomes higher, reaching the shares observed for those who identified as neither trans or cis, while the penalty previously observed for trans individuals increases and becomes significant.

Figure 2.1. Predicted Shares of Post-secondary Education Completion, by Gender trajectory



Note. Predicted shares are taken from OLS Linear Probability Models explaining secondary education, controlling for age and sex assigned at birth. Ref: Cis. Error bars indicate 95% confidence intervals.

Table 2.5. Linear Regression Models' Coefficients Estimating the Likelihood of Post-Secondary Education Attainment Across Gender Identities (Cis, Trans, and Queer Groups)

	Model 1	Model 2	Model 3
(Intercept)	0.452*** (0.00)	0.472*** (0.00)	0.162*** (0.01)
Trans	-0.069 (0.05)	-0.071 (0.05)	-0.158*** (0.04)
Queer, gender fluid, or something else	0.141*** (0.03)	0.139*** (0.03)	0.017 (0.03)
Assigned Female At Birth		-0.037*** (0.00)	-0.034*** (0.00)
<u>Age group</u>			
19-24			0.556*** (0.01)
25-29			0.479*** (0.01)
30-34			0.434*** (0.01)
35-39			0.385*** (0.01)
40-44			0.315*** (0.01)
45-49			0.272*** (0.01)
50-54			0.251*** (0.01)
55-59			0.223*** (0.01)
60-64			0.150*** (0.01)
65-69			0.109*** (0.01)

*Note. Ref. Category: Cis; Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

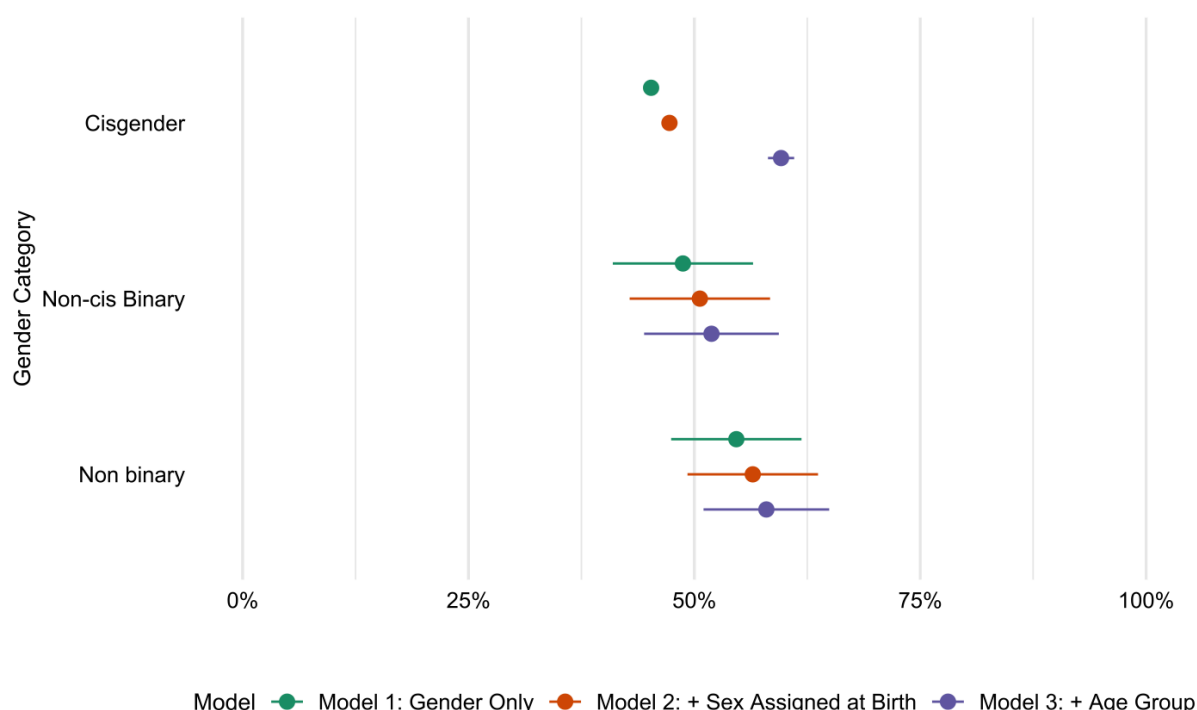
2. 4. 3. Second categorization: Binary Dimension

For the second categorization, 3 groups are studied: cis binaries, non-cis binaries, and non-cis non-binaries. Men and women were coded as binary, while those who describe themselves as both man and woman, none of these, or in a different way, were coded as non-binary.

Once again, three OLS models were conducted, adding sex assigned at birth and age sequentially. As can be seen in Figure 2.2, the effects of this way of categorization are almost non-significant for all the different categories. Comparing cis individuals with their non-cis binary peers, the differential becomes significant once controlling for age, with non-cis binary individuals performing worse than cis binaries. Non-binary individuals show no clear difference with the cisgender group in any of these models. Results are presented in detail in Table 2.6.

Figure 2.2. Predicted Shares of Post-secondary Education Completion, by Binary

Dimension (Cis Binary, Non-cis Binary, and Non-Cis Non-Binary)



Note. Predicted shares are taken from OLS Linear Probability Models explaining secondary education, controlling for age and sex assigned at birth. Ref: Cis. Error bars indicate 95% confidence intervals.

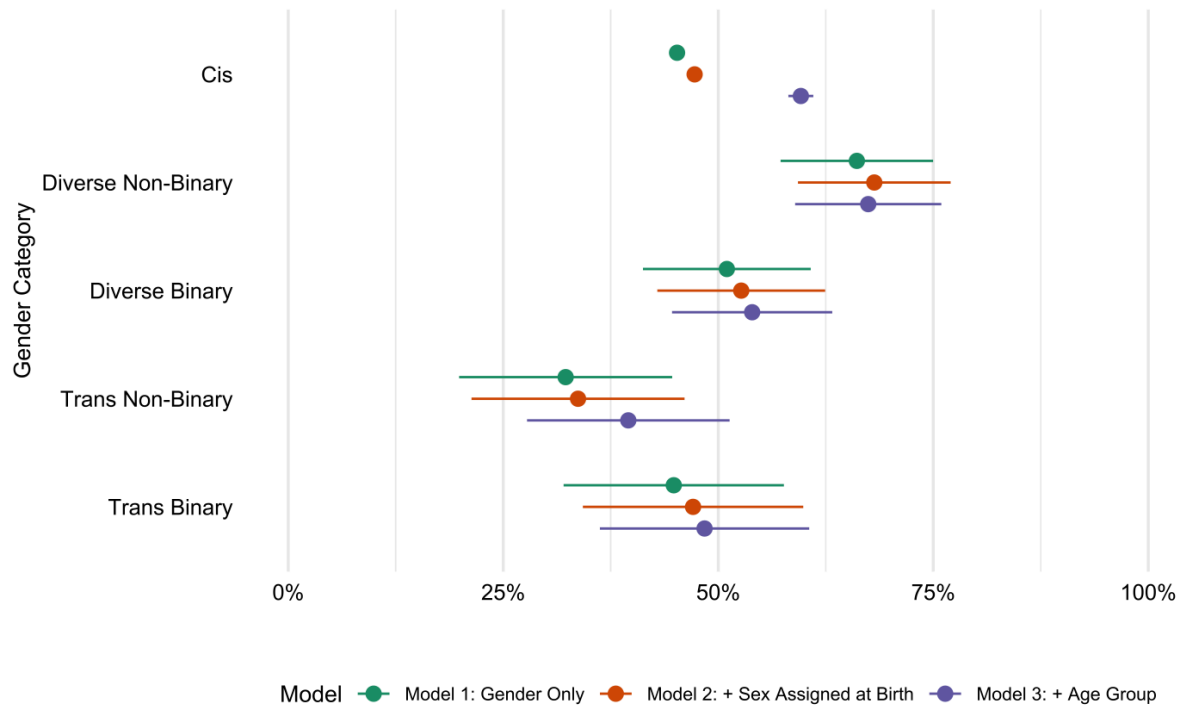
Table 2.6. Linear Regression Models' Coefficients Estimating the Likelihood of Post-Secondary Education Attainment Across Gender Identities (Cis Binary, Non-Cis Binary, and Non-Cis, Non-Binary Groups)

	Model 1	Model 2	Model 3
(Intercept)	0.452*** (0.00)	0.472*** (0.00)	0.162*** (0.01)
Non-cis Binary	0.035 (0.04)	0.034 (0.04)	-0.077* (0.04)
Non-cis Non Binary	0.094* (0.04)	0.092* (0.04)	-0.016 (0.03)
Assigned Female At Birth		-0.037*** (0.00)	-0.034*** (0.00)
19-24			0.557*** (0.01)
25-29			0.479*** (0.01)
30-34			0.434*** (0.01)
35-39			0.385*** (0.01)
40-44			0.314*** (0.01)
45-49			0.272*** (0.01)
50-54			0.251*** (0.01)
55-59			0.223*** (0.01)
60-64			0.150*** (0.01)
65-69			0.110*** (0.01)

*Note. Ref. Category: Cis; Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

2. 4. 4. Third categorization: Combined Gender Identification

Figure 2.3. Predicted Shares of Post-secondary Education Completion, by Multidimensional Gender Categorization



Note. Predicted shares are taken from OLS Linear Probability Models explaining secondary education, controlling for age and sex assigned at birth. Ref: Cis. Error bars indicate 95% confidence intervals.

Adopting a multidimensional understanding of gender—one that considers both gender trajectory and the binary/non-binary dimension—offers a clearer picture of predicted educational outcomes and potential inequalities. In the second categorization method, which considered only the binary/non-binary dimension, non-binary individuals—as a single aggregated group—showed no statistically significant effect.

However, when gender trajectory is incorporated in combination with binary identification, a different pattern emerges. As shown in Figure 2.3 and detailed in Table 2.7, only one gender group appears to be significantly penalized: trans non-binary individuals. In contrast, those who identify as non-binary but not as trans or cis show a significant positive effect, which

disappears once age is accounted for—likely due to the younger age profile of this group. Among those with binary identifications, both trans binary and diverse binary individuals show negative but non-significant effects.

Table 2.7. Linear Regression Models' Coefficients Estimating the Likelihood of Post-Secondary Education Attainment Across Gender Identities

	Model 1	Model 2	Model 3
(Intercept)	0.452*** (0.00)	0.473*** (0.00)	0.162*** (0.01)
Diverse Binary	0.058 (0.05)	0.054 (0.05)	-0.057 (0.05)
Diverse Non-Binary	0.209*** (0.05)	0.209*** (0.05)	0.078 (0.04)
Trans Binary	-0.004 (0.07)	-0.002 (0.07)	-0.112 (0.06)
Trans Non-Binary	-0.130* (0.06)	-0.136* (0.06)	-0.201*** (0.06)
Assigned Female at Birth		-0.037*** (0.00)	-0.034*** (0.00)
19-24			0.556*** (0.01)
25-29			0.479*** (0.01)
30-34			0.434*** (0.01)
35-39			0.385*** (0.01)
40-44			0.314*** (0.01)
45-49			0.272*** (0.01)
50-54			0.251*** (0.01)
55-59			0.223*** (0.01)
60-64			0.150*** (0.01)
65-69			0.109*** (0.01)

*Note. Ref. Category: Cis; Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

Additional analyses were conducted separately for individuals assigned female at birth (AFAB) and those assigned male at birth (AMAB). The same three categorization methods were applied, and overall the patterns were consistent with those obtained for the full sample. However, some differences emerged. The educational disadvantages among trans non-binary individuals remained significant only for AFAB, while for AMAB the effect lost

significance although the coefficient remained negative. Moreover, a significant penalty not observed in other models is present for trans binary AMAB individuals, whereas diverse non-binary AMAB individuals showed a relative premium. Detailed results for each categorization are available in the Supplementary Materials.

Muñoz et al. (2024) found that AFAB individuals as a group displayed lower levels of post-secondary completion compared to cisgender individuals. Yet, inequalities differ when examined by gender identification, as the subgroup analysis reveals variation across gender groups. For AMAB individuals, Muñoz et al. reported no significant difference compared to the cisgender population. However, our disaggregated analysis shows that several inequalities do emerge, most notably the penalty affecting trans binary AMAB individuals. This highlights the value of looking at non-cis experiences by gender location rather than treating them as a homogeneous group.

2. 5. Conclusions

This study has examined educational inequalities through a multidimensional lens of gender, using a nationally representative sample from Mexico. By comparing three different categorization strategies—based on gender trajectory, binary identification, and their combination—it shows that the way gender is measured and grouped deeply influences empirical findings.

The results reveal that cis individuals consistently show higher rates of post-secondary education attainment than trans individuals, echoing previous findings (Carpenter et al., 2020), and partially in concordance with Muñoz et al. (2024), who find this disadvantage only for trans individuals who were registered female at birth, but not for those assigned male at birth. However, this pattern becomes more complex when non-binary identification is introduced. In the binary vs. non-binary comparison, no significant effects were found for non-binary individuals. Yet, when both gender trajectory and binary identification are considered simultaneously, important differences emerge: trans non-binary individuals face

the strongest educational disadvantage, while diverse non-binary individuals (i.e., those who identify outside both cis and trans categories) show a positive effect—though this appears to be driven by a younger age structure of this group. These findings demonstrate the value of a multidimensional gender framework for uncovering intra-group variation, particularly within the broad category of non-binary.

Several limitations must be acknowledged. First, the survey design forced respondents to choose between mutually exclusive gender identity categories—transgender/transsexual vs. queer/gender-fluid/other—limiting the possibility of capturing overlapping identifications, which may have led to some ambiguity or misclassification. Second, while the ENDISEG survey offers rich and nationally representative data, the sample sizes for gender-diverse groups remain small, especially for those who identify as both trans and non-binary. This requires cautious interpretation of statistical significance and limits the generalizability of findings. Finally, the study could not account for key variables such as socioeconomic background, school trajectory, or timing of gender realization and social transition, which may condition both educational outcomes and how gender is expressed or disclosed. However, we consider descriptive differences between groups to be an important first step towards understanding gender differences in education.

These findings are partly consistent with existing literature. As in Carpenter et al. (2020) and Kolk et al. (2025), cis individuals tend to show educational advantages compared to trans individuals. However, previous studies based on administrative data in the United States found an advantage in educational attainment for transgender individuals, which may reflect selection biases, as they typically capture those who access legal or medical transition (Carpenter et al., 2024). This study, by contrast, includes a broader range of gender-diverse individuals—many of whom may not be institutionally recognized—highlighting hidden inequalities.

Wilkinson et al. (2021) found that binary trans individuals face disadvantages in educational outcomes, while non-binary individuals do not. Yet, this study shows that such a distinction may obscure important variation within the non-binary category. As Darwin (2020) and Beischel et al. (2023) argue, gender identities are not easily mapped into dichotomous frameworks, and recognizing their multidimensionality is crucial for empirical accuracy. Gender trajectories, on one hand, and binary/non-binary relation to this trajectory on the other hand, capture distinct dimensions of gender and are related to education in different ways.

This study offers one of the first empirical applications of a multidimensional gender framework to educational inequalities in a national survey. It demonstrates that how individuals relate to both cisnormative and binary gender expectations has meaningful consequences for their access to education. Future research could explore: The timing of gender realization or transition, and how it intersects with school experiences; Contextual factors, such as urban vs. rural settings, which may offer more or less space for non-normative gendered individuals; The role of socioeconomic origin, and whether it mediates or amplifies the penalties and advantages associated with gender diversity; And the inclusion of intersecting dimensions such as ethnicity, disability, and sexuality, which remain underexplored in this field. As Shuster (2020) mentions, taking into consideration multiple axes of oppression is highly needed, considering not only identification but how non-cis individuals move through life, if they can express their identities freely, and how their relation to different institutions and networks is, among other aspects of social life.

This study highlights the importance of treating gender as a multidimensional construct when analyzing social inequalities. Rather than a single identity marker, gender intersects with two axes of normativity: alignment with sex assigned at birth and conformity to binary identification. Individuals who challenge these norms—whether by identifying as trans, non-binary, or both—often face institutional barriers, particularly in educational contexts where cisnormative and binary logics prevail. The findings show that different ways of

categorizing gender lead to significantly different empirical results, underscoring the need for more nuanced methodologies in population research that move beyond rigid categories and reflect the complexity of lived gender experiences.

2. 6. References

Badgett M, Carpenter C, and Sansone D. (2021) LGBTQ Economics. *Journal of Economic Perspectives*, 35 (2): 141-70. DOI:10.1257/jep.35.2.141

Badgett, M. L., Carpenter, C. S., Lee, M. J., & Sansone, D. (2024). A review of the economics of sexual orientation and gender identity. *Journal of Economic Literature*, 62(3), 948-994.

Baumle, A. K., & Nordmarken, S. (Eds.). (2022). Demography of Transgender, Nonbinary and Gender Minority Populations. *Springer*.

Beischel, W.J., Schudson, Z.C., Hoskin, R.A., Van Anders, S.M., (2023). The gender/sex 3×3: Measuring and categorizing gender/sex beyond binaries. *Psychology of Sexual Orientation and Gender Diversity* 10, 355–372. <https://doi.org/10.1037/sgd0000558>

Boertien, D., Perales, F., & Pessin, L. (2024). Does intergenerational educational mobility vary by sexual identity? A comparative analysis of five OECD countries. *European Sociological Review*, 40(2), 226-241.

Butler J. (1990). Gender Trouble: Feminism and the Subversion of Identity. *New York: Routledge*

Carpenter, C.S., Eppink, S.T., Gonzales, G. (2020). Transgender Status, Gender Identity, and Socioeconomic Outcomes in the United States. *ILR Review*, 73(3), 573–599. <https://doi.org/10.1177/0019793920902776>

Carpenter, C.S., Lee, M.J., Nettuno, L., (2022). Economic Outcomes for Transgender People and Other Gender Minorities in the United States: First Estimates from a Nationally

Representative Sample. *IZA Discussion Papers, No. 15116, Institute of Labor Economics (IZA), Bonn*

Carpenter, C. S., Goodman, L., & Lee, M. J. (2024). Transgender earnings gaps in the United States: Evidence from administrative data (No. w32691). *National Bureau of Economic Research*.

Compton D R., Meadow T., Schilt, K. (Eds.), (2018). Other, please specify: queer methods in sociology. *University of California Press, Oakland, California*.

Darwin, H., (2020). Challenging the Cisgender/Transgender Binary: Nonbinary People and the Transgender Label. *Gender & Society* 34, 357–380.
<https://doi.org/10.1177/0891243220912256>

Fausto-Sterling, A. (2019). Gender/sex, sexual orientation, and identity are in the body: How did they get there?. *The Journal of Sex Research*, 56(4-5), 529-555.

Flores A, Herman J, Gates G, and Brown T. (2016). “How many adults identify as transgender in the United States?” *The Williams Institute*.

GATE (2013) Identidad de Género y Derechos Humanos. *Nueva York: GATE*. Available at: <https://gate.ngo/es/identidad-de-genero-y-derechos-humanos-en-2-paginas/>

Gutierrez, E., & Rubli, A. (2024). LGBT+ persons and homophobia prevalence across job sectors: Survey evidence from Mexico. *Labour Economics*, 87, 102500.

Guyan, K. (2022). Queer data.

Hausman, R., Tyson, L. D., & Zahidi, S. (2012). The global gender gap report. In *World Economic Forum* (pp. 11-23).

ILGA. (2020) Trans Legal Mapping Report 2019: Recognition before the law, *Geneva: International Lesbian, Gay, Bisexual, Trans and Intersex Association*.

ILGA Database (2025). <https://database.ilga.org/en> Accessed 05 March 2025.

INEGI (2021) Encuesta Nacional sobre Diversidad Sexual y de Género (ENDISEG).
<https://www.inegi.org.mx/programas/endiseg/2021/>

Kolk, M., Tilley, J. L., von Essen, E., Moberg, Y., & Burn, I. (2025). The Demography of Sweden's Transgender Population: A Research Note on Patterns, Changes, and Sociodemographics. *Demography*, 62(2), 349-363.

Lagos D. (2022). Has There Been a Transgender Tipping Point? Gender Identification Differences in U.S. Cohorts Born between 1935 and 2001. *American Journal of Sociology* 128, 94–143. <https://doi.org/10.1086/719714>

Mittleman, J. (2022). Intersecting the academic gender gap: The education of lesbian, gay, and bisexual America. *American Sociological Review*, 87(2), 303-335.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European sociological review*, 26(1), 67-82.

Muñoz, E., Saavedra, M., & Sansone, D. (2024). Socio-economic disparities by sexual orientation and gender identity in Mexico.

Ortiz, M. Á. L. (2019). Avances en el reconocimiento del derecho humano a la identidad de género auto-percibida en México.

Miller K., & Willson S. (2023) Development and Evaluation of a Single, Non-Binary Gender Question for Population-Based Federal Health Surveys. *NCHS*

Ramirez, J., & Munar, A. M. (2022). Hybrid gender colonization: The case of muxes. *Gender, Work & Organization*, 29(6), 1868-1889.

Richards, C., Bouman, W. P., Seal, L., Barker, M. J., Nieder, T. O., & T'Sjoen, G. (2016). Non-binary or genderqueer genders. *International Review of Psychiatry*, 28(1), 95–102.
<https://doi.org/10.3109/09540261.2015.1106446>

Ruberg, B., & Ruelos, S. (2020). Data for queer lives: How LGBTQ gender and sexuality identities challenge norms of demographics. *Big Data & Society*, 7(1), 2053951720933286.

shuster, S. (2020). Carving Out a Niche or Finding a Place at the Table? The Sociology of Transgender Studies.

Stryker, S. (2004). Transgender studies: Queer theory's evil twin. *GLQ: A journal of lesbian and gay studies*, 10(2), 212-215.

Suárez, M. I., Marquez-Velarde, G., Compton, D. L. R., & Poston, D. L. (2022). Demographic and Socioeconomic Characteristics of the Transgender and Cisgender Populations in the US. In *Demography of Transgender, Nonbinary and Gender Minority Populations* (pp. 73-90). Cham: Springer International Publishing.

Thomsen, M. K., Andersen, M., & Greve, J. (2024). Transgender lives at the population level: Evidence from Danish administrative data. *Social Science & Medicine*, 358, 117182.

Torche, F., (2015). Intergenerational Mobility and Gender in Mexico. *Social Forces* 94,563–587. <https://doi.org/10.1093/sf/sov082>

Valfort M. A. (2017) LGBTI in OECD Countries. Paris: OECD

Valencia, Á. R. V., Flores, J. A. N., Avalos, C. L. M., & Sánchez, J. C. A. (2023). Estudio estadístico multivariante sobre las actitudes hacia las minorías sexo-genéricas en Colima, México. *Estudios sobre las Culturas Contemporáneas*, 19(57), 131-158.

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & society*, 1(2), 125-151.

Westbrook L, Saperstein A. (2015). New categories are not enough: rethinking the measurement of sex and gender in social surveys. *Gend. Soc.* 29:534–60

Wilkinson, L., Shifrer, D., & Pearson, J. (2021). Educational Outcomes Of Gender-Diverse Youth: A National Population-Based Study. *Gender & Society*, 35(5), 806–837. <https://doi.org/10.1177/08912432211038689>

2. 7. Appendix (Online Supplementary Materials)

In these supplementary materials, results on post-secondary education by gender identity are presented, disaggregated by sex assigned at birth. All three categorization methods were examined separately for individuals assigned female at birth (AFAB) and those assigned male at birth (AMAB). For each categorization, we present models including only gender categories as well as models additionally controlling for age (Tables 2.7.1, 2.7.2, and 2.7.3).

First categorization: Gender trajectory

For this categorization, the results obtained when separating the analysis by sex assigned at birth are consistent with those for the overall sample in terms of statistical significance.

Table 2.7.1. Linear Regression Models' Coefficients Estimating the Likelihood of Post-Secondary Education Attainment Across Gender Identities (Cis, Trans, and Queer Groups)

	AFAB		AMAB	
	Null Model	+Age group	Null Model	+Age group
(Intercept)	0.435*** (0.00)	0.123*** (0.01)	0.472*** (0.00)	0.169*** (0.01)
Trans	-0.046 (0.06)	-0.171** (0.06)	-0.095 (0.06)	-0.146* (0.06)
Queer, gender fluid, or something else	0.132** (0.05)	-0.002 (0.04)	0.146** (0.05)	0.035 (0.05)
19-24		0.570*** (0.01)		0.540*** (0.02)
25-29		0.475*** (0.01)		0.485*** (0.02)
30-34		0.445*** (0.01)		0.420*** (0.02)
35-39		0.397*** (0.01)		0.371*** (0.02)
40-44		0.327*** (0.01)		0.300*** (0.02)
45-49		0.280*** (0.01)		0.262*** (0.02)

50-54	0.265*** (0.01)	0.234*** (0.02)
55-59	0.208*** (0.01)	0.241*** (0.02)
60-64	0.151*** (0.02)	0.148*** (0.02)
65-69	0.106*** (0.02)	0.114*** (0.02)

Note. Ref. Category: Cis; Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.

Second categorization: Binary dimension

Following the second categorization method, based on the binary/non-binary dimension, a negative and significant effect was found for individuals assigned male at birth (AMAB) who identify within the binary, as was observed for the sample as a whole. This effect was not observed among individuals assigned female at birth (AFAB).

Table 2.7.2. Linear Regression Models' Coefficients Estimating the Likelihood of Post-Secondary Education Attainment Across Gender Identities (Cis Binary, Non-Cis Binary, and Non-Cis, Non-Binary Groups)

	AFAB		AMAB	
	Null Model	+Age group	Null Model	+Age group
(Intercept)	0.435*** (0.00)	0.123*** (0.01)	0.472*** (0.00)	0.169*** (0.01)
Non-cis Binary	0.090 (0.06)	-0.032 (0.05)	-0.024 (0.06)	-0.123* (0.05)
Non-cis Non Binary	0.053 (0.05)	-0.086 (0.05)	0.130* (0.05)	0.049 (0.05)
19-24		0.570*** (0.01)		0.540*** (0.02)
25-29		0.475*** (0.01)		0.485*** (0.02)
30-34		0.445*** (0.01)		0.420*** (0.02)
35-39		0.397*** (0.01)		0.370*** (0.02)
40-44		0.327*** (0.01)		0.299*** (0.02)

45-49	0.280*** (0.01)	0.262*** (0.02)
50-54	0.265*** (0.01)	0.234*** (0.02)
55-59	0.208*** (0.01)	0.240*** (0.02)
60-64	0.151*** (0.02)	0.148*** (0.02)
65-69	0.106*** (0.02)	0.115*** (0.02)

*Note. Ref. Category: Cis; Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

Third categorization: Combined gender identification

The third categorization method reveals some differences compared to the full sample results. The disadvantage for trans non-binary individuals remains significant among those assigned female at birth (AFAB). However, among individuals assigned male at birth (AMAB), the effect for trans non-binary individuals is no longer significant, although the coefficient remains negative. For trans binary AMAB individuals, a significant penalty is observed, while a premium is found for diverse non-binary AMAB individuals.

