Intimate Partner Violence in Mexico

by

Helena Saenz de Juano Ribes

Submitted in fulfilment of the requirements for the

Degree of

Doctor of Philosophy in Economics

to

The Adam Smith Business School

College of Social Sciences

University of Glasgow

September, 2023

This page is intentionally left blank.

To my family .....

This page is intentionally left blank.

Abstract

This thesis contributes to our understanding of Intimate Partner Violence (IPV) within the field of Economics, with a specific focus on the context of Mexico, a country notorious for violence against women. Chapter 1 investigates the associations between IPV and factors such as educational attainment, employment status, and income contribution within couples. It distinguishes between two forms of IPV based on the perpetrator's intentions - Situational Couple Violence (SCV) and Intimate Terrorism (IT). The findings reveal significant positive associations between women's employment status with IPV victimization, while the link with relative education and income within couples appears to be minimal. Chapter 2 addresses a common challenge in empirical research, which is the need to combine data from multiple sources to answer research questions. It provides an extensive overview of existing statistical matching methods and introduces cutting-edge machine-learning techniques to improve and refine the matching process. Chapter 3 applies this technique to merge data from two separate surveys: the Mexican National Survey on the Dynamics of Household Relationships and the Mexican Time Use Survey. The aim is to understand how the allocation of time across different activities relates to the incidence of IPV. The results uncover disparities in intimate partner violence perpetration concerning the allocation of time across various activities such as childcare, particularly among men working fewer than 40 hours per week. In Chapter 4, the final chapter, the thesis explores the impact of the COVID-19 pandemic on IPV, with a specific focus on geographical variations in the severity of COVID-19 at the municipality level. The findings suggest that as the COVID-19 death rate increased, more households experienced heightened tensions and conflicts related to financial and employment uncertainty. However, due to the fear of infection, perpetrators exhibited less physically harmful violence.

Contents

List of Figures

List of Tables

Acknowledgements

This thesis would not have been possible without the constant support and encouragement I have been fortunate to receive over the past several years. I wish to take this opportunity to extend my sincere gratitude to all those who, in their unique ways, have contributed to my ability to complete this thesis today.

First and foremost, my deepest appreciation goes out to my advisors, Dr. Tanya Wilson and Prof. Michele Battisti. Your patience in meticulously reviewing my work, addressing my countless inquiries, and suggesting references has been invaluable. Beyond that, I am profoundly thankful for your empathy, for being patient with me, and for offering support and guidance through every setback I encountered over these four years.

I would also like to express my gratitude to Lidia Farre and Mario Berrios for introducing me to the world of research and for igniting the spark that initiated this incredible journey. Without your initial guidance, none of this would have been possible.

In addition, I extend my sincere appreciation to all the individuals who provided me with job opportunities during these research-intensive years: Wenya Cheng, Anthoulla Phella, Amy Thomas, Michele Battisti, Cecilia Poggi, Héctor Rufrancos, and Eva-María Egger. The varied experiences I gained from these roles have profoundly enriched and shaped this thesis.

I would like to thank my cherished friends in Glasgow - Claudia, Melina, Darío, Francesca, Rohan, Laura, Nikolas, Damiano, Elizaveta, Foivos, Natalia, Ahn, Gaspar, Maksym, Dev, Vladimir, Prateek, Ashraful, and Zhou - your camaraderie has transformed this experience and city into something truly memorable. Thank you for bringing balance into my work hours and for making this challenging journey more enjoyable. Particularly, I am grateful for your understanding during the most stressful periods.

I would like to give special recognition to my friend, Gema. Ever since we crossed paths a decade ago during our master's studies, you have guided and supported me on this unique journey. Thank you for encouraging me to venture to Geneva, encouraging me to pursue the PhD, and more importantly, for extending your hand when I needed it most.

I am thankful to all my friends in Valencia, especially my three dear Martas, Clara, and Alba. Your persistent presence, regardless of where I am or where I go, is a source of tremendous comfort. You all are my

terreta, and I cherish you deeply.

Thank you Rohan for being an indispensable part of this thesis. Your love, encouragement, and patience have been invaluable. I really appreciate you asking those probing questions that made me a better researcher and person. Thank you for being by my side every day, celebrating my victories, and comforting me during defeats. Above all, thank you for your patience in enduring the stress and insecurities that come with the most challenging moments.

Finally, this thesis, and indeed any project in my life, would not have been feasible without the unconditional support of my parents, Víctor and Amparo, and my sister, Mara. Thank you for supporting me through all the adventures I have embarked upon, from my determination to complete my degree in Barcelona, to my pursuit of job opportunities in Geneva, and my aspiration to pursue a PhD in Glasgow. Your constant belief in me, along with your financial support during times of need, has consolidated all my life achievements. I proudly assert that this thesis is not just my own but also a product of your guidance and support. Thank you.

Affidavit

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: *Helena Saenz de Juano Ribes*

Signature:

This page is intentionally left blank.

Chapter 0

Introduction

Intimate Partner Violence (IPV) encompasses various forms of abuse and aggression that occur within current or former romantic relationships. It is a deeply concerning and prevalent form of violence against women that transcends socioeconomic backgrounds, educational levels, racial or ethnic backgrounds, age, and religious affiliations. According to statistics from the World Health Organization statistics, roughly 30% of women worldwide have experienced physical or sexual violence from an intimate partner or former partner at some point in their lives. Furthermore, IPV is not limited to isolated incidents; it can manifest as a single episode of violence, or persist as a chronic and severe abuse spanning multiple years, with fatal consequences in some cases. Alarmingly, approximately 38% of femicides, which refers to the killing of women, are committed by individuals who share an intimate relationship with the victim.

The existing literature on IPV primarily focuses on identifying key factors that either increase women's vulnerability to IPV or provide protection against it. These factors span multiple levels, including individual characteristics, the dynamics of relationships, community-level influences, and broader societal norms. Researchers investigate these factors across diverse settings and regions, seeking a deeper understanding of the elements associated with variations in IPV prevalence. Several risk factors consistently emerge across studies conducted in numerous countries. These universal risk factors include prior experiences of violence perpetrated by intimate partners, alcohol consumption, financial strain, and elevated levels of violence within specific regions or communities. In contrast, some factors are context-specific and exhibit variability both among and within countries. These contextual factors can encompass gender roles, employment status, financial autonomy, the strength of family networks, and societal levels of tolerance towards acts of violence.

This thesis investigates several of these factors within the specific context of Mexico, a country that has garnered international attention for its concerning levels of violence against women. Currently, 22 out of Mexico's 32 states have declared a Violence Against Women Alert, mandating a series of emergency government actions when a region reports a concerning number of femicides. However, this situation only represents the tip of the iceberg, as Mexico faces more systemic and intricate forms of violence against women. Data from 2021 reports that 70% of Mexican women aged 15 or older have experienced any form of violence during their lifetime either in educational, workplace, community, familial or intimate environments. Moreover, a staggering 40% of Mexican women reported encountering some form of IPV at some point in their lives, equating to nearly 19 million Mexican women affected. Despite these alarming statistics, IPV remains one of the least disclosed forms of violence against women. Only 27% of the affected women do not disclose their experiences because they perceive them as irrelevant. The majority refrain from sharing their victimization due to fear, shame, or the hope that their perpetrators might change their behaviour.

While both the Mexican government and civil society have made consistent efforts to combat IPV, Mexico's cultural landscape is heavily influenced by three key ideologies: machismo, marianismo, and familismo. These ideologies shape social norms and prescribe behavioural expectations for individuals within society. Machismo promotes the ideal of manhood, emphasizing that men should primarily serve as providers for their families. Marianismo enforces the notion that a woman's sole purpose in life should be to become a dutiful wife and mother, confining her to the domestic sphere. Familismo emphasizes the importance of family needs over individual needs. In a society where these gender roles are deeply entrenched, any deviation from these norms can lead to increased tensions and, in some cases, violence. Understanding how these cultural factors intersect with individual, relationship, and community-level variables is crucial to comprehending and addressing the complex issue of IPV in Mexico.

This thesis comprises four chapters, each contributing to the understanding of IPV in Mexico. Chapters 1 and 4 are independent papers that delve into the determinants and mechanisms influencing IPV against women. In Chapter 2, I introduce a statistical technique that serves as the foundation for the analysis conducted in Chapter 3, where the objective is to uncover differences in behaviour within couples living with intimate partner violence. Each of these chapters is self-contained, featuring its own literature review that explores its relevance to existing research, highlights its unique contributions to the field and discusses policy implications of the research undertaken. A brief overview of the chapters is provided below:

In Chapter 1, I delve into the relationship between IPV and women's empowerment. I examine existing empowerment theories that link these two variables. The first theory, known as "marital dependency", posits that women who rely on their husbands for economic support are at a higher risk of experiencing IPV. In contrast, the second theory, termed "resource theory and male backlash," suggests that men, when threatened by a potential loss of economic power in the relationship, may resort to violence to maintain control. This research aims to determine which of these two theories holds more sway in the context of Mexico, and it investigates the relevance of specific empowerment indicators. Among these indicators, I focus on educational attainment, employment status, and economic resources. I make a unique contribution by considering a relative perspective on these empowerment indicators. Moreover, I delve into the intentions of the perpetrators behind each act of violence. The results of Chapter 1 reveal that women who are employed face a higher risk of IPV compared to those who are not employed. Notably, this risk is associated with employment status, and neither education nor financial resources exhibit a significant correlation with IPV.

In Chapter 2, I tackle a common challenge encountered in empirical research: the unavailability of all necessary information in a single dataset. To overcome this issue, I delve into statistical matching techniques. This chapter provides a comprehensive overview of the origins, evolution, and framework of statistical matching, along with its current implementations. Furthermore, I contribute to the existing arsenal of statistical matching tools by incorporating propensity score estimation and machine learning techniques. Using a carefully prepared synthetic dataset and the matching techniques outlined by D'Orazio (2017) as the baseline, I introduce more sophisticated matching variables. These variables combine the propensity score matching technique with random forests and boosted models of classification and regression trees, alongside the sum of relative weights derived from the genetic matching algorithm. The results of this analysis reveal that while propensity score matching does not outperform the baseline method from D'Orazio (2017), the sum of weights obtained from the genetic algorithm significantly enhances the overall matching process.

In Chapter 3, I put into practice the technique introduced in Chapter 2 to address a research question that would have otherwise remained unanswered. This chapter seeks to understand how individuals allocate their time based on whether they experience/exert IPV or not. Building on the findings from Chapter 1, where we learnt that employed Mexican women experience higher rates of IPV compared to their non-working counterparts and that this discrepancy is not linked to their earnings, I turn the focus to how they distribute their time. Chapter 3 begins by investigating the time allocation of both men and women in various activities such as work, housework, caregiving, and leisure separately. This study goes beyond analysing women's time alone (comparing victims with non-victims) and extends its scope to comprehend their husbands' time (comparing perpetrators to non-perpetrators). Therefore, this research contributes to the IPV literature by taking this dual approach to exploring the intimate sphere. The evidence reveals that victims of IPV, on average, dedicate more time to work in comparison to non-victims. Subsequently, the study explores the distribution of time among activities related to housework, caregiving, and leisure, taking into account the total weekly hours dedicated to work. The results demonstrate that, regardless of the amount of time spent working, there is no clear difference in time allocation towards housework between violent and non-violent settings. Instead, distinctions in time allocation are more prominent in areas related to childcare and family connections.

Finally, Chapter 4, the final chapter of this thesis, investigates the impact of the COVID-19 outbreak on IPV. Given the significant economic disparities among various regions in Mexico, the country experienced one of the highest death tolls, surpassing 300,000 casualties by the end of the pandemic. Previous pandemics had already demonstrated an increase in violence against women, prompting concerns about the potential impact of this virus on IPV. Most prior research primarily focused on the lockdown period, considering its direct influence on IPV, relying on police reports due to the unavailability of specific survey data at that time. This research contributes by examining the severity of COVID-19, as measured by age-standardized deaths, which extends the study's timeframe. Additionally, this chapter adds to the literature by using a survey dataset specifically designed for studying IPV. This approach allows for the inclusion of many unreported cases typically absent from police reports, enabling a deeper analysis of the predominant forms of IPV during the pandemic. The findings from this analysis reveal that in regions most severely affected by COVID-19, characterized by higher COVID-19-related deaths, more women experienced IPV. This increase was particularly notable in emotional, economic, and sexual IPV. However, the research also indicates a decrease in the frequency of physical IPV during this time.

This page is intentionally left blank.

Chapter 1

Intimate Partner Violence and Women's Empowerment in Mexico

Abstract

In this research, I explore the potential link between women's empowerment and Intimate Partner Violence (IPV) victimization in Mexico. Using the 2016 National Survey on the Dynamics of Households Relationships (ENIDREH), I undertake a comparative analysis of educational attainment, employment status and the income contribution within couples to discern their correlations with IPV. Moreover, this research distinguishes between two forms of violence depending on the controlling intentions of the perpetrator - Situational Couple Violence (SCV) and Intimate Terrorism (IT). The findings derived from this analysis show that women's education and employment exhibit a significantly higher association with IPV victimization, while the correlation with relative income appears to be notably minimal.

1.1 Introduction

During the last decades, intimate partner violence (IPV) has been recognized as a major public health issue in need of serious attention. Statistics from the World Health Organization (WHO) show that approximately 30% of women worldwide have been subjected to physical and/or sexual violence by an intimate partner or ex-partner, and 38% of femicides are committed by an intimate person (WHO, 2021). Reports from the WHO identify that the main factors that protect or put women at risk of IPV are education, financial autonomy, previous victimization, female empowerment, social support, and the inter-generational transmission of violence (WHO, 2012, 2021). As a result, understanding how these factors relate to the incidence of IPV in different regions or countries has naturally been an active area of research.

In this paper, I investigate the determinants of IPV in the context of Mexico. Mexico is an interesting country of choice to study theories that relate women's education level, employment status and wages with their risk of intimate violence victimization for the following reasons. Firstly, as a part of the Latino-American heritage, Mexico's culture still has instilled in it three ideologies relevant in the current context — *familismo*, *machismo* and *marianismo* (Terrazas-Carrillo and McWhirter, 2015). *Familismo* values the family's needs above the individuals, guaranteeing the family's harmony and unity. *Machismo* represents an exaggerated sense of manhood in which men are expected to be the head of the household and the main breadwinners. The *marianismo* ideal is the acceptance of women in a submissive role and with full dedication to their husband's and children's needs. Therefore, the combination of these three ideals may become a burden for women who seek out independence or try to escape from an abusive relationship.

Secondly, the Mexican government takes an active interest in implementing public policies to encourage women's empowerment. Programs have been implemented to reduce the share of women not in employment, education, or training at their working age. As a consequence, women's labour force participation has increased during the last three decades (from 34% during the '90s to 47% in 2016). However, traditional gender roles, motherhood and long working hours may become a disadvantage for women to reach gender equality in labour force participation (OECD, 2017).

This chapter contributes to the literature by deepening our understanding of intimate violence depending on the intentions of the perpetrator. Most empirical studies concentrate on the *type* of violence experienced: any violence, or categorization such as physical, emotional, economic and sexual violence. The present analysis also considers the context under which violence occurs by considering the perpetrator's intentions. I follow the framework of Johnson (1995) and Johnson and Ferraro (2000), who define specific contexts in which violence can occur. Situational Common Violence (SCV) emanates from sporadic heated arguments between partners, whereas Intimate Terrorism (IT) is exerted to maintain dominance and control (Johnson and Ferraro, 2000). The findings of the paper indicate that the increased risk of IPV victimization amongst working women occurs predominantly in situations of Intimate Terrorism (IT). This is consistent with theories where perpetrators perceive a symbolic threat to their manhood and where there are coercive controlling incentives in the use of violence (Macmillan and Gartner, 1999).

Furthermore, this research adds a distinct dimension by encompassing and contrasting the partner's characteristics. This approach aims to address potential biases that have hitherto remained little explored. Additionally, it allows for examination of the IPV association with the relative women's education and income in comparison to their husbands. I employ a Least Absolute Shrinkage and Selection Operator (LASSO) technique to forecast individuals' wages, as concerns had been raised about the accuracy of the reported income. The results reveal a positive linear correlation between the education gap and IPV, though this relationship does not attain statistical significance. This implies that heightened conflicts tend to arise when a woman possesses more education than her husband. In contrast, relative wages show negative associations, in which a greater earning potential for women corresponds to a reduced, though not significant, likelihood of experiencing IPV.

The chapter is organized as follows. Section 1.2 presents the institutional context, discussing Mexican cultural characteristics salient to the analysis, such as education, labour force participation, and household expenditures. Section 1.3 reviews the previous literature on Intimate Partner Violence in women's education, employment, and monetary contribution. Section 1.4 describes the data and defines key variables. Section 1.5 presents the empirical strategy and its main findings. Section 1.6 presents a discussion of the findings and concludes.

1.2 Institutional Context

1.2.1 Mexican Cultural Characteristics

One of the main values that historically relate to Latino American culture overall, and Mexican culture in particular, with IPV, is the ideal of *machismo* (Flake and Forste, 2006). During the Spanish conquest, indigenous men were subjugated and humiliated which created an inferiority complex with a need to be compensated (Riding, 2011). This resulted in an exaggerated sense of manhood where a proper "macho" needs to be masculine, strong, fearless, sexually aggressive, and able to consume large amounts of alcohol without getting drunk (Giraldo, 1972). As it is expressed in Terrazas-Carrillo and McWhirter (2015), in more recent times the *machismo* ideal alludes to a hyper-masculinity concept of men as the head of the household, in charge of providing for the family, proud of being hard workers, and always able to apply effective violence as a way of showing their physical power and dominance (Orozco, 1995).

The female counterpart to reinforce the culture of *machismo*, comes from the concept of *marianismo*. *Marianismo* is the expectation of women to become good wives and mothers dedicating all their time and effort to fulfilling their husband's and children's needs. Under this ideal, women should be submissive, dependent, and sexually faithful to their husbands, as well as, take care of the household and children (Flake and Forste, 2006). This is reflected in women's sacrifice of their individual aspirations, and any deviation from these norms justifies a response of violence or abuse. In particular, studies based in Mexico, show that women accept partner violence if they do not fulfil their obligations as wives (INEGI, 2007), their feminine roles (Agoff et al., 2006) or they trespass social norms on relationships, dressing code, defending their children, or being bad housekeepers (Alvarado-Zaldivar et al., 1998). For instance, Fuchsel et al. (2012) find testimonies of women who accept violence to keep their families together and blame themselves when they misdirect their submissive role as wives and mothers.

The above social roles are supported by the standard concept of Latino families. *Familismo* is the term used to emphasize family values, relationships, reciprocity, loyalty, and solidarity, above individual needs (Marin and Marín, 1991). Mexican families are characterized by either living with nuclear and extended family all together or having frequent visits (Terrazas-Carrillo and McWhirter, 2015). All members are expected to place responsibilities and obligations to their immediate family members ahead of the individual's interests (Ingoldsby, 1991). In particular, women become the centrepiece that guarantees family harmony and unity (Terrazas-Carrillo and McWhirter, 2015). On one hand, this ideal brings strong social networks, which have positive effects on mental health (Ayón et al., 2010). On the other hand, it enables a perfect scenario for partner abuse and becomes a risk factor for women (Flake and Forste, 2006). Women face social barriers when they try to escape from a tormented relationship. For instance, Castro (2019)'s evaluation of violence against women in Mexico, explained that single-parent families where women are the household head are socially considered dysfunctional families since there is a lack of male values instilled. Moreover, female-headed households are more likely to experience higher poverty risks and instability (OECD, 2017). Finally, García-Ramos (2021) finds that years after unilateral divorce legislation was introduced[1], IPV increased as a persuasion method of not getting divorced.

Finally, Mexico is one of the countries with high rates of homicide, robbery, extortion, kidnapping, and missing people (Almanza-Avendaño et al., 2018; Bergman et al., 2021).[2] Decades of organised crime and drug-related violence significantly increased after Felipe Calderón launched a war in 2007 using the military and police services to combat them (Cutrona et al., 2022; Carpenter, 2014; Castañeda, 2010). Calderón's strategy of militarizing the drug war caused an increase in criminal organizations who fought not only among themselves for territory control, but also among each other for the trafficking routes, leading to extremes levels of insecurity (Cutrona et al., 2022; Bagley, 2012; Shirk, 2011). In a context of high community violence, unaccompanied women become more vulnerable in the streets, facing additional barriers to their outside options. Au Yong Lyn (2021) finds that the impact of the gold mining boom in the North States of Mexico increased the perceptions of safety outside of the home which increased women's intra-household decision-making power. However, she did not find any significant impact on their labour force participation. Moreover, Jewkes (2002) contemplates that in societies where violence becomes usual in conflict situations, IPV becomes more frequent, especially sexual violence (Peterman et al., 2011; Swiss et al., 1998).

1.2.2 Women's indicators of empowerment in Mexico

1.2.2.1 Mexican gender education characteristics

Throughout the last decades, the education of Mexican boys and girls has experienced an upward trend regardless of the student's gender. According to data from the World Bank shown in Figure 1.1, in 2016 approximately 81% of the Mexican population had secondary education and 38% undergraduate or postgraduate studies. In

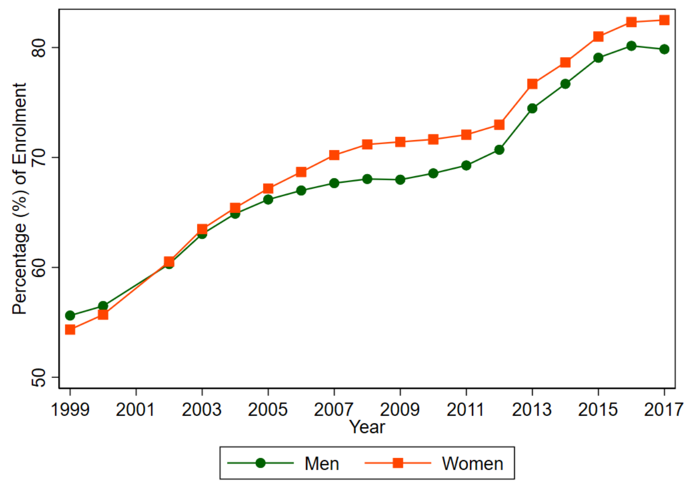
[1] The first state to implement the unilateral divorce was Mexico City in 2008, followed by other states until 2016.

[2] According to data from the UNODC (2019) Mexico ranks as the fourth country in Latin America in terms of homicide rates, with 15.7 deaths per 100,000 people, following Colombia (44.4), Venezuela (41.7) and Brazil (25.1).

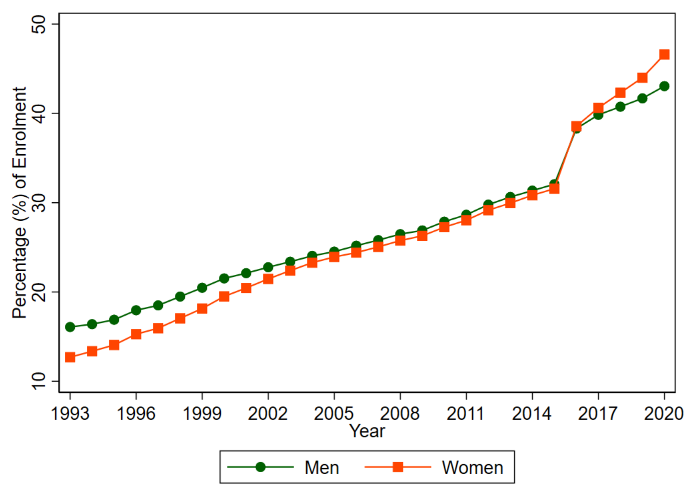
addition, in both cases the level of enrollment for women is slightly higher than that of men, thus, Mexico has managed to end the education gender gap, achieving equality in enrollment, level of education, and literacy (Zamudio Sánchez et al., 2014). Garay and Valle-Díaz-Muñoz (2011) discuss that only remaining gender differences in tertiary education are found in the initial degree choice, where STEM studies have the least female representation.

Figure 1.1: Mexican Education Enrolment over time

(a) Secondary Studies



(b) Graduate Studies



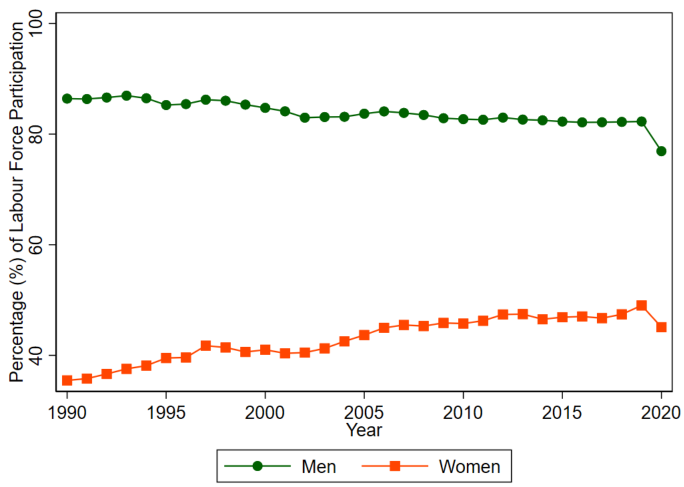
Note: Percentage of enrolment of men and women in Mexico for (a) Secondary studies and (b) Graduate studies over time. Source: The World Bank Data.

Achieving higher levels of education has allowed women to bring theoretical and practical knowledge that will help them to make informed decisions about their professional and personal aims (García and Adame, 2020). However, this is not reflected when we look at the national labour force statistics. In fact, the study from Parrado and Zenteno (2002) show that the improvement in women's education in Mexico did not change significantly the marriage timing and the after-marriage gender roles.

1.2.2.2 Mexican gender employment status differences

As seen in Figure 1.2, Female labour force participation has increased over the last decades in Mexico, from 34%, on average, in 1990 to 47% in 2016 (OECD, 2017). However, this number does not reach half of the total female population, and it is 35 percentage points below the average for Mexican working-age men (82%), and far below the OECD average for women (67%). Among the causes of the low female labour force participation, motherhood, social stereotypes, and working conditions play a major role. Mexican mothers are very much prone to leaving school or the workforce after the first child to dedicate full time to housework and childcare. This becomes a serious disadvantage for young mothers to reincorporate in the labour market since they are missing crucial time to settle their career prospects (OECD, 2017).

Figure 1.2: Mexican Labour Force Participation over time



Note: Percentage of Mexican men and women aged between 15 and 64, active in the labour force market across time. Source: The World Bank Data, estimates from the International Labour Organization.

Moreover, Mexico is one of the countries with the highest share of workers spending over 40 hours per week on the job (OECD, 2017). As mentioned above, Mexican men have huge respect for having a job and earning money to provide for their families. However, not only does overwork have negative effects on fathers, who lose well-being time spent with their family, but also this becomes an additional burden for women to reconcile work and family care. For instance, women face high discrimination from employers in the hiring process, since employers claim that during their childbearing age, women will leave their company to care for children.

Therefore, women's jobs in Mexico are highly polarized, the vast majority either work in the public sector or the informal economy. Women with easier access to the labour market prefer public sector jobs (51% of workers in the public sector are women), since these positions include job stability, better benefit packages, and the existence of anti-harassment policies, although they have difficult access to management positions and they earn much less than they would in the private sector. However, Mexican women also represent high percentages of informal jobs as domestic workers, or street vendors, in addition to unpaid informal jobs as supporters of their husband's businesses. Although women, in general, have a higher preference for being in the formal economy than men, the lack of job opportunities, access to bank credit, or a difficult business environment, becomes a greater barrier. Therefore, male self-employed workers in Mexico (27% of total male workers) are more likely to have their business registered with the government, compared to self-employed female workers (25% of a total of female workers) who are more likely to have informal businesses.

To mitigate the low levels of women's labour participation, in 2006, Mexico established an institutional framework for gender equality through the General Law on Equality between Women and Men, and the National Program for Equality and Non-Discrimination (PROIGUALDAD), which are legal frameworks established in different governments and agencies to lay down gender units in all public institutions to promote equality initiatives (OECD, 2017). From this legal framework, it created Mexico's National Women's Institute (INMUJERES) foundation to promote awareness of gender-based discrimination and foster equality of opportunity and treatment in the political, cultural, economic and social areas. Moreover, it implemented the "Estancias infantiles" (Child care centres), which brings childcare coverage for the young children of working mothers. Finally, a cash transfer program PROSPERA provided income subsidies, or scholarships, to women living in poor households for schooling their children. Although Maldonado et al. (2005); Bobonis et al. (2013) claim that the program does not increase intra-household violence, since the money is directly intended for children, Angelucci (2008) states that beneficiaries of this program, actually see increased exposure of IPV.[3] In addition, Canedo and Morse (2021) find that partnered women receivers of cash transfers in urban Mexican areas, increase their likelihood of experience physical and sexual IPV.

1.2.2.3 Economic Resources in Mexican Households

According to the National Household Income and Expenditure Survey, the main expenses in Mexican households are related to day-to-day life: food, transportation, education, and housing. These expenses are mainly supported by income from labour and each household distribute them differently (ENIGH, 2020). Despite the Mexican collective imagery that associates men with being the ones in charge of all household expenses, in some instances, both members of the couple contribute equally, or it is the woman who bears 100% of the responsibility. However, the Government of Mexico finds itself incapable of presenting specific statistics on this subject as it faces a persistent taboo in society whereby couples do not discuss their income and what they use it for (Government, 2020; CONDUSEF, 2018). In fact, there are those who hide bonuses, salary increases and even credit cards from their partner to enjoy them themselves. Therefore, little is known about how Mexican married women having their own salary affects the overall household finances and consequently, their real bargaining power which in its turn, may be related to greater or lesser marital conflicts. Some researchers have already addressed this issue in other countries as Anderberg and Rainer (2013); Dalal (2011), however, it is not clear how this can be exemplified in the specific case of Mexico.

[3] In both cases the analysis is used with the Oportunidades program, the precursor of Pospera

1.3 Literature Review

1.3.1 IPV and Coercive Controlling Attitudes

The seriousness of any act of violence is measured by the prevalence and nature of the criminal event (Sellin and Wolfgang, 1964). However, the intentions behind the crime determine whether this event can be assumed as an accident or a deliberate act of aggression (Sebba, 1984). In the intimate sphere, Johnson (1995)'s work attributes these intentions to the desire for control over the partner and proposes it as the central issue in differentiating types of violence. Following this line of research, Johnson and Ferraro (2000) defined four types of Intimate Partner Violence depending on the incentives and goal to exert the violence: Situational Couple Violence, Intimate Terrorism, Violent Resistance, and Mutual Violent Control. Most of the literature, either only considers physical assaults or lumps all types of intimate violence together. However, the failure to distinguish between them can lead to inaccurate policies that do not provide the expected results to eradicate them. The root of the differences comes from the controlling behaviour that the violent partners exert over their victims. These coercive controlling attitudes include emotional abuse, isolation, threats, using children, using male privilege, possessiveness, intimidation, and blaming (Pence et al., 1993; Johnson, 1995). In this regard, there are two distinct groups of individuals: high-controllers and low-controllers (Johnson and Leone, 2005) depending on the amount and recurrence that they exert on their partners. In heterosexual couples, it can also bring light to those gender symmetries and asymmetries discussed in the literature (Johnson, 2006).

Situational Couple Violence (SCV), previously named Common Couple Violence in (Johnson, 1995), arises when a certain couple's argument gets "out of hand", leading to physical or emotional aggression. This type of violence is not related to a general pattern of control (low-controllers), and therefore, is not likely to be frequent or escalate over time. In fact, it is endemic to family life, and, depending on the cultural context, can be considered acceptable as a conflict that sometimes leads to violence. In heterosexual couples, women are equally likely to utilize this type of violence as men, bringing a gendered symmetry to this violence. This violence is the most common to be measured in general surveys.

Intimate Terrorism (IT) is motivated by total domination over the relationship. One member of the couple manifests power and control over the other member, i.e. a high controller, apart from the use of violence. The effectiveness of these coercive controlling attitudes is derived from their relationship with the use of violence, which escalates over time and is likely to involve serious injuries. This type of violence is asymmetric in heterosexual couples, with men as the main perpetrators.[4]

Violence Resistance occurs when a victim of Intimate Terrorism decides to exert violence as an act of response or "self-defence".[5] Battered women are usually the actors of this type of violence, which becomes an

[4] In fact, due to the nature of this violence, it used to be called Patriarchal Terrorism Johnson (1995).