Table 2.7.3. Linear Regression Models' Coefficients Estimating the Likelihood of Post-Secondary Education Attainment Across Gender Identities by Sex registered at birth.

	AFAB		AMAB	
	Null Model	+Age group	Null Model	+Age group
(Intercept)	0.435*** (0.00)	0.123*** (0.01)	0.472*** (0.00)	0.169*** (0.01)
Diverse	0.098 (0.07)	-0.013 (0.07)	0.018 (0.07)	-0.092 (0.06)
Binary				
Diverse	0.156* (0.06)	0.005 (0.06)	0.273*** (0.07)	0.163* (0.06)
Non-Binary				
Trans Binary	0.079 (0.08)	-0.057 (0.08)	-0.125 (0.10)	-0.198* (0.10)
Trans	-0.227* (0.10)	-0.337*** (0.10)	-0.078 (0.08)	-0.115 (0.08)
Non-Binary				
19-24		0.570*** (0.01)		0.540*** (0.02)

25-29	0.475*** (0.01)	0.485*** (0.02)
30-34	0.445*** (0.01)	0.420*** (0.02)
35-39	0.397*** (0.01)	0.371*** (0.02)
40-44	0.327*** (0.01)	0.299*** (0.02)
45-49	0.281*** (0.01)	0.262*** (0.02)
50-54	0.265*** (0.01)	0.234*** (0.02)
55-59	0.208*** (0.01)	0.240*** (0.02)
60-64	0.151*** (0.02)	0.148*** (0.02)
65-69	0.106*** (0.02)	0.115*** (0.02)

*Note. Ref. Category: Cis; Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

Chapter 3: Gender Trajectory Milestones from a Life-Course Perspective: A Comparison across Cohorts

Abstract

Gender identification and affirmation develop through key milestones like realization, disclosure, and, when desired, medical transition. The timing and sequencing of these milestones are life-course processes that are closely related to, and shaped by, historical context, specially for individuals who challenge gender norms. While existing research has documented cohort differences in gender milestones, no study has modeled complete gender trajectories as holistic sequences. Using microdata from the 2019 EU Fundamental Rights Agency LGBTQ+ Survey (n=7,621 across 30 countries), this study applies sequence analysis and cluster methods to identify common trajectory typologies and examines how these patterns vary by birth cohort. Three distinct trajectories emerged: childhood realization with gradual disclosure (15%), adolescent realization with rapid disclosure (43%), and prolonged non-disclosure (43%). Multinomial regression models reveal substantial differences between cohorts: younger individuals (born 1985-1994) were twice as likely to follow rapid-disclosure trajectories compared to those born before 1985, who predominantly experienced extended periods between realization and external milestones. These findings demonstrate that gender trajectories are not uniform but responsive to the social context, with implications for healthcare planning, socioeconomic outcomes, legal frameworks, and demographic measurement in increasingly diverse populations.

Keywords: Life course, gender, milestones, birth cohort

3. 1. Introduction

Embodying and expressing one's gender is a fundamental aspect of our lives, and, when gender identity does not align with sex assigned at birth, this development is quite different. The *timing* and *process* of this journey, from self-perception to disclosure, or seeking affirmative medical and legal procedures, are not trivial; they have direct and profound implications for an individual's life.

There are many outcomes that have been studied in relation to gender milestones. Access to gender-affirming care is robustly associated with improved well-being. A systematic review by Cornell University's What We Know Project (2021) found that gender transition improves overall quality of life, including significant reductions in anxiety, depression, and suicidality. Similarly, the context of one's transition influences educational attainment. Supportive school environments are critical, as experiences of discrimination during these formative years can negatively impact academic trajectories (Wilkinson et al., 2018; Bercovich, 2025). This impact extends to socioeconomic outcomes and relationships. Demographic research using Swedish register data, for example, finds that trans individuals have lower legal union rates and substantially worse socioeconomic outcomes when compared to cisgender individuals (Kolk, 2025). This topic is increasingly relevant, as recent studies indicate a growing number of individuals are identifying outside of binary cisgender norms and moving beyond the gender/sex they were assigned at birth (Lagos, 2022). Younger generations in particular embrace diverse gender identities more frequently than older cohorts.

Previous studies have taught us that gender transition does not occur in a vacuum (Puckett et al., 2022; Wilkinson et al., 2018). Realizing that one's gender differs from the one assigned at birth, openly sharing this identity, accessing legal recognition, or undergoing body modifications, when desired, are all milestones deeply shaped by the social, legal, and medical context surrounding the individual. A recent systematic review on the timing of these milestones (Baams et al., 2025) provides a general illustration of this. The review found that while the age of internal realization (self-awareness of one's gender) has remained relatively

stable across generations, the ages of social and medical milestones have decreased dramatically. For instance, the average age of first disclosure dropped from 26.8 for Gen X to 17.3 for Generation Z, and the age of seeking gender-affirming healthcare plummeted from 50.9 for Baby Boomers to 17.6 for Generation Z. This indicates that while realization may be an individual process, the pacing of disclosure and affirmation is highly dependent on societal acceptance, visibility, and access to care.

Existing literature has focused on these milestones in isolation or highlighted attitudinal differences across birth cohorts (Puckett et al, 2022; Scandurra et al., 2021; Wilkinson et al., 2018). However, what is often missing is a comprehensive life-course perspective. Existing studies often fall short of connecting these dots into a temporal sequence, leaving questions unanswered about the *timing* and *pacing* of key gender milestones. It is still unknown how the *duration* between them relates to an individual's developmental context, particularly the social climate of their birth cohort. This is a critical gap in the literature, as the duration of these processes, especially a prolonged gap between internal realization and disclosure, is a measurable proxy for periods of internal and external incongruence, which are linked to well-being, minority stress, and, potentially, socioeconomic outcomes.

To address this gap, the present study aims to explore gender trajectories through a life-course lens. It seeks to answer the following research questions: Are there common patterns in gender trajectories? Do these trajectories and patterns vary across cohorts? To answer these questions this study focuses on three key milestones: the realization that one's gender does not align with the sex assigned at birth, the age of first disclosure, and, when applicable, the age at which body interventions occurred. A key aspect, in which this study goes beyond previous studies, is to document the time that elapses between different gender milestones and how this changed over time. By analyzing microdata from the 2019 European Union Fundamental Rights Agency (FRA) LGBTQ+ Survey, covering 30 countries, this research constructs 7,621 individual life-course sequences. Common trajectory typologies are identified based on milestone timing, the duration of spells between

milestones, and their resulting sequences. These typologies are then related to birth cohort using multinomial regression models.

This article contributes to the literature on the life course of trans, non-binary, and gender diverse populations by identifying different trajectories, focusing on the timing and spacing of key milestones. The relevance of this approach to the literature is twofold. First, it allows us to better understand the impact of social context on gender development, as we compare two different cohorts and test for country differences. Second, at a time when legal and political frameworks surrounding trans rights are under constant debate, understanding how individuals actually navigate their trajectories is vital. The findings show that there are clear cohort differences in the way gender is developed during the life course, with, in general, shorter paths to affirmation in younger cohorts. These findings can help inform the design of public policy and gender-affirmative pathways that are more responsive to the real-world needs and timing of trans, non-binary, and gender-diverse people.

3. 2. Theoretical Background

3. 2. 1. Gender trajectories and milestones

When studying the lived experiences of individuals whose gender identity differs from the sex they were assigned at birth, it is essential to understand their life stories holistically. Adopting a life course perspective (Elder et al., 2003) on gender trajectories emphasizes the trajectory as a process. This perspective focuses not only on the key milestones, such as self-realisation, disclosure, or expression, but also on the *pacing* and duration between them. The life course perspective also considers that the timing of these trajectories consist of three dimensions: chronological timing (age), social timing (socially constructed expectations related to age), and historical timing (Wilkinson et al., 2018). Gender trajectories are then understood as being influenced not only by internal factors but also as highly dependent on their historical timing.

The study of gender trajectory milestones focuses on the diverse pathways individuals take when their gender identity diverges from the sex they were registered with at birth. Existing literature identifies key events and processes in the lives of non-cis individuals; including self-recognition, identity disclosure and expression, social and medical transitions, and legal changes (Wilkinson et al., 2018; Puckett et al., 2022; Klein, 2023). These milestones are shaped by both personal and contextual factors, reflecting the individuality and potential complexity of each transition process. To organise these milestones, the literature can be grouped thematically into a process that often begins with internal recognition, followed by social expression, and subsequently, medical or legal affirmation.

Across the literature, a consistent starting point for gender trajectories is the *internal* milestone: the questioning of the gender assigned at birth. This "*recognition of self as different*" (Wilkinson et al., 2018) or "*coming-into-identity*" (Nordmarken, 2023) is an internal cognitive and emotional process. Wilkinson et al. (2018) understand the process of self-realisation to include two separate parts: the recognition of self as different due to gender and the identification of oneself as transgender. Nordmarken (2023) describes this process as one of self-reflection often sparked by exposure to diverse gender conceptualizations, experimentation with gender, emotionally significant events, or conversations with others.

Following this internal realization, the next critical milestone is often social. The disclosure of one's gender identity and expressing it in alignment with that identity are fundamental steps to well-being and personal development (Puckett et al., 2022). This milestone represents the first bridge between the internal self and the social world. However, this step is not immediate; the literature documents that many individuals go through "prolonged periods of internal identification before disclosing their gender to others" (Factor, 2008). This temporal gap between internal realization and social disclosure is theorized to be strongly influenced by the external context (Nordmarken, 2023).

Gender pathways often include subsequent steps of medical, legal, and social affirmation. Scholars have categorized these steps in detail. For example, Klein (2023) identifies 11 transition milestones, while Barrientos (2021) categorizes affirmation steps into seven domains. Both frameworks cover the key areas of internal realization, social transition (e.g., changing expression), legal affirmation (e.g., legal name and gender rectification), and medical affirmation (e.g., initiation of hormone therapy, and various gender-affirming surgeries). Klein (2023) notes that legal and medical affirmation steps are particularly significant as they contribute positively to psychological well-being.

Finally, it is crucial to note that the timing and sequencing of these milestones are not uniform. Factor (2008) highlights variations in the timing of these events across groups. For example, male-to-female (MTF) individuals often recognize their gender identity earlier but delay public presentation compared to female-to-male (FTM) individuals. FTMs, in contrast, are more likely to engage in conversations about their experiences before identifying fully as transgender.

While all these milestones, from internal realization to social, legal, and medical affirmation, constitute key events of gender trajectories, this study narrows its focus to the gap between self-realisation and disclosure. The theoretical frameworks presented (e.g., Nordmarken; Wilkinson et al., 2018) clearly distinguish between an *internal* process of realization and *external* actions of social disclosure. The temporal gap between these two specific milestones is of central importance in this paper. The following section will argue that this *timing* is not random, but is profoundly shaped by the historical and social context of an individual's cohort.

3. 2. 2. Contextual factors

An individual's journey towards living in their authentic selves is an experience profoundly shaped by the world around them. The specific historical, social, and geographical context a person is born into can either create barriers or, conversely, provide the freedom to thrive. As

established in the previous section, these factors shape the timing, pace, and accessibility of key milestones, shaping the accessibility and social risk of acts like disclosure and medical or legal affirmation. Previous studies have highlighted how variations in societal visibility, legal frameworks, and generational contexts contribute to differences in trans individuals' experiences (Puckett et al., 2022).

Wilkinson et al. (2018) highlights three dimensions of timing that shape these trajectories. Chronological timing refers to developmental stages such as puberty. Social timing relates to societal expectations tied to age and roles, while historical timing reflects broader societal shifts and cultural norms. For instance, Millennial transgender individuals (born after 1981) who came of age in the 1990s or later benefited from growing societal awareness and increased visibility of LGBTQ+ issues. In contrast, Generation X (born between 1965 and 1981) and Baby Boomers (born before 1965) faced adolescence during less supportive periods, when transgender activism was emerging. The impact of this changing historical context is not merely theoretical; it is empirically documented.

A recent systematic review by Baams et al. (2025) on the timing of gender and sexual identity milestones provides direct evidence of how the age of internal realization has remained relatively stable and early across generations. Wilkinson et al. (2018), using a large U.S. sample of 3,796 transgender and gender-nonconforming adults, found that self-realisation occurred at around age 8 across Baby Boomers (1946–1964), Gen X (1965–1981), and Millennials (1982–1996). Other studies with smaller samples report similar ages: Puckett et al. (2022), based on 695 transgender and gender-diverse participants in the United States, place this milestone in late childhood or early adolescence, between ages 9 and 13, and Scandurra et al. (2021), drawing on a sample of 197 transgender and gender-nonconforming individuals in Italy, identify comparable patterns. Together, these studies suggest that the age of internal realization is relatively consistent across cohorts; this existing evidence comes mostly from single-country studies with diverse sample sizes, highlighting the need for large, cross-national analyses such as the one presented here.

These studies also found that the ages for social and medical milestones have dropped dramatically across cohorts. The average age of first disclosure fell from 26.8 for Generation X to 17.3 for Generation Z (Scandurra et al. 2021), and the age of seeking gender-affirming healthcare plummeted from 50.9 for Baby Boomers to 17.6 for Generation Z (Puckett et al., 2022). This empirically shows a shrinking gap between internal realization and external action, relating the pacing of trajectories to one's historical cohort.

These cohort differences illustrate the mechanisms of historical change. Nordmarken (2023) emphasizes that gender identity formation is a fundamentally social process that occurs through interactions with others. Exposure to diverse gender models and narratives, or, conversely, exposure to social stigma, can accelerate or delay the “coming into identity” process. In this light, the hard-won gains in visibility, the advancement of legal rights, and generational shifts toward acceptance (Puckett et al., 2022) are key contextual factors. Consequently, younger cohorts have benefited from greater access to affirmative narratives and resources (Kosciw et al., 2015; Levitt and Ippolito, 2014), facilitating an earlier alignment between their internal and social identities.

Context is, however, not only historical: it is also geographical. The social, legal, and medical landscapes that shape gender trajectories vary significantly not just over time but also over space. Within the European Union, there is vast divergence in contexts, from countries with progressive legal gender recognition based on self-determination to those with significant legal and medical barriers (ILGA-Europe, 2024). Analyzing data across these 30 diverse European countries, as this study does, allows for a more robust test of the link between context and trajectory timing.

As the life course perspective highlights, our individual stories are always embedded in our collective history, which explains why these lived experiences vary across regions and cohorts. Building on this life-course framework, this article aims to fill a critical gap by showing the diversity gender of trajectories. While existing literature has clearly established

that context matters and has examined cohort shifts in individual milestones (e.g., Baams et al., 2025; Puckett et al., 2022), to our knowledge, no study has yet applied a life-course sequence approach to model how these complete gender trajectories, as holistic sequences of events, vary across birth cohorts.

Therefore, the present study seeks to answer the following research questions: Are there emerging patterns when looking at gender trajectories? Do these trajectories vary across cohorts? To address these questions, the analysis proceeds in two stages. First, I employ sequence analysis and cluster analysis as an exploratory approach to identify whether distinct typologies of gender trajectories emerge from the data. Second, once typologies are established, we test theory-driven hypotheses about how these trajectories vary across birth cohorts.

The first empirical goal is mainly exploratory: to identify distinct typologies of gender trajectories using sequence and cluster analysis. Based on previous work on gender identity milestones (Baams et al., 2025; Puckett et al., 2022), I expect that if distinct trajectory typologies exist, they will be primarily differentiated by two factors: (a) the timing (chronological age) of each milestone and (b) the pacing (duration) of the sequence, specifically the gap between internal realization and the first disclosure milestones. However, the specific number and characteristics of trajectory types that emerge will be determined empirically through the clustering procedure.

The second empirical goal is to test whether trajectory typology membership varies systematically across birth cohorts. Building on the exploratory identification of trajectory typologies, and grounded in theoretical and empirical evidence of cohort differences in gender milestone timing (Baams et al., 2025; Puckett et al., 2022), the following hypotheses are tested:

H1: Younger cohorts (born between 1985 and 1994) will be significantly more likely to belong to trajectory typologies characterized by early external milestones (disclosure

and/or medical steps) and a shorter duration between realization and these external steps.

H2: In contrast, older cohorts (born on or before 1985) will be more likely to belong to trajectory typologies characterized by (a) delayed external milestones (disclosure/medical steps) or (b) an "internal-only" trajectory (showing only the "realization" milestone, with no external milestones appearing before age 25).

While the second empirical goal provides specific, testable predictions, the overall research design remains fundamentally exploratory in nature (in line with the first empirical goal). Beyond simply testing these hypotheses, this study makes use of sequence analysis to identify a broader map of common pathways. This exploratory approach seeks to find emerging typologies in the lived experiences of gender identification and to provide a rich, data-driven understanding of how belonging to different birth cohorts shapes these life-course trajectories.

3. 3. Research Design

3. 3. 1. Data

This analysis draws on data from the European Union Agency for Fundamental Rights' LGBTQ+ survey. The primary goal of this survey was to address the lack of data reflecting the lived experiences of LGBTQ+ individuals across European countries. This survey was first conducted in 2012 and included around 93,000 individuals. It was repeated in 2019, reaching nearly 140,000 respondents. Data from the 2019 iteration is the focus of the present study, which was collected through an online survey between 27 May and 22 July 2019. Participants were recruited through a multi-channel campaign that included LGBTQ+ organisations, national media, social media, dating apps, and offline outreach to reach hard-to-access groups, and the questionnaire was completed through an open online survey available across 30 European countries.

Although the sample is non-probabilistic, it provides a uniquely large sample of individuals and captures detailed information on gender-related milestones that would not typically be included in general population surveys. All individuals, gender diverse or not, can experience gender questioning and reflection, and in this LGBTQ+ addressed survey, this was included. Currently, probabilistic samples with sufficient numbers of respondents and detailed milestone data are not available. Thus, non-probabilistic data are currently the only available source for quantitative analysis of gender trajectories and milestones. Previous analyses of the FRA LGBTQ survey have examined transgender discrimination and experiences across Europe (TGEU, 2021; ILGA-Europe & TGEU, 2023), and Caprinali and Vitali (2025) recently used this data to conduct event history analyses of sexual identity development, demonstrating its value for studying LGBTQ+ life course trajectories.

Two self-identification questions were considered to capture gender identity. The first asked, *“How would you describe yourself today?”* with response options including *woman, trans woman, man, trans man, non-binary (non-binary, genderqueer, agender, polygender, or gender-fluid)*, and *something else*. Among the 7,621 participants providing gender milestone data, 1,182 identified as trans women, 959 as trans men, and 3,519 as non-binary. Participants who identified as men, women, or something else were given a follow-up question to clarify whether they identified as trans, using an inclusive definition: *“Are/were you a trans person? In this survey, ‘trans’ is used as a broad umbrella term, encompassing transgender, non-binary, gender variant, polygender, agender, gender-fluid, cross-dressers, transsexual individuals, men and women with a transsexual past, and other related identities.”* Based on these responses, an additional 1,011 women, 714 men, and 236 other individuals were included in the sample.

As the main aim was to explore gender trajectories, the analysis focused on individuals with valid information related to key gender milestones. Questions related to these milestones were asked to individuals whose gender identification differed from their sex assigned at birth and to those who explicitly mentioned having a trans history. The questionnaire

included four questions addressing gender milestones: age of realization was assessed by asking, *“At what age did you first realize that your feelings about your gender did not match the gender assigned to you at birth?”*. Age of disclosure was determined through the question, *“How old were you when you first told someone about this?”*. Body interventions were explored with two questions: the first asked, *“Have you undergone any kind of intervention to change your body so it better matches your gender identity?”* If respondents answered affirmatively, a follow-up question, *“At what age did this take place for the first time?”*, helped establish the age of intervention.

Two cohorts were created to study change over time: one for individuals born between 1985 and 1994, and another for those born before 1985. Due to the age grouping structure in the FRA LGBTQ survey data, 1985 was used as the cohort cut-off point rather than 1982, as employed in previous studies (e.g., Wilkinson et al., 2018). These cohorts experienced distinct social contexts during their formative years: those born between 1985 and 1994 came of age during a period of increasing LGBTQ+ visibility and rights expansion in Europe, while those born before 1985 navigated adolescence and early adulthood in less supportive environments with limited public discourse on transgender issues (Wilkinson et al., 2018)

To ensure comparability across cohorts, trajectories were studied through sequences that go from 0 to 25 years. Therefore, individuals younger than 25 were excluded from the sample. Variables related to gender, educational level, and sex assigned at birth are used as covariates in the analysis. Educational level was recoded into three categories: Low (*“No formal education,” “Primary education,”* or *“Lower secondary education”*), Mid (*“Upper secondary education”* or *“Post-secondary education other than college/university”*), and High (*“Bachelor or equivalent,” “Master or equivalent,”* or *“Doctoral or equivalent”*).

3. 3. 2. Methods

As the main goal of this article is to gain insights into gender trajectories from a life course perspective, sequence analysis is employed as the primary analytical approach. Sequence

analysis is particularly well-suited for studying life course phenomena because it allows for a holistic examination of entire trajectories rather than focusing on isolated events (Billari & Piccarreta, 2005; Aisenbrey & Fasang, 2010). Unlike event history analysis, which examines the timing of single transitions, sequence analysis treats the life course as an integrated unit, capturing the timing, sequencing, and duration of multiple states (Abbott & Tsay, 2000). This is especially important for understanding gender trajectories, as they are characterized by heterogeneous patterns that unfold over extended periods (Aboim & Vasconcelos, 2021).

The main variables under analysis are the age of realization (when individuals first recognized their gender identity as something different from their sex assigned at birth), the age of sharing (when they first disclosed their identity to others), and, if applicable, the age at which body interventions occurred (such as hormone treatments or surgeries). Importantly, these milestones are linear in time: self-recognition precedes sharing, which precedes body interventions, and no alternative orderings of events are observed. For each age of a person's life, four possible states were coded: initial, self-recognition, disclosed, and body intervention. Sequences were created based on these states to represent each individual's gender trajectory (Gabadinho et al., 2011).

Figure 3.1. Example sequences of gender trajectories.

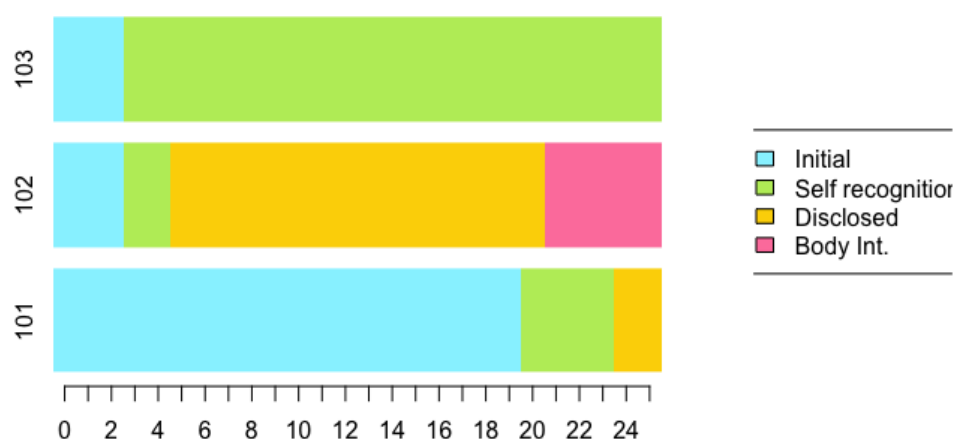
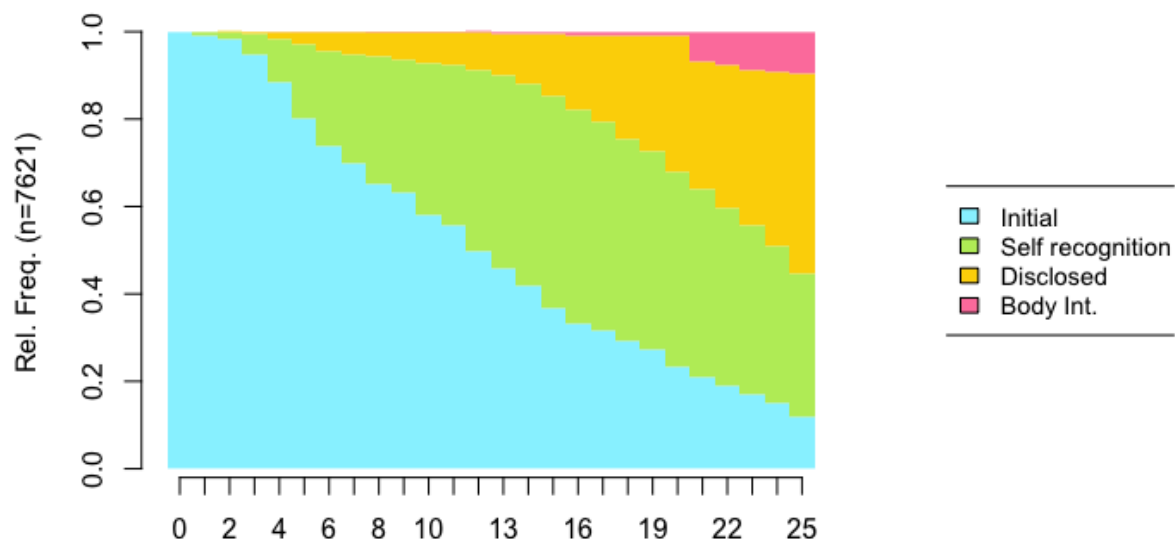


Figure 3.1 presents three example sequences. Sequence 101 illustrates the trajectory of an individual who self-recognized a gender different from the one assigned at birth by age 19

and remained in this state until disclosing it at age 23. Sequence 102 presents a different story: this individual realized their gender identity as early as age 3 and revealed it just a year later. Additionally, this sequence includes a body intervention at age 20. Sequence 103 offers the simplest scenario among the examples, showing an individual who recognized their gender identity difference at age 3 but, at least until age 25, never disclosed it to anyone. As can already be seen from these example sequences, gender trajectories can vary greatly across individuals. Figure 3.2 presents the state distribution by age for the entire sample.

Figure 3.2. State distribution by age for individuals 25 and older (n = 7,621).



To gain a better understanding of trajectory trends and address the research question, underlying patterns across the sample were analyzed. First, Optimal Matching (OM) distances were calculated for each pair of sequences. These distances reflect the cost of transforming one sequence into another (Abbott & Tsay, 2000). Using the calculated distances, hierarchical agglomerative clustering was performed following Ward's method to group sequences into clusters based on dissimilarity. This approach allowed the identification of similar patterns across the sample.

Cohort change was examined using multinomial regression models to predict the probability of belonging to each cluster by cohort. Controls for education, gender, and sex assigned at birth were included in the models. All analyses were conducted in R, using the following packages: *TraMineR* for sequence analysis, *cluster* for clustering, and *nnet* for multinomial modeling.

3. 4. Results

3. 4. 1. Descriptives

This descriptive analysis examines key characteristics across two cohorts: individuals born between 1985 and 1994, and those born before 1985; as presented in Table 3.1. The mean age at realization of gender identity is similar across cohorts (14.49 vs. 14.31 years). However, the mean age at disclosure is notably higher in the older cohort (28.16 years) than in the younger cohort (21.16 years). The mean age at medical intervention also follows this pattern, occurring earlier in the younger cohort (25.82 years) than in the older cohort (39.85 years).

In terms of gender identification, the majority in both cohorts identified as non-binary, 52.70% in the younger group vs. 37.21% in the older, which points to this identity being more prevalent in younger generations. A greater share of trans women was found in the older cohort (20.41% vs. 11.94%), while trans men were more prevalent in the younger cohort (14.37% vs. 10.13%), this is in line with findings in other countries such as Sweden (Kolk, 2025). The proportion of those who identified as men and women was generally low in both groups. A higher share of individuals in the older cohort (12.50%) had never disclosed their gender identity compared to the younger cohort (9.56%). However, the rate of medical intervention was higher in the older cohort (43.22%) than in the younger cohort (34.16%). Similarly, legal gender change was more common in the older cohort (36.27% vs. 27.67%). Education level was relatively similar between cohorts, with most individuals having a high level of education (55.26% in the younger cohort and 57.71% in the older cohort). Additional

sociodemographic information, including civil status, occupation, urban/rural residence, and involvement in LGBTQ+ collectives, is available in Appendix A, Table 3.A1 and provides complementary background for interpreting cohort-based differences.

Table 3.1. Descriptives by Cohort

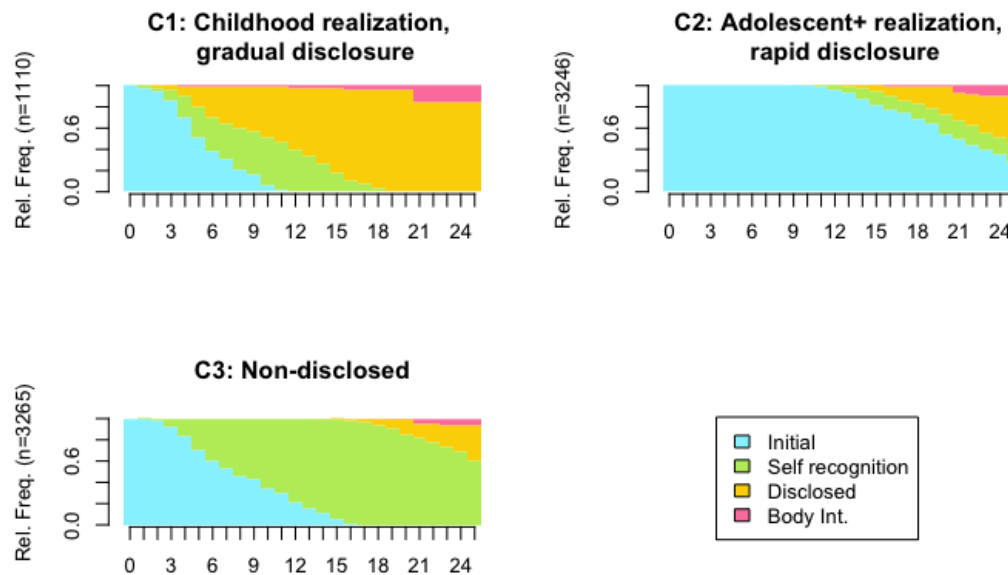
Variable	Cohort		Type
	1985-1994	Before 1985	
Mean Age	28.70	46.01	Mean
Mean Realization Age	14.49	14.31	Mean
Mean Disclosure Age	21.16	28.16	Mean
Mean Age Intervention	25.82	39.85	Mean
Non Disclosed	9.56	12.50	Share
Intervention	34.16	43.22	Share
Legal gender changed	27.67	36.27	Share
<u>Gender</u>			
Woman	9.77	18.07	Share
Man	9.25	9.54	Share
Trans Woman	11.94	20.41	Share
Trans Man	14.37	10.13	Share
Non-Binary	52.70	37.21	Share
Please specify...	1.97	4.64	Share
<u>Education</u>			
Low	7.50	9.82	Share
Mid	37.24	32.47	Share
High	55.26	57.71	Share

3. 4. 2. Gender trajectory typologies

The use of clusters to represent gender trajectory typologies allows for the grouping of individuals based on sequence similarity in terms of self-realisation, disclosure, and body interventions. This approach helps simplify the identification of similar trajectories across individuals, facilitating a clearer understanding of the complexity and diversity of gender identity development over time. To capture these distinct patterns, a solution with three

clusters was chosen, as presented in Figure 3.3. The number of clusters was determined by combining empirical interpretation with technical criteria⁸.

Figure 3.3. Cluster typologies of gender trajectories.



Cluster 1 consisted of individuals who recognized their gender identity during childhood. This group includes 1,110 individuals (14.6% of the sample). The trajectories within this cluster varied, with some individuals taking longer to disclose their gender identity while most did so at earlier stages of their lives. Cluster 2 includes individuals who realized their gender identity during their teenage years and disclosed it shortly afterward. This cluster consisted of 3,246 individuals (42.6%). Cluster 3 comprised individuals who, for various reasons, did not disclose their gender identity, either at all or after a prolonged period of time. This group consisted of 3,265 individuals (42.9%). These individuals spent a significant amount of time in a state of non-disclosure, reflecting potential social, personal, or cultural factors that may have delayed or constrained disclosure.

⁸ The silhouette score, a measure of cluster quality, indicated that three clusters provided the best solution in terms of balancing both the internal cohesion of the clusters and their separation from one another.

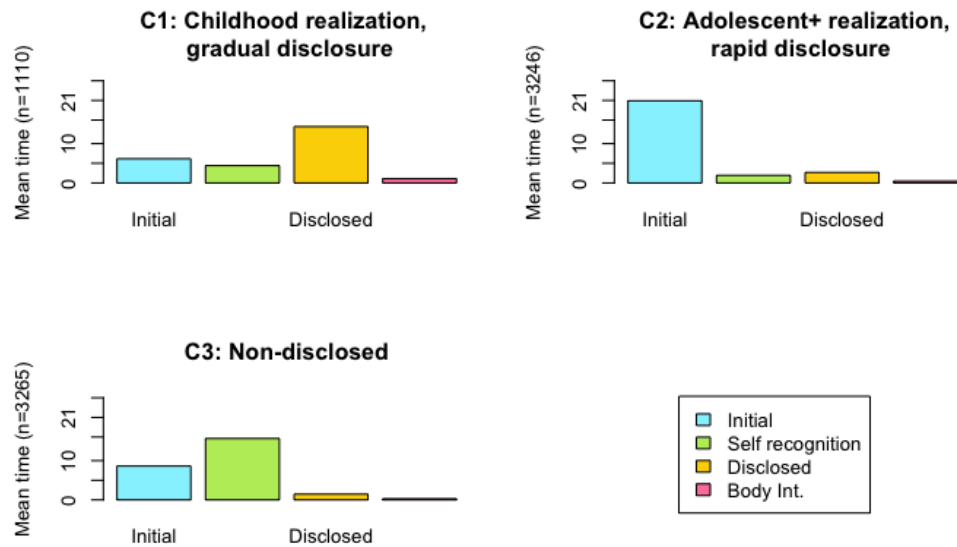
Figure 3.4. Mean time spent on each state, by cluster.

Figure 3.4 presents the mean time spent in each state by cluster. As shown, individuals in Cluster 1 spent the majority of their time in the disclosed state, after a short initial period before self-recognition, followed by disclosure. In Cluster 2, the longest mean time was observed in the initial state, reflecting that recognition of their gender identity occurred at a later stage compared to other clusters, followed by relatively brief periods from realization to disclosure. In Cluster 3, the longest duration was spent between self-recognition and disclosure, indicating a prolonged period of non-disclosure before, and if, disclosure occurred.

Table 3.2. Descriptives by cluster.

Variable	Cluster			Type
	C1: Childhood realization, gradual disclosure	C2: Adolescent+ realization, rapid disclosure	C3: Non-disclosed	
Mean Age	36.88	33.88	37.79	Mean
Mean Realization	6.15	23.09	8.61	Mean
Mean Disclosure age	10.24	24.99	28.64	Mean
Mean Age Intervention	31.13	30.03	34.83	Mean

Non Disclosed	0.00	9.09	16.17	Share
Body Intervention	45.77	30.44	42.82	Share
Legal gender changed	42.70	23.26	35.41	Share
<u>Gender</u>				
Woman	14.95	9.67	16.26	Share
Man	12.34	9.37	8.36	Share
Trans Woman	15.77	11.89	19.02	Share
Trans Man	16.04	11.18	12.80	Share
Non-Binary+	36.67	55.27	40.37	Share
Please specify...	4.23	2.62	3.19	Share
<u>Educational level</u>				
Low	10.09	7.98	8.42	Share
Mid	36.13	34.29	35.87	Share
High	53.78	57.73	55.71	Share

Table 3.2 presents descriptive statistics by cluster. Mean realization age differs notably across clusters, occurring earliest in Cluster 1 (6.15 years), later in Cluster 2 (23.09 years), and at an intermediate stage in Cluster 3 (8.61 years). Disclosure age follows a similar pattern: individuals in Cluster 1 disclosed their gender identity at a mean age of 10.24, while those in Cluster 2 did so at 24.99, and those in Cluster 3 at 28.64. Non-disclosure is concentrated in Cluster 3, where 16.17% of individuals never disclosed their gender identity. In contrast, 9.09% in Cluster 2 and no individuals in Cluster 1 remained undisclosed.

The mean age at medical intervention was highest in Cluster 3 (34.83 years), followed by Cluster 1 (31.13) and Cluster 2 (30.03). The proportion of individuals who underwent body interventions is highest in Cluster 1 (45.77%), followed by Cluster 3 (42.82%) and Cluster 2 (30.44%). Similarly, legal gender change was more frequent among Cluster 1 participants (42.70%), compared to 35.41% in Cluster 3 and 23.26% in Cluster 2.

In terms of gender identification, non-binary individuals represent the largest share in Cluster 2 (55.27%), compared to 36.67% in Cluster 1 and 40.37% in Cluster 3. Trans women are more prevalent in Cluster 3 (19.02%), whereas trans men are slightly more represented in Cluster 1 (16.04%). Education level is relatively similar across clusters, with most

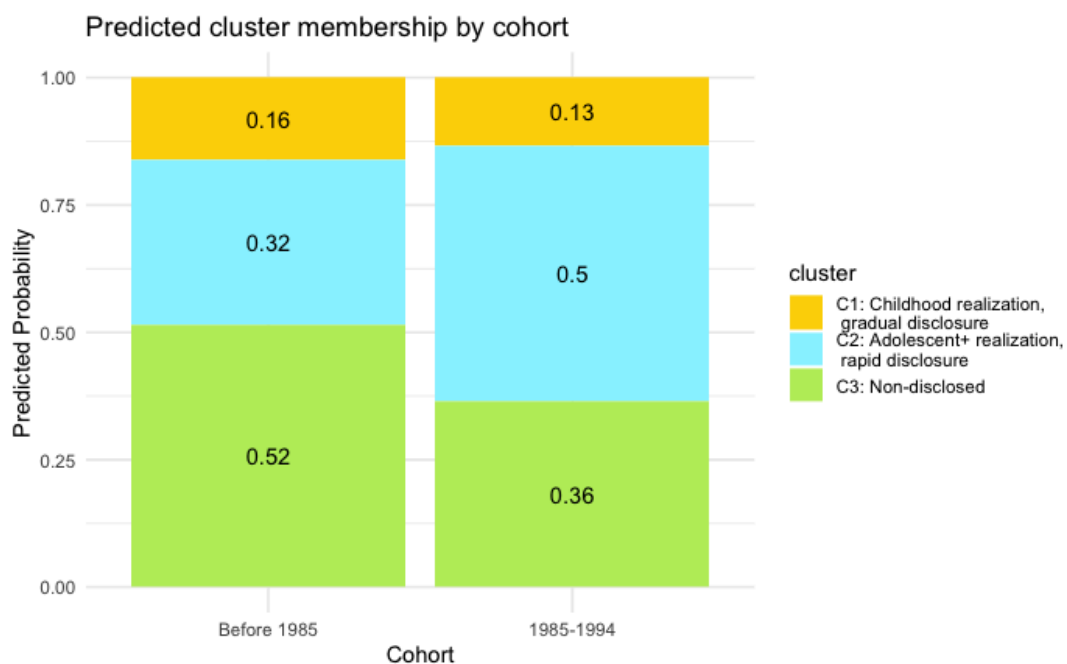
participants holding a high level of education (ranging from 53.78% in Cluster 1 to 57.73% in Cluster 2). Additional descriptive variables by cluster are provided in Appendix Table A2.

3. 4. 3. Cohort distribution across clusters

Figure 3.5 indicates that individuals born between 1985 and 1994 were most likely to belong to Cluster 2, characterized by realization occurring at a later stage (compared to Cluster 1) and a short interval between realization and disclosure. The predicted probability of membership in Cluster 2 for this cohort is 0.50, compared to 0.13 for Cluster 1 and 0.36 for Cluster 3.

For individuals born before 1985, the probability of belonging to Cluster 1 was nearly the same as in the younger cohort (0.16). However, their likelihood of being in Cluster 2 was substantially lower, with a predicted probability of 0.32. Notably, this older cohort was most strongly associated with Cluster 3, with a predicted probability of 0.52, suggesting that the majority of this cohort experienced trajectories characterized by extended intervals between gender identity realization and disclosure.

Figure 3.5. Predicted probability for cluster membership by birth cohort.



As shown in Table 3.3, cohort membership significantly shaped the relative likelihood of belonging to one cluster versus another. Individuals born between 1985 and 1994 were 1.8 times more likely to be in Cluster 2 rather than Cluster 1 ($p < 0.01$) and twice as likely to be in Cluster 2 rather than Cluster 3 ($p < 0.01$) compared to those born before 1985. This suggests that younger participants were more frequently represented in trajectories characterized by later realization and rapid disclosure.

Table 3.3. Relative risk ratio from multinomial regression analysis for cluster membership.

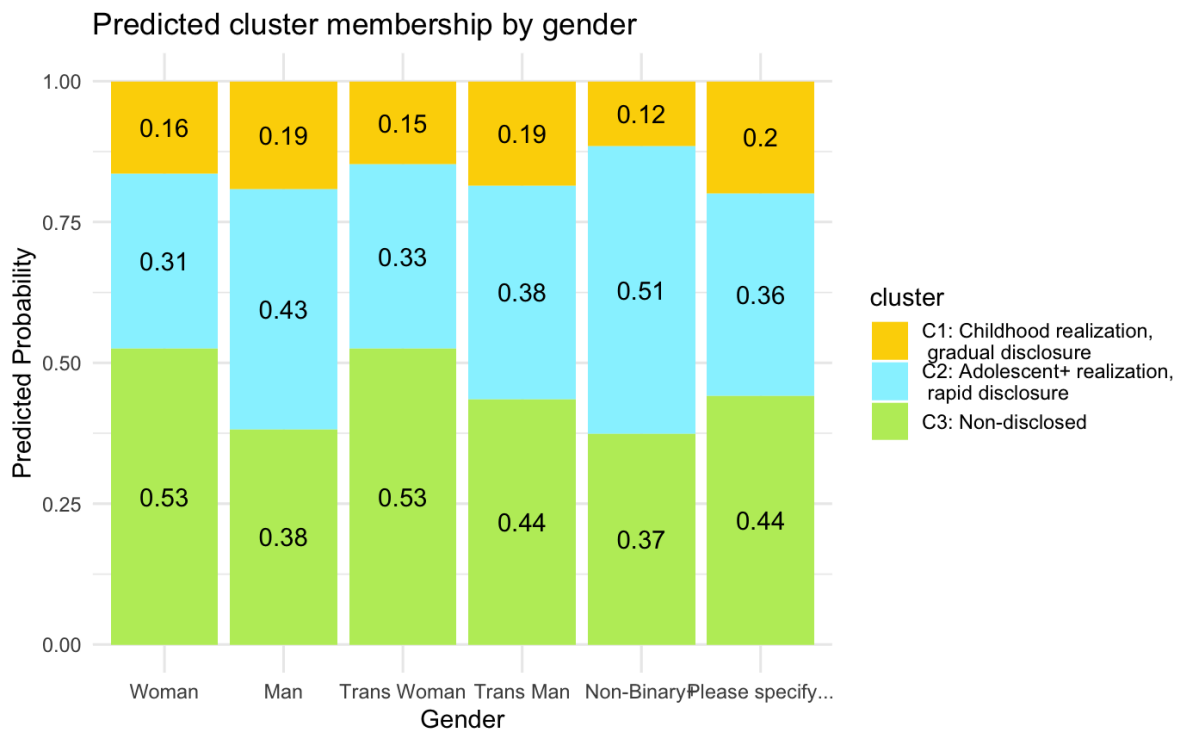
	C2 (ref: C1)	C3 (ref: C1)	C2 (ref: C3)
(Intercept)	0.84	1.88***	0.45***
cohort: 1985-1994	1.79***	0.88*	2.03***
<u>Gender</u>			
Man	1.28*	0.73**	1.76***
Trans Woman	1.05	1.02	1.024
Trans Man	1.35**	1.04	1.30**
Non-Binary+	2.71***	1.30**	2.09***
Please specify...	1.29	0.85	1.52**
<u>Sex assigned at birth</u>			
Male	1.82***	1.70***	1.07
Other, please specify	0.95	1.09	0.864
<u>Education</u>			
Mid	1.08	1.23	0.88
High	1.18	1.28**	0.919

Notes: Ref. Categories: Cohort before 1985, gender woman, SAAB woman, education low. Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.

Gender identity was also associated with cluster membership (Figure 3.6). Non-binary respondents most often belonged to Cluster 2 (Teen+ realization, fast disclosure) (predicted probability ≈ 0.51), with comparatively lower probabilities for Cluster 3 (Non-disclosed) (0.37) and Cluster 1 (Early realization, disclosed) (0.12). Women and trans women showed the highest probabilities of non-disclosure (Cluster 3 ≈ 0.53 in both groups). Men displayed a relatively higher probability of Cluster 2 (0.43) and a lower probability of Cluster 3 (0.38).

Trans men fell in between, with Cluster 3 (0.44) and Cluster 2 (0.38) as the most common outcomes; the “please specify” group showed a similar split (Cluster 3 0.44, Cluster 2 0.36). Across all groups, Cluster 1 remained comparatively small (≈ 0.12 – 0.20).

Figure 3.6. Predicted probability for cluster membership by gender identity.



3. 4. 4. Additional analysis

To complement the main analysis based on cohort differences, the distribution of clusters was examined by country to assess whether national-level contexts were associated with differences in gender trajectory typologies. Countries were compared according to their legal and policy environments, using ILGA World Legal Gender Recognition Database (2023) scores as a proxy for the degree of legal protection and social recognition of gender-diverse individuals. The results, which can be found in Appendix B, did not reveal systematic differences in cluster representation across countries with higher versus lower legislative scores.

Also, an event history analysis of disclosure was conducted to further examine cohort differences in the timing of this gender milestone. Restricting the sample to individuals who had experienced self-realisation by age 25 to increase comparability, Kaplan–Meier estimates confirm that the younger cohort had a significantly higher probability of disclosure by age 25 (see Appendix C). Whereas most people in the younger cohort who realized before age 25 also disclosed before age 25, this was the case for a bit more than half of people in the older cohort. An event-history model directly testing whether the time between self-realisation and disclosure is shorter for younger cohorts confirmed this result (i.e. an event-history model explaining disclosure, where the clock starts at the age at self-realisation; see Appendix C).

This additional analysis reinforces the robustness of the main findings, indicating that the three identified trajectory typologies are broadly stable across legislative contexts, and that cohort differences remain the primary source of variation in the timing and pacing of gender milestones, in particular the time that elapses between self-realisation and disclosure.

3. 5. Conclusions and Discussion

This study applied a life-course perspective to move beyond the analysis of singular gender milestones and, for the first time, model complete gender trajectories across cohorts to look at lived experiences in a more holistic way. In doing so, we confirmed that there is not a single "transgender trajectory" but rather a set of distinct, but common, pathways. We identified three different paths: Cluster 1 (14% of the sample) represents individuals who recognized their gender identity during childhood and presented diverse disclosure timings. Cluster 2 (42%) is characterized by later realization, around or after adolescence, with a short period between internal and external identification. Cluster 3 (43%) includes individuals who experienced prolonged periods of non-disclosure, regardless of the timing of realization.

Our findings on the older, pre-1985 cohort (who were more likely to present non-disclosed trajectories) empirically confirm the "classic narrative" of gender transition long-established

in the literature (Baams et al., 2025). In this sense, self-realisation happened around late childhood (mean age 8.6), but was followed by a long, latent period with no external milestones before age 25. These results coincide with previous findings for Generation X, where the mean self-realisation age varied between 7.89 (Puckett et al., 2022) and 9.35 (Wilkinson et al., 2018), while the coming out happened later in life, around age 26.78 (Scandurra et al., 2021). The decades-long gap that these individuals experience between internal and external identification is probably not a personal choice but a reflection of societal barriers, fear of rejection, and a lack of access to affirmative care.

The younger cohort (1985-1994) is mostly present in the "Adolescent+ realization, rapid disclosure" path, which is characterized not by an earlier realization, but by a later one (mean age 23.1) followed by a very short gap to disclosure (mean age 25.0). This finding challenges the simple assumption that younger cohorts just "start the same path, but earlier", as found in previous studies (Puckett et al., 2022; Wilkinson et al., 2018; Scandurra et al., 2021). It suggests that for the majority of individuals in this generation the very nature of "realization" has changed.

As Nordmarken (2023) argues, identity formation is a social process. With increased societal visibility (Lagos, 2022) and the available language to describe non-binary and trans identities, younger individuals may not "realize" in isolation at age 8. Instead, they "come into identity" as adolescents or young adults with the concepts and models already available, leading to a short period between realization and disclosure. At the same time, when looking at cluster membership by gender identification, it is valuable to notice that non-binary individuals also are mostly represented by this same cluster. This is a major addition to the discussion, suggesting that realization age is not as stable as previously thought (Wilkinson et al., 2018), but also dependent on the context and gender identification of each individual.

It is also possible that recent increase in trans, non-binary, and gender diverse identification (Lagos, 2022) made these identities more available. For individuals who are going through

such formative periods as adolescence, this might have had a bigger impact on their identity development than for older cohorts, who went through this “transgender tipping point” at a later, more stable, point in life. This could point to a period effect that was more consequential for the younger cohort.

Despite the contributions of this study, some limitations must be acknowledged. First, the data is retrospective, which introduces a risk of recall bias. This is particularly relevant for the timing of internal milestones like realization, which may be reinterpreted in light of subsequent events. In this sense, earlier ages of realization in the older cohort could be related to recall bias, as the time span between this milestone and survey completion was longer.

Methodologically, the sequence states are also a necessary simplification. "Disclosure" (telling one person) is a specific event that happens at a point in time, and is not the same as the process of "living openly" (social transition). Similarly, "Body Intervention" groups both hormone treatment and various surgeries, which are, in reality, distinct pathways with different timings and barriers. Also, these sequences fall short to represent other events and processes associated with gender trajectories and identification, such as legal rectification or gender expression.

Beyond these measurement limitations, the non-probabilistic nature of the sample needs to be considered. It is plausible that individuals who have recently realized or disclosed their gender identity might be more motivated to participate in an LGBTQ+ survey than those for whom these events are in the distant past. This could introduce a selection bias, particularly within the younger cohort, potentially overrepresenting those currently navigating these specific "Adolescent realization with rapid disclosure". However, despite this limitation, the FRA survey remains the most comprehensive dataset currently available to study gender trajectories. Crucially, our findings are robust to alternative specifications that restrict the sample to people who experienced realization before age 25: this complementary

event-history analysis confirmed that the duration between realization and disclosure has significantly shortened across cohorts.

A final limitation that opens up the possibility for future research is the absence of retrospective information on other life-course processes. The FRA survey, while providing excellent data on identity milestones, does not include longitudinal information on partnership formation, entry into the labor market, or other key transitions to adulthood, such as leaving the parental home or completing education. As the Life Course Perspective (Elder et al., 2003) posits, life domains are intertwined. Our findings suggest a link, for instance, with the "Rapid disclosure" aligning with the transition to adulthood, but we cannot test it. In this sense, future research could move towards multi-domain sequence analysis, as this would need new data efforts. For example, how does progressing through the educational system (e.g., attending university) provide new social networks or resources that facilitate a rapid alignment between internal and external identification? How does being in a supportive partnership influence the pacing of affirmation, perhaps by providing the emotional resources to shorten the gap? And crucially, looking at the inverse relationship, how do these different trajectory typologies subsequently shape educational attainment, socioeconomic outcomes, and long-term labor market integration? Understanding this two-way relationship, how life events provide contexts for affirmation, and how transition trajectories, in turn, shape future life outcomes, would be a great next step.