[5] It is not called self-defence violence since does not always meet the legal definition of self-defence.

indicator of the incentives of leaving the abusive relationship.

Mutual Violent Control emerges from both partners battling for control of the relationship, and therefore, using coercive controlling tactics and violence. It could be defined as an IT coming from both members. This type of violence is very rare since it is difficult to compare measures of violence or control where a victim also becomes a perpetrator (Howard-Bostic, 2013; Miller and Meloy, 2006).

As a contribution to the previous analysis of IPV in Mexico, this research will distinguish between Situational Couple Violence and Intimate Terrorism, since their implications are very different. The important matter in this analysis is not the act of abuse or aggression per se, but the incentives behind it. Therefore, the relevance will be to understand whether any type of exerted violence (either emotional, economic, physical or sexual) is driven by authoritarian behaviours or not. For instance, given the high rates of violence in Mexico, one could consider that most of the reported violence is due to passionate arguments rather than controlling and dominating incentives from the perpetrators. The other two types of violence cannot be considered, since the data contains only information about women's victimization and not perpetration.

1.3.2 IPV and Women's Empowerment: Marital dependency, Resource theory and Male backlash

Several theories have tried to substantiate the origin of intimate violence regarding women's economic status, however, they lead to different conclusions. The marital dependency theory suggests that women's economic independence reduces their risk of suffering IPV (Kalmuss and Straus, 1990; Anderberg et al., 2015; Villarreal, 2007). Some theorists claim that the lack of sufficient economic resources, in particular being inactive in the labour market, increases women's incidence of spousal violence (Anderberg et al., 2015; Villarreal, 2007; Capaldi et al., 2012). However, there is a complimentary version in which there is a selection effect coming from women: working women are more likely to leave abusive relationships (Vyas and Watts, 2009; Bhattacharya et al., 2009). Therefore, from this point of view, men under a spousal leaving threat, prefer not to use physical violence against them. Empirical evidence that supports this theory has been found in the UK (Anderberg et al., 2015), and in Mexico (Villarreal, 2007).

A different theoretical tradition suggests that violence is caused by an imbalance in economic resources between spouses. The resource theory is consistent with the use of violence for those partners who cannot derive power from their economic resources or employment status (Fox et al., 2004; Villarreal, 2007; Macmillan and Gartner, 1999). However, some authors have found that in heterosexual couples, the use of violence as the "ultimate resource" to maintain their domain is exerted exclusively by men (Anderson, 1997; Macmillan and Gartner, 1999; Fox et al., 2004; Villarreal, 2007; Guarnieri and Rainer, 2021; Bhalotra et al., 2021b). In her research, Anderson (1997) examines the use of violence by men and women in heterosexual couples, finding that the earning capacity provided by men has a symbolic value tied to their construction of male identity. This conclusion was also supported by Macmillan and Gartner (1999)'s research in which they identify coercive tactics from unemployed men against their employed wives. This response behaviour to sabotage women's empowerment is accommodated by the theories of male backlash (Guarnieri and Rainer, 2021).

As Vyas and Watts (2009) claim in their study, before implementing any policy or experimental program that develops women's empowerment it is essential to understand which of these above theories is the most predominant in each country or region. Two factors need to be considered: gender inequalities in society and in the intimate sphere, and the normative use and acceptance of violence to solve conflicts (Jewkes, 2002; Lawoko, 2008). This research contributes to Kessler and McRae Jr (1982) findings, which state that women's employment status, rather than their income or occupational status, has effects on husbands' mental health, which may result in intimate violence. For men, and in particular Mexican men, having a job and being the main provider of economic resources becomes essential for their construction of masculinity (Macmillan and Gartner, 1999; Ramirez, 2011).

1.3.2.1 IPV and Women's Education

Education is the most powerful instrument to improve socioeconomic status and guarantee better financial security. Therefore, the inclusion of women in higher education levels has given them the opportunity of achieving financial independence with greater-paying jobs and economic alternatives beyond marriage. In the framework of an abusive relationship, women with education become less vulnerable and more likely to leave a harmful marriage with the safety that singlehood would not be a financial problem for them (Vyas and Watts, 2009). In the US, Kreager et al. (2013) found that, overall, educated women tend to have more stable marriages than uneducated women, and a greater tendency to leave the marriage in the presence of violence.

Similar results were found in India (Boyle et al., 2009), where at the individual level, women's education prevented the risk of IPV. However, this violence protection from education varied depending on the community area of study. In those communities with higher acceptance of mistreatment, the education shielding was weaker than in those communities where there were averse attitudes towards the existence of IPV. Therefore, the individual-collective fit in the occurrence of IPV needs to be evaluated more carefully, since women's education may challenge the established roles, and so be associated with an increase in intimate violence.

For instance, Guarnieri and Rainer (2021), by comparing the French and British colonial regimes in Cameroon and their differences in the education system, brings evidence of how the employment potential that higher levels of education offer become a threat to men, who use violence to restrain the power balance. The authors find that women born in the British territories, beneficiaries of a universal education system, were 30% more likely to become victims of domestic violence compared to the French territories, where education was limited to only the elite and the regional industry was male-dominated.

In the context of Mexican education, most of the research agrees that girls' schooling is a protective instrument against IPV, specifically in the lowest social strata. In the 25 indigenous regions that Valdez-Santiago et al. (2013) focused their research on, the average woman had a low level of education and was employed in unskilled jobs such as farm labourers, housemaids, or street vendors. The authors found that for those women, each one-year increment of education functioned as a protector against severe domestic violence. Similar results were obtained in Avila-Burgos et al. (2009) where women with low education and socioeconomic status were at greater risk of victimization. Finally, Rivera-Rivera et al. (2004) contributes to this line of analysis by introducing the husband's education, since the risk of IPV decreased when both members of the couple were educated.

As women's education is a relevant asset for women's empowerment, this research will compare the abusive dynamics among the different couples with regard to their education differences. The main goal is to understand whether changes in the education gap between partners are associated with changes in intra-household bargaining power and whether it plays a part in IPV. This will contribute to the previous literature by not only measuring the absolute level of education but also the relative level compared to their partner.

1.3.2.2 IPV and Women's Employment Status

Previous evidence that focused on the relationship between women's employment status and intimate violence has yielded unclear results which highly depend on the socioeconomic context, as it manifests in Vyas and Watts (2009) revision of women's economic empowerment and its relation with IPV literature from different countries. They found a decrease in women's violence related to economic participation in Ecuador, Egypt, India, Haiti, South Africa, Bangladesh, Ukraine, and the United States (Koenig et al., 2003; Bates et al., 2004; Yount, 2005; Dude, 2007; Kishor and Johnson, 2004; Yount and Carrera, 2006; Bhattacharya et al., 2009; Capaldi et al., 2012; Gage, 2005; Hadi, 2005; Peterman et al., 2015; Kim et al., 2007; Schuler et al., 1996). In this line, Anderberg et al. (2015) also find that only unemployed women with a working husband are the ones at higher risk of suffering IPV in the UK, correlated with the marital dependency theory. Moreover, Aizer (2010) researches the income effect, by using gender wage gap and assaulting data in the US. Calculating measures of local market income in male and female-dominated industries for low-skilled workers, she found that decreases in the wage gap were correlated with a reduction of IPV.

In contrast, Vyas and Watts (2009) review other evidence for Bangladesh, India and the United States as well as in Peru, the Dominican Republic, and Nicaragua that relates positively employment, or micro-finance, with IPV (Capaldi et al., 2012; Dalal, 2011; Flake and Forste, 2006; Kishor and Johnson, 2004; Koenig et al., 2003; Naved and Persson, 2005; Rahman et al., 2011). More recently, Gage and Thomas (2017) compare risks of suffering IPV among employed women with paid work, unpaid work, and non-employed in Nigeria. Their results show that any working activity, regardless of its payment, higher the risk of IPV, increasing for non-cash work. In addition, Guarnieri and Rainer (2021) compare the Anglo-French historical border in Cameroon where women living on the British side, and with better labour opportunities, were also at higher risk of suffering spousal violence than women on the French side. Finally, Anderberg and Rainer (2013) develop an intrahousehold sabotage model which states that a woman's salary has a U-shaped IPV consequence bringing higher cases when income is low, but not low enough to dissuade employment.

There is an additional body of research that compares not only women's employment status but also their male partner's employment conditions. Macmillan and Gartner (1999) using data from the Violence Against Women Survey in Canada, found that the risk of women's labour force participation decreased only when their male partner was also employed, but substantially increased when they were unemployed. Another similar result is found in Dalal (2011), where Indian working women at higher risk of victimization were the ones with unemployed husbands. These results are congruent with Bhalotra et al. (2021b) research-based in different developing countries, men reacted adversely to women's employment prospects when they were unemployed; and in Brazil, the loss of the unemployment insurance was associated with an increase in domestic violence perpetration (Bhalotra et al., 2021a). The same results are found in European countries Costa et al. (2016), in particular in Spain, where Tur-Prats (2021) shows the different relationship between employment status and victimization depending on the family structure from the potential threat of the traditional breadwinner role from men.

For the specific case of Mexico, depending on the employment approach, the existent evidence is mixed. Villarreal (2007) using instrumental variables to address of women's employment status endogeneity, found that employed women had a lower risk of IPV. In contrast, Terrazas-Carrillo and McWhirter (2015) concluded that women's employment status was positive and significant when husbands' controlling behaviour was excluded from the regression, indicating that coercive controlling attitudes may become a mediator between employment status and IPV. In addition, Angelucci (2008) and Canedo and Morse (2021) found a positive association between receiving cash transfers from OPORTUNIDADES/PROSPERA programs and IPV. Canedo and Morse (2021) specify that IPV risk was higher among benefited women who were also employed than those who were not employed. More recently, Gupta et al. (2018) found that low-income women workers in Mexico City experienced high-level of work-related disruption such as having to change jobs, missing work or school, or losing their jobs due to IPV abuse. However, this literature based in Mexico lacks the partner's employment status comparison, nor identifies the type of intimate violence to which this research will contribute.

1.3.2.3 IPV and Economic Resources in the Household

Besides employment status, many authors have agreed that a relevant variable to consider is men's income and their ability to provide enough for the whole family. There is evidence that disadvantageous income, men's low education and low occupational status are significantly associated with female victimization (Costa et al., 2016; Anderberg and Rainer, 2013; Dalal, 2011; Babcock et al., 1993; Melzer, 2002). Fox et al. (2004) use data from the US National Survey of Families and Households (NSFH) to examine different household economic indicators on the risk of IPV. They found that those women who wanted an increase in their partner's work, men's feelings of job exhaustion or irritability, and prior experiences of violence were salient factors in increasing the risk of intimate violence. Contrary, men with a larger household income-to-need ratio, lower number of debts, or larger share of the couple's earnings (more common in households with higher incomes), and perceptions of financial well-being, reduced significantly the man-to-woman violence.

As Jewkes (2002) mentions, poorer men have fewer resources to bring the bare minimum into the house and fewer resources to reduce stress. This is intensified when social expectations of manhood are unattainable, and they can no longer control or economically support their wives. Thus, men may choose to express their dominance and power over women using violence. To give evidence of this, Anderberg and Rainer (2013) presents a model that predicts the intra-household employment sabotage of men to women depending on women's relative earnings. The authors state a non-monotonic relationship between the couple's wage gap and the incidence of economic abuse. Women's low and high relative wages decrease the likelihood of victimization since in the lowest extreme the husbands can financially support their wives to fully specialize at home, and in the highest extreme husbands are too dependent on their wives' wages to interfere with violence. However, in the intermediate relative wages women's work becomes a threat to their husband's preferred pre-established roles, therefore, they seek the increase of their financial support while making use of economic sabotage. Similar results are found in Dalal (2011), where Indian women with lower or equal wages as their partners experienced less spousal abuse compared to those who earned more.

Some of the Mexican literature uses cash transfers to analyse the level of IPV (Angelucci, 2008; Canedo and Morse, 2021). However, according to the National Household Income and Expenditure Survey of Mexico, those cash transfers represent, on average, 13% of the total income, wages being the predominant income contribution to the household. Therefore, this research will contribute to tackling specifically how the absolute and relative salary contributions may incentivise the husband's perpetration of violence.

1.4 Data

This chapter uses data from the 2016 Mexican National Survey on the Dynamics of Households Relationships (Encuesta Nacional Sobre la Dinamica de las Relaciones en Los Hogares, ENIDREH). The ENDIREH is a national representative survey administrated by the National Institute of Statistics and Geography in Mexico (INEGI). This survey includes three different datasets. One with household characteristics, another with demographic and employment characteristics for each member of the household, and a third dataset, which is specialized in collecting information on women aged 15 years or more who experienced violence in school, work, community, family and intimate settings. With the aim of avoiding under-reporting cases, ENDIREH ensures the safety of every respondent with face-to-face interviews in isolated home spaces and offers reschedules when these circumstances are not satisfied.

The individuals selected for this analysis are working-age women (between 15 and 64 years old), who are married or living in a couple with a man, and whose partners are also of working age. Using the variables of marital status and household relationship in the demographic dataset, I kept in the sample those households whose interviewed women and their partners were able to be identified, and thus, also their individual characteristics. From a total of 111,256 surveyed women, 72,855 were married or cohabiting. Moreover, I dropped those couples where one of the members was not included in the working-age interval, and with missing values in any of the key variables for this research leaving a total of 60,028 couples.

Although ENDIREH records monthly income variables for women and for their partners, it's important to acknowledge that these responses are approximations and, as such, might not accurately reflect the true values.[6] Therefore, I employ information from the 2016 year of the Mexican Labour Force Survey (Encuesta Nacional de Ocupación y Empleo, ENOE) where I estimate the potential value of the hourly wages of both members within each couple. The ENOE survey is a nationally rotating panel, tracking the same individuals for five consecutive quarters. For instance, an individual initially interviewed in the first quarter of 2016, would continue to be surveyed until the first quarter of 2017 (included). Therefore, I created a sample by pooling those non-repeated individuals for each of the quarters of 2016.[7] Overall, the dataset comprises a total of 597,039 observations from the year 2016, featuring a subset of 42,968 individuals who were both married and employed (consisting of 14,656 women and 28,312 men).

1.4.1 Measurement

Intimate Partner Violence (IPV). The ENDIREH survey covers a collection of 36 questions starting with "In the past 12 months, has your partner..." and continuing with a list of different tactics related to physical, emotional, sexual, and economic violence. As examples of *physical violence*, the survey includes questions such as "In the past 12 months has your partner pushed you? or has your partner slapped you?". Other examples of *emotional violence* comprise questions such as "In the past 12 months, has your partner shamed you? or has your partner humiliated you?". Regarding the questions related to *sexual violence*, ENDIREH includes among other questions, "In the past 12 months, has your partner used physical force to have sexual intercourse?". Finally, examples of *economic violence* tactics are considered in the questions "In the past 12 months, has your partner taken over your money? or has he spent your money without your consent?". The Appendix subsection 1.A.1.1 details and classifies all the 36 questions included in the survey. For each question, respondents can answer whether they experienced that violent tactic "one time, several times, or never". In this analysis, the IPV variable has been constructed as a dichotomous variable with a value of 1 for respondents who ever experienced any type of violence at least once during the past year, and 0 otherwise. The intention behind this definition is to evaluate whether any tactic of violence derives from a controlling pattern regardless of the number of tactics exerted and their frequency.

[6]The ENDIREH survey collects partner's income data from surveyed women which could potentially lead to inaccuracies.

[7]For example, to ensure distinct observations, individuals interviewed in the second quarter of 2016 who had already participated in the first quarter were excluded from the dataset.

Coercive Controlling Attitudes. This variable measures the degree of coercive controlling tactics employed by surveyed women's partners. It was constructed as an index in which a unit was added to the variable if respondents answered affirmatively to each of the 29 questions (Appendix subsection 1.A.1.2) such as "Does your husband/partner get mad because you work, you make more money, or he thinks you are cheating". The Cronbach's alpha coefficient for this variable is 0.896[8].

The combination of the coercive controlling attitude index with the intimate partner violence variable allows distinguishing whether any act of violence recorded may be originated from a controlling behaviour exerted by the partner. In other words, the combination of these two variables enables the construction of the two main types of violence (Johnson, 1995; Johnson and Ferraro, 2000; Johnson, 2006): Situational Couple Violence (SCV) and Intimate Terrorism (IT). Following the previous research of Johnson and Ferraro (2000); Graham-Kevan and Archer (2003) and Frye et al. (2006) I use a k-mean cluster analysis with a two-cluster solution to create low and high controlling profiles.