These findings raise critical questions about the responsiveness of current policy and legal systems to the lived realities of younger trans, non-binary, and gender-diverse individuals. The analysis revealed consistent cluster membership patterns across diverse national contexts (see Appendix B), with limited variation driven by country of residence and legislative score. This suggests that the generational acceleration of gender milestones might be a transnational phenomenon, driven by shared global factors, such as digital connectivity and cultural shifts, that transcend borders. Consequently, it appears that legal and healthcare systems across Europe, regardless of their specific national frameworks, are

collectively struggling to keep pace with the accelerated timing and needs of this younger generation.

The trajectory characterized by realization occurring in late adolescence or early adulthood, followed almost immediately by disclosure reflects a generation for whom increased societal visibility has fundamentally transformed the process of gender identification. Unlike older cohorts who spent years, sometimes decades, between internal awareness and external expression, younger individuals are coming into their identities with less time spent in what might be called a state of incongruence between internal and external identification. This compression of the timeline creates new needs: when realization and disclosure nearly overlap, individuals might opt for immediate access to documentation changes, gender-affirming healthcare, and legal recognition to maintain congruence across all domains of their lives. However, many European healthcare and legal systems remain structured around gatekeeping models that require psychiatric diagnoses, multiple assessments of "mental readiness," and prolonged waiting periods, in some countries extending from 6 months to over 7 years (TGEU, 2022; Verbeek et al., 2022). These frameworks, rooted in pathologizing approaches that position mental health professionals as arbiters of who is "trans enough" to access care (Verbeek et al., 2022), are structurally misaligned with how younger people actually experience and navigate their gender identities. The policy implication is clear: systems must evolve toward models of informed consent and self-determination that accommodate not just one trajectory, but multiple pathways, with particular attention to reducing barriers for external gender affirmation.

In conclusion, this study advances knowledge on gender identity trajectories by identifying common patterns and examining cohort-based variations. The findings highlight how younger individuals in general experienced shorter intervals between gender realization and disclosure, whereas older cohorts often underwent prolonged periods of non-disclosure. These insights underscore the importance of historical and social contexts in shaping gender development processes. This study provides empirical, life-course evidence for what

transgender and gender-diverse communities have long known: there is no single path to gender affirmation.

3. 6. References

- Abbott, A., & Tsay, A. (2000). Sequence analysis and optimal matching methods in sociology: Review and prospect. *Sociological Methods & Research*, 29(3), 3-33.
- Aboim, S., & Vasconcelos, P. (2021). Gender trajectories and the production of inequalities from a life course perspective. In M. Nico & G. Pollock (Eds.), *The Routledge Handbook of Contemporary Inequalities and the Life Course* (pp. 338-348). Routledge.
- Aisenbrey, S., & Fasang, A. E. (2010). New life for old ideas: The 'second wave' of sequence analysis bringing the 'course' back into the life course. *Sociological Methods & Research*, 38(3), 420-462.
- Baams, L., Caprinali, A., & Fischer, M. (2025). Navigating time: A systematic review of the timing of sexual and gender identity milestones. In E. D. Rothblum (Ed.), *Sexually and gender diverse adolescents* (pp. 145–163). Amsterdam University Press.
- <https://library.oapen.org/bitstream/handle/20.500.12657/100842/9781040353974.pdf>
- Barrientos Delgado, J., Saiz, J. L., Guzmán-González, M., Bahamondes, J., Gómez, F., Castro, M. C., ... & Giami, A. J. (2021). Sociodemographic characteristics, gender identification, and gender affirmation pathways in transgender people: a survey study in Chile. *Archives of Sexual Behavior*, 50(8), 3505-3516.
- Bercovich Szulmajster, S. (2025). Categorizing gender beyond the binary: inequalities in education from a multidimensional gender perspective. *Genus*, 81(1), 1-20.
- Billari, F. C., & Piccarreta, R. (2005). Analyzing demographic life courses through sequence analysis. *Mathematical Population Studies*, 12(2), 81-106.

Caprinali, A., & Vitali, A. (2025). The demography of sexual identity development and disclosure among LGB people in Europe. *Demographic Research*, 52, 125-140.

Cornell University. (2021). What we know: The research on transgender health. Cornell University, What We Know Project.

<https://whatweknow.inequality.cornell.edu/topics/lgbt-equality/what-does-the-scholarly-research-say-about-the-well-being-of-transgender-people/>

Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the Life Course* (pp. 3-19). Springer.

European Union Agency for Fundamental Rights (FRA) (2021). The EU LGBTI II Survey, 2019. GESIS Data Archive, Cologne. ZA7604 Data file Version 1.1.0, <https://doi.org/10.4232/1.13733>.

Factor, R., & Rothblum, E. (2008). Exploring gender identity and community among three groups of transgender individuals in the United States: MTFs, FTMs, and genderqueers. *Health Sociology Review*, 17(3), 235-253.

Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of statistical software*, 40, 1-37.

ILGA-Europe. (2024). Annual Review of the Human Rights Situation of Lesbian, Gay, Bisexual, Trans and Intersex People in Europe and Central Asia. ILGA-Europe.

ILGA-Europe & TGEU. (2023). Intersections: The LGBTI II Survey – Trans and Non-binary Analysis. Brussels: ILGA-Europe.

ILGA World. (2023). Legal gender recognition: Data and resources on laws, policies and practices. ILGA World. <https://database.ilga.org/legal-gender-recognition>

Klein, H., & Washington, T. A. (2023). Transition milestones and psychological distress in transgender adults. *Urban Social Work*, 7(1), 66-84.

Kolk, M., Tilley, J. L., von Essen, E., Moberg, Y., & Burn, I. (2025). The Demography of Sweden's Transgender Population: A Research Note on Patterns, Changes, and Sociodemographics. *Demography*, 62(2), 349-363.

Kosciw, J. G., Palmer, N. A., & Kull, R. M. (2015). Reflecting resiliency: Openness about sexual orientation and/or gender identity and its relationship to well-being and educational outcomes for LGBT students. *American journal of community psychology*, 55(1), 167-178.

Lagos, D. (2022). Has there been a transgender tipping point? Gender identification differences in US cohorts born between 1935 and 2001. *American Journal of Sociology*, 128(1), 94-143.

Levitt, H. M., & Ippolito, M. R. (2014). Being transgender: The experience of transgender identity development. *Journal of homosexuality*, 61(12), 1727-1758.

Nordmarken, S. (2023). Coming into identity: How gender minorities experience identity formation. *Gender & Society*, 37(4), 584-613.

Puckett, J. A., Tornello, S., Mustanski, B., & Newcomb, M. E. (2022). Gender variations, generational effects, and mental health of transgender people in relation to timing and status of gender identity milestones. *Psychology of Sexual Orientation and Gender Diversity*, 9(2), 165.

Scandurra, C., Carbone, A., Baiocco, R., Mezzalana, S., Maldonato, N. M., & Bochicchio, V. (2021). Gender identity milestones, minority stress and mental health in three generational cohorts of Italian binary and nonbinary transgender people. *International Journal of Environmental Research and Public Health*, 18(17), 9057.

TGEU (Transgender Europe). (2021). Trans discrimination in Europe: A TGEU analysis of the FRA LGBTI Survey 2019. Berlin: TGEU.

TGEU (Transgender Europe). (2022). Trans Health Map 2022: The State of Trans Healthcare in the EU. Berlin: TGEU. <https://tgeu.org/trans-health-map-2022/>

Verbeek, W., Baici, W., MacKinnon, K. R., Zaheer, J., & Lam, J. S. H. (2022). "Mental readiness" and gatekeeping in trans healthcare. *The Canadian Journal of Psychiatry*, 67(11), 828-830. <https://doi.org/10.1177/07067437221102725>

Wilkinson, L., Pearson, J., & Liu, H. (2018). Educational attainment of transgender adults: Does the timing of transgender identity milestones matter?. *Social Science Research*, 74, 146-160.

3. 7. Appendix

3. 7. 1. Appendix A. Additional Descriptive Characteristics

Appendix Table 3.A1 presents additional descriptive characteristics by cohort, including civil status, occupational status, place of residence (urban/rural), and involvement in LGBTQ+ collectives. These variables are not directly analyzed in the main text but provide useful background information about the sample composition. Appendix Table 3.A2 presents the same variables by cluster.

Appendix Table 3.A1. Additional descriptive variables by cohort

Variable	Cohort		Type
	1985-1994	Before 1985	
<u>Occupational Status</u>			
Employed	56.75	68.15	Share
Other	15.28	16.11	Share
Unemployed	10.27	7.26	Share
Student	17.29	1.99	Share
Retired	0.41	6.48	Share
<u>Urban/Rural</u>			
Urban	90.80	86.04	Share
Rural	9.13	13.93	Share
<u>Civil Status</u>			
Single	82.41	48.92	Share
Married/ registered partnership (same-sex)	8.23	15.14	Share
Married/ registered partnership (different sex)	6.69	16.61	Share
Divorced	1.43	12.06	Share
Separated	1.11	5.42	Share
Widowed	0.14	1.84	Share
<u>Involved in LGBTQ+ Collective</u>			
Yes	47.78	53.01	Share
No	52.22	46.99	Share

Appendix Table 3.A2. Additional descriptive variables by cluster

Variable	Cluster			Type
	C1: Childhood realization, gradual disclosure	C2: Adolescent+ realization, rapid disclosure	C3: Non-disclosed	
<u>Occupation</u>				
Employed	62.25	60.41	62.45	Share
Other	16.49	14.57	16.39	Share
Unemployed	7.57	9.18	9.31	Share
Student	10.27	13.34	8.58	Share
Retired	3.42	2.50	3.28	Share
<u>Location</u>				
Urban	87.48	90.26	87.78	Share
Rural	12.43	9.64	12.22	Share
<u>Civil status</u>				
Single	65.86	72.89	64.59	Share
Married/ registered partnership with same-sex partner	12.52	10.35	11.45	Share
Married/ registered partnership with different sex partner	11.17	9.98	11.64	Share
Divorced	7.03	4.31	7.11	Share
Separated	2.61	1.94	4.01	Share
Widowed	0.81	0.52	1.19	Share
<u>Involved in LGBTQ+ Collective</u>				
Yes	53.69	49.01	49.68	Share
No	46.31	50.99	50.32	Share

3. 7. 2. Appendix B. Country-level variation in cluster representation

To explore whether national legal frameworks are associated with differences in gender trajectory typologies, additional analyses were conducted at the country level. Each country was assigned a legal recognition score ranging from 0 to 4, reflecting the extent to which gender recognition procedures are available and accessible.

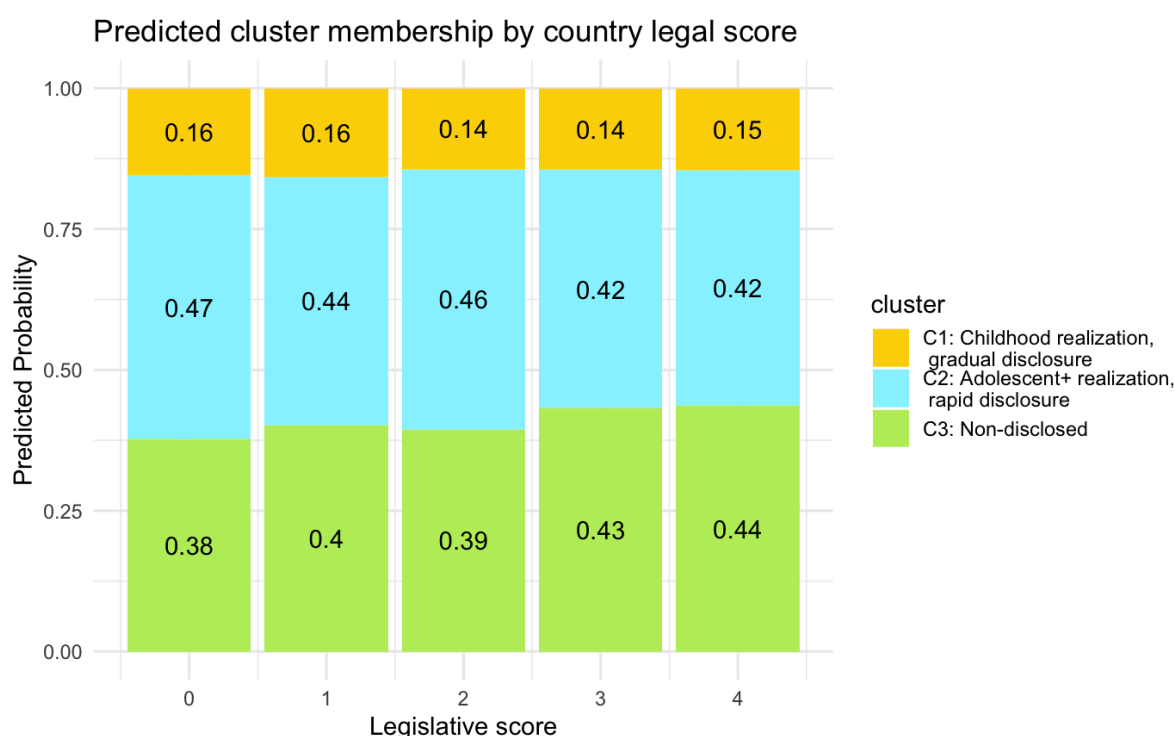
This score was derived from the ILGA World Legal Gender Recognition database (2023) and is based on the presence of four key legal provisions:

1. the possibility of name change,
2. the possibility of gender marker change,
3. the existence of an established legal procedure for gender recognition, and
4. whether such a process is based in self-identification.

Countries fulfilling all four conditions, such as Belgium, Denmark, Finland, Germany, Ireland, Luxembourg, Malta, Portugal, and Spain, scored 4, representing the most inclusive frameworks. Countries meeting three of the four conditions (e.g., France, Italy, the Netherlands, Sweden, United Kingdom) scored 3, typically lacking full self-determination. Intermediate cases (score = 2) including Austria, Poland, and Romania have partial procedures in place, while North Macedonia (1), Bulgaria (0), and Hungary (0) lack comprehensive legal mechanisms for gender recognition.

Overall, the proportions remain similar across the entire range of legal recognition scores, suggesting that the availability of legal gender recognition mechanisms does not substantially alter the distribution of gender trajectory typologies. Within the European context captured by the 2019 FRA LGBTQ+ Survey, cohort and individual-level factors appear to play a more decisive role than national legislation in shaping the timing and sequence of gender milestones.

Figure 3.B1. Predicted cluster membership by country legal recognition score (0–4)



Source: Author's calculations based on the ILGA World Legal Gender Recognition Database (2023).

Predicted probabilities of cluster membership were then calculated for each score level (Figure 3.B1). The results show only minor variation across groups: individuals in countries with the lowest score (0) had a 0.38 probability of belonging to Cluster 3 (Non-disclosed), 0.16 for Cluster 1 (Childhood realization, gradual disclosure), and 0.47 for Cluster 2 (Adolescent+ realization, rapid disclosure). Among countries with the most comprehensive legal systems (score = 4), the corresponding probabilities were 0.44, 0.15, and 0.42, respectively. Detailed cluster membership prediction by country can be found in Table 3.B1.

Appendix Table 3.B1. Cluster membership prediction by country.

Legislative Score	Country	Childhood realization	Rapid disclosure	Non-Disclosed
4	Belgium	0.14	0.41	0.45
	Denmark	0.15	0.40	0.45
	Finland	0.15	0.43	0.42
	Germany	0.15	0.40	0.45
	Ireland	0.14	0.41	0.44
	Luxembourg	0.16	0.39	0.45
	Malta	0.16	0.39	0.45
	Portugal	0.14	0.44	0.42
	Spain	0.13	0.45	0.41
3	Croatia	0.15	0.43	0.42
	Czech Republic	0.15	0.44	0.41
	Estonia	0.14	0.44	0.42
	France	0.14	0.46	0.41
	Greece	0.15	0.43	0.42
	Italy	0.14	0.46	0.41
	Netherlands	0.15	0.41	0.45
	Serbia	0.14	0.45	0.41
	Slovakia	0.15	0.42	0.43
	Slovenia	0.13	0.47	0.40
	Sweden	0.14	0.44	0.42
	United Kingdom	0.15	0.40	0.45
2	Austria	0.15	0.42	0.43
	Cyprus	0.15	0.49	0.35
	Latvia	0.14	0.48	0.38
	Lithuania	0.13	0.50	0.38
	Poland	0.14	0.48	0.38
	Romania	0.14	0.45	0.40
1	North Macedonia	0.16	0.44	0.40
0	Bulgaria	0.15	0.51	0.35
	Hungary	0.16	0.46	0.38

3. 7. 3. Appendix C. Event History Analysis

To complement the cluster-based results, an event history analysis was conducted to examine the timing of disclosure among individuals who experienced self-realisation by age 25. Since individuals who did not yet realize they are transgender or non-binary are less likely to be part of the sample, we excluded individuals whose self-realisation occurred after age 25 and like that make the cohorts more comparable. Since the younger cohort has not yet reached older ages (e.g., 40+) at the time of the survey, they have simply not had the opportunity to experience a late realization. Including late-life realizations, which are only observable for the older cohort, would introduce a bias. Restricting both groups to the same 0–25 age window allows to compare their trajectories on equal ground.

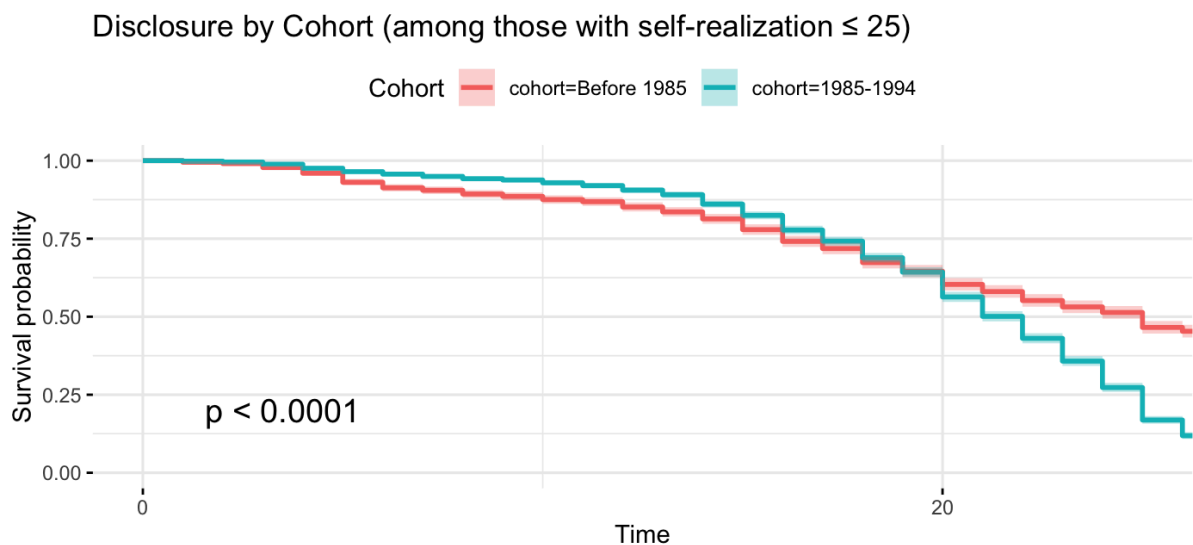
Kaplan–Meier survival curves (Figure 3.C1) show clear cohort differences in the probability of disclosure across the life course. By age 25, the younger cohort (born 1985–1994) displayed a higher cumulative probability of having disclosed their gender identity compared to the older cohort (born before 1985). These findings are consistent with the broader pattern observed in the main analyses, indicating that younger generations tend to disclose their gender identity earlier in life, probably reflecting changing social and cultural contexts that increasingly enable earlier expression.

When interpreting these results, it is important to note that the event history analysis was restricted to individuals who had experienced self-realisation by age 25. This group represents 91.5% of the younger cohort (born 1985–1994) and 83.7% of the older cohort (born before 1985). Accordingly, 8.5% of the younger cohort and 16.3% of the older cohort were excluded from this analysis because they had not yet reported self-realisation by that age.

This means that the curves reflect the pace of disclosure among those who already recognized their gender identity relatively early in life, rather than the overall probability of disclosure across the population. Because younger cohorts include individuals who may not

yet have reached this stage, the comparison focuses on the timing of disclosure, rather than on absolute rates. Within this subset, the younger cohort shows a steeper decline in non-disclosed probability, indicating that once realization occurs, disclosure tends to follow more rapidly among younger individuals than among older ones.

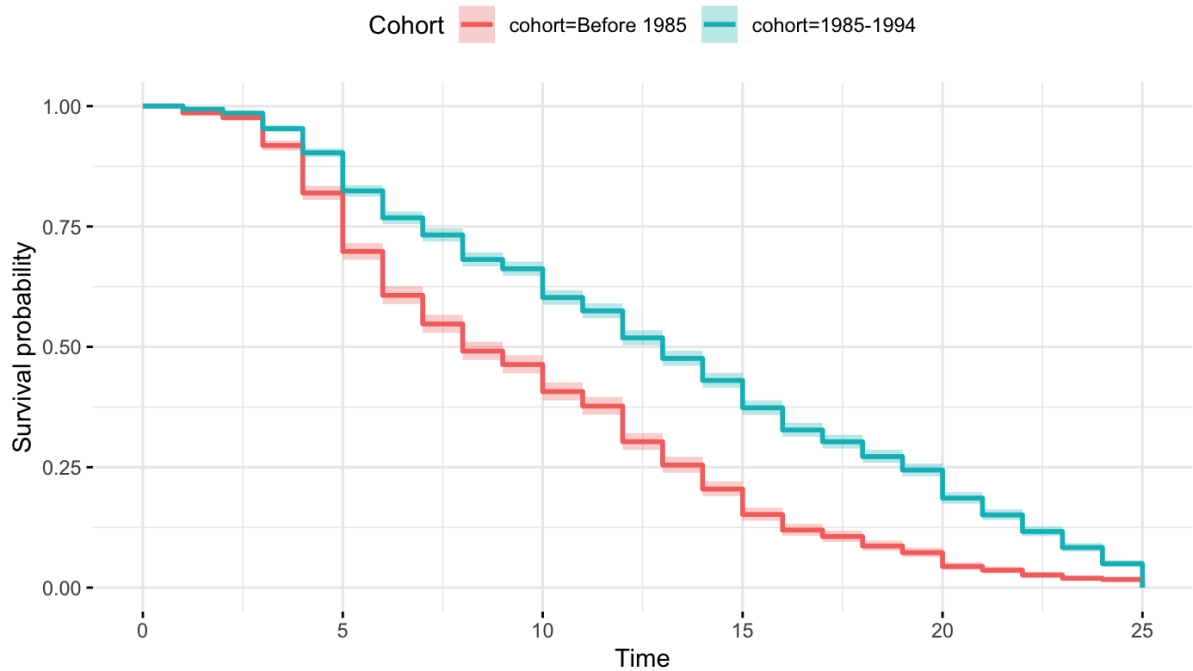
Figure 3.C1. Kaplan–Meier survival curves for disclosure by cohort (restricted to individuals with self-realisation ≤ 25)



An additional Kaplan–Meier estimation was performed for self-realisation, once again restricted to individuals who had experienced this milestone by age 25 (Figure 3.C2). Since this restriction implies that all events occurred within the observed window, the curves represent the cumulative distribution of realization ages rather than survival probabilities. The curves show that the older cohort (born before 1985) reached self-realisation earlier than the younger cohort (born 1985–1994). At any given age below 25, a higher proportion of younger respondents had not yet recognized their gender identity.

Figure 3.C2. Kaplan–Meier survival curves for self-realisation by cohort (restricted to individuals with self-realisation ≤ 25)

Self-realization by Cohort (≤ 25 years)



This difference should be interpreted with caution. Potential recall bias may have influenced older respondents' retrospective reporting of when they first recognized their gender identity. Individuals in older cohorts, having had more time to reconstruct their life narratives, may recall this milestone as having occurred earlier than it actually did. Therefore, differences between cohorts in self-realisation timing likely reflect a combination of age at observation and retrospective reinterpretation, rather than purely generational change.

Figure 3.C3. Kaplan–Meier survival curves for lapse between self-realisation and disclosure by cohort (restricted to individuals with self-realisation ≤ 25)

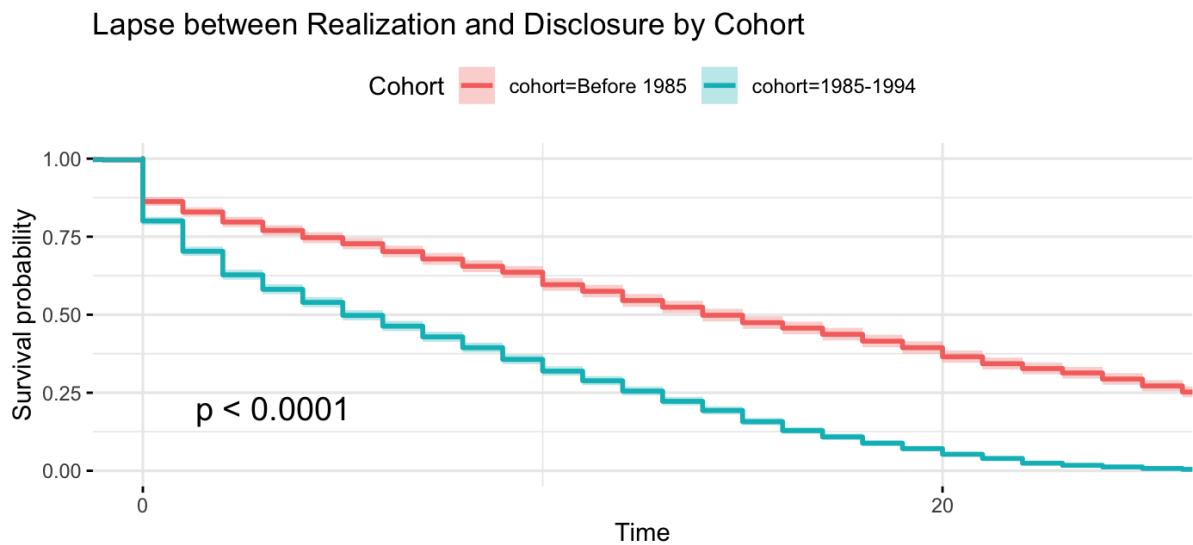


Figure 3.C3 presents the Kaplan-Meier curves for the duration in years between self-realisation and disclosure. Once again, it can be directly observed that for the younger generation the lapse is in general shorter, with half of individuals disclosing around or before 5 years after they realized. In contrast, the older cohort (born before 1985) exhibits a much more prolonged duration between these milestones. For this group, the median time to disclosure extends to approximately 13 years. Moreover, the curve for the older cohort flattens out at a significantly higher level, indicating that around 20% of individuals had not yet disclosed their identity even 20 years after first realizing it.

Chapter 4: Beyond the Swipe: Gender Diversity and Interactions in Dating Apps

Co-authored with Francesco Rampazzo & Diederik Boertien

Abstract

Dating apps have become a central space for meeting romantic partners, yet little quantitative research examines how transgender, non-binary, and other gender-diverse individuals navigate these platforms. Using behavioural data from 34,139 users of Zoe, an LGBTQ+ dating app for queer womxn in the United Kingdom (2023-2025), we investigate two key questions: Do gender diverse users show different levels of engagement compared to other groups? And do they receive different returns from their engagement in the form of matches? Negative Binomial regression models reveal that gender diverse users engage substantially more with the platform, with trans women giving approximately three times as many likes as women, and non-binary and other gender-diverse users also showing elevated activity. However, this heightened engagement does not translate into proportionate matches. Ordinary Least Squares regression and Coarsened Exact Matching analyses demonstrate that gender diverse users experience systematically lower like-to-match conversion rates, even after accounting for age, profile characteristics, and liking patterns. These disparities persist within an explicitly LGBTQ+-designed platform, suggesting that structural exclusion and potentially algorithmic bias operate even in self-described inclusive spaces. Our findings show how digital dating reproduces inequalities affecting gender diverse populations, with implications for understanding relationship formation disparities, well-being outcomes, and questions of algorithmic justice in platforms that increasingly mediate intimate life.

Keywords: Gender diversity, dating, digital trace data, inequalities

4. 1. Introduction

Dating and relationships are important parts of life for large parts of the population, and are related to many facets of well-being (Reczek, 2020; Rodríguez Sánchez & Boertien, 2025). Understanding differences in dating experiences is therefore an important goal towards understanding inequalities in well-being. In this regard, people who challenge norms in society might experience specific obstacles when dating. In this article, we focus on people who challenge binary cisgender norms and aim to better understand their dating experiences. With dating apps becoming a central space for meeting romantic partners (Hobbs, Owen, & Gerber, 2017), we use data from the LGBTQ+ dating app Zoe to understand how gender diversity relates to activity on dating apps and the number of matches people receive on apps.

Existing research has shown how transgender, non-binary and other gender diverse people experience particular challenges when dating. For instance, a Canadian study showed how only a small share of the population stated to be open towards dating a transgender person (Blair & Hoskin, 2019). In addition, gender diverse people have to deal with more violence and safety issues once dating than binary cisgender individuals (Webster et al., 2025). Online dating has become the most important way in which certain marginalized groups date and find partners (Rosenfeld & Thomas, 2012; Sironi & Kashyap, 2022). Even though obstacles experienced by gender diverse people are present in both offline and online dating, online dating can make it easier to find dating partners and to block or ignore people who behave in a hostile manner (Sharabi et al., 2023). Some of the issues related to online dating, such as the possibility to disclose an identity or feeling safe, are arguably less present in apps specifically designed for dating among LGBTQ+ people (Fernandez & Birnholtz, 2019; Sharabi et al., 2023).

One of the questions that arises from these observations is the extent to which gender diverse people experience obstacles within LGBTQ+ dating apps. To our knowledge, there

are no quantitative studies that have aimed to quantify how app use and the number of matches received differ between gender diverse and binary cisgender people (both in general, nor specifically within LGBTQ+ dating apps). To fill this gap in the literature, we use data from the app Zoe which is designed for use by primarily LGBTQ+ women. We aim to answer the following questions: Do transgender, non-binary and other gender diverse people show higher levels of engagement in LGBTQ+ apps as compared to cisgender binary women? And, do gender diverse people receive higher or lower “returns” from engagement in the form of matches compared to cisgender binary women?

On the one hand, the opportunities that LGBTQ+ dating apps provide can increase the engagement of gender diverse people in such apps, as alternative ways of meeting people can be less available or unsafe (Griffiths & Armstrong, 2024; Webster et al., 2025). On the other hand, gender diverse individuals also report unpleasant and hostile experiences within LGBTQ+ dating apps (Fernandez & Birnholtz, 2019; Sharabi et al., 2023), and even though more LGBTQ+ people are open towards dating transgender people, large shares also are not (Blair & Hoskin, 2019). In addition, LGBTQ+ dating apps can limit the available options of potential dating partners. This can discourage gender diverse people from engaging with LGBTQ+ dating apps, or lead to lower “success rates” in terms of matches they receive. An important question is also to what extent these processes differ between transgender, non-binary and other gender diverse groups. An issue about which there is very little research beyond a few studies on the challenges that transgender women experience compared to transgender men when dating (Blair & Hoskin, 2019; Taylor, 2022).

To look into these questions and issues, we use data from Zoe, a dating app for queer womxn, collected between January 2023 and June 2025. The dataset includes detailed user-level information on gender identity, sexual orientation, age, location, and profile characteristics (such as photo count, bio completion, and verification status), as well as behavioural indicators capturing user engagement (number of likes given) and outcomes (number of matches received). These variables allow us to analyze both activity levels and

returns to engagement across gender groups. Recent work using Zoe data has, for example, examined global patterns in sexual orientation among app users (Rampazzo, Winer, & Kovačič, 2025), but did not focus on gender diversity or returns to engagement.

We employ count data models, specifically Negative Binomial regression, to examine engagement patterns, and Ordinary Least Squares (OLS) models to assess the like-to-match conversion rate, controlling for demographic and profile characteristics to isolate the effects of gender diversity. One of the advantages of using data from dating apps, is that most research has relied on stated preferences rather than actual behavior within dating (Griffin & Fingerman, 2018), which arguably reveals more clearly people's preferences and experienced obstacles in dating (Potârca, Mills & Neberich, 2015).

4. 2. Theoretical Background

4. 2. 1. Gender Diversity and Dating

There is little research on the dating life and relationships of gender diverse individuals. Survey data shows that transgender women, transgender men, and gender non-conforming individuals are less likely to be in a relationship than cisgender men and cisgender women in the US (Badgett et al., 2024). Given the challenges that gender diverse people encounter once dating (Blair & Hoskin, 2019; Webster et al., 2025), these lower rates of partnering are likely to result from obstacles rather than preferences.

In this article, we focus on dating experiences, which is a topic of importance in itself, but also a starting point for understanding differences in relationship outcomes between binary cisgender and gender diverse people. Our two outcomes of interest are engagement with dating apps and the “returns” people get from engagement in dating apps, i.e. the number of matches people make relative to the likes they give. In the following two sections, we discuss the determinants of engagement and returns, consecutively.

4. 2. 2. Engagement with Dating Apps

The first outcome we consider in this paper is engagement with dating apps, engagement encompasses both the likelihood of using a dating app, and once using one, how active people are on these apps. Empirically, we will only be able to look at this latter aspect of engagement among app users, more specifically the number of likes people give and the amount of information they disclose on their profiles.

Engagement with dating apps is shaped by its attractiveness compared to other ways of meeting people, such as meeting in bars, through friends, or in other social settings (Rochat et al., 2019; Konings, 2024; Lampard, 2007). The relative attractiveness of dating apps depends on the relative availability of dating partners and how experiences during the dating process compare to those of other ways of dating. Dating apps provide opportunities and challenges in both regards for gender diverse people.

Online dating provides the possibilities to meet people beyond “traditional” pathways such as friends, family, work, bars, or other places. The large pool of potential dating partners that dating apps offer make it easier to find a partner that has specific characteristics that might be harder to find or observe in offline settings. However, the gamified design of dating apps, which emphasizes swiping and instant matching, has been argued to commodify intimacy and transform romantic pursuit into a marketplace dynamic (Hobbs et al., 2017). Within this context, online dating has become instrumental for many marginalized groups to meet people (Potârcă, Mills & Neberich, 2015; Sironi & Kashyap, 2022). Given widespread discrimination against gender diverse people, such as the low openness towards dating a transgender person (Blair and Hoskin, 2019), having a large pool of potential dating partners could be more important for gender diverse people as compared to cisgender people, increasing the use of and engagement with dating apps.

Marginalization can also affect the various uses that people give to dating apps. Byron and colleagues (2021) showed that online dating apps are also used as ways to make friends or

meet people without romantic intentions. Previous research has shown that people who seek same-sex relationships are more likely to state looking for shared experiences rather than romance in areas where attitudes towards LGBTQ+ rights are unfavorable (Griffin & Fingerman, 2018). If marginalization leads people to use dating apps not only for romance but also for meeting people more generally, gender diverse people might show higher levels of engagement with dating apps than cisgender people. Qualitative research indeed indicates that dating apps are an important tool for gender diverse people to connect with other people (Griffiths & Armstrong, 2024).

The engagement of gender diverse people with dating apps is likely to also depend on experiences on those platforms. Here we discuss how dating apps relate to two major challenges when dating: safety and disclosing transgender status. Important shares of the gender diverse population report experiences of discrimination, fetishization, violence and abuse once dating (Perez & Pepping, 2024). In this regard, the physical separation and anonymity of the online environment can be a safety advantage (Fernandez & Birnholtz, 2019). However, gender diverse people also report unsafety in online dating (Griffiths & Armstrong, 2024; Perez & Pepping, 2024). Experiences of safety depend on the features that apps offer such as the possibility to report or block people. Some users have reported that geolocalization can decrease safety as it increases the possibility of users' physical location to be tracked down (Griffiths & Armstrong, 2024).

Besides physical safety, many transgender people, in particular women, express anxiety about dating partner's responses to their identity (Riggs et al., 2015). Online dating can reduce the barriers towards disclosing gender diverse status (Sharabi et al., 2023), as people generally report that anonymous meetings allow people to disclose more of their inner self compared to face-to-face meetings (Sironi & Kashyap, 2022). Transgender daters indeed report physical separation and the anonymity of online environments as motivations to disclose transgender status on online platforms (Fernandez and Birnholtz, 2019).

In our empirical analysis, we look at engagement within a dating app aimed at LGBTQ+ women which can further affect the attractiveness for gender diverse people to use it. First of all, many dating apps have excluded gender diverse identities from their design, including options and filters, and LGBTQ+ apps might be more inclusive in that regard (Griffits & Armstrong, 2024). Secondly, gender diverse people have reported to feel more safe interacting with other people from the LGBTQ+ community and that LGBTQ+ dating apps are experienced as providing safer spaces (Sharabi et al., 2023). Hence, gender diverse people might be particularly active on LGBTQ+ dating apps. However, some gender diverse people have also reported to not use LGBTQ+ dating apps because they restrict the dating pool by focusing on certain genders or identities, or because they do not feel included in LGBTQ+ environments (Sharabi et al., 2023).

In short, if online dating is relatively safer than other dating routes, and if online environments provide more authentic dating experiences, gender diverse people might engage more with dating apps than cisgender people. This could particularly be the case once studying engagement with LGBTQ+ apps, as is the case for this paper. We therefore expect that:

H1a: Users who indicate a trans, non-binary, or gender-diverse identity show higher levels of engagement within LGBTQ+ dating apps than cisgender people.

However, the presence of discrimination and harassment also within online dating apps, and the limited dating options that LGBTQ+ dating apps sometimes provide, could also reduce engagement among gender diverse people. This would then lead to:

H1b: Users who indicate a trans, non-binary, or gender-diverse identity show lower levels of engagement within LGBTQ+ dating apps than cisgender people.

4. 2. 3. Returns to Engagement in Dating Apps

Engagement with dating apps is only one side of the connection process. Even more important is the extent to which engagement “pays off” in the form of mutual connections. We conceptualize this through the like-to-match conversion rate: how many matches are obtained relative to the number of likes given. For trans, non-binary, and gender-diverse users, returns to engagement are shaped by both structural and platform-specific mechanisms.

Widespread discrimination in the dating market reduces opportunities for gender diverse people to form romantic or sexual relationships. Surveys show that the majority (87%) of cisgender individuals explicitly exclude transgender people from their dating pool, with particularly high levels of exclusion for trans women (Blair & Hoskin, 2019). This exclusion suggests that, even when gender diverse users actively engage with dating apps, they may receive fewer matches than cisgender users. Online transphobia and harassment remain widespread, and experiences of fetishization (e.g., “trans chasers” seeking out transgender users in exploitative ways) complicate whether matches reflect genuine interest or objectifying dynamics (Perez & Pepping, 2024). More broadly, research on online dating in Europe has documented strong patterns of partner exclusion along racial lines (Potârncă & Mills, 2015), suggesting that digital platforms can reproduce and even sharpen existing desirability normativities.

At the same time, some features of LGBTQ+ dating apps may create more favorable conditions for gender diverse people. Compared to mainstream apps, LGBTQ+ platforms may provide safer and more inclusive environments, with greater openness toward gender diversity and more tools to manage disclosure. Within these apps, T4T connections, where trans people seek to match with other trans people, may represent a positive mechanism of community-building and affirmation, potentially resulting in higher conversion rates (Griffiths & Armstrong, 2024).

Recent scholarship has shown that dating apps do not simply mirror users' pre-existing identities but actively shape how gender, sexuality, and desire are experienced and expressed. As Guyan (2025) argues, the classificatory systems embedded in dating app design both enable and constrain queer users' possibilities for self-presentation and recognition, revealing how inclusivity efforts can simultaneously reproduce exclusionary norms. Dating apps often privilege images and profile formats that align with hegemonic gender norms (MacLeod et al., 2019). Recent research and legal cases show how algorithmic systems can reproduce and even legitimize discrimination. Beyond individual preferences, online dating platforms structure visibility and choice in ways that can reinforce social boundaries. For example, racialized sorting in European dating sites has been documented even in the absence of explicit exclusion criteria (Potârcă & Mills, 2015). The Breeze case in the Netherlands revealed that a dating app's matching algorithm likely discriminated against dark-skinned users, leading the Dutch Institute for Human Rights to mandate corrective action (de Jonge et al., 2024). Similarly, analyses of Bumble's recommendation system indicate that its algorithms inherit and reinforce the biases embedded in the data they are trained on, perpetuating patterns of exclusion toward users with non-normative gender identities (Kalra et al., 2023). From a data feminism perspective (D'Ignazio & Klein, 2020), these examples show that data and algorithms are not neutral, rather, they reflect and even amplify existing inequalities. The design of matching algorithms, and what counts as "desirable," "visible," or "normal", plays a central role in reproducing inequalities within online dating spaces. These biases could impact the visibility of gender diverse individuals in these apps, potentially leading to less matches.

In short, the returns to engagement for gender diverse users are shaped by contradictory forces: structural exclusion and algorithmic bias may reduce conversion rates, while LGBTQ+ spaces and intra-community dynamics may enhance them. This leads us to two competing hypotheses:

H2a: Users who indicate a trans, non-binary, or gender-diverse identity have a higher like-to-match conversion rate.

H2b: Users who indicate a trans, non-binary, or gender-diverse identity have a lower like-to-match conversion rate.

4. 2. 4. United Kingdom context

The United Kingdom provides a relevant context for studying gender diverse dating experiences given its particular legal and social landscape around gender recognition. The Gender Recognition Act 2004 established a legal framework for transgender people to obtain legal recognition of their gender, though the process has been critiqued for being medicalised, lengthy, and bureaucratic (Hines, 2020; Raha, 2017). More recently, policy debates around gender recognition reform have intensified public discourse about transgender rights, creating what some scholars describe as a particularly hostile media environment for gender diverse people in the UK compared to other Western European countries (Pearce et al., 2020).

Survey research indicates that whilst public attitudes towards LGB rights in the UK have become increasingly accepting over recent decades, attitudes towards transgender people remain more ambivalent, with significant portions of the population still expressing uncertainty or discomfort (Clements & Field, 2014; Bachmann & Gooch, 2018). Experimental evidence also suggests that opposition to trans-inclusive policies is often driven less by attitudes toward trans people per se and more by generalized concerns about male violence (Morgenroth, Axt, & Westgate, 2024), highlighting how public debates frequently frame trans rights through broader anxieties about safety. Within this context, gender diverse individuals report high rates of discrimination in various life domains, including healthcare, employment, and public spaces (Government Equalities Office, 2018). The extent to which these broader societal attitudes shape intimate relationship formation and partner selection, particularly

within LGBTQ+ community spaces such as dating apps, remains an empirical question that our analysis seeks to address.

4. 3. Research Design

4. 3. 1. Data

This project uses data from Zoe, a dating app for queer womxn, collected between January 2023 and June 2025. The data were obtained through a data-sharing agreement between the app developers and the University of Oxford. Zoe is available on both the App Store and Google Play, where it has received over 5 million downloads and maintains high user ratings (4.3 stars on Google Play based on more than 104,000 reviews, and 4.5 stars on the App Store with 3,400 ratings). The platform reports over 7.2 million users worldwide, indicating broad adoption and stable engagement within the LGBTQ+ community. Upon registration, users provide information on their gender identity, age, sexual orientation, photos, and bio. The anonymized dataset includes user-level statistics on demographic and profile characteristics (such as photo count, bio completion, and verification status), as well as behavioural indicators of activity and interaction within the app, including the number of likes given and matches received. The data were provided in CSV format, with each observation representing an individual user. Our study builds on and complements previous work using Zoe data to analyze sexual orientation patterns in a global sample of users (Rampazzo et al., 2025), by focusing specifically on gender-diverse users in the United Kingdom and on their levels of engagement and returns.

We focus on two primary metrics: the number of likes given and the number of matches. Likes capture how often a user expresses interest in other profiles, while matches indicate mutual interest between users. To account for variation in the length of time users have been active on the platform, we include the variable exposure months, measuring the number of months between account creation and June 2025. However, we cannot know how long

people have been active, as we do not have information on when the likes were given or when users were last active.

Users self-identify their gender as *female* (n=12,384), *non-binary* (n=619), *trans female* (n=397), *other gender (fluid, bigender, androgynous, agender, pangender)* (n=379), or opt not to disclose, in which case the entry is recorded as *not reported (NA, n=20,360)*. In the text, we will use these exact terms to refer to the different groups of people. In other words, when we refer to females, we refer to people who gave that answer option and not all women. Respondents are also asked to select one sexual orientation from the following options: *asexual, bisexual, gay, lesbian, pansexual, or queer*. Missing responses are recorded as not reported. These variables allow us to analyze user's gender and sexual identification across the Zoe database.

Additional variables capture aspects of user profiles and demographic characteristics. User profile characteristics include the number of profile photos, the "about" section length, and whether the user has verified their account. Related to demographic characteristics, age and location are used. Age is grouped into four categories: *18–30, 31–45, 46–60, and 60+*. As the database provided information about the geolocation of the user at the moment of profile creation, we also coded a binary indicator for whether the user is located in an urban area based on NUTS3 and their urbanization level.

The initial dataset includes over 4 million accounts created between January 2023 and June 2025. We restrict the analysis to users located in the United Kingdom (n = 85,941), who gave at least one like on the platform. After applying these filters, the final analytic sample includes 34,139 users.

4. 3. 2. Methods

To analyze the engagement, three different variables are considered. Picture count and bio length provide an indicator of profile completion and are evaluated through a descriptive analysis of means. To study the engagement in terms of the number of likes users give on

the platform, we employed count data models due to the discrete and non-negative nature of the outcome variable. We compared three approaches: Ordinary Least Squares (OLS), Poisson regression, and Negative Binomial regression. Model comparison was based on Akaike Information Criterion (AIC) and log-likelihood values.

The Poisson model, while appropriate for count data, assumes equidispersion—that the mean and variance of the dependent variable are equal. This assumption was violated in our case, as the variance of likes substantially exceeded the mean. As a result, the Negative Binomial regression model was preferred, as it accounts for overdispersion by including an additional parameter that models excess variance (Cameron & Trivedi, 2013; Hilbe, 2011). The Negative Binomial model provided a better fit to the data based on both AIC and log-likelihood comparisons.

In addition to modeling the count of likes, we examined the likes-to-matches ratio as a continuous outcome variable. For this, we applied OLS regression and evaluated model performance using the R-squared statistic. However, users in different gender identity groups may systematically differ across other characteristics that also predict matching success (e.g., age, profile completeness, liking patterns, or sexual orientation). These differences could confound the direct relationship between gender identity and matching returns. To address this and ensure our comparisons are more robust, we supplemented our OLS models by implementing Coarsened Exact Matching (CEM). CEM is a preprocessing method designed to match individuals with similar characteristics between the groups being compared. By matching users across gender categories on key observable characteristics (such as age group, liking patterns, and profile characteristics), CEM generates a set of weights for each observation. Running our OLS models on this matched sample allows us to better isolate the net association of gender identity on the likes-to-matches ratio, reducing the bias that may arise from simply comparing dissimilar groups of users.

All models were conducted in a stepwise manner, adding the full set of covariates described in the data section sequentially, including gender identity, sexual orientation, age group, urban status, verification status, profile completeness, and exposure time on the app.

4. 4. Results

The results section is organised into three parts. First, we present descriptive findings on demographic characteristics and profile completion, broken down by gender identity. Second, we report the results of Negative Binomial regression models predicting the number of likes given, with a particular focus on differences across gender groups. Finally, we analyze the match-to-like ratio using OLS regression models to examine patterns of “returns” to engagement. This analysis is conducted in two steps: first using the full unweighted sample, and then applying a Coarsened Exact Matching approach to account for engagement differences across groups.

4. 4. 1. Descriptives

To contextualize the analyses that follow, we first describe the demographic and identity composition of the sample. Sexual orientation and age are two key variables shaping users’ experiences and behaviors on dating apps, and they are therefore included as controls in the regression models. In the case of the Coarsened Exact Matching (CEM) models, age is also used as one of the matching criteria to ensure comparability across gender groups. The distributions of sexual orientation and age by gender identity are presented in Table 4.1.

Table 4.1. Age and sexual orientation distributions by gender identity

	Not reported	Female	Non binary	Trans female	Other gender
Obs. (n)	20,360	12,384	619	397	379
<u>Age group (share per group)</u>					
18-30	0.61	0.64	0.86	0.74	0.70
31-45	0.34	0.31	0.13	0.18	0.25
46-60	0.05	0.04	0.01	0.06	0.04
60+	0.00	0.00	0.00	0.01	0.01
<u>Sexual Orientation</u>					

Asexual	0.00	0.00	0.03	0.01	0.03
Bisexual	0.03	0.38	0.18	0.26	0.26
Gay	0.00	0.02	0.04	0.01	0.04
Lesbian	0.03	0.43	0.36	0.41	0.25
Pansexual	0.00	0.04	0.17	0.15	0.17
Queer	0.00	0.03	0.18	0.07	0.12
Not reported	0.92	0.09	0.06	0.08	0.12

In terms of age distribution, the majority of users across all gender groups are between 18 and 30 years old, especially among non-binary users (86%). Users aged 31–45 make up the second largest share, with highest representation among users who did not report sexual orientation (34%). Older users (46+) constitute a small fraction across all groups.

Regarding sexual orientation, the most common identification among female users is lesbian (43%), followed by bisexual (38%). Trans female users also primarily identify as lesbian (41%) or bisexual (26%). Among non-binary and other gender-diverse users, identities are more varied, with higher shares of queer, pansexual, and bisexual identification when compared to their cis counterparts. Notably, a relatively high proportion of users in the not reported group did not report a sexual orientation (92%).

Table 4.2. Profile completion descriptives by gender identity

	Not reported	Female	Non binary	Trans female	Other gender
Obs. (n)	20,360	12,384	619	397	379
About filled (Share)	0.49	0.83	0.90	0.94	0.84
About size (Mean characters)	76.31	133.51	156.20	172.82	157.45
About size (Mean words)	14.47	25.07	28.48	31.72	28.78
Photo count (mean)	2.93	4.11	4.34	4.07	3.99
Verified (Share)	0.19	0.47	0.58	0.50	0.50

Table 4.2 reports descriptive statistics on profile completion by gender identity, offering a better understanding in terms of how individuals construct and present their profiles within the app. Users who reported their gender are generally more likely to have filled out the “About” section of their profile (ranging from 0.83 to 0.94), with the highest completion among trans female users (0.94). In contrast, only about half of those who did not state a

gender provided an “About” section (0.49). These patterns are also similar for the length of the “About” section. Trans female users write the longest bios on average (173 characters; 32 words), followed by non-binary users and those identifying with another gender. Females filled, on average, shorter bios, while users who did not state their gender write the briefest ones overall.

Verification and photo count patterns follow the same direction: users who self-identify their gender show higher verification rates (0.47–0.58) and upload more photos (around 4 on average for all groups) than users with missing gender information (0.19 verification rate; 2.93 photos). Among all groups, non-binary users stand out with the highest average number of photos (4.34) and the highest verification rate (0.58).

4. 4. 2. Likes given

This section presents results from a series of Negative Binomial regression models predicting the number of likes given by dating app users. Model 1 includes only gender identity; Model 2 adds sexual orientation; Model 3 incorporates age group; and Model 4 includes the full set of controls (photo count, urban/rural location, account verification, bio presence, and time since registration). Across all models, Female serves as the reference category for gender, Lesbian for sexual orientation, and 18–30 for age group. The complete set of effects is reported in Table 4.3.