The k-mean clustering is an unsupervised machine learning algorithm that partitions the sample data into k disjoint clusters. In each cluster, the data points are aggregated together for their similarities with the nearest mean (or cluster centre). These similarities are usually the solution to minimization of variances using squared Euclidean distances[9]. Therefore, in this particular case, I create two different clusters (low and high controlling profiles) using the Euclidean distance, and with randomly selected means as starting points. The low profile covers a range from 0 to 9 (max) of different controlling attitudes and the high profile covers from 5 to 29 controlling attitudes. This technique allows objectively to determine whether a partner delivers many coercive controlling tactics, whether those tactics are of a similar nature, and thus, whether he will become a high controlling profile or not.

Therefore, I partition Intimate Partner Violence (IPV) into two disjoint groups depending on the profile: ***Situational Couple Violence (SCV)*** equals 1 for those cases of IPV with low profile controlling attitudes, while ***Intimate Terrorism (IT)*** equals 1 for those cases of IPV with high controlling attitudes. In other words, SCV is defined as any act of violence ($IPV = 1$) experienced by the surveyed victim that has been committed by a non-controlling partner, and thus, there are no controlling incentives behind it. Contrary, IT is defined as any act of violence perpetrated by a high-controlling partner with the clear intention of bolstering the power of the relationship.

The sample of analysis in ENDIREH 2016 shows that 29.6% of the women experienced Intimate Partner Violence (IPV) during the previous year. The prevalent form of this violence is Situational Couple Violence

[8] The Cronbach's alpha coefficient measures the internal consistency of a set of surveyed questions, i.e., whether this set of questions measures the same characteristic. With a 0 to 1 standardized scale, high Cronbach's alpha values indicate that the responses of surveyed individuals are consistent, compared to low values, being 0.7 as the benchmark value (Frost, 2023)

[9] For a full description of the k-means algorithm, see Appendix 1.A.2.1

(SCV), which comprises 17%, while the remaining 12.6% falls under the category of Intimate Terrorism (IT). The distribution of each type of violence is outlined in Table 1.1 which is interpreted as follows: for each type of violence (IPV, SCV or IT) I compute the percentage of women who had experienced physical, emotional, sexual or economic violence along with their corresponding standard deviations. Therefore, the cumulative numbers may not add up to 100%, as some victims could have encountered multiple forms of violence. Table 1.1 reveals noteworthy patterns. Overall, when women are subjected to IT, there are considerably higher levels of physical, emotional, sexual, or economic violent experiences compared to instances of SCV or the broader category of IPV. Indeed, the percentages for physical and sexual violence exhibit a substantial increase compared to both IPV and SCV. Specifically, the figures rise from 15% and 26% to 40% for physical violence, and from 2% and 7% to 14% for sexual violence. These findings align with prior research suggesting that IT is associated with more severe injuries and poses a greater threat to victims' lives (Johnson and Leone, 2005).

Table 1.1: Violence

	Intimate Partner Violence Percentage (SD)	Situational Couple Violence Percentage (SD)	Intimate Terrorism Percentage (SD)
Physical	26.41% (0.44)	15.43% (0.36)	41.21% (0.49)
Emotional	86.27% (0.34)	81.83% (0.39)	92.26% (0.27)
Sexual	7.60% (0.26)	2.41% (0.15)	14.60% (0.35)
Economic	53.73% (0.5)	42.01% (0.49)	69.53% (0.46)
Observations	17754	10193	7561

Note: Types of Intimate Partner Violence distribution of married couples where both members are in working age. Each column represents the form of violence coming from the perpetrators' intentions (i.e., column 1 is for Intimate Partner Violence (IPV), column 2 is for Situational Couple Violence (SCV), and column 3 is for Intimate Terrorism (IT). Each row represents the percentage of physical, emotional, sexual and economic violence for each form of violence (i.e., those who suffer from IPV, 26% suffered from physical violence, 86% from emotional, etc.). Source: ENDIREH 2016

In addition to data related to IPV, the ENDIREH survey includes various demographic and determinative factors pertinent to this analysis. For instance, it encompasses information on individual and household characteristics, indicators reflecting gender role attitudes and intra-couple bargaining power, as well as intergenerational transmission of violence. Detailed statistical descriptions for all variables concerning both genders (women and their husbands) are presented in Table 1.2.

Individual and Household Characteristics. Among the set of demographic variables, I have included those particularly relevant to IPV based on (Rennison and Welchans, 2002). These encompass age, education level, ethnicity, rural, poverty status, children present in the household, health insurance coverage, and employment status. As detailed in Table 1.2, the mean age of married women interviewed for the survey is 38.2 years, while the mean of their partners is 41.7, three years older. The variable for ethnicity is an indicator of whether the individual considers her/himself indigenous or not, I find that both genders have a similar percentage (around 27.5%). The variable education level is divided into six main categories: No education, primary education, secondary education, preparatory (high school), university studies, and postgraduate studies. Table 1.2 show that there is a slightly higher percentage of men with higher education. As the OECD (2017) report claims, the gender educational gap is narrow enough to consider that both genders have the same educational opportunities. However, Table 1.2 presents an evident gender inequality in labour force participation (OECD, 2017). In this sample, only 40.7% of women are employed compared to 92.8% of their male counterparts, and from the non-employed category group, the majority of women identify as homemakers (55.9%).

Furthermore, Table 1.2 presents household common variables that take a value of 1 when surveyed women answer affirmatively to questions of whether they live in rural areas, have given birth to living children, are covered with health insurance IMSS or are covered with Seguro Popular[10]. Moreover, the ENDIREH survey includes housing characteristics that determine whether the family lives under the poverty line. The National Council for the Evaluation of Social Development Policy (CONEVAL) established different indicators that measure the level of poverty in Mexico. Among those indicators, CONEVAL includes housing quality and space and access to basic housing services such as "the material of the floors of the house is ground; the ratio of people per room is greater than 2.5; the water is obtained from a well, river, lake, stream or pipe, or by hauling from another dwelling, or from the public or hydrant wash; they do not have a drainage service, the drain is connected to a pipe that leads to a river, lake, sea, ravine or crack; and does not have electricity".

Therefore, from Table 1.2 it can be observed that 27% of the surveyed married women live in rural areas, and nearly half (46%) of them live under poverty conditions[11], almost every one of them has given birth to at least one living child (93%), less than a third of the households (32%) have at least one of the members working in a formal job, and another 38% prefer to have health coverage from Seguro Popular.

Gendered Role Attitudes (GRA). ENDIREH collects 9 questions that relate to gender role differences (Appendix subsection 1.A.1.3). Therefore, the GRA index is created by adding one unit for an affirmative answer to questions such as "Do you agree that women should be responsible for caring, men should earn more money or have better jobs than women?", or rather a negative answer to questions such as "Do you agree that men should be equally responsible for house chores and caring, or do women have the same right to party alone?". By employing this method, the index captures the level of gender inequality beliefs ranging from 0 (indicative of more gender-equal beliefs) to 9 (indicative of less gender-equal beliefs). This index has a Cronbach's alpha of 0.65, indicating a moderate level of internal consistency. The sample's average of the GRA index stands at 2.47, providing insight into the women's low endorsement of more traditional gender role attitudes.

Asking for permission index. This index controls women's dependency on their partners in decision-making from 7 different questions such as "Should you ask permission to work, go shopping or visit relatives

[10]The health coverage IMSS is also commonly used in the literature as a proxy for formal employment since workers need to be registered to be covered. It only needs one member of the family to be registered to cover the rest. Seguro Popular, however, only includes informal workers who are registered for basic health coverage

[11]This percentage is not far from the national 42% of CONEVAL statistics

Table 1.2: Individual Characteristics

	Women	Men
Demographic Characteristics		
Age	38.16 (11.07)	41.67 (11.25)
Indigenous	27.54%	27.49%
Household Head	5.47%	89.34%
Education Characteristics		
No Education	3.64%	3.84%
Primary	24.36%	26.14%
Secondary	37.49%	33.22%
Preparatory	19.38%	19.60%
University	13.82%	15.46%
Post Graduate	1.30%	1.73%
Employment Characteristics		
Employed	40.74%	92.77%
Unemployed	0.22%	1.53%
Student	0.61%	0.11%
Retired	0.82%	1.84%
Housekeeper	55.89%	0.22%
Other	1.72%	3.53%
Inter-generational transmission of violence		
Childhood IPV	3.31 (1.66)	0.49 (0.71)
She/He beats kids	32.93%	15.00%
Income		
Monthly Income	5706.27 (16525.84)	6683.9 (7781.75)
Predicted Hourly Income	39.15 (19.16)	42.84 (20.76)
Household Characteristics		
Rural	26.18%	
Poverty Status	46.74%	
Children	93.18%	
IMSS	31.50%	
Seguro Popular	37.99%	
Social Norms		
GRA	2.47 (1.91)	
Ask	0.47 (1.17)	
Observations	60029	

The distribution of individual and household characteristics pertains to married couples where both members fall within the working age bracket. The presentation features percentages for categorical variables, while continuous variables are depicted as Mean (Standard Deviation). The table also encompasses the six primary tiers of education and employment status, where "Other" refers to individuals responding affirmatively to circumstances such as having had a job but not working (due to holidays, leave, etc.), experiencing physical or mental limitations that hindered work, or simply not working for various reasons. The inter-generational transmission of violence is captured through indicators related to childhood experiences for both men and women, along with whether they perpetrate violence against their own children. Social norms are gauged through the Gender Role Attitudes (GRA) index and the index reflecting wives' practice of seeking their husband's permission (Ask). Source: ENDIREH 2016.

and friends?" (Appendix subsection 1.A.1.4 details all questions). This means that the higher the value of the index, the more dependent is the surveyed woman to her partner's judgements. Cornbach's alpha for this variable is 0.8, with a sample index average of 0.47. Therefore, this sample exhibits a strong tendency to endorse less dependent behaviours within the relationship.

Inter-generational transmission of violence. This index determines the extent of violence encountered or witnessed by each respondent during their childhood including physical, emotional or sexual violence. ENDIREH questionnaire covers questions like "In the place where you lived until you were 15 years old, do you remember if, between the cohabiting adults, there were beatings or curses? Did those adults beat you? Did someone abuse you sexually?". Appendix 1.A.1.5 details the 10 questions about women's childhood experiences. To formulate this index, I added one unit value for each affirmative response to these questions. The resulting Cronbach's alpha for this index is 0.73.

Moreover, the same surveyed women are also queried about their awareness of their partners' childhood experiences involving IPV. Specifically, they are asked if their husbands endured or/and witnessed violence before reaching the age of 15. To measure it, I created a categorical variable that gives a value equal to one when they either witnessed or experienced IPV, and a value equal to two when both things happened. Additionally, I incorporate controls for the transmission of violence to the next generation, by including two identifiers of whether she or her partner resorts to physical aggression against their children when provoked. These variables may serve as a proxy for evaluating the presence of violence within the household as a means of conflict resolution.

As highlighted in Table 1.2, the mean value for women's childhood IPV exposure is 3.3 (ranging from a minimum of 0 to a maximum of 10), and the average of the two categories for their partner's experience is 0.49. Moreover, Table 1.2 presents that 32.7% of the sampled women physically harmed their child after becoming angry compared to 14.8% of their husbands. While this might suggest that women are twice as likely as their husbands to resort to violent behaviour, it is important to note that these variables do not quantify the intensity or trauma caused by such actions. Instead, they provide a good insight into the normalization of violence within the household.

Monthly and Hourly wages. Interviewed women answer the following questions in which the monthly income for her and her partner is derived: "Approximately, how much money do you earn from your job?" and "Approximately, how much money does your husband earn from his job?"[12]. Since these values are approximated values and, therefore, they could be imprecise, I calculated hourly wages through a LASSO computation in ENOE and applied the obtained coefficients in the labour force ENDIREH variables.

[12]Income records valued as 999,997 were dropped from the analysis. Despite this value means that the individual earns more than that amount, it becomes an outlier biasing the income mean.

LASSO is a regression analysis method that improves the prediction accuracy of linear regression models by including a regularization term, which is optimized after cross-validation processes. LASSO modelling allows for a great number of covariates, selecting those more relevant for the predictions, and thus, it is also suited for models with high levels of multicollinearity[13].

The explanatory variables that were common in both the ENDIREH and ENOE samples were: age, education level, gender, marital status, state, urban/rural, IMSS registration, and occupation. Although the LASSO process is a precise predictor, key variables are missing to refine this calculation. In fact, I compute the hourly wages as the outcome variable since ENDIREH misses the information on how many hours the individuals work. The LASSO was computed for married women and married men separately using both the same covariates mentioned above. For the predicted wage for women, the LASSO model used 59 non-zero covariates of the total 65 covariates in a sample of 14,656 observations. With an estimation of 72 lambdas the optimally selected lambda value is 0.077. To predict the hourly wage of a sample of 28,312 married men, 62 non-zero covariates were used with a selected optimal lambda valued at 0.043 of 71 lambdas estimated. Table 1.2 presents the average and the standard deviation of both, the monthly reported income and the hourly predicted income for each surveyed woman and her husband. In both scenarios, the average income of men surpasses the average income of women.

1.5 Empirical Strategy and Results

The first regression I consider is the association between the different control variables using a probit model. The dependent variable $Violence_i$ are the three different kinds of intimate violence — general Intimate Partner Violence (IPV), Situational Couple Violence (SCV), and Intimate Terrorism (IT), for each woman i. Among the independent variables X_k, I include demographic variables (age, belonging to an indigenous minority, rural/urban setting, living under the poverty line, the age of marriage, parenthood, health coverage), social norm variables (being head of the household, believing in gender roles, women asking permission from her husband) and violence experience (childhood experience for both members and resort violence to their children).

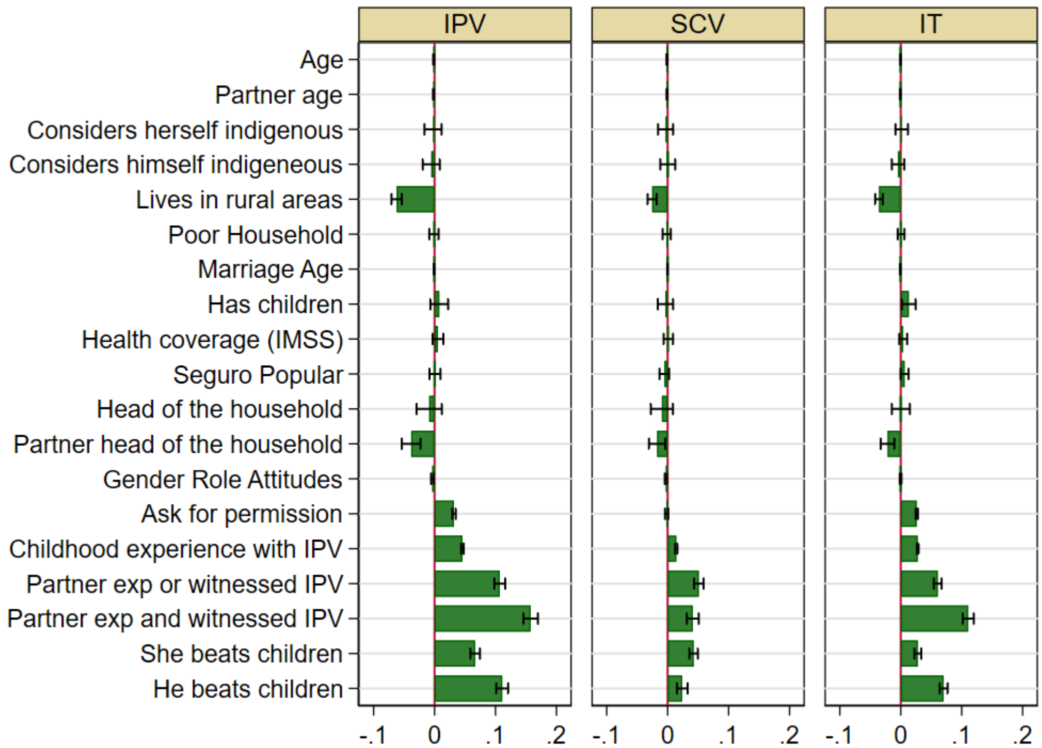
$$Violence_i = \beta_k X_{ki} + u_i \tag{1.1}$$

Figure 1.3 displays the first results. Each column represents each dependent variable for each of the different types of violence. There is a negative association between the different forms of IPV and variables age, partner's age, living in rural areas, the partner being head of the household, and agreement on gender role norms. Contrarily, variables that affect positively the risk of IPV are parenthood, demanding the husband's permission before any decision, childhood IPV experiences for both members and if they use violence against

[13]For a full description of the LASSO methodology and application, see Appendix 1.A.2.2 section

their children. These associations are in line with previous studies in the literature which focused on these types of variables.