Table 4.3. Negative Binomial models predicting number of likes given: Incidence Rate Ratios (IRR) with 95% CIs

Block	Term	Model 1	Model 2	Model 3	Model 4
Intercept	Intercept	167.40*** [162.67, 172.31]	182.50*** [175.32, 190.07]	164.25*** [157.39, 171.50]	37.23*** [34.25, 40.48]
Gender	Trans Female	2.99*** [2.55, 3.53]	3.07*** [2.62, 3.64]	3.23*** [2.75, 3.82]	2.82*** [2.41, 3.33]
		1.27*** [1.11, 1.45]	1.39*** [1.21, 1.60]	1.45*** [1.26, 1.66]	1.35*** [1.19, 1.55]
Gender	Non Binary	1.15 [0.98, 1.37]	1.33*** [1.13, 1.59]	1.36*** [1.15, 1.61]	1.30** [1.11, 1.53]

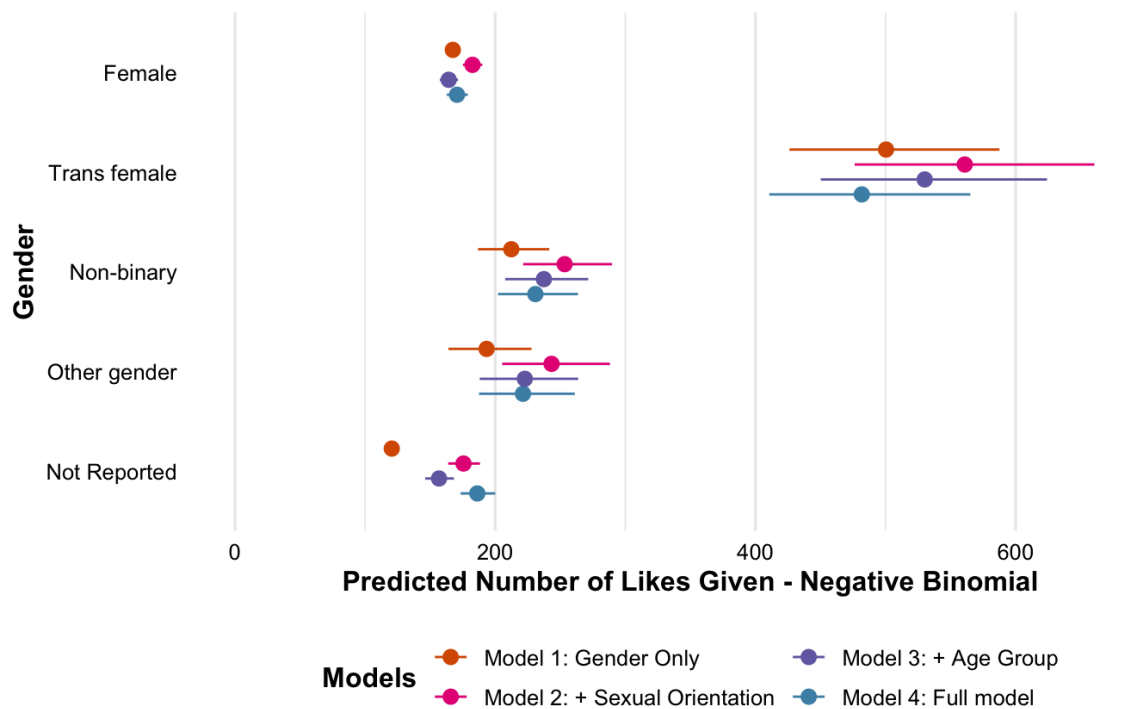
Gender	Not reported	0.72*** [0.69, 0.75]	0.96 [0.90, 1.02]	0.95 [0.90, 1.02]	1.09** [1.02, 1.16]
Sexual orientation	Asexual	—	0.61** [0.44, 0.87]	0.61** [0.44, 0.87]	0.51*** [0.37, 0.72]
Sexual orientation	Bisexual	—	0.92** [0.87, 0.98]	0.96 [0.91, 1.02]	0.92** [0.87, 0.98]
Sexual orientation	Gay	—	0.60*** [0.50, 0.74]	0.61*** [0.50, 0.74]	0.65*** [0.54, 0.79]
Sexual orientation	Not reported	—	0.66*** [0.62, 0.71]	0.68*** [0.63, 0.73]	0.77*** [0.72, 0.82]
Sexual orientation	Pansexual	—	0.92 [0.82, 1.04]	0.94 [0.84, 1.06]	0.83*** [0.74, 0.93]
Sexual orientation	Queer	—	0.54*** [0.47, 0.62]	0.56*** [0.49, 0.64]	0.53*** [0.46, 0.60]
Age group	31-45	—	—	1.34*** [1.29, 1.40]	1.56*** [1.51, 1.62]
Age group	46-60	—	—	0.72*** [0.66, 0.78]	0.73*** [0.67, 0.79]
Age group	60+	—	—	0.40*** [0.30, 0.54]	0.41*** [0.31, 0.55]
Profile & context	About filled	—	—	—	2.12*** [2.04, 2.20]
Profile & context	Exposure months	—	—	—	1.03*** [1.02, 1.04]
Profile & context	Exposure months ²	—	—	—	1.00*** [1.00, 1.00]
Profile & context	Photo count (per photo)	—	—	—	1.05*** [1.04, 1.06]
Profile & context	Rural (vs Urban)	—	—	—	0.90*** [0.86, 0.93]
Profile & context	Verified account	—	—	—	0.88*** [0.85, 0.92]

*Notes: IRR = exp(coef.), 95% confidence intervals in brackets. IRR predictors indicate multiplicative change per one-unit increase. Across all models, Female is the reference category for gender, Lesbian for sexual orientation, and 18–30 for age group. Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$*

Negative Binomial models reveal clear differences in engagement by gender identity. Relative to women, trans women give roughly three times as many likes (IRR 2.82–3.23 across models; 2.82 [2.41, 3.33] in the full model). Non-binary users also show higher activity (IRR 1.35 [1.19, 1.55]), as do users identifying with other genders (IRR 1.30 [1.11, 1.53]). In the initial model, users who did not report a gender are less active (IRR 0.72 [0.69,

0.75]), but this difference reverses after adjusting for sexual orientation, age, and profile features (IRR 1.09 [1.02, 1.16]). Overall, these findings confirm higher engagement among gender diverse users. Figure 4.1 presents a summary of the predicted number of likes given by gender for each model specification.

Figure 4.1. Predicted Number of Likes Given by Gender Identification (Female, Trans Female, Non-Binary, Other Gender, and Not Reported) based on Negative Binomial Models. The reference categories are “Female,” “Lesbian,” and “18–30.”



We also estimated Ordinary Least Squares (OLS) models as a sensitivity check. These results are presented in Appendix A and are consistent with the main findings reported here. Given that the number of likes given is a count variable characterized by overdispersion, Negative Binomial models provide a more robust and appropriate specification than OLS as this approach accounts for the high variability in user activity levels and provides clearer estimates of group differences, particularly among highly active users. Finally, we examined interactions with urban–rural locations and with the share of gender diverse users in each region (Appendix B). Although these contextual effects are not straightforward to interpret, the core gender differences in likes given are robust.

4. 4. 3. Like-to-match ratio

To examine how engagement translates into successful connections, we estimate a series of Ordinary Least Squares (OLS) models predicting the ratio of matches to likes. This indicator provides a measure of “returns” to engagement on the app. The analysis focuses again on differences across stated gender identities. We begin with a null model (Model 1) and sequentially add controls: sexual orientation (Model 2), age group (Model 3), and finally all profile and contextual characteristics (Model 4).

Table 4.4. OLS regression models predicting the ratio of likes-to-matches

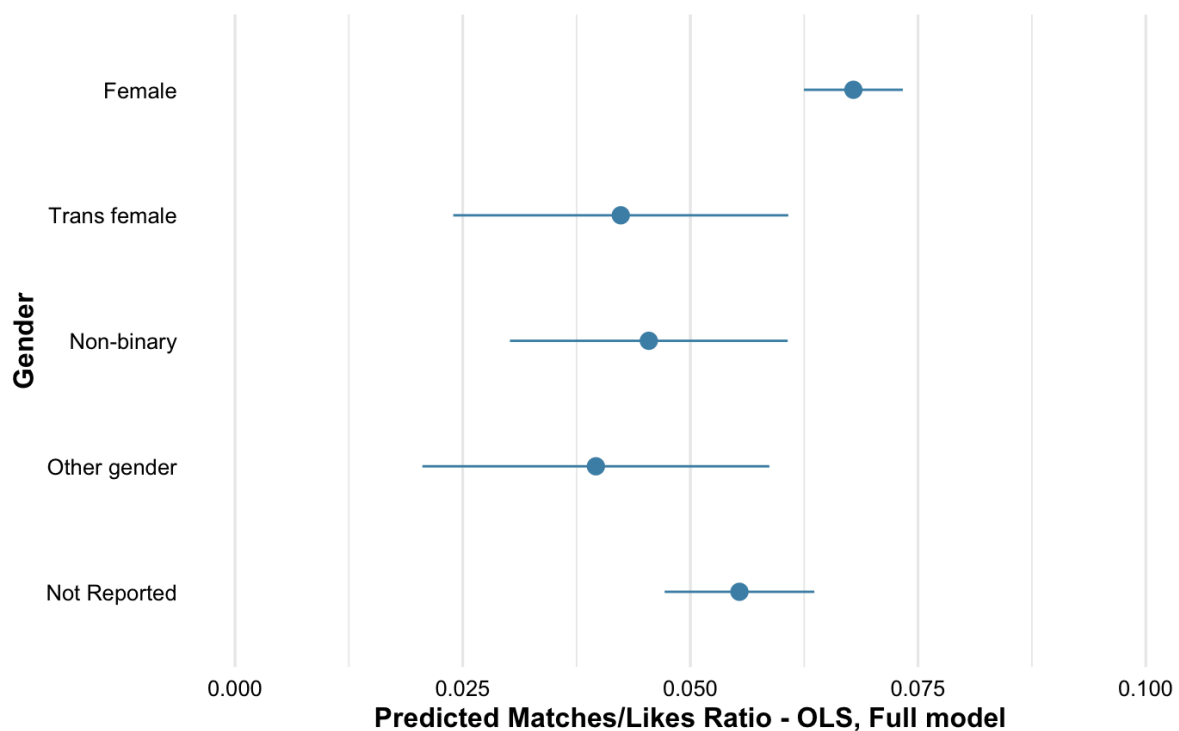
		Model 1	Model 2	Model 3	Model 4
	(Intercept)	0.08*** (0.00)	0.08*** (0.00)	0.08*** (0.00)	0.02** (0.00)
Gender	non_binary	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02** (0.01)
	trans_female	-0.03** (0.01)	-0.02* (0.01)	-0.02* (0.01)	-0.03** (0.01)
	other	-0.03** (0.01)	-0.03** (0.01)	-0.03** (0.01)	-0.03** (0.01)
	NA	-0.03*** (0.00)	-0.02*** (0.00)	-0.02*** (0.00)	-0.01*** (0.00)
	asexual		-0.04* (0.02)	-0.04* (0.02)	-0.05* (0.02)
Sexual Orientation	bisexual		0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	gay		0.00 (0.01)	0.00 (0.01)	0.00 (0.01)
	pansexual		0.00 (0.01)	0.00 (0.01)	-0.01 (0.01)
	queer		0.01 (0.01)	0.01 (0.01)	0.00 (0.01)
	NA		-0.01** (0.00)	-0.01** (0.00)	0.00 (0.00)
Age Group	31-45			0.01* (0.00)	0.01* (0.00)
	46-60			0.02*** (0.00)	0.02*** (0.00)
	60+			-0.02 (0.02)	-0.01 (0.02)
	Photo count				0.01*** (0.00)
	Rural				-0.01*** (0.00)
	Verified account				0.02*** (0.00)
	About filled				0.00 (0.00)
	exposure months				0.00*** (0.00)
	exposure months sqr				0.00*** (0.00)

*Notes: Across all models, Female is the reference category for gender, Lesbian for sexual orientation, and 18–30 for age group. Standard errors are reported in parentheses. Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$*

As can be noted in Table 4.4, regression models consistently show that users who identify as trans, non-binary, or gender-diverse experience lower returns to engagement compared to cisgender women. Across specifications, non-binary, trans female, and “other gender” users display significantly lower like-to-match conversion rates, even after controlling for sexual

orientation, age group, profile completeness, and app exposure. The negative coefficients remain robust in Models 1–4, indicating that these groups need to give more likes to achieve the same number of matches as cisgender women. The strongest disadvantage is observed among trans female users, who systematically exhibit reduced conversion rates, highlighting the structural barriers and forms of exclusion they face in the online dating environment. Figure 4.2 provides a visual summary of the effects for the full model.

Figure 4.2. Predicted Number of Match/Likes ratio, by Gender Identification, OLS models



4. 4. 4. Like-to-match ratio - Coarsened Exact Matching

We implement Coarsened Exact Matching (CEM) to probe whether the lower returns to engagement observed for gender diverse users reflect genuine differences in how interest is reciprocated, rather than compositional differences across groups. Descriptives and negative binomial models show sizeable variation in liking intensity and profile completeness by gender. CEM mitigates this by restricting comparisons to users with similar characteristics, improving the robustness of the comparisons.

Concretely, we match users on age, liking behavior (likes given), and key profile features (photo count, “About” section filled, and verification). After matching on these covariates, we test whether differences in the match-to-like conversion rate persist when comparing users with similar engagement and visibility. We run separate matched comparisons for each gender-diverse group (trans female, non-binary, other) against (a) females and (b) users who did not report a gender.

Table 4.5. CEM results comparing gender diverse users to females

	n	Matched	Unmatched	Treatment Coef.	Std. Error
Trans Female	397	394	3	-0.019**	0.0070
Non-binary	619	618	1	-0.012	0.0065
Other genders	379	378	1	-0.022**	0.0081

*Notes: Each model run separately for each comparison, matching on likes count, age group, picture count, about completeness, and verification status. Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

Table 4.6. CEM results comparing gender diverse users to those who did not report their gender

	n	Matched	Unmatched	Coef.	Std. Error
Trans Female	397	394	3	-0.019**	0.0070
Non-binary	619	618	1	-0.025***	0.0069
Other genders	379	378	1	-0.031***	0.0079

*Notes: Each model run separately for each comparison, matching on likes count, age group, picture count, about completeness, and verification status. Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$.*

Results from the Coarsened Exact Matching (CEM) models confirm that differences in returns to engagement persist even when comparing users with similar profile completion, age, and levels of activity. After matching on age group, number of likes given, photo count, profile completeness, and verification status, trans, non-binary, and other gender-diverse users continue to show significantly lower match-to-like ratios than their comparison groups. When compared to women (Table 4.5), users identifying with other genders display the largest disadvantage, followed by trans women. The gap remains when comparing gender

diverse users to those who did not report a gender (Table 4.6), with consistently negative and statistically significant effects for all gender-diverse groups. These results indicate that, even among users with equivalent age profiles and similar engagement or profile visibility, gender-diverse users still experience lower chances of having their likes reciprocated, suggesting enduring inequalities in how interest is returned on the platform.

Taken together, these results reveal a consistent pattern of inequality in dating outcomes. While users who state their gender as trans women, non-binary, or other gender-diverse identities tend to engage more actively by giving a higher number of likes, this effort does not translate into equivalent matches. Even after matching users with comparable profiles, age, and liking patterns, these groups receive fewer matches relative to the likes they give. This suggests that, beyond behavioural or profile presentation differences, there are differentials in how expressions of interest are reciprocated on the platform, representing the unequal dynamics faced by gender-diverse users in online dating environments.

4. 5. Conclusions and Discussion

This study provides new quantitative evidence on the dating experiences of trans, non-binary, and gender-diverse people within an LGBTQ+ dating app. Our results show that gender diverse users engage more actively with the platform, particularly trans women who give substantially more likes than cisgender women. However, this higher engagement does not translate into proportional benefits: when considering the like-to-match conversion rate, gender diverse users consistently experience lower returns to engagement. These findings highlight the persistent barriers that gender-diverse individuals face in digital dating spaces, even in those that are explicitly designed for LGBTQ+ communities.

Our finding that gender diverse users engage more actively on the Zoe app, with trans women giving approximately three times as many likes as other women, aligns with theoretical expectations about how gender diverse groups navigate dating markets. This heightened engagement might reflect three different, but interconnected, mechanisms. First,

the larger pool of potential partners that dating apps provide becomes particularly valuable for gender diverse individuals who face widespread exclusion in offline dating contexts (Potârcă, Mills & Neberich, 2015; Sironi & Kashyap, 2022). Second, online environments may offer relative safety advantages compared to traditional meeting venues such as bars or social gatherings, where gender diverse people report experiencing higher rates of discrimination and violence (Webster et al., 2025). The physical separation inherent to app-based dating allows users greater control over disclosure timing and the ability to block hostile interactions (Fernandez & Birnholtz, 2019; Sharabi et al., 2023). Third, dating apps also operate as spaces for friendship formation and community connection, going beyond romantic partnering. These purposes may be particularly important for gender diverse individuals who face social isolation (Griffiths & Armstrong, 2024). The elevated engagement we observe may therefore reflect not only romantic seeking behaviour, but also the use of LGBTQ+ dating platforms as crucial sites of social connection and community participation. A final possible interpretation of the higher number of likes given by gender diverse individuals, relates to our next result: gender diverse individuals receive lower returns in terms of matches from their likes, which can lead to people giving more likes to achieve a given number of matches.

Our results showed that despite their higher engagement, gender diverse users consistently experience lower like-to-match conversion rates, a pattern that persists even after accounting for age, profile completeness, verification status, and liking intensity through Coarsened Exact Matching. This finding provides quantitative confirmation of the structural exclusion in dating that qualitative research has documented (Blair & Hoskin, 2019; Griffiths & Armstrong, 2024; Perez & Pepping, 2024). The fact that 87% of cisgender individuals explicitly exclude transgender people from their dating pool (Blair & Hoskin, 2019), with particularly high rates of exclusion for trans women, manifests in our data as systematically lower returns to engagement. Even within an LGBTQ+ dating app designed for queer womxn, gender-diverse users face reduced reciprocity of interest, suggesting that

transphobia and cisnormative preferences operate across both mainstream and community-specific dating platforms. This is particularly striking given that LGBTQ+ spaces are often theorized as more inclusive and affirming environments (Sharabi et al., 2023). Our results indicate that whilst these platforms may provide safer spaces for disclosure and community connection, they do not fully insulate gender diverse users from discrimination in partner selection.

Our results complement prior work on sexual orientation patterns in Zoe (Rampazzo et al., 2025) and on racial and gendered sorting in online dating more broadly (Potârcă & Mills, 2015), by foregrounding the specific dynamics faced by gender-diverse users. A major contribution of this study is comparing not only cisgender vs non-cisgender groups, but to include different gender-diverse groups, comparing the behaviour and returns of trans, non-binary, and other gender diverse identities alongside female users. As previous studies show (Darwin, 2020; Bercovich, 2025), when studying gender diverse populations, it is relevant to compare across gender identifications, as treating all non-cis individuals as a single group can shield mixed results. In this analysis we demonstrate the value of considering each gender identity as a single group, as trans females present a clear divergence in terms of liking patterns when compared to any other gender group. At the same time, it is notable to mention that regardless of their gender identity, all gender diverse groups present lower returns to their participation.

The persistent disparities in match outcomes raise important questions about the role of algorithmic systems and platform design in reproducing inequalities. Recent research demonstrates that dating app algorithms are not neutral arbiters of connection, but rather encode and amplify existing biases present in user behaviour and training data (D'Ignazio & Klein, 2020; Kalra et al., 2023). The Breeze case in the Netherlands, where a dating app's algorithm was found to discriminate against dark-skinned users (de Jonge et al., 2024), illustrates how matching systems can institutionalize exclusion under the guise of technical optimization. Similarly, MacLeod and McArthur's (2019) analysis reveals how interface

design privileges hegemonic gender presentations, potentially reducing the visibility of users with non-normative gender identities. Whilst our data cannot directly measure algorithmic effects, the fact that conversion rate disparities persist after matching users on observable profile characteristics suggests that factors such as differential visibility in recommendation systems or lower prioritization in matching queues may contribute to the disadvantages gender diverse users experience. From a data feminism perspective, this underscores the need for dating platforms to critically examine how their algorithmic systems may reproduce structural inequalities.

Several limitations should be considered when interpreting our findings. First, our data come from a single LGBTQ+ dating app designed specifically for queer womxn, which limits generalizability to other platforms and dating contexts. The user base of Zoe may differ systematically from users of mainstream apps such as Tinder or Bumble, or other LGBTQ+ platforms such as Her or Lex. Future research should examine whether the patterns we observe—higher engagement but lower conversion rates—hold across different platform types, particularly in mainstream apps where gender diverse users may face additional barriers.

A second major limitation related to the app studied, is that we only capture gender diverse individuals who are interested in queer womxn. It is therefore unclear whether our results for, for instance, transgender women are specific to transgender women interested in queer womxn, or whether transgender women interested in men experience similar disadvantages in LGBTQ+ dating apps. Similarly, a major question is what the experiences are of gender diverse people interested in queer womxn whose gender identity was not part of the options offered by the Zoe app. Future research on other apps can give further insight into this issue.

The third limitation relates to the use of behavioural data and more generally encompasses the connected limited generalizability of our results. Since our data are not collected through surveys based on a probabilistic sampling frame, it is impossible to know to what extent we

capture a representative part of the population. However, the advantage of using behavioural data, compared to survey data, is that we observe preferences as expressed through actual behavior, rather than relying on people's stated preferences, which might be more subject to desirability bias. Future work could leverage emerging probability-based datasets with dedicated LGBTQ+ samples (Fischer et al., 2022) to assess how patterns observed in dating apps map onto broader population-level inequalities.

Lastly our sample is restricted to users in the United Kingdom, and dating norms, LGBTQ+ community structures, and levels of transphobia vary substantially across national contexts (FRA, 2020; ILGA-Europe, 2024; TGEU, 2023). Cross-national comparative work would help identify which aspects of our findings reflect universal dynamics of digital dating and which are specific to particular socio-cultural environments.

There are specificities in terms of behavior that would be important to study further. First, we cannot observe who is liking whom and which specific combinations of gender identities are more or less likely to result in matches. Understanding whether gender diverse users primarily direct their likes toward women, other gender diverse individuals, or a mix would illuminate whether lower conversion rates reflect general exclusion or patterns of T4T (trans-for-trans) seeking that may have different dynamics (Griffiths & Armstrong, 2024). Second, we are measuring engagement solely through likes given, but other indicators such as messaging frequency, time spent on the app, or profile updating behaviour might reveal additional dimensions of platform use. More generally, an important aspect of engagement is whether people use dating apps in the first place. It is possible that the high levels of engagement among gender diverse people observed in our study reflect selection into app-use based on characteristics that increase engagement (e.g. gender diverse people might only use apps when most actively looking for dates or a relationship). Third, and perhaps most importantly, we cannot trace what happens after matches occur. Do matches translate into conversations, dates, relationships, or community connections? Or do they sometimes lead to negative experiences such as harassment, discrimination, or

disappointment? Understanding post-match outcomes is crucial for understanding whether dating apps ultimately serve the relationship and community needs of gender diverse users, or primarily expose them to additional discrimination within ostensibly safe spaces. Finally, there is a substantial number of users ($n=20,360$) who did not report their gender identity. Are these users gender diverse individuals who choose not to disclose, cisgender women seeking privacy, or those uncertain about identity categories offered by the platform? Understanding patterns of non-disclosure could reveal additional information related to dating apps' uses and behaviour.

The patterns documented in this study carry significant implications for understanding social stratification and well-being among gender-diverse populations. Beyond the experiences within the apps, and less access to dates. Lower conversion rates in dating apps may contribute to the documented disparities in relationship formation, with gender diverse individuals being substantially less likely to be partnered than their cisgender counterparts (Badgett et al., 2024). Given the well-established links between romantic relationships and multiple dimensions of well-being, including mental health, economic security, and social support (Reczek, 2020; Rodríguez Sánchez & Boertien, 2025), systematic disadvantages in dating markets represent more than mere romantic disappointment. They constitute a mechanism through which structural discrimination translates into tangible inequalities in life chances and quality of life. Reduced access to partnerships may have cascading effects on access to parenthood, even considering multiple pathways such as co-parenting arrangements, assisted reproduction, or adoption, many of which presume or favor partnered individuals. The fact that these disparities persist even within LGBTQ+ dating spaces/platforms explicitly designed to provide safer and more inclusive environments suggests that addressing them will require more than simply creating community-specific platforms. It demands a deeper examination of how prejudices, algorithmic systems, and platform design choices shape who gets to connect with whom in digital spaces that increasingly mediate intimate life.

This study reveals a paradox at the heart of digital dating for transgender, non-binary, and gender-diverse individuals: heightened engagement that does not translate into proportionate connection. Gender diverse users invest considerably more effort into seeking partners on LGBTQ+ dating platforms yet receive systematically lower returns in the form of matches. Our findings point to enduring structural inequalities in how expressions of interest are reciprocated, reflecting both individual-level prejudices and potentially algorithmic mechanisms that disadvantage gender-diverse users. These disparities matter not only for understanding the dating experiences of gender diverse individuals, but also for illuminating broader processes of social stratification and digital inequality. The implications extend beyond individual happiness to fundamental questions of algorithmic justice, platform governance, and whether digital spaces can serve as sites of liberation for those who challenge gender norms—or whether they inevitably replicate the exclusions of the offline world in new ways.

4. 6. References

- Bachmann, C. L., & Gooch, B. (2018). LGBT in Britain: Trans report. Stonewall, YouGov.
- Badgett, M. L., Carpenter, C. S., Lee, M. J., & Sansone, D. (2024). A review of the economics of sexual orientation and gender identity. *Journal of Economic Literature*, 62(3), 948-994.
- Blair, K. L., & Hoskin, R. A. (2019). Transgender exclusion from the world of dating: Patterns of acceptance and rejection of hypothetical trans dating partners as a function of sexual and gender identity. *Journal of Social and Personal Relationships*, 36(7), 2074-2095.
- Byron, P., Albury, K., & Pym, T. (2021). Hooking up with friends: LGBTQ+ young people, dating apps, friendship and safety. *Media, Culture & Society*, 43(3), 497-514.
- Cameron, A. C., & Trivedi, P. K. (2013). Regression analysis of count data (No. 53). Cambridge university press.