Figure 1.3: Regression with control variables



Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Point estimates (end of bars) along with their 95% confidence intervals (line intervals) are displayed. Number of Observations: 60028 in all regressions. Number Coefficients are represented in Tables A1 and A2 in Appendix. Base Line Categories: Does not consider herself indigenous; Does not consider himself indigenous; Lives in urban areas; Does not live in a poor household; Does not have children; Does not have IMSS health coverage; Does not have Seguro Popular health coverage; She is not the head of the household; Her husband is not the head of the household; The husband never experienced or witnessed IPV during infancy; She does not exert violence against her children; He does not exert violence against his children.

1.5.1 IPV and Education Gaps

In the initial phase of the analysis, I explore to what extent the education levels attained by women compared to those of their husbands can be related to the occurrence of abuse. To this end, I have classified the level of education of both men and women into three distinct groups: (1) primary or non-education, (2) secondary or preparatory (high school) education, and (3) university education (encompassing both graduate and postgraduate studies). Based on these categories, I have derived the variable Education Gap (as defined in equation 1.2) which encapsulates the difference between the education level received by the men (husbands of the surveyed women) and the women themselves.

$$EducationGap = Partner's\ Education\ Level - Women's\ Education\ Level \tag{1.2}$$

Table 1.3 provides the sample distribution of the five different education gap categories. These numbers suggest a normal distribution pattern, where approximately two-thirds of women (66%) are married to partners with the same educational backgrounds. Additionally, there is an equivalent representation of a one-level education gap (16%) for couples with either men or women with higher education attainment, and notably a lower occurrence of two-level gap couples. Moreover, Table 1.3 offers insights into the corresponding violence percentages. In couples where a woman has a higher education level than her husband higher percentages of violence emerge. This indicates that in cases where the wife possesses more advanced education than her husband, increases the incidence of conflict, suggesting a response to the perceived shift in traditional gender roles.

Table 1.3: Education Gap between Couples

	Total Percentage	Total Observations	Percentage of IPV	SCV	IT
Education Gap = -2	0.45%	271	35.42%	22.51%	12.92%
Education Gap = -1	16.39%	9837	31.49%	17.41%	14.08%
Education Gap = 0	66.44%	39882	29.37%	16.97%	12.41%
Education Gap = 1	16.06%	9639	28.53%	16.52%	12.01%
Education Gap = 2	0.67%	400	24.00%	15.25%	8.75%

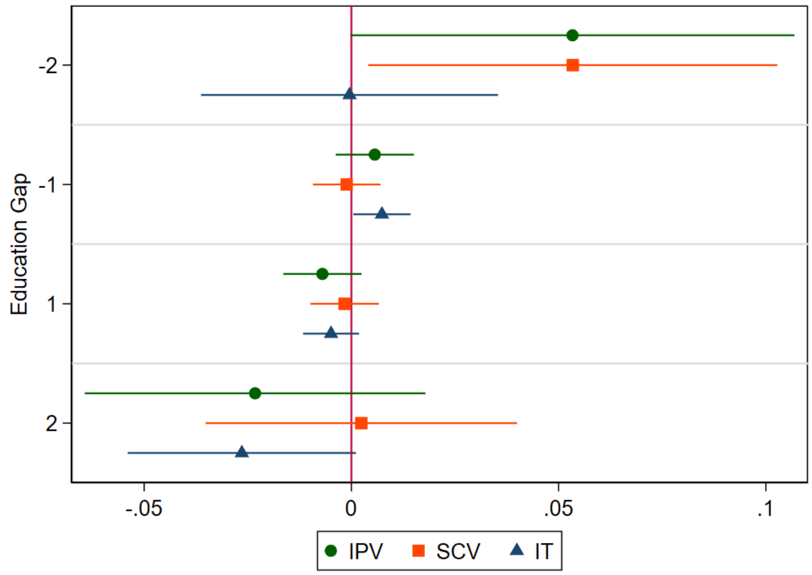
Note: Education gap distribution of married couples. Education levels are classified into three categories: primary or less education; secondary education and tertiary education (graduate and post-graduate studies). The negative Education Gap relates to those couples where women have a higher level of education than their husbands. The table presents the Percentages of women who suffer from Intimate Partner Violence (IPV), Situational Couple Violence (SCV) and Intimate Terrorism (IT) for each Education Gap category.

To investigate the significance and nature of this behaviour, I proceed to conduct a probit model by regressing the variable $Violence_i$ comprising Intimate Partner Violence (IPV), Situational Couple Violence (SCV), and Intimate Terrorism (IT) using equation 1.3. In this regression, the variable $EducationGap_i$ represents the education gap level within couples. Additionally, the vector x_{ki} encompasses the principal demographic, gender role attitudes, and generational transmission of violence controls listed in equation 1.1.

$$Violence_i = \mu EducationGap_i + \beta_k x_{ki} + u_i \tag{1.3}$$

Figure 1.4 visually represents the outcomes derived from the probit regression. One of the primary observations is the negative linear correlation between the IPV coefficients and the education gap existing between spouses. This negative relationship attains statistical significance at the 10% level for the extreme categories of 2, and -2 (for detailed values, see Table A3 in Appendix). This indicates that the husband's educational advantage over the wife decreases there appears to be an increase in conflicts within the household. Specifically when compared to couples with equal levels of education ($EducationGap = 0$), those where the wife possesses an educational advantage of two levels[14] exhibit a significant increase in the incidence of Situational Couple Violence (SCV).

Figure 1.4: Women's Education on IPV



Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Point estimates (dots) along with their 95% confidence intervals (lines) are displayed. Number of Observations: 60028 in all regressions. Controlled by demographic characteristics, gender norm variables, and generational transmission of violence. Baseline comparison: Education Gap = 0 (same level of education). Education levels are classified into three categories: primary or less education; secondary education and tertiary education (graduate and post-graduate studies). The negative Education Gap relates to those couples where women have a higher level of education than their husbands. Positive numbers refer to couples where the partner has a higher level of education. Coefficients' numbers are represented in Table A3 in Appendix.

1.5.2 Employment Status and IPV

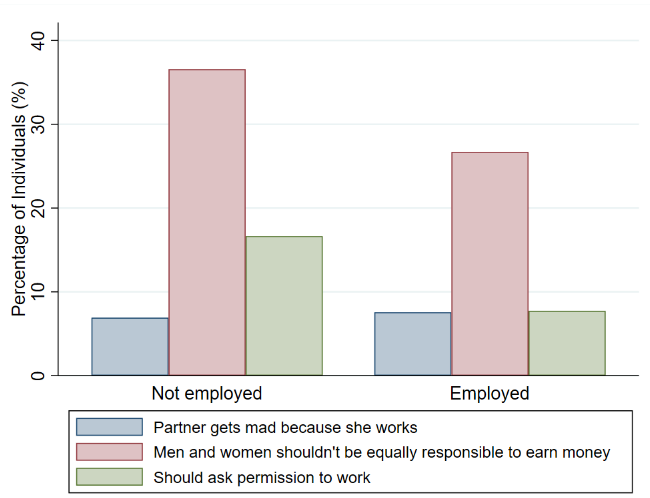
In the ENDIREH survey, individuals are defined as employed if they have engaged with work during the preceding week either as employees, self-employed, employers or unpaid workers. Conversely, those not fitting this criterion are classified as not employed (including both the unemployed and non-active population)[15]. Moreover, the ENDIREH survey incorporates questions that link gender attitudes with women's roles in employment which I selected the three following ones: *"Does your partner get mad because you work", "Do you think men should and women shouldn't be equally responsible to earn money", and "Should you ask permission to work?"*. These attitudes are measured on a dichotomous variable of 0 and 1. Figure 1.5 portrays the percentages of these attitudes towards the role of women's employment comparing employed women to those not employed. On average, a smaller proportion of respondents agree that men and women should assume different responsible roles about who should be the main provider when women are employed (27% agree to

[14] This case refers to those couples where the woman has at least graduated studies and her spouse primary or non-education.

[15] The Appendix subsection B.2.1 shows the determinants of being employed for the surveyed women

this statement) compared to when they are not (37%). In addition, fewer employed respondents hold the belief that they should seek their husband's permission to work (8% compared to 17% of non-working women). However, there is a slight increase in the share of partners who become upset when their wives work if women respondents are employed than when they are not (8% of men become mad when their wives work compared to 7% when they are not employed). This suggests that despite the more progressive beliefs of employed women concerning their work status, they might encounter a less favourable reaction from their partners.

Figure 1.5: Gender Role Beliefs' of Women's Employment



Note: Share of women asked about their perceived ideals over gender roles in work comparison by their employment status. The higher the percentage the more women believe in that specific statement. Source: ENDIREH 2016

To evaluate the association of women's and men's employment statuses on IPV, I classify couples into four distinct types. In the first category, both spouses are employed. In the second, only women are employed while men are not employed. The third category includes couples where only men are employed, and in the fourth category, neither spouse is employed. Table 1.4 presents the distribution of these couples across the sample. Approximately 55% of the sample is composed of traditional gender norms couples, where men are the primary earners and women undertake homemaking roles (couple type 3). Following this, around 40% of the couples are both spouses employed (type 1). Subsequently, in nearly 5% of the couples, neither of the partners is employed (type 4), and finally, in only 2% of the couples, the woman is employed with a not employed spouse (type 2). Furthermore, this table presents insights into the disparities in experiencing IPV. Notably, households where women are the primary earner show a higher percentage of IPV, with a difference of almost 9 percentage points with the lowest couple, where neither spouse is employed. Similar variations are observed in cases of intimate terrorism (IT), and situational couple violence (SCV). Therefore, it is plausible to expect a male backlash in response to women's engagement in the workforce.

To explore the potential existence and significance of male backlash against women's employment, I implement a probit model targeting the three types of *Violence*: Intimate Partner Violence (IPV), Situational Couple Violence (SVC), Intimate Terrorism (IT) over the different types of couples (variable *Couple*). Moreover, the

Table 1.4: Types of Couples

	Total Percentage	Total Observations	Percentage of IPV	SCV	IT
Type 1: Both Employed	38.34%	23014	32.06%	17.95%	14.11%
Type 2: Only Women Employed	2.40%	1442	33.43%	17.75%	15.67%
Type 3: Only Men Employed	54.44%	32678	28.03%	16.52%	11.51%
Type 4: Both Not Employed	4.82%	2895	25.32%	14.06%	11.26%

Note: Employment distribution of married couples compared to their husbands. The table presents the Percentages of women who suffer from Intimate Partner Violence (IPV), Situational Couple Violence (SCV) and Intimate Terrorism (IT) for each Type of Couple. Source: ENDIREH 2016.

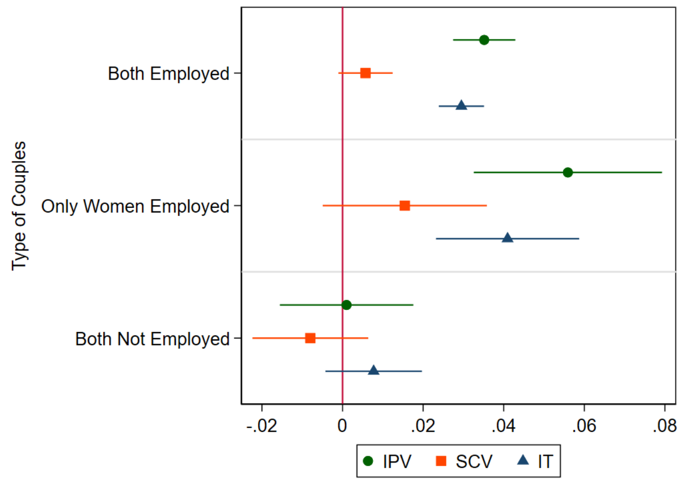
model controls for demographic characteristics, social norms and inter-generational transmission of violence, applied in regression 1.1 (X_{ki} in the regression)[16]. The variable *Couple* compares the traditional type of couple (where men are employed and their wives are not employed, type 3) with the rest of the couples. This approach investigates whether exists a male backlash effect when couples deviate from the established traditional couple.

$$Violence_i = \alpha Couple_i + \beta_k X_{ki} + u_i \tag{1.4}$$

Figure 1.6 visually represents the marginal effects of each type of IPV for different couple categories in comparison to the traditional couple (type 3). Remarkably, within couples where both spouses are employed (type 1), the likelihood of experiencing Intimate Partner Violence (IPV) exhibits a significant average increase of 3.6%. Specifically, as illustrated in the graph, this significant increase is attributed to Intimate Terrorism (IT) with a marginal effect of 3%. Moreover, for the type 2 couples, where only women are employed, the coefficients surpass type 1 couples (increasing the likelihood by approximately 5%). However, it is important to acknowledge that due to the limited sample size of type 2 couples (where only women are employed), the marginal effects are less precisely estimated, and are not significantly different from type 1 couples (where both partners are employed). These results affirm the hypothesis that regardless of men's employment status if their spouses are employed, they are more prone to suffer IPV compared to non-employed spouses. Specifically, this escalation of violence is predominantly linked to high levels of coercive controlling attitudes characterized by intimate terrorism (IT). Appendix Figure A5 illustrates the outcomes for various couples categories (type 1, type 2, and type 3) by different employment types (paid worker, employer, self-employed, unpaid worker). The findings suggest limited statistical significance across most categories. Only in the case of couples of type 1, where both spouses are employed, there is a significant increase in both IPV and IT when husbands hold the role of employers compared to being paid employees.

[16]The controls of these regressions includes women's and their spouses' education level.

Figure 1.6: Type of Couple depending on their Employment Status on IPV



Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Point estimates (dots) along with their 95% confidence intervals (lines) are displayed. Number of Observations: 60028 in all regressions. Controlled by demographic characteristics, gender norm variables, and generational transmission of violence. Baseline comparison: Couple Type 3: Only Men Employed (Traditional employment standards). Coefficients' numbers are represented in Table A4 in Appendix.

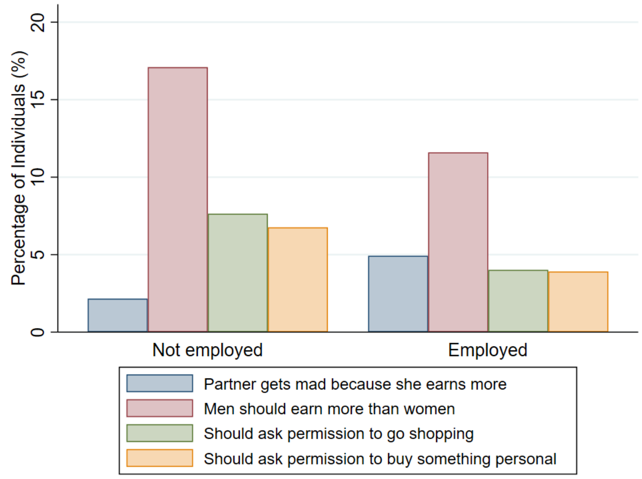
1.5.3 Income and IPV

Addressing the financial contribution to the household, the women surveyed in ENDIREH were posed questions about their viewpoints regarding their beliefs on who should be the primary household earner and their autonomy on how to spend their own income, and their spouses' reaction towards their wages. I selected the following questions: *" Does your partner gets mad at you because you earn more money?", "Do you think men should earn more than women?", "Should you ask permission from your partner to go shopping, or to buy something personal?".* Figure 1.7 illustrates the distribution of these beliefs among both employed and non-employed women. Overall, working women tend to disagree more with the notion of men being the breadwinner and are less inclined to seek permission for their personal expenditures than their non-employed counterparts[17]. However, akin to the results from the employment status beliefs in Figure 1.5, there is a higher proportion of spouses who become more aggressive when their wives earn more money than them, particularly when the women are employed. Specifically, approximately 5% of husbands react negatively when their employed wives earn higher incomes, whereas this percentage drops to 2% for husbands with non-employed wives).