Clements, B., & Field, C. D. (2014). Public opinion toward homosexuality and gay rights in Great Britain. *Public opinion quarterly*, 78(2), 523-547.

D'ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT press.

Darwin, H. (2020). Challenging the cisgender/transgender binary: Nonbinary people and the transgender label. *Gender & Society*, 34(3), 357-380.

de Jonge, T., & Borgesius, F. Z. (2024). Mitigating Digital Discrimination in Dating Apps--The Dutch Breeze case. *arXiv preprint arXiv:2409.15828*.

European Union Agency for Fundamental Rights (FRA). (2020). A long way to go for LGBTI equality.

Fernandez, J. R., & Birnholtz, J. (2019). " I Don't Want Them to Not Know" Investigating Decisions to Disclose Transgender Identity on Dating Platforms. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-21.

Fischer, M. M., Kroh, M., De Vries, L., Kasproski, D., Kühne, S., Richter, D., & Zindel, Z. (2022). Sexual and Gender Minority (SGM) Research Meets Household Panel Surveys... *European Sociological Review*, 38(2), 321–335.

Government Equalities Office. (2018). *National LGBT Survey: Summary report*. UK Government.

Griffin, E. M., & Fingerman, K. L. (2018). Online dating profile content of older adults seeking same-and cross-sex relationships. *Journal of GLBT Family Studies*, 14(5), 446-466.

Griffiths, D. A., & Armstrong, H. L. (2024). "They were talking to an idea they had about me": A qualitative analysis of transgender individuals' experiences using dating apps. *The Journal of Sex Research*, 61(1), 119-132.

Guyan, K. (2025). *Rainbow Trap: Queer Lives, Classifications and the Dangers of Inclusion*. Bloomsbury Publishing.

Hilbe, J. M. (2011). *Negative binomial regression*. Cambridge University Press.

Hines, S. (2020). Sex wars and (trans) gender panics: Identity and body politics in contemporary UK feminism. *The Sociological Review*, 68(4), 699-717.

Hobbs, M., Owen, S., & Gerber, L. (2017). Liquid love? Dating apps, sex, relationships and the digital transformation of intimacy. *Journal of sociology*, 53(2), 271-284.

ILGA-Europe. (2024). *Rainbow Map 2024*.

Kalra, R. A., Gupta, P., Varghese, B., & Rangaswamy, N. (2023). Exploring Gender Disparities in Bumble's Match Recommendations. *arXiv preprint arXiv:2312.09626*.

Konings, F., Sumter, S. R., & Vandenberg, L. (2024). A linkage study investigating sexualized self-presentation on mobile dating apps and user traits. *Body Image*, 51, 101781.

Lampard, R. (2007). Couples' places of meeting in late 20th century Britain... *European Sociological Review*, 23(3), 357–371.

MacLeod, C., & McArthur, V. (2019). The construction of gender in dating apps: An interface analysis of Tinder and Bumble. *Feminist Media Studies*, 19(6), 822-840.

Morgenroth, T., Axt, J. R., & Westgate, E. C. (2024). What underlies the opposition to trans-inclusive policies?... *Personality and Social Psychology Bulletin*, 50(4), 533–549.

Pearce, R., Erikainen, S., & Vincent, B. (2020). TERF wars: An introduction. *The Sociological Review*, 68(4), 677-698.

Perez, J., & Pepping, C. A. (2024). Relationship experiences of transgender and non-binary adults: exploring dating goals, relationship structures, minority stress, sexual fetishization, and relationship victimization. *International Journal of Transgender Health*, 1-20.

Potârcă, G., & Mills, M. (2015). Racial preferences in online dating across European countries. *European Sociological Review*, 31(3), 326-341.

Potârcă, G., Mills, M., & Neberich, W. (2015). Relationship preferences among gay and lesbian online daters: Individual and contextual influences. *Journal of Marriage and Family*, 77(2), 523-541.

Raha, N. (2017). Transfeminine brokenness, radical transfeminism. *South Atlantic Quarterly*, 116(3), 632-646.

Rampazzo, F., Winer, C., & Kovačič, M. (2025). Sexual orientation on Zoe: A global analysis of a lesbian dating app. *Demographic Research*, 53.

Reczek, C. (2020). Sexual-and gender-minority families: A 2010 to 2020 decade in review. *Journal of Marriage and Family*, 82(1), 300-325.

Riggs, D. W., Ansara, G. Y., & Treharne, G. J. (2015). An evidence-based model for understanding the mental health experiences of transgender Australians. *Australian Psychologist*, 50(1), 32-39.

Rochat, L., Bianchi-Demicheli, F., Aboujaoude, E., & Khazaal, Y. (2019). The psychology of “swiping”: A cluster analysis of the mobile dating app Tinder. *Journal of Behavioral Addictions*, 8(4), 804-813.

Rodríguez Sánchez, A., & Boertien, D. (2025). Decomposing Gender Identity Differences in Subjective Well-Being Using Two Colombian Surveys. *Socius*, 11, 23780231251315577.

Rosenfeld, M. J., & Thomas, R. J. (2012). Searching for a mate: The rise of the internet as a social intermediary. *American Sociological Review*, 77(4), 523-547.

Sharabi, L. L., Ryder, C. V., & Niess, L. C. (2023). A space of our own: Exploring the relationship initiation experiences of lesbian, gay, bisexual, transgender, queer, intersex, and asexual dating app users. *Journal of Social and Personal Relationships*, 40(7), 2277-2297.

Sironi, M., & Kashyap, R. (2022). Internet access and partnership formation in the United States. *Population Studies*, 76(3), 427-445.

Taylor, A. (2022). "But where are the dates?" Dating as a central site of fat femme marginalisation in queer communities. *Psychology & Sexuality*, 13(1), 57-68.

Transgender Europe (TGEU). (2023, May 11). Trans Rights Index & Map 2023: Continued Progress Amid Anti-trans Backlash in Europe & Central Asia.

Webster, A., Brooker, J., Misajon, R., & Wilson, C. (2025). Online dating experiences of transgender and gender-diverse people: A systematic review and thematic synthesis of qualitative evidence. *Psychology of Sexual Orientation and Gender Diversity*

4. 7. Appendix

4. 7. 1. Appendix A. Ordinary Least Squares (OLS) regression results predicting number of likes given

This appendix presents the full OLS regression models predicting the number of likes given. Models include progressively added controls for sexual orientation, age group, and profile characteristics. Table 4.A1 shows the complete set of effects, while Figure 4.A1 presents the predicted number of likes given by gender identification. Significant differences emerge across gender identities. In Model 1, trans women give substantially more likes than women, a pattern that remains remarkably stable even after controlling for sexual orientation, age, and other variables. Similarly, non-binary users consistently give more likes than women across all models. The category “Other gender” does not significantly differ from the reference group, while users who did not report a gender do not show significant differences once covariates are included. These findings suggest that trans women engage more actively in liking behavior, even when accounting for multiple user characteristics.

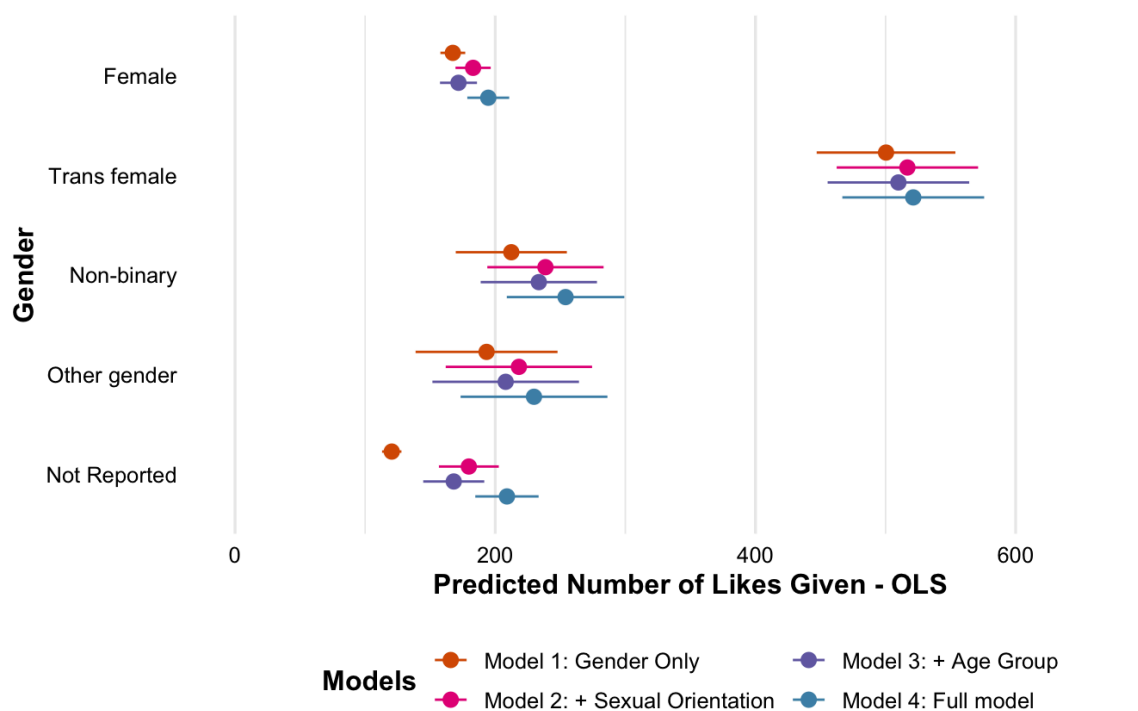
Table 4.A1. Ordinary Least Squares (OLS) regression models predicting number of likes given

		Model 1	Model 2	Model 3	Model 4
	(Intercept)	167.40*** (4.87)	183.00*** (6.90)	171.75*** (7.25)	-4.44 (14.64)
	Non-binary	44.96* (22.32)	55.60* (22.70)	61.75** (22.71)	59.42** (22.55)
Gender	Trans woman	332.99*** (27.63)	333.80*** (27.75)	338.15*** (27.73)	326.65*** (27.53)
	Other gender	25.94 (28.26)	35.20 (28.48)	36.29 (28.46)	35.15 (28.24)
	Not reported	-46.93*** (6.18)	-3.28 (10.97)	-3.64 (10.96)	14.36 (10.99)
	Asexual		-93.41 (57.70)	-95.04 (57.64)	-104.31 (57.21)
	Bisexual		-16.05 (9.74)	-15.57 (9.75)	-16.96 (9.68)
Sexual Orientation	Gay		-73.95* (32.84)	-74.08* (32.81)	-72.05* (32.56)
	Pansexual		3.29 (19.65)	2.24 (19.63)	-0.98 (19.50)
	Queer		-86.27*** (21.89)	-85.58*** (21.88)	-93.82*** (21.73)
	Not reported		-63.38*** (11.90)	-64.37*** (11.89)	-38.80** (11.94)
Age Group	31-45			45.41*** (6.46)	50.93*** (6.45)

46-60	-39.26** (14.49)	-32.77* (14.41)
60+	-77.95 (50.39)	-71.63 (50.01)
Photo count		3.21 (1.82)
Rural		-3.51 (7.09)
Verified account		-29.52*** (7.00)
About filled		107.41*** (6.90)
Exposure months		2.70* (1.37)
Exposure months sqr		0.13** (0.05)

Notes: Across all models, Female is the reference category for gender, Lesbian for sexual orientation, and 18–30 for age group. Standard errors are reported in parentheses. Significance levels: *** for $p < 0.001$; ** for $p < 0.01$; * for $p < 0.05$

Figure 4.A1. Predicted Number of Likes Given by Gender Identification (Female, Trans Female, Non-Binary, Other Gender, and Not Reported) based on Ordinary Least Squares (OLS) Models. The reference categories are “Female,” “Lesbian,” and “18–30.”



4. 7. 2. Appendix B. Geographical variations

To explore whether gender differences in the number of likes given vary across local contexts, we estimated two additional negative binomial models with interaction terms (Table B1). In both specifications, the outcome is the number of likes given and coefficients are

presented as incidence rate ratios (IRR). The reference categories are women for gender, lesbian for sexual orientation and ages 18–30 for age group.

The first model adds interactions between gender group and a binary urban-rural indicator based on the Eurostat DEGURBA classification. The main gender coefficients can be read as the IRR for users living in urban areas, while the interaction terms show how these gender gaps change in rural regions. In urban settings, non-binary users and those in the “other gender” category give more likes than women (IRR $\approx$ 1.22 and 1.42, respectively), and trans women give around three times as many likes as women (IRR $\approx$ 3.06), net of controls. The interaction terms below 1 for non-binary (IRR = 0.65) and other gender users (IRR = 0.63) indicate that these urban advantages are attenuated in rural areas: their like-giving rates converge towards those of women. For trans women, the rural interaction (IRR = 0.86) suggests that their very large differential in urban areas is somewhat smaller in rural contexts, although it remains substantial in absolute terms. In contrast, users who did not report their gender show a positive interaction (IRR = 1.39), meaning that the gap between “not reported” and women is larger in rural areas than in cities. Overall, these results point to some contextual moderation of gender gaps, but the pattern is heterogeneous.

The second interaction model examines whether gender gaps vary with the local concentration of gender-diverse users in each NUTS-3 region (Table 4.B1). Here we include the share of users whose gender is non-binary, trans female or “other” (pct_not_female) and interact this continuous measure with gender groups. The main effect of pct_not_female (IRR = 0.95) suggests that, among women (the reference), living in areas with a higher proportion of gender-diverse users is associated with slightly fewer likes given. For users who did not report gender information, the interaction above 1 (IRR = 1.19) implies that their like-giving increases more steeply with local gender diversity than it does for women. For the “other gender” group, the interaction below 1 (IRR = 0.88) indicates the opposite: their advantage in likes is smaller in more gender-diverse regions. For trans women, the

interaction term is slightly above 1 (IRR = 1.03) and statistically significant, but very close to the null value, so the substantive change in the slope is modest. The interaction for non-binary users is not statistically different from 1, suggesting little evidence of contextual moderation for this group.

Taken together, these robustness checks show that the large gender differences in like-giving documented in the main text are still there, even considering local context. Some interaction terms are statistically significant, indicating that gender gaps are somewhat smaller in rural areas for several gender-diverse groups and that the association between local gender diversity and like-giving is not identical across gender categories. However, most IRR values for the interactions are close to 1 and the direction of the effects is not uniformly reinforcing or dampening. We therefore interpret these contextual patterns as suggestive rather than central to our argument, and we retain the simpler specification without interactions as our main model.

Table 4.B1. Negative Binomial models including interactions predicting number of likes given: Incidence Rate Ratios (IRR) with 95% CIs

Block	Term	Full model + Urban/Rural interaction	Full model + % not female interaction
Intercept	Intercept	36.27*** [33.34, 39.47]	45.56*** [40.85, 50.83]
Gender	Not reported	1.13*** [1.05, 1.20]	0.98 [0.87, 1.09]
Gender	Non Binary	1.22** [1.05, 1.43]	0.59** [0.40, 0.86]
Gender	Other	1.42*** [1.18, 1.71]	2.43*** [1.50, 3.92]
Gender	Trans Female	3.06*** [2.57, 3.69]	3.15*** [2.05, 4.90]
Gender × Urban–rural	Not reported × Rural	1.39* [1.03, 1.88]	—
Gender × Urban–rural	Non Binary × Rural	0.65* [0.45, 0.97]	—
Gender × Urban–rural	Other Gender × Rural	0.63* [0.43, 0.95]	—
Gender × Urban–rural	Trans Female × Rural	0.86*** [0.79, 0.94]	—
Gender × % not-female	Not reported × % not-female	—	1.19*** [1.11, 1.29]
Gender × % not-female	Non Binary × % not-female	—	0.98 [0.90, 1.07]
Gender × % not-female	Other Gender × % not-female	—	0.88** [0.80, 0.97]
Gender × % not-female	Trans Female × % not-female	—	1.03** [1.01, 1.05]
Sexual orientation	Asexual	0.53*** [0.38, 0.74]	0.51*** [0.37, 0.73]

Sexual orientation	Bisexual	0.93* [0.88, 0.98]	0.92** [0.87, 0.98]
Sexual orientation	Gay	0.64*** [0.54, 0.78]	0.64*** [0.53, 0.77]
Sexual orientation	Na	0.77*** [0.72, 0.83]	0.76*** [0.71, 0.81]
Sexual orientation	Pansexual	0.84** [0.75, 0.95]	0.83** [0.74, 0.94]
Sexual orientation	Queer	0.53*** [0.47, 0.60]	0.52*** [0.46, 0.60]
Age group	31-45	1.57*** [1.51, 1.63]	1.56*** [1.50, 1.62]
Age group	46-60	0.73*** [0.67, 0.80]	0.72*** [0.67, 0.79]
Age group	60+	0.41*** [0.31, 0.56]	0.41*** [0.31, 0.56]
Profile & context	About filled	2.12*** [2.04, 2.20]	2.11*** [2.03, 2.19]
Profile & context	Exposure months	1.03*** [1.02, 1.04]	1.03*** [1.02, 1.04]
Profile & context	Exposure months ²	1.00*** [1.00, 1.00]	1.00*** [1.00, 1.00]
Profile & context	Photo count (per photo)	1.05*** [1.04, 1.06]	1.05*** [1.04, 1.06]
Profile & context	Rural (vs Urban)	0.98 [0.91, 1.04]	0.91*** [0.87, 0.95]
Profile & context	Verified account	0.89*** [0.85, 0.92]	0.88*** [0.84, 0.91]
Other	pct_not_female	—	0.95*** [0.93, 0.96]

Chapter 5: Conclusions

This thesis began with a simple idea: if we want to understand the lives of transgender, non-binary, and gender-diverse people, we need demographic tools that reflect the reality of their experiences. As feminist demographers previously argued that demography should move beyond treating sex as a demographic variable, to understanding gender as a social process (Watkins, 1993; Riley, 1997), this work seeks to criticize current approaches to gender operationalization. Standard approaches in population studies often assume that gender is binary and fixed, and as a result, they overlook important forms of inequality and miss the processes through which identities develop (Westbrook & Saperstein, 2015; Compton et al., 2018; Baumle & Nordmarken, 2022). The three studies in this dissertation show that new patterns and dynamics become visible when we rethink how we measure gender, when we allow for fluidity and change, and when we work with different types of data.

How are outcomes across the life course shaped by experiences that depart from or challenge dominant gender norms? What role does gender identity play in shaping educational and relational outcomes? In this work, inequalities related to educational outcomes were found, with a clear penalty for trans non-binary individuals. Also, in terms of relationship formation, our results based on behavioural data from a dating app show that gender-diverse individuals are much more prone to giving likes, but the return in terms of matches is much lower. Both studies show that gender diversity is linked to life outcomes, and in both cases, we found worse outcomes for those who challenge cisgender norms, as found in previous literature (Rodríguez Sánchez & Boertien, 2025; Carpenter et al., 2022).

If gender diversity relates to people's outcomes in life, what are the relevant gender dimensions that we should consider? How can demography include gender diversity in a way that does justice to people's lives? Knowing that gender diversity is related to life outcomes, a focus was made on how we are measuring and studying gender. In the first article, we present how a multidimensional gender dimension provides a clearer picture of inequalities. We also consider gender to be a process that unfolds across life, and not a

fixed category. In the second study, a life course perspective is applied to gender development, providing an innovative perspective on how gender can be studied in a dynamic way. These questions engage with previous critics made to gender categorization and operationalization, as can be found on the “New categories are not enough” (Westbrook & Saperstein, 2015) and “Other, please specify” edited by Compton et al., (2018).

What are the data sources and methods that can be used to study gender diversity? In this thesis, there is a strategic choice of complementary data sources. Each data source provides a particular value in terms of possibilities for research, but at the same time presents limitations. The first article uses a representative survey that retrieves gender in a three-step manner, allowing for the study of a specific characterization of gender in a multidimensional way. The second study is based on a specialized survey for the LGBTQ+ population that was collected through a collective effort between the EU Fundamental Rights Agency and LGBTQ+ communities. As this survey was designed to provide specific answers related to the LGBTQ+ population, questions related to sexual and gender identity development were included, allowing for a life course perspective. The third article leverages digital trace data to better understand real-world behaviour, providing evidence of the use that new data sources, such as big data, can provide to demographic research. As no data source is perfect, the strategy of complementing sources aims to provide a more comprehensive picture of gender diversity in demographic research.

The conclusions of this thesis are structured by first introducing an overview of the general findings of the thesis chapter by chapter. Second, there is a section dedicated to the different contributions, focusing on the main research questions of the thesis. Later on, a section dedicated to the discussion and implications of these findings and contributions is presented, accounting for limitations and presenting potential future lines of research. This thesis ends with some final remarks.

5. 1. Overview of Findings

The first study examines how the categories in which we operationalize gender shape our interpretation of educational inequalities. Although many studies have documented inequalities in various SES outcomes among trans and non-binary individuals (Carpenter et al., 2022; Badgett et al., 2021), most rely on broad classifications that collapse very different gender groups into a single category, commonly named *trans*, that is then compared to cis individuals. This chapter shows that once we move beyond the man/woman binary, or overly simplified cis versus trans distinctions, a diversity of experiences emerge. Using a multidimensional classification that reflects both gender trajectory (cis, trans, or something else) and the gender's relation to the binary (binary/non-binary), the analysis reveals educational penalties that remain hidden under commonly used gender categorizations. These patterns resonate with earlier work that reflects on how gender is measured crucially affects what we see and the results we get on our studies (Westbrook and Saperstein, 2015; Compton et al., 2018), and with Mittleman (2022) intersectional analysis of educational attainment, considering sexual identification, gender, and predicted gender.

This study is also developed in conversation with research on gender-diverse educational inequalities, where cis individuals in general present higher educational outcomes when compared with their trans counterparts (Carpenter, 2020; Kolk, 2025). Wilkinson, in their 2021 study, found that while binary trans individuals faced disadvantages compared with cis individuals, non-binary individuals did not. With aims to add to the findings of these studies, the first study empirically shows that binary/non-binary and gender trajectory represent two intersecting axes of gender identification, and relate to educational outcomes in different ways. The main finding is that trans individuals, particularly those who identified outside the man/woman binary, face consistent penalizations in educational attainment. Individuals who fall outside both the cisgender and the binary gender norms face more barriers in educational settings, affecting life-long outcomes. Choosing how we capture gender may highlight or completely obscure these inequalities. The main message is simple but

consequential: measurement is never neutral. It shapes the population we think we are studying, and therefore the inequalities we are able to document.

After looking at the multidimensionalities of gender, and the ways we can operationalize these in our studies, the second article turns to gender identity development as a dynamic life-course process. By applying Elder et al.'s (2003) life course perspective to gender identity development, this study demonstrates that gender trajectories can be studied in a dynamic way, and as fundamentally shaped by both internal and external processes. Previous research has shown that, when looking at gender development, self-realisation and disclosure are key processes, and that they occur at different ages for different cohorts, often in response to shifting social norms, visibility and legal recognition (Puckett, 2022; Wilkinson et al., 2018; Baams, 2025). Using sequence analysis, this chapter follows the timing and duration of key milestones rather than treating them as a single point in time.

The findings reveal clear cohort differences: Older cohorts tend to spend a long period between self-realisation and disclosure, while younger cohorts tend to realise their gender during or after adolescence, and they move to disclosing much more quickly than older cohorts. These results echo prior work that highlights the importance of social context in shaping LGBTQ+ identity trajectories (Nordmarken, 2023), but they add new detail by showing how these processes unfold for different generations, presenting demographic effects that could be potentially related to both cohort or period. The chapter shifts the focus from static labels, such as assuming that the sex assigned at birth is the life-long gender with which an individual will identify, toward a more dynamic view of gender identity as something lived and developed across time. It shows that gender development is not simply a matter of whether a person is cis, trans, non-binary, but also of when realisation occurs, how disclosure unfolds and how these experiences differ across historical periods. If demographers seek to deepen the understanding of the diversity of gendered experiences, then this study shows why and how gender could be considered a dynamic characteristic, and not solely a fixed variable based on sex assigned at birth.

Beyond developing the study of gender as multidimensional and dynamic, the third article of this thesis moves from surveys to digital trace data to present innovative data sources to study this population. This chapter explores intimate relationships by using behavioural data from a dating application and offers a look at actual interaction patterns rather than self-reported preferences or hypothetical choices. By using these data alongside demographic reasoning, the thesis connects with ongoing work in computational demography and shows how digital environments can be used to study social processes in new ways.

Prior qualitative research on dating has documented inequalities in dating experiences and relationship patterns across gender groups, as well as biases in dating apps (Blair & Hoskin, 2019; Griffiths & Armstrong, 2024; Perez & Pepping, 2024). Complementing these qualitative findings, this chapter provides a quantitative approach, and shows that transgender, non-binary and gender-diverse users are highly active and tend to give more likes than others, yet they receive fewer matches (lower reciprocity) in return. This finding speaks directly to Noble's (2018) concept of 'algorithms of oppression,' demonstrating how platform design and recommendation systems can encode cisnormative biases even in LGBTQ+ spaces. From a data feminism perspective (D'Ignazio & Klein, 2023), the lower conversion rates experienced by gender-diverse users reveal how digital infrastructures are not neutral but actively reproduce structural inequalities through both user behavior and algorithmic mediation.

The chapter expands what demographic research can say about intimate life and invites us to consider how digital environments shape the possibilities for connection and partnership formation, as being partnered has been shown to have an impact in well-being and life opportunities (Reczek, 2020; Rodríguez Sánchez & Boertien, 2025). This study also expands the possibilities in terms of data sources that we can use to study LGBTQ+ populations.

5. 2. Contributions

This thesis makes several contributions to demographic research on gender diversity. Does gender diversity relate to outcomes in people's lives? What are the relevant dimensions of gender we can consider? And what data sources and methods can be used? Although each chapter addresses a different empirical question, the studies come together to show how demography can better account for the lived experiences of transgender, non-binary, and gender-diverse people.

5. 2. 1. What role does gender diversity play in shaping educational and relational outcomes?