This analysis narrows the focus exclusively to couples where both members are employed (Couple Type 1). The main goal of this is to scrutinize whether the root correlation of intimate violence is associated with economic resources or solely employment status. Therefore, I exclude couples in which only women are employed (Couple Type 2) as it signifies a dual threat: not only she is employed but also the sole (hence, the primary)

[17] 12% of working women agree that men should be the main provider compared to 17% of non-working women. 4% of employed women usually seek permission to go shopping or buy something personal compared to 8 and 7% of their non-employed counterparts.

Figure 1.7: Perceptions on the Monetary Contribution by Women's employment status



Share of women asked about their perceived ideals over gender roles in their wages. The higher the percentage the more women believe in that specific statement. Source: ENDIREH 2016

income provider for the household. Table 1.5 details the distribution of the reported monthly income and the predicted hourly income for these types of couples. Consistent with the trends observed in Table 1.2, women's median and average monthly wages are lower than their male counterparts. Similarly, this trend extends to the predicted hourly wages obtained from the Mexican Labour Force Survey (ENOE, 2016).

Table 1.5: Distribution of the Monthly and Hourly income for Couples Type 1: Both working

	Mean	SD	Median	Observations
Women's Monthly Income	5893.62	(17479.89)	4300	17569
Men's Monthly Income	7501.14	(8285.315)	5590	17538
Predicted Women's Hourly Income	39.37	(19.33)	34.63	14320
Predicted Men's Hourly Income	46.42	(21.31)	39.97	14320

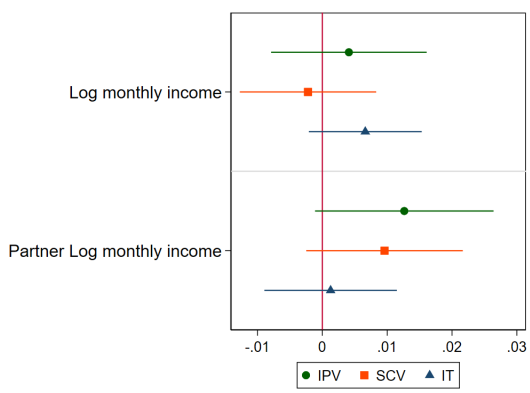
Note: Women and men's monthly income of married couples where both members are working (Couple Type = 1). Mean and Standard Deviations (SD) are displayed. Source: ENDIREH 2016. Women and men's LASSO-predicted hourly income of married couples where both members are working (Couple Type = 1). Means and Standard Deviations (SD) are displayed. Source: ENOE 2016.

In order to discern the potential effect of an individual's income on intimate violence, I conduct probit regression analyses involving the three types of violence: Intimate Partner Violence (IPV), Situational Couple Violence (SVC) and Intimate Terrorism (IT) over the logarithm of the women's and their husband's income ($logIncome_i$ and $logPartnerIncome_i$ respectively). These regressions also control for demographic characteristics, social norms indicators and inter-generational transmission of violence (X_{ki}). the same controls from regression 1.1 and the log of income. I initiate the analysis by regressing the aforementioned outcome variables against recorded monthly incomes obtained from the ENDIREH survey. Following this, I replicate the regressions using the predicted hourly incomes derived from the ENOE survey data.

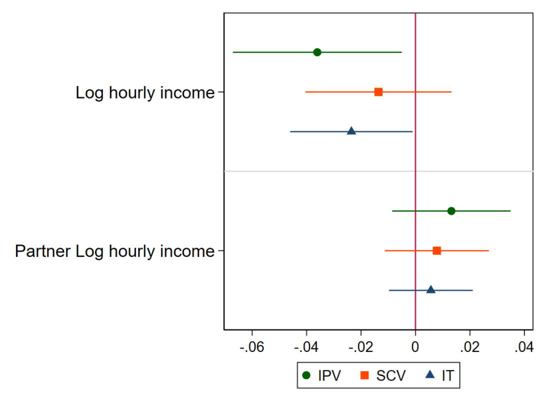
$$Violence_i = \gamma_1 logIncome_i + \gamma_2 logPartnerIncome_i + \beta_k X_{ki} + u_i \tag{1.5}$$

Figure 1.8 presents the results for both the log monthly income (Figure 1.8 (a)) and the log hourly income (Figure 1.8 (b)). When focusing on the reported monthly income (Figure 1.8 (a)), the marginal effects for both women and their partners are not statistically significant and exhibit a small magnitude, albeit they reflect a positive association. Turning to Figure 1.8 (b), which features the predicted log hourly income extracted from the Mexican Labour Force Survey (ENOE), the results unveil a distinct pattern. Specifically, while there is a positive yet statistically insignificant association in men's salary, a noteworthy (although relatively small) negative and significant correlation emerges between the overall IPV and the rise in women's predicted hourly wage.

Figure 1.8: Suffered IPV the previous 12 months



(a) Reported Monthly Income



(b) Predicted Hourly Income

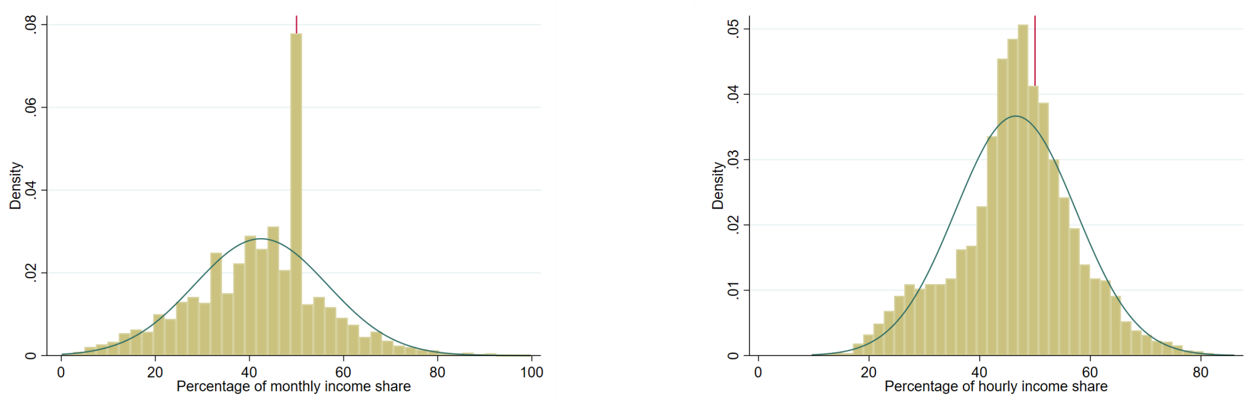
Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Point estimates (dots) along with their 95% confidence intervals (lines) are displayed. Number of Observations of Regression (a) Reported Monthly Income: 14746; Number of Observations of Regression (b) Predicted Hourly Income: 14320. Controlled by demographic characteristics, gender norm variables, and generational transmission of violence. Coefficients' numbers are represented in Table A5 for Regression (a) and Table A6 for Regression (b) in Appendix.

To examine the potential influence of income contribution proportions on intimate partner violence, I calculate the women's income share as the percentage of wage she earns relative to her spouse's wage. This is represented by Equation 1.6. I proceed to estimate this equation twice: firstly, using the reported monthly salaries from ENDIREH, and subsequently, employing the predicted hourly wages sourced from ENOE. Although this predicted hourly wage does not directly encompass the actual income contributed by each member of the household, it remains a pertinent measure as it determines the relative job positions of women compared to their spouses. For instance, consider the scenario where a woman earns a higher hourly wage than her husband, yet she is employed part-time while her husband works full-time. In this context, the husband's income most probably will surpass hers. However, given her superior hourly wage, if she were to transition to a full-time job, her income would exceed her husband's, thus, elevating the rank of her job compared to his. Furthermore, Wolf (2002) presents a model that justifies the basis for why individuals working fewer hours typically earn less per hour. Finally, the utilization of hourly wages is empirically grounded by ENOE due to its significant positive correlation with the number of working hours and income per hour, applicable to both married men and women.

$$Income\ Share = \frac{Women's\ Incmone}{Women's\ Incmone + Partner's\ Income} * 100 \tag{1.6}$$

Figure 1.9 displays the distribution of both income share percentages, featuring a vertical line of 50% mark. The first histogram represents the percentage of monthly income, while the second pertains to the predicted hourly wage. A noteworthy observation for both distributions is that there is a concentration at the 50% line. In fact, in Figure 1.9 (a), there is a noticeable accumulation of surveyed women who declare that they earn the same as their partners, making a high pick before the 50% mark. The rationale behind this phenomenon is unclear. It is plausible that some respondents may intentionally declare to earn the same or less than their husbands, while others may genuinely lack information on their spouse's salaries. As mentioned in the literature, there is a persistent taboo in Mexico regarding the couple's finances.

Figure 1.9: Women's Income Share



(a) Reported Monthly Income

(b) Predicted Hourly Income

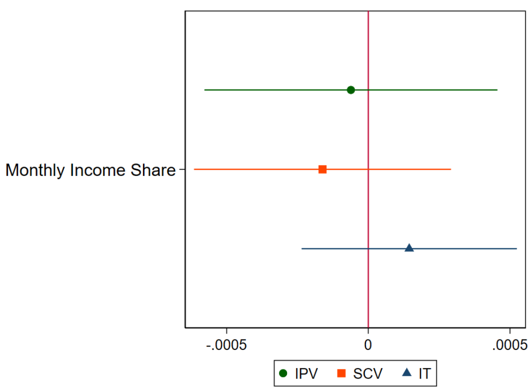
Note: Histogram of the percentage of income that women contribute in the household relative to their spouses. Figure (a) represents the reported monthly income (SOURCE: ENDIREH 2016), and Figure (b) represents the LASSO predicted hourly income (SOURCE: ENOE 2016).

The final regression in this analysis studies the association between the relative income with violence (Intimate Partner Violence, Situational Couple Violence and Intimate Terrorism) experienced in the household represented in Equation 1.7. Aligned with the resource theory and male backlash, this probit regression aims to ascertain whether the risk of violence experienced by women is influenced by their monetary contribution to the household. This phenomenon is measured by the variable $IncomeShare_i$. Moreover, the model controls for demographic characteristics, social norms and inter-generational transmission of violence indicators (X_{ki}).

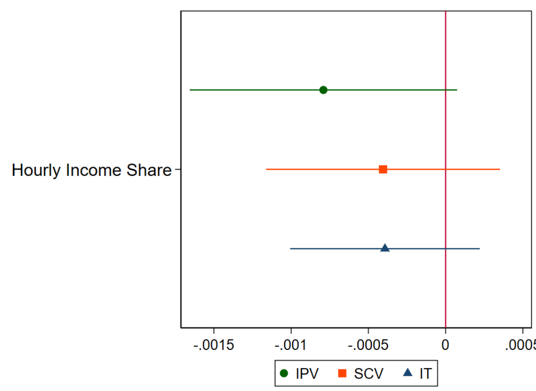
$$Violence_i = \delta IncomeShare_i + \beta_k X_{ki} + u_i \tag{1.7}$$

Figure 1.10 presents the principal findings stemming from Equation 1.7 for both the relative monthly income, recorded in the ENDIREH survey, and the relative hourly income, predicted from the ENOE (the Mexican Labour Force Survey). Notably, the marginal effect of each outcome is very close to the zero value in both instances. As a matter of fact, while the relative hourly income of women, predicted using LASSO, results in a negative association significant to the 10% level between the relative hours and the overall Intimate Partner Violence (IPV), the magnitude of this effect is almost negligible (see Table A8). Therefore, it cannot be deemed a substantial impact. In the robustness checks section of the Appendix (section 1.A.3.3), Figure A6 shows that the significant negative association between IPV and the hourly income share is solely sustained for those who live over the Poverty line.

Figure 1.10: Suffered IPV the previous 12 months



(a) Reported Monthly Income Share



(b) Predicted Hourly Income Share

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Point estimates (dots) along with their 95% confidence intervals (lines) are displayed. Number of Observations of Regression (a) Reported Monthly Income Share: 14746; Number of Observations of Regression (b) Predicted Hourly Income Share: 14320. Controlled by demographic characteristics, gender norm variables, and generational transmission of violence. Coefficients' numbers are represented in Table A7 for Regression (a) and Table A8 for Regression (b) in Appendix.

1.6 Conclusion and Discussion

Intimate Partner Violence is a global human rights issue that is still prevalent despite many efforts from governments and international organizations to eradicate it. This research explores in which form Mexican women's empowerment increases the risk of suffering IPV. It employs three main empowerment indicators: education level, labour force participation, and income contribution, all of them relative to their spouses. By controlling for individual and partner characteristics, the obtained results indicate that Mexican men see their wives' education and employment as a threat to their male identity rather than what they actually earn. In particular, men are more likely to exert violence with highly coercive controlling tactics when their spouse is employed. Moreover, this study shows that despite the more progressive views on gender roles and the place of women in modern Mexican society, men's preferences of being the sole provider are still a hindrance to achieving it without violence. Further research is needed to properly apply policy programs that support women's empowerment without the risk of increasing IPV.

This analysis contributes to the literature on IPV and its economic determinants in three different ways. Firstly, all five regressions not only control for women's demographic characteristics and employment status and childhood characteristics, but they also integrate their partner's counterpart, mitigating a possible bias. For instance, the first regression that solely incorporates control variables already exhibits significant results regarding gender differences in social norms. Strong beliefs on gender role differences are linked to a reduced likelihood of experiencing Intimate Partner Violence (IPV) and Situational Couple Violence (SCV). Moreover, the regression demonstrates that when women are required to seek permission from their husbands for certain activities, their likelihood of IPV increases primarily driven by Intimate Terrorism (IT). This examination of gendered social standards is looked into further in the second regression of the analysis which shows that the lower the education level of women compared to their spouses, the fewer intra-household conflicts. This may be interpreted as women with more years of schooling than their spouses may differ from behaving in a passive and compliant manner, while less educated women are less willing to enter into arguments that may lead to major conflicts. Furthermore, in Appendix Section 1.A.3.1, Figure A3 displays women's reactions to IPV depending on their level of education. On the whole, the more educated, the greater share of women who communicate their abuse situation, and know where to ask for help, but fewer actually take the necessary action of asking for institutional help or reporting the abuse.

Secondly, this research contributes to knowing which type of violence is affecting the most Mexican women. Johnson (1995) and Johnson and Ferraro (2000) sustain that there is a difference in the exerted violence depending on whether perpetrators get violent under a heated argument (SCV), or want to demonstrate power and control against their victims (IT). The results from the classification of couples depending on their employment status show that regardless of their partner's employment status, working women are at a higher risk of being victims of IT than their non-employed counterparts. These outcomes differ from Villarreal (2007) and Terrazas-Carrillo and McWhirter (2015) findings, however, although their use of ENDIREH previous surveys, they are not exactly comparable in measurement differences. Both studies agree that coercive controlling attitudes are strongly related to IPV and also may affect women's job performance. However, both use them as predictors, and not part of the outcome, as this research does. These findings are consistent with Canedo and Morse (2021) results in which the only type of Mexican women's employment that didn't increase the risk of IPV after receiving cash transfers, was unpaid household employment. These results infer that although working women have a more progressive vision of the role of women in society, they find themselves systematically rejected by their spouses, who would prefer that they stay at home.

Finally, this study delves into whether absolute income and women's income share could be potential explanatory variables. Aizer (2010) already found that with the tightening of the gender wage gap in the US, there were fewer physical assaults that required hospitalization. However, in the case of Mexico, this can only be confirmed with the predicted absolute hourly income of women, and the relative hourly income share with almost zero magnitude. Neither the reported absolute monthly income nor the relative monthly income shows any significant association with any form of IPV. These findings pose further more sophisticated methods to study the real effect of women's monetary contribution on IPV. Figure 1.7 provides an insight into how there was a greater share of women who faced their spouse's disapproval of earning more when they were employed. However, this was not reflected in the probit models. These findings may show that it is women's employment rather than how much she provides to the household, that has an increasing effect on violence intimate terrorism victimization. Although they are consistent with the resource theory and male backlash, they are more in line with Kessler and McRae Jr (1982) argument in which they proved that men's mental health was harmed when their wives were active in the labour force, and with Pyke (1996) qualitative analysis in which some men in her study sample affirmed using coercive controlling tactics and being violent against their wives when they ceased being the sole earner of the household. Therefore, as previously discussed, Mexican men link their male identity and worth to their ability to be the main provider of the family (Macmillan and Gartner, 1999; Jewkes, 2002; Pyke, 1996). Moreover, the income share coefficients are so low, that it cannot be analysed whether it follows a U-shape as Anderberg and Rainer (2013) modelled.

Future policies in Mexico regarding women's increase in labour force participation that take into consideration these results may focus their strategies in the workplace, and not consider this violence as only an intimate matter. Given those victims of IPV are more likely to be employed, the workplace is an important area for awareness, intervention and protection. In fact, targeting Mexican immigrants in the US, Galvez et al. (2011) identify common tactics used by IPV perpetrators to disrupt their partner's job: on-the-job surveillance tactics, such as constant messages and repeated calls; or on-the-job harassment tactics, like physically appearing in her workplace and causing problems. Furthermore, since almost 90% of the male population is employed, gender equality programs may contribute to changing *machismo* and *marianismo* views, allowing women to participate in the labour force without the burden of social norms.

1.A Appendix Chapter 1

1.A.1 Definition of variables

1.A.1.1 Suffered IPV the previous 12 months

In the last 12 months, has your partner...
Physical violence:
- Pushed you or pulled your hair?
- Slapped you?
- Tied you up?
- Kicked you?
- Thrown an object at you?
- Beaten you with his hands or an object?
- Tried to choke you or suffocate you?
- Hurt you with a knife?
- Fired a weapon on you?

Emotional violence:
- Shamed on you, or humiliated you?
- Ignored you, didn't take you into account, or didn't give affection?
- Told you that you are cheating on him?
- Made you feel fearful?
- Threatened to leave you, hurt you, take your children?
- Kept you at home and prohibited you to leave or receive visits?
- Stalked you, spies on you?
- Constant calls and messages to know where are you?
- Threatened you with a weapon?
- Threatened to kill you. hurt you or your children?
- Destroyed, thrown away, or hidden your things away?
- Stopped talking to you?
- Demands your passwords and checks your emails and phone?
- Turned your children and relatives against you?

Sexual violence:
- Threatened or blackmailed to have sexual intercourse?
- Made you do things you do not like when you have sexual intercourse?
- Used physical force to make you have sexual intercourse?
- Forced you to watch porn?
- Forced sexual intercourse without protection?

Economic violence:

- Have been mad because the house chores are not perfect?
- Forbidden you to work or study?
- Taken over your money or spent it without your consent?
- Taken over your goods or properties?
- Spent the money you need to pay bills?
- Not given you money to pay bills or has threatened not to give you money for bills?
- Not given you money for home expenses, even if he has money?
- Demanded about how you spend your money?

1.A.1.2 Coercive Controlling Attitudes

Does your husband get mad at you because...

- You work or study?
- You make more money than he does?
- You don't spend with him enough time and attention?
- You dedicate too much time to your job or because of your schedule?
- You are jealous?
- You are too possessive?
- You cry for everything?
- He thinks you cheat on him?
- He does not like the way you dress?
- You visit or are visited by friends and family?
- He doesn't like your friends?
- You go out without asking permission?
- Other men talk to you?
- You agree on something and you don't respect it?
- You are not obedient?
- You make your own decision in matters he believes are his responsibility?
- You have different or disagree in front of others?
- You consume alcohol or drugs?
- You don't want to have sexual intercourse?
- You use contraceptives?
- He doesn't like how you raise your kids?
- He considers you are a bad mother?
- A relative intervenes in the way you raise your kids?
- You tell him his responsibilities?
- You can not have kids?

- He wants more kids?
- You are sick and he needs to take care of you?
- Without any reason?
- Other?

1.A.1.3 Perseverance of Gender Roles

- Women should be responsible for caring for children, elders and ill people?
- Men should earn more than women?
- Men and women are equally responsible to earn money?
- Men should be equally responsible for house chores and caring?
- Women have the same right to night party alone?
- Men should have better occupations?
- Working women do not care properly for their children?
- Women shouldn't show their necklines to avoid molestation.
- It is a woman's duty to have sexual intercourse with her husband even if she does not want too

1.A.1.4 Ask for permission

Should you ask for permission...
- To work?
- To go shopping?
- To visit relatives and friends?
- To buy something personal?
- To participate in a community activity?
- To make new friends?
- To vote for a party?

1.A.1.5 Inter-generational Transmission of Violence

When you were a child...
- Do you remember if, between the cohabiting adults, there were beatings...?
- Do you remember if, between the cohabiting adults, there were curses...?
- Did those adults beat you...?
- Do you remember if those adults cursed you...?
- Were you touched or forced to touch intimate parts without your consent?
- Were you forced to show/watch your intimate body parts?

- Were you forced to watch porn?
- Did someone abuse you sexually?
- Were you sexually abused under threat?
- Were you sexually abused in exchange for money?

When your partner was a child (before his 15's)...

- Do you know if he was beaten or insulted?...
- Do you know if his mother suffered physical aggression from her husband?

1.A.2 Machine Learning Techniques

1.A.2.1 K-mean clustering

K-mean clustering is an unsupervised machine learning algorithm with a very simple objective: group similar data points together and discover underlying patterns (Dabbura, 2018). The algorithm partitions the data into k disjoint clusters that include all data points aggregated together with the nearest mean (a centroid or cluster centre). Therefore, the objective of the algorithm is to make the intra-cluster data points as similar (close) as possible while also keeping the clusters as different (far) as possible.

In other words, the k-means clustering solves a two-step minimization problem whose objective function J formulated as follows:

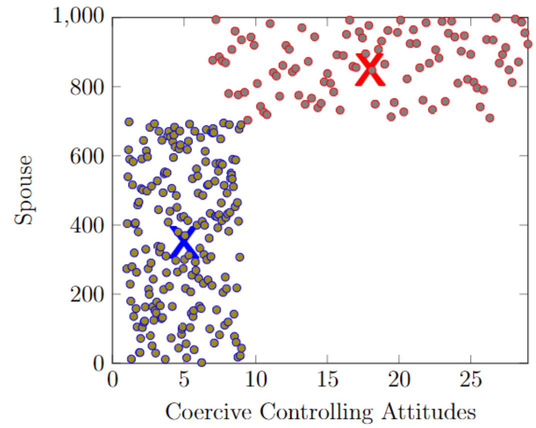
$$J = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} ||x^i - \mu_k||^2$$

where w_{ik} is an indicator of cluster inclusion (i.e., $w_{ik} = 1$ if x^i is in cluster k, and $w_{ik} = 0$ otherwise), and μ_k is the centroid of x^i's cluster. The differences are commonly measured with the Euclidean distance since it equates to the intra-cluster variance measurement. Therefore, the first step of the optimizing problem assigns the data points to the closest cluster by minimizing J with respect to w_{ik} having μ_k fixed. Then, the second step computes the optimal centroid of each cluster by minimizing J with respect to μ_k having wik fixed.

To achieve the creation of clusters in its practical application, the algorithm of the k-means clustering is the following:

1) Define a target number k, which is the total number of clusters that is willing to be formed (i.e. define the number of total centroids).
2) The k centroids are randomly selected across the data.
3) Every data point is allocated to each of the clusters by reducing the in-cluster distance:
    3.1) Compute the sum of the squared distance between a data point and a centroid.
    3.2) Assign each data point to the closest centroid (cluster).

Figure A1: K-means Clustering ($k = 2$) Representation



Fictional representation of the resulting k-mean clustering for the spouses' coercive controlling attitudes. The X-axis measures the sum of the spouses' coercive controlling attitudes while the Y-axis represents each spouse. The dots represent the data points in the sample and the cross represents the centroid.

4) Compute a new centroid of each cluster by calculating the average of all the clusters.
5) Repeat steps 3 and 4 until the position of the clusters is optimized.
6) The optimization of the problem is achieved when one of these two conditions is satisfied
   6.1) The centroids are stabilized (i.e., they don't change their position after step 4), and thus the clustering has been successful.
   6.2) The pre-determined number of iterations (repetitions) has been achieved.

Figure A1 is a fictional representation of the k-mean clustering applied to the analysis data. It illustrates how the k-mean clustering would look after the application of the algorithm. The X-axis measure the sum of the coercive controlling attitudes from the respondent's spouses, and the Y-axis represents each of them (here are represented only 1000 of them). The coloured dots represent the data points, i.e. each coercive controlling value for each spouse, and the crosses represent the centroids. It can be noticed that after the algorithm is applied, two disjoint clusters are created, one that will represent those individuals with low numbers of coercive controlling attitudes, and the other with a higher number of coercive controlling attitudes.

1.A.2.2 LASSO Algorithm

The Least Absolute Shrinkage and Selection operator (LASSO) is a supervised machine learning method that improves the prediction accuracy of linear regression models (Tibshirani, 1996). It solves the minimization problem of Equation (1.8):

$$\sum_{i=1}^{n}\left(Y_i - \left(\beta_0 + \sum_{j=1}^{p} X_{ij}\beta_j\right)\right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \tag{1.8}$$

The LASSO regression is composed of two main parts as shown in Equation (1.8). The first part

$\sum_{i=1}^{n}\left(Y_i - \left(\beta_0 + \sum_{j=1}^{p} X_{ij}\beta_j\right)\right)^2$ represents the standard linear regression model where Y is the dependent variable, X the regressors (or independent variables) and β the parameters to be estimated. Furthermore, $\lambda \sum_{j=1}^{p} |\beta_j|$, is the L1 regularization part in which $\sum_{j=1}^{p} |\beta_j|$ is the sum of the absolute value of the none intercept beta coefficients, and λ is a tuning parameter, also named penalty parameter, or regularization parameter.

λ controls the strength of the penalty in the regression equal to the absolute value of the magnitude of the coefficients. This means that some coefficients can become 0 and be eliminated from the model if they are not strongly associated with the outcome variable. λ denotes the shrinkage in the following way:

- When $\lambda = 0$, no parameters are eliminated. The estimate is equal to the linear regression model.
- As λ increases, more and more coefficients are set to 0. For instance, when $\lambda = \infty$ all coefficients are eliminated.
- As λ increases, the bias of the estimators $\hat{\beta}$ increases
- As λ decreases, the variance of the estimators $\hat{\beta}$ increases.

In order to find the optimal λ a range of λ's are tested using cross-validation. For instance, the data sample is separated into a training set and a test set. Then the LASSO model is built in the training set and estimated with the mean squared error calculated in the test set. This is also known as a supervised machine learning algorithm.

The LASSO regression minimization problem of Equation (1.8) is also equivalent to Equation (1.9) which is also seen in the literature (Tibshirani, 1996).

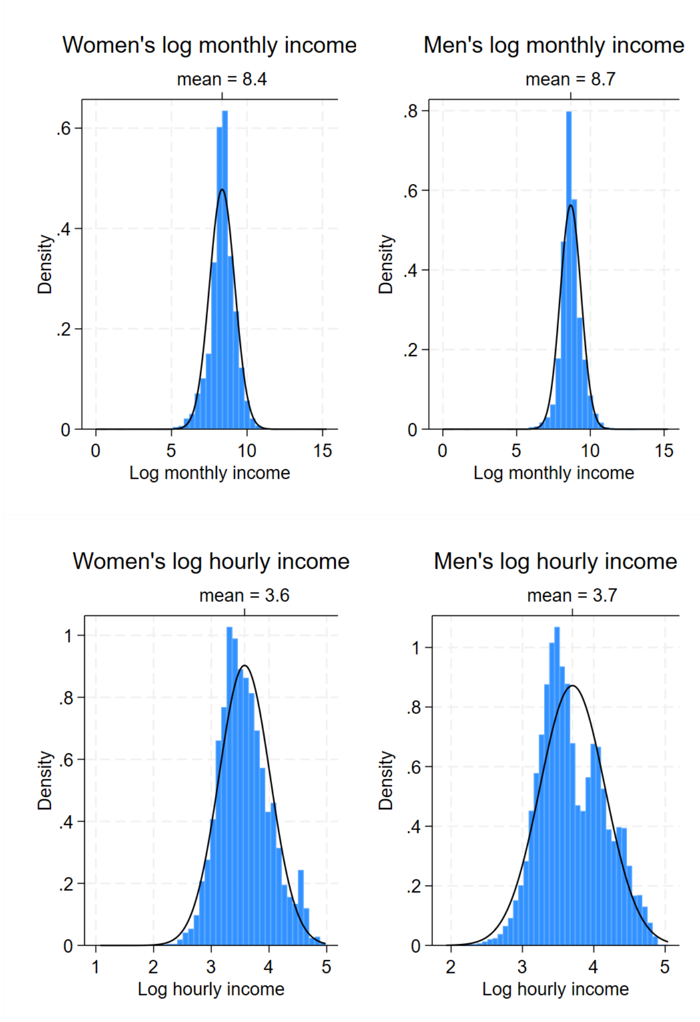
$$\min \sum_{i=1}^{n}\left(Y_i - \left(\beta_0 + \sum_{j=1}^{p} X_{ij}\beta_j\right)\right)^2 \quad subject \quad to \quad \sum_{j=1}^{p} |\beta_j| \leq t \tag{1.9}$$

in which t is now the tuning parameter. Similarly, as t decreases towards 0, the β coefficients shrink towards 0 forcing the least associated β coefficients to decrease all the way to 0. As a result, those β coefficients that are not strongly associated with the outcome variable are shrunk to 0, i.e., are removed from the regression.

Therefore, the LASSO modelling becomes a suitable method for prediction or variable selection. It performs well in high-dimensional datasets and allows for a great number of covariates, thus, it is also suited for models with high levels of multicollinearity. With its shrinkage feature, it balances simplicity and accuracy managing the risk of overfitting (Kumar, 2023).

The outcomes obtained through the application of the LASSO method to derive income values are visually depicted in Figure A2. The initial row of histograms portrays the logarithm of monthly income as reported by employed women who were interviewed, along with their estimations of their spouses' monthly income. The subsequent row of the figure displays the predicted hourly income of both women and men using the LASSO technique. The primary distinctions apparent within the histograms stem from the inherent dissimilarity in the income variable itself, where one is represented on a monthly basis (first row) and the other on an hourly basis(second row).

Figure A2: Logarithm of the income measurements

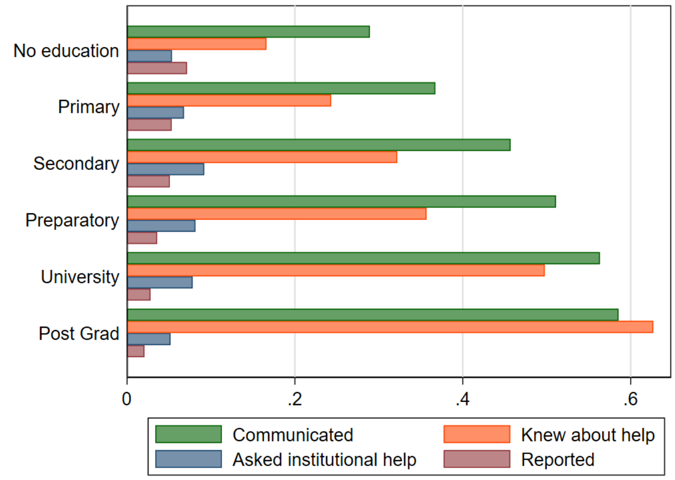


Histogram of the Logarithm of the income distribution for men and women. The first figure represents the reported monthly income (SOURCE: ENDIREH 2016), and the second figure represents the LASSO predicted hourly income (SOURCE: ENOE 2016).

1.A.3 Robustness Checks

1.A.3.1 Women's response against IPV over Education

Figure A3: Women's Response of IPV by Level of Education

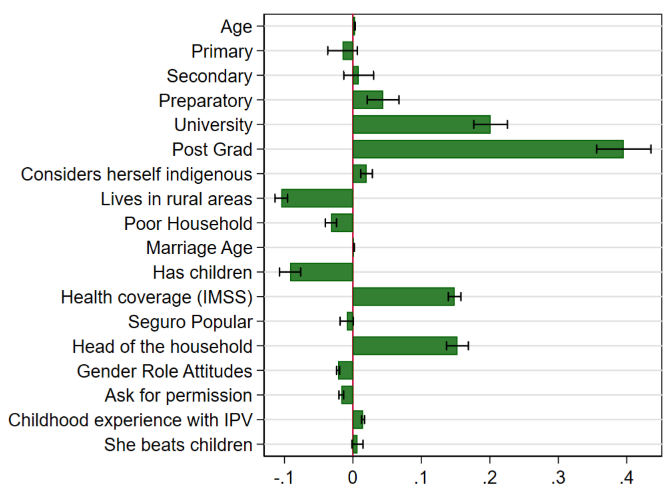


Percentage of response for communicating, searching for help and reporting by women's level of education.