The thesis offers several empirical contributions that deepen our understanding of structural inequality among transgender, non-binary and gender-diverse populations. Demographic research can be expanded to include a gender diverse perspective to study life outcomes, and this work does so to relate gender diversity to education, dating, and developmental possibilities. These findings extend Badgett et al.'s (2024) economic review of SOGIE populations by documenting how gender diversity shapes not only labor market outcomes and earnings but also foundational life outcomes, such as education and dating, that structure long-term socioeconomic trajectories. By demonstrating educational penalties for trans non-binary individuals and relational inequalities in digital dating, this thesis fills critical gaps identified by Baumle and Nordmarken (2022) in the demographic study of transgender and gender diverse populations, particularly in domains beyond health and employment.

The first article provides quantitative evidence of educational disadvantage that becomes visible only when gender is measured in a multidimensional way. This helps clarify inconsistent findings in the literature and shows which groups are most affected. As an example, Wilkinson (2018) had found no evidence of educational inequalities between cisgender and non-binary individuals, while in this chapter we find strong disadvantages for non-binary individuals when they also identify as trans. Also, and in line with previous research (Carpenter et al., 2020; Kolk et al., 2025) trans individuals presented lower

educational outcomes than cis individuals. It is very much needed to understand what might be the mechanisms that are leading to these inequalities in the educational system for those who move further from gender norms.

The third article documents relational inequality in digital intimate life. It shows that gender-diverse users are active participants in online dating environments but receive fewer reciprocal responses. Previous research has shown how digital spaces, mediated by algorithms, are reproducing offline inequalities and prejudices (de Jonge et al., 2024; D'Ignazio & Klein, 2020; Kalra et al., 2023). This chapter provides new empirical knowledge on how inequality operates in digital dating apps, even those aimed at the LGBTQ+ population, and raises questions about the roles of social norms and platform design. It also contributes to the conversation about transgender exclusion from the world of dating (Blair & Hoskin, 2019; Griffiths & Armstrong, 2024; Perez & Pepping, 2024), and provides evidence on one of the potential mechanisms behind the lower percentages of partnership found for the trans population when compared to their cisgender counterparts (Badgett et al., 2024).

The second study offers empirical insight into how gender identity unfolds across the life course. It identifies cohort-specific trajectories in self-realisation, disclosure and affirmation, and shows that these processes are shaped by social context and historical change. Younger cohorts spent in general shorter periods between self-realisation and disclosure when compared to older cohorts, which could be understood to be beneficial in terms of wellbeing (Klein, 2023), and which could be related to more beneficial contexts (Nordmarken, 2023). Deepening our understanding of how gender develops could better inform other life outcomes, as the timing of these outcomes are probably closely related.

When viewed together, the chapters reveal a pattern that repeats itself across very different areas of life. Educational outcomes, the process of understanding and sharing one's gender, and interactions in digital spaces all show signs of unequal treatment for gender-diverse people. Taken as a whole, the results make clear that gender shapes people's opportunities

and constraints across the life course. This helps justify the question of whether we should incorporate gender diversity, showing that it captures meaningful variation in life outcomes.

5. 2. 2. If gender diversity relates to people's outcomes in life, what are the relevant gender dimensions that we should consider?

An important contribution of the thesis is methodological, and it cuts across all three studies. This involves expanding and testing the way gender is considered and studied in demographic work. Each chapter goes beyond the usual man/woman or cis/trans classifications that dominate demographic research, showing that the field can adopt richer and more flexible ways of representing gender. Using data that include detailed gender identity questions, the thesis compares several ways of gender categorization. This methodological contribution directly addresses calls from queer methods scholarship (Compton et al., 2018; Westbrook & Saperstein, 2015) to move beyond current approaches that simply include transgender respondents as a single category. By operationalizing gender as multidimensional, following Beischel et al.'s (2023) framework, and as processual, through a life course perspective and sequence analysis of identity milestones, the thesis demonstrates that demographic methods can be adapted to capture the complexity of gender-diverse lives without flattening them into binary comparisons.

The first article follows a multidimensional perspective to gender to test for gender inequalities. Considering identification as trans, cis or something else on one axis, and the binary/non-binary relation to gender on another axis, a multidimensional approach to gender was conducted. Having several dimensions to gender is highly valuable to find deeper inequalities even inside the plurality of gender expansive lives. Despite the fact that adding these specific questions to general surveys and instruments dedicated to population-level measurement can be complex and expensive, rethinking gender as a multiple-dimension characteristic of individuals can provide very rich perspectives. Embracing gender dimensions such as an individual's expression, gender roles, sexual identity, as relational perspectives to gender (West & Zimmerman, 1987; Butler, 1990). Even cisgender individuals

might challenge gender norms in subtle ways, and in evolving populations it is meaningful to keep an open perspective to gender, as presented in Mittleman's (2022) work.

All the chapters of this thesis include, at least, the categories of cis, trans, non-binary, and other gender-diverse identities in its analysis. The first study intersects gender categories from a multidimensional perspective, as mentioned above. The second article studies the patterns of gender development, and their commonalities, by gender groups; and the last study also uses a unidimensional variable based on gender identity, understood in an expansive way. This allows for a more detailed picture of the commonalities and divergences inside the gender-diverse population.

The thesis provides evidence that gender categories are not neutral containers but active forces that shape what becomes visible in research (Westbrook & Saperstein, 2015; Compton et al., 2018). Different ways of defining gender produce different populations and, by extension, different patterns of inequality. Taking gender as a multidimensional axis can shed light on groups that might not be seen in a simplistic approach, such as the experiences in education for those who challenge not only cisnormativity, but the binary man/woman system altogether. The results show that categorization choices matter a lot. Depending on how gender is measured, inequalities can appear, disappear or change in size. This supports a broader theoretical point: demographic knowledge is never independent from the classification systems it relies on. The ways we opt to measure populations' processes and characteristics has a direct impact on the results we can interpret from our analyses.

Understanding gender as dynamic can provide tools to deepen our studies and how we interpret the relation between gender diversity and life processes. The application of a life course perspective to gender incorporated a dynamic lens to gender experiences. In addition to previous studies that study gender transitions (Baams, et al., 2025); Wilkinson et al., 2018; Puckett, 2022), the thesis shows that it can develop through sequences of events that

unfold over time and are deeply influenced by historical and social context. Thinking about gender in terms of timing, transitions and trajectories expands the conceptual toolkit of demography and opens new ways of understanding how identities form, change and interact with structural conditions.

As can be noted, it is hard to introduce, or suggest, a detailed measure of gender that complies with all the characteristics explained above. It is important first to question which populations we are seeking to study, and which dimensions of gender are relevant for the outcome we are studying. Binary versus non-binary has been shown to provide nuances to the cis/trans dichotomy (Darwin, 2020), which could be further explored. Variations in gender development and trajectories could also be considered when studying the gender diverse population, as differences in processes and timing can lead to inequalities in other outcomes.

The thesis develops a conceptual framework in which gender is understood as a lived process that shapes opportunities across multiple domains of social life. It demonstrates that demography can engage with gender diversity in ways that respect complexity rather than flatten it, and that demographic concepts such as trajectories, exposure, timing and context can be meaningfully applied to questions of gender identity and lived experience. Taken together, the studies show that demographic analysis needs to approach gender not as a fixed attribute but as something dynamic, multidimensional, and shaped by time and context. Queer and feminist perspectives help to question assumptions behind gender categories and to think critically about how data infrastructures shape who appears in the data and whose experiences are captured, which leads to the last main question of the thesis:

5. 2. 3. Which data sources and methods help us understand how these lives unfold over time?

Another contribution comes from using several types of data rather than relying on a single source. The thesis combines a nationally representative survey from Mexico, a large community-based survey from Europe and behavioural digital trace data from a dating app in

the United Kingdom. Working with these three datasets shows how different forms of data make different experiences visible, and how each of them helps answer different kinds of questions. This multi-source strategy responds to what Guyan (2022) identifies as the 'queer data paradox': gender-diverse populations are simultaneously over-surveilled in clinical and administrative contexts yet under-represented in population-level research. By leveraging nationally representative surveys, community-based convenience samples, and behavioural digital trace data, the thesis demonstrates how different data sources make different aspects of gender-diverse lives visible, and, crucially, how researchers must critically examine who is included and excluded in each source.

Two out of three studies are based on survey data. One of them uses representative data, while the other is based on a convenience sample. Both surveys have in consideration the fact that the respondents may be gender diverse, and go beyond the man/woman and cis/trans binaries. The main strengths of these surveys for gender diversity demographic research is the inclusion of questions dedicated to expansive gender experiences, such as self-realisation and disclosure timelines, gender expression, affirmation steps, among gender identification questions that go beyond two-step questions.

The third article focuses on the use of digital trace data extending demographic inquiry to social processes that are increasingly shaped by digital infrastructures. By analysing behavioural data from a dating app, the thesis shows how gendered experiences are also mediated by algorithms, platform design, and online norms. Despite the sample being highly selected to individuals who opt to use this specific app, digital trace data have been proven useful to study population phenomena (Rampazzo et al., 2025). This brings demographic theory into conversation with work on digital sociology and computational demography, recognising that intimate life and social connection are increasingly formed in environments that blend personal agency with technological structures.

The selection of these data sources, and their limitations, are related to broader challenges in demographic research on gender diversity, where population-based data on gender diverse groups remain scarce and where methodological innovation often requires working with imperfect but meaningful sources. Selection bias can be present in all three studies, as their data sources are not representative. The second and third studies rely on non-probabilistic samples, which limits their representativeness and the generalisability of the results. The life-course analysis one chapter uses retrospective self-reports, which may be subject to recall bias, particularly for older respondents. The first study relies on cross-sectional survey data, even though the data is based on a probabilistic sample, there are still constraints on the ability to make causal claims about educational inequality. Individuals may opt not to disclose their gender identity, something that can be related to the very high prevalence of non-response to gender identification in the last study. At the time of this thesis development, these non-probabilistic data sources are our most valuable input for studying gender diversity, as representative sources rarely exist, and when they do, the specific outcomes under study are not present. Even with these limitations in mind, these imperfect data sources still provide highly valuable information on gender processes, life outcomes, and inequalities.

The thesis also expands the methodological toolkit used to study gender by incorporating inductive approaches that allow patterns to emerge from the data rather than being imposed in advance. In the analysis of gender identity development, sequence analysis and clustering are used to look at how identities unfold over time. This demonstrates that demographic methods commonly used in other areas, such as family transitions and morbimortality, can be adapted to follow dynamic processes such as gender realisation, disclosure and affirmation, which are difficult to capture with more traditional, static research designs.

This thesis demonstrates how working in an innovative way can make use of imperfect but meaningful data sources, and to understand the knowledge that can be generated from a specific source, regardless of their limitations. This relates closely to the research design,

the research questions we seek to answer on our studies, and the methodologies we opt to use on our analyses.

5. 3. Discussion and implications

The findings of this thesis show that gender diversity shapes people's experiences across different areas of life, from education to identity development and intimate relationships. Although legal and policy advances have expanded recognition and rights for transgender and even non-binary people in many settings (ILGA-Europe, 2020), the results here confirm that formal rights do not always translate into equal lived experiences and opportunities. Educational penalties, slower or more constrained pathways of identity development for older cohorts, and reduced reciprocity in digital intimate spaces point to structural inequalities that remain embedded in everyday institutions and interactions. These results echo existing work documenting discrimination, well-being gaps, and unequal access to resources among gender diverse populations (Rodríguez Sánchez & Boertien, 2025; Carpenter et al., 2022; Westbrook & Saperstein, 2015), highlighting areas that have received far less attention in demographic research.

The implications of these findings are broad. For education, they highlight the importance of inclusive school environments and anti-discrimination policies that recognise non-binary identities. For identity development, the results underline how social visibility, information and access to support networks shape the timing and pacing of self-realisation and disclosure (Baams, 2025; Wilkinson et al., 2018). For digital environments, the findings suggest that platform design and algorithmic systems can reinforce exclusion, even within LGBTQ spaces. These findings underscore a critical disconnect between legal and policy advances (ILGA World, 2024; TGEU, 2021), whether guided by the Yogyakarta Principles or national reforms, and their translation into substantive equality. The gaps identified here show that many obstacles persist in everyday settings: in schools, in relationships, and in digital environments, which suggests that policy interventions must extend beyond legal

recognition to address institutional practices, cultural norms, and algorithmic systems that reproduce inequality.

Several limitations cut across the three studies and matter for how the findings should be interpreted. First, sampling and selection: two of the three empirical chapters rely on non-probability samples (FRA LGBTQ+ Survey II and Zoe), which likely over-represent people who are more “out”, more digitally connected, and more engaged with LGBTQ+ communities or platforms. Even in the representative survey (ENDISEG), the analytical *n* for gender-diverse groups is small, which reduces statistical power, makes estimates less stable for specific subgroups (especially at intersections), and limits how far the results can be disaggregated.

Second, measurement constraints vary by data source. The life-course chapter is based on retrospective self-reports of ages at milestones, which are vulnerable to recall error (particularly for older respondents and for long intervals between events). The education chapter is constrained by survey categories, which can flatten identities and may undercount people who opt not to disclose. The dating-app chapter captures behaviour up to likes and matches, but has no post-match information (e.g., messaging, meeting, relationship formation), so observed inequalities in reciprocity cannot be directly linked to downstream outcomes or to the quality of interactions.

Third, generalisation and inference: two chapters are anchored in single-country contexts (Mexico for ENDISEG; the UK for Zoe), and even the EU-wide survey chapter reflects the specific social, legal, and community infrastructures of Europe. This matters because gender categories, disclosure practices, and discrimination mechanisms vary across institutional settings. Across chapters, designs are primarily observational and cross-sectional (or cross-sectional with retrospective timing), so the thesis is not positioned to make causal claims about the origins of inequalities; the results document patterned disparities and trajectories, but do not isolate mechanisms.

Finally, missing information limits the scope of explanation: the FRA data do not provide rich measures of socioeconomic background for modelling selection into different trajectories; the life-course chapter cannot link identity trajectories to later life outcomes in a longitudinal way; and the Zoe data do not contain match outcomes beyond the match itself. Taken together, these constraints mean the thesis likely provides a conservative and partial picture: it identifies robust inequalities and temporal patterns, but leaves open key questions about mechanisms, heterogeneity across contexts, and how early trajectories translate into later-life outcomes. This leads to several lines to continue these conversations forward.

First, there is a need for comparative work that systematically links the trajectories and outcomes of gender-diverse populations to differences in legal recognition and policy environments. Existing evidence suggests that institutional contexts shape both identity development and access to resources (Flores et al., 2016), but comparative demographic work in this area remains limited. The findings of this thesis underline why this gap matters: across the empirical chapters, results vary substantially depending on how gender is measured and categorised, as can be noted specially in the first article.

This leads to a broader comment for comparative research: we need to think about gender measurement in a way that is linked to its context and the specific gender identifications present in each population, while seeking to understand what this means for comparative research. The way people name themselves can mean different things in different contexts, so comparative research on how gender is conceptualized by people is highly needed. The divergent findings between administrative studies (e.g., Kolk, 2025; Carpenter et al., 2024) and survey-based research (e.g., this thesis; Wilkinson et al., 2021) underscore the importance of comparative methodological work. Administrative data capture a highly selected population of individuals who complete legal or medical transitions, while surveys may include broader gender-diverse groups. Comparative research must therefore account not only for cross-national policy variation but also for how different data sources produce different populations.

Second, future research should move beyond documenting inequalities to identifying mechanisms. This includes integrating qualitative and quantitative approaches, linking survey and administrative data, or studying the effects of specific contextual changes, such as legislative or public policies, on identity development and well-being. We know that inequalities exist, but we are not sure about the causes that lead to them. To produce meaningful change and equality of opportunities, we should know why inequalities in education and dating are still present. Other gender dimensions could also be included to get a more comprehensive perspective, such as the relation of legal gender rectification, medical affirmation steps, and expression, to study and understand mechanisms. As Shuster (2020) argues, understanding gender inequalities requires moving beyond single-axis analyses to examine how gender diversity also intersects with race, class, disability, and other axes of marginalization. Future research should integrate intersectional frameworks with life course methods to identify the specific institutional mechanisms that produce the educational and relational penalties documented in this thesis.

Third, research on digital environments should work to disentangle user-level bias from algorithmic bias. As relational life increasingly unfolds on digital platforms, understanding how algorithms classify, rank and present gender-diverse users becomes essential for addressing inequality (Noble, 2018). The Breeze case documented by de Jonge et al. (2024) demonstrates that algorithmic discrimination in dating apps is both measurable and legally actionable. Many studies report inequalities that are magnified by the use of algorithms, and the increasing advance in AI mediated environments call for a reevaluation of who is counted and privileged in these settings (D'Ignazio & Klein, 2020; Guyan, 2022). Following Seaver's (2019) call to study algorithms as cultural objects embedded in social practices, future research should examine not only user preferences but also how recommendation systems, ranking algorithms, and interface design shape which gender-diverse users become visible and to whom. This requires collaboration between demographic researchers,

computer scientists, and platform designers to develop methods for auditing algorithmic fairness in digital domains.

Finally, the thesis highlights a broader need within demography: to distinguish between studying transgender and gender-diverse populations as a demographic group, and applying a gender-aware lens that questions binary assumptions across demographic research. Both approaches are necessary. This thesis contributes to each by documenting structural inequalities and by showing how demographic research can be adapted to study gender beyond the binaries, and with an increasing number of surveys and data sources where gender can be studied with a diverse perspective, future lines of inquiry are ready to be pursued.

5. 4. Final Remarks

The studies presented in this thesis demonstrate that demographic research can meaningfully engage with gender diversity when it allows for complexity, fluidity and change. By combining multiple data sources, inductive methods and a life-course perspective, the thesis made that clear. The aim of this research was not to provide a final answer, but to show what we can learn when we treat gender as something lived, dynamic, and shaped by context.

Understanding populations requires recognising all the people who compose them, and by sticking to a binary understanding of gender we are not doing so. There is a part of our population that we, as demographers, know very little about due to assumptions that do not fit current societies. There remains much to do: Better data, comparative work, and closer attention to mechanisms will help refine the understanding of the lived experiences of increasingly diverse populations.

The studies presented in this thesis show that demography can expand in ways that make room for these complexities: incorporating gender diversity strengthens demographic research, both analytically and ethically. When doing so, we can observe inequalities that

remain obscured when gender is treated as a fixed, binary, characteristic, particularly among individuals whose lives divert from dominant gender norms, and help inform public policies that acknowledge diverse lived experiences.

As Lagos (2022) documents, we are living through a 'transgender tipping point' characterized by unprecedented visibility and population-level increases in gender-diverse identification, particularly among younger cohorts. Demographic research must evolve to meet this moment. The analytical and ethical call by Compton et al. (2018) that demography must develop methods capable of studying queer lives without erasing their complexity is not just an invitation for methodological innovation but a fundamental requirement for producing knowledge about increasingly diverse populations. We need to continue incorporating better tools and perspectives that are useful for observing diversity: By combining different data sources, methods and perspectives, this thesis illustrates how demographic research can engage with gender diversity more rigorously, while opening paths for further work that continues to push the field forward.

5. 5. References

Baams, L., Caprinali, A., & Fischer, M. (2025). Navigating time: A systematic review of the timing of sexual and gender identity milestones. In E. D. Rothblum (Ed.), *Sexually and gender diverse adolescents* (pp. 145–163). Amsterdam University Press. <https://library.oapen.org/bitstream/handle/20.500.12657/100842/9781040353974.pdf>

Badgett, M., Carpenter, C., & Sansone, D. (2021). LGBTQ economics. *Journal of Economic Perspectives*, 35(2), 141–170. <https://doi.org/10.1257/jep.35.2.141>

Badgett, M. L., Carpenter, C. S., Lee, M. J., & Sansone, D. (2024). A review of the economics of sexual orientation and gender identity. *Journal of Economic Literature*, 62(3), 948–994.

Baumle, A. K., & Nordmarken, S. (Eds.). (2022). *Demography of transgender, nonbinary and gender minority populations*. Springer.

Beischel, W. J., Schudson, Z. C., Hoskin, R. A., & Van Anders, S. M. (2023). The gender/sex 3×3: Measuring and categorizing gender/sex beyond binaries. *Psychology of Sexual Orientation and Gender Diversity*, 10(3), 355–372. <https://doi.org/10.1037/sgd0000558>

Blair, K. L., & Hoskin, R. A. (2019). Transgender exclusion from the world of dating: Patterns of acceptance and rejection of hypothetical trans dating partners as a function of sexual and gender identity. *Journal of Social and Personal Relationships*, 36(7), 2074–2095.

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. Routledge.

Carpenter, C. S., Eppink, S. T., & Gonzales, G. (2020). Transgender status, gender identity, and socioeconomic outcomes in the United States. *ILR Review*, 73(3), 573–599. <https://doi.org/10.1177/0019793920902776>

Carpenter, C. S., Lee, M. J., & Nettuno, L. (2022). Economic outcomes for transgender people and other gender minorities in the United States: First estimates from a nationally representative sample (IZA Discussion Paper No. 15116). Institute of Labor Economics (IZA).

Carpenter, C. S., Goodman, L., & Lee, M. J. (2024). Transgender earnings gaps in the United States: Evidence from administrative data (NBER Working Paper No. w32691). National Bureau of Economic Research.

Compton, D. R., Meadow, T., & Schilt, K. (Eds.). (2018). *Other, please specify: Queer methods in sociology*. University of California Press.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press.

Darwin, H. (2020). Challenging the cisgender/transgender binary: Nonbinary people and the transgender label. *Gender & Society*, 34(3), 357–380.
<https://doi.org/10.1177/0891243220912256>

de Jonge, T., & Borgesius, F. Z. (2024). Mitigating digital discrimination in dating apps—The Dutch Breeze case (arXiv:2409.15828). arXiv. <https://arxiv.org/abs/2409.15828>

Elder, G. H., Jr., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 3–19). Springer.

European Union Agency for Fundamental Rights. (2020). A long way to go for LGBTI equality. Publications Office of the European Union.

Flores, A., Herman, J., Gates, G., & Brown, T. (2016). How many adults identify as transgender in the United States? The Williams Institute.

Griffiths, D. A., & Armstrong, H. L. (2024). “They were talking to an idea they had about me”: A qualitative analysis of transgender individuals’ experiences using dating apps. *The Journal of Sex Research*, 61(1), 119–132.

Guyan, K. (2022). *Queer data: Using gender, sex and sexuality data for action*. Bloomsbury Academic.

ILGA World. (2024). Laws on us: Pride and protest. ILGA World.

Kalra, R. A., Gupta, P., Varghese, B., & Rangaswamy, N. (2023). Exploring gender disparities in Bumble’s match recommendations (arXiv:2312.09626). arXiv. <https://arxiv.org/abs/2312.09626>

Klein, H., & Washington, T. A. (2023). Transition milestones and psychological distress in transgender adults. *Urban Social Work*, 7(1), 66–84.

Kolk, M., Tilley, J. L., von Essen, E., Moberg, Y., & Burn, I. (2025). The demography of Sweden's transgender population: A research note on patterns, changes, and sociodemographics. *Demography*, 62(2), 349–363.

Lagos, D. (2022). Has there been a transgender tipping point? Gender identification differences in U.S. cohorts born between 1935 and 2001. *American Journal of Sociology*, 128(1), 94–143. <https://doi.org/10.1086/719714>

Mittleman, J. (2022). Intersecting the academic gender gap: The education of lesbian, gay, and bisexual America. *American Sociological Review*, 87(2), 303–335. <https://doi.org/10.1177/00031224221075776>

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Nordmarken, S. (2023). Coming into identity: How gender minorities experience identity formation. *Gender & Society*, 37(4), 584–613.

Perez, J., & Pepping, C. A. (2024). Relationship experiences of transgender and non-binary adults: Exploring dating goals, relationship structures, minority stress, sexual fetishization, and relationship victimization. *International Journal of Transgender Health*, 1–20.

Puckett, J. A., Tornello, S., Mustanski, B., & Newcomb, M. E. (2022). Gender variations, generational effects, and mental health of transgender people in relation to timing and status of gender identity milestones. *Psychology of Sexual Orientation and Gender Diversity*, 9(2), 165.

Rampazzo, F., Winer, C., & Kovačič, M. (2025). Sexual orientation on Zoe: A global analysis of a lesbian dating app. *Demographic Research*, 53.

Reczek, C. (2020). Sexual- and gender-minority families: A 2010 to 2020 decade in review. *Journal of Marriage and Family*, 82(1), 300–325.

Riley, N. E. (1998). Research on gender in demography: Limitations and constraints. *Demography*, 35(1), 5–19. <https://doi.org/10.2307/3004024>

Rodríguez Sánchez, A., & Boertien, D. (2025). Decomposing gender identity differences in subjective well-being using two Colombian surveys. *Socius*, 11, 23780231251315577.

Scandurra, C., Carbone, A., Baiocco, R., Mezzalana, S., Maldonato, N. M., & Bochicchio, V. (2021). Gender identity milestones, minority stress and mental health in three generational cohorts of Italian binary and nonbinary transgender people. *International Journal of Environmental Research and Public Health*, 18(17), 9057.

Seaver, N. (2019). Knowing algorithms. In J. Vertesi & D. Ribes (Eds.), *Digital STS: A field guide for science & technology studies* (pp. 412–422). Princeton University Press.

shuster, S. M. (2020). Carving out a niche or finding a place at the table? The sociology of transgender studies. *Contemporary Sociology*, 49(1), 2–9.

Transgender Europe. (2021). *Trans discrimination in Europe: A TGEU analysis of the FRA LGBTI Survey 2019*. TGEU.

Watkins, S. C. (1993). If all we knew about women was what we read in *Demography*, what would we know? *Demography*, 30(4), 551–577. <https://doi.org/10.2307/2061815>

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society*, 1(2), 125–151.

Westbrook, L., & Saperstein, A. (2015). New categories are not enough: Rethinking the measurement of sex and gender in social surveys. *Gender & Society*, 29, 534–560.

Wilkinson, L., Pearson, J., & Liu, H. (2018). Educational attainment of transgender adults: Does the timing of transgender identity milestones matter? *Social Science Research*, 74, 146–160.

Wilkinson, L., Shifrer, D., & Pearson, J. (2021). Educational outcomes of gender-diverse youth: A national population-based study. *Gender & Society*, 35(5), 806–837.