1.A.3.2 Robustness Checks of Women's Employment Status

Determinants of being employed for women

Figure A4: Women's Employment Status

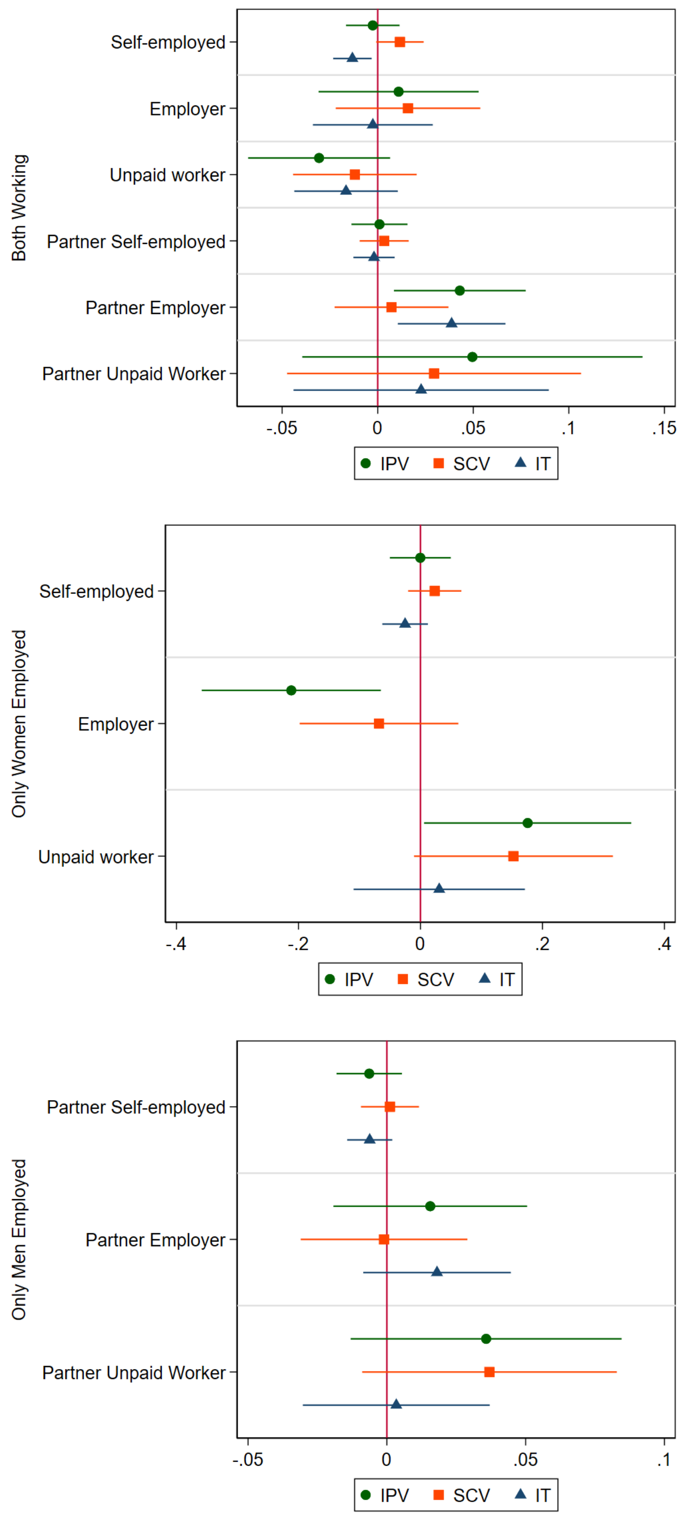


This graph Source: ENDIREH 2016

1. Intimate Partner Violence and Women's Empowerment in Mexico

Differences in Type of Employment for the different types of Couples

Figure A5: Women's Employment Status

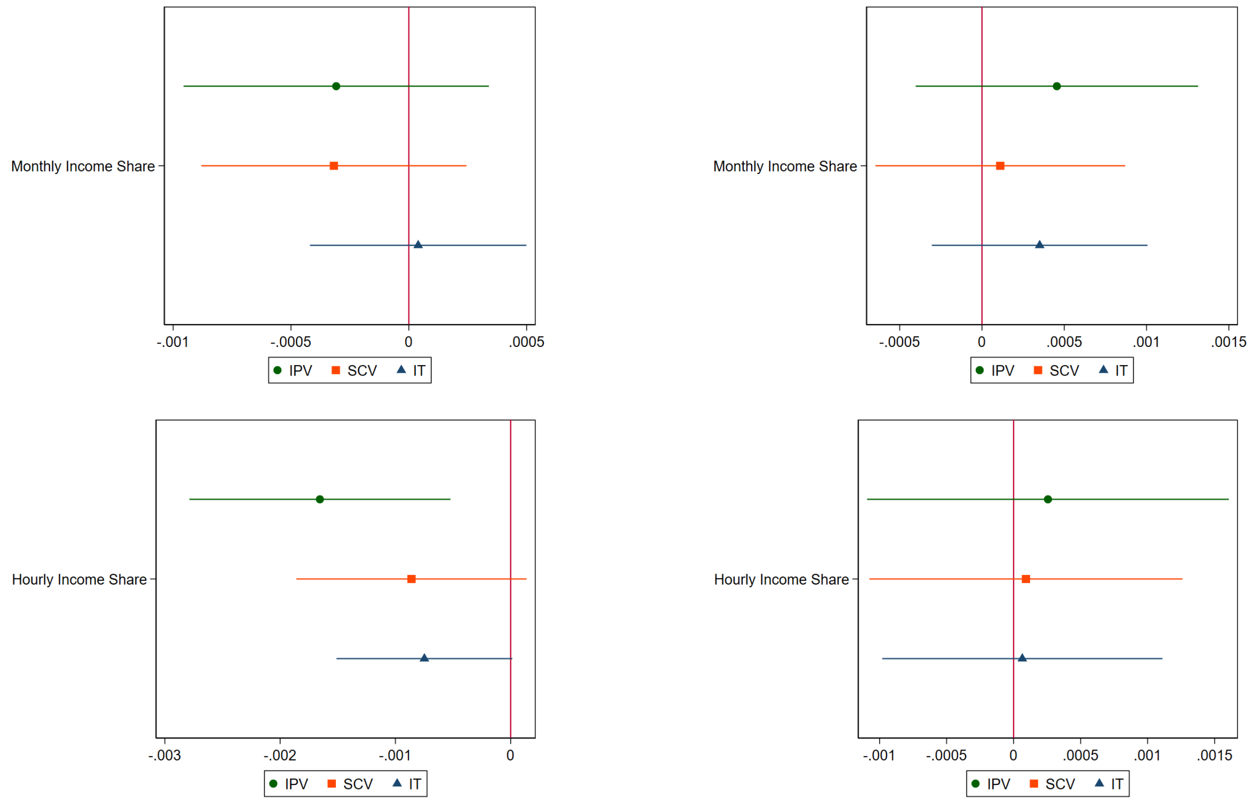


This graph Source: ENDIREH 2016

1.A.3.3 Robustness Checks on Women's Income Contribution

Relative Monetary Contribution Depending on the Poverty Conditions

Figure A6: Relative Monetary Contribution Depending on the Poverty Conditions



(a) Over the Poverty Line

(b) Under the Poverty Line

1.A.4 Tables of the main graphs

Table A1: Suffered Violence the previous 12 months: Control Variables

	IPV	SCV	IT
Age	-0.00144***	-0.00123***	-0.000214
	(0.000368)	(0.000322)	(0.000267)
Partner age	-0.00180***	-0.00111***	-0.000685***
	(0.000346)	(0.000302)	(0.000252)
Considers herself indigenous	-0.00249	-0.00337	0.00162
	(0.00717)	(0.00628)	(0.00522)
Considers himself indigeneous	-0.00531	0.000222	-0.00437
	(0.00712)	(0.00624)	(0.00523)
Lives in rural areas	-0.0622***	-0.0254***	-0.0357***
	(0.00433)	(0.00378)	(0.00318)
Poor Household	-0.000983	-0.00155	0.000557
	(0.00389)	(0.00340)	(0.00280)
Marriage Age	-0.000946***	-0.000336**	-0.000705***
	(0.000188)	(0.000160)	(0.000145)
Has children	0.00780	-0.00360	0.0132**
	(0.00741)	(0.00639)	(0.00576)
Health coverage (IMSS)	0.00565	0.00112	0.00406
	(0.00450)	(0.00390)	(0.00330)
Seguro Popular	0.000727	-0.00544	0.00631*
	(0.00458)	(0.00400)	(0.00332)
Observations	60028	60028	60028

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Base Line Categories: Does not consider herself indigenous; Does not consider himself indigenous; Lives in urban areas; Does not live in a poor household; Does not have children; Does not have IMSS health coverage; Does not have Seguro Popular health coverage.

Table A2: Suffered Violence the previous 12 months: Control variables

	IPV	SCV	IT
Head of the household	-0.00877	-0.00943	0.000216
	(0.0106)	(0.00923)	(0.00753)
Partner head of the household	-0.0383***	-0.0173**	-0.0216***
	(0.00785)	(0.00680)	(0.00570)
Gender Role Attitudes	-0.00371***	-0.00298***	-0.000219
	(0.000994)	(0.000869)	(0.000717)
Ask for permission	0.0320***	-0.00168	0.0264***
	(0.00147)	(0.00136)	(0.000922)
Childhood experience with IPV	0.0458***	0.0142***	0.0279***
	(0.00105)	(0.000933)	(0.000706)
Partner exp or witnessed IPV	0.107***	0.0512***	0.0607***
	(0.00456)	(0.00396)	(0.00325)
Partner exp and witnessed IPV	0.158***	0.0413***	0.111***
	(0.00602)	(0.00500)	(0.00463)
She beats children	0.0667***	0.0427***	0.0281***
	(0.00398)	(0.00353)	(0.00287)
He beats children	0.111***	0.0240***	0.0704***
	(0.00498)	(0.00447)	(0.00330)
Observations	60028	60028	60028

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Base Line Categories: She is not the head of the household; Her husband is not the head of the household; The husband never experienced or witnessed IPV during infancy; She does not exert violence against her children; He does not exert violence against his children.

Table A3: Suffered Violence the previous 12 months: Education Gap

	IPV	SCV	IT
-2 Education Gap	0.0533*	0.0534**	-0.000452
	(0.0273)	(0.0252)	(0.0183)
-1 Education Gap	0.00564	-0.00113	0.00737**
	(0.00482)	(0.00416)	(0.00352)
1 Education Gap	-0.00700	-0.00163	-0.00491
	(0.00481)	(0.00422)	(0.00345)
2 Education Gap	-0.0232	0.00239	-0.0264*
	(0.0210)	(0.0192)	(0.0141)
Demographic controls	YES	YES	YES
Gender Role Characteristics	YES	YES	YES
Transmission of violence	YES	YES	YES
Observations	60028	60028	60028

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Base Line Categories: Education Gap = 0 (Same Level of Education). Education levels are classified into three categories: primary or less education; secondary education and tertiary education (graduate and post-graduate studies). The negative Education Gap relates to those couples where women have a higher level of education than their husbands.

Table A4: Suffered Violence the previous 12 months: Employment Status

	IPV	SCV	IT
Both Employed	0.0352***	0.00571*	0.0295***
	(0.00394)	(0.00344)	(0.00286)
Only Women Employed	0.0559***	0.0154	0.0410***
	(0.0119)	(0.0104)	(0.00907)
Both Not Employed	0.00100	-0.00799	0.00773
	(0.00845)	(0.00733)	(0.00611)
Demographic controls	YES	YES	YES
Gender Role Characteristics	YES	YES	YES
Transmission of violence	YES	YES	YES
Observations	60028	60028	60028

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). Baseline comparison: Couple Type 3: Only Men Employed (Traditional employment standards).

Table A5: Suffered Violence the previous 12 months: Absolute Monetary Contribution

	IPV	SCV	IT
Log Monthly Income	0.00444	-0.00189	0.00674
	(0.00602)	(0.00530)	(0.00436)
Log Partner Monthly Income	0.0108	0.00784	0.00139
	(0.00684)	(0.00599)	(0.00508)
Demographic controls	YES	YES	YES
Gender Role Characteristics	YES	YES	YES
Transmission of violence	YES	YES	YES
Observations	14746	14746	14746

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). The logarithm of the Reported Monthly Income obtained directly from ENDIREH 2016.

Table A6: Suffered Violence the previous 12 months: Absolute Monetary Contribution

	IPV	SCV	IT
Log Hourly Income	-0.0361**	-0.0136	-0.0235**
	(0.0158)	(0.0137)	(0.0115)
Log Partner Hourly Income	0.0132	0.00787	0.00567
	(0.0111)	(0.00975)	(0.00784)
Demographic controls	YES	YES	YES
Gender Role Characteristics	YES	YES	YES
Transmission of violence	YES	YES	YES
Observations	14320	14320	14320

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). The logarithm of the Predicted Hourly Income was calculated from ENOE 2016.

Table A7: Suffered IPV last year

	IPV	SCV	IT
Monthly Income Share	-0.0000424	-0.000145	0.000147
	(0.000264)	(0.000231)	(0.000194)
Demographic controls	YES	YES	YES
Gender Role Characteristics	YES	YES	YES
Transmission of violence	YES	YES	YES
Observations	14746	14746	14746

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). The logarithm of the Reported Monthly Income obtained directly from ENDIREH 2016.

Table A8: Suffered IPV last year

	IPV	SCV	IT
Hourly Income Share	-0.000791*	-0.000405	-0.000393
	(0.000441)	(0.000387)	(0.000313)
Demographic controls	YES	YES	YES
Gender Role Characteristics	YES	YES	YES
Transmission of violence	YES	YES	YES
Observations	14320	14320	14320

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Probit regressions with dependent variables Intimate Partner Violence (IPV), Situational Couple Violence (SVC), and Intimate Terrorism (IT). The logarithm of the Predicted Hourly Income was calculated from ENOE 2016.

This page is intentionally left blank.

Chapter 2

Statistical Matching with Propensity Score and Machine Learning

Abstract

Statistical matching, also known as data merging, data fusion, or synthetic matching, constitutes a technique employed to integrate two or more datasets originating from the same population without a common identifier. This approach offers a noteworthy advantage by amalgamating complementary information from existing data sources that do not overlap. In the present paper, I present a comprehensive overview of the prevalent statistical matching methods while delving into the introduction of emerging machine learning tools. The primary focus lies in optimizing the selection of matching variables, wherein I introduce the utilization of propensity scores and the genetic matching algorithm. Subsequent to the application of these distinct matching methodologies, the results reveal that, on the whole, the propensity score does not yield a discernible enhancement in matching quality. Conversely, the genetic matching algorithm exhibits significant improvements. Thus, the integration of advanced technology into the original statistical matching framework results in a heightened quality of synthetic datasets, facilitating more robust inferential analyses.

2.1 Introduction

Oftentimes, economic analysts encounter limitations due to the availability of information from a single data source, restricting their ability to gain deeper insights into the research problem at hand. Consider, for instance, the investigation of Intimate Partner Violence. While surveys may provide data on victim (perpetrator) demographics and crime characteristics, they might lack information on specific employment attributes relevant to understanding the commission of such crimes. This scenario exemplifies a broader issue that spans various economic branches, necessitating the integration of data from diverse resources.

One solution to this problem is for researchers to collect new data through a bespoke survey or through randomized control trials, the advantage of the latter being they give unbiased estimates of the causal effect of a certain treatment. However, practical, financial, or ethical constraints might render this approach unfeasible. An alternative solution is to link existing observational data through various statistical techniques. Commonly employed methods fall under two categories: exact matching, which relies on a unique identifier to link individual respondents across distinct datasets[1], and statistical matching, which leverages a set of common variables among different data sources to link respondents with the same characteristics.

This work presents a thorough review of the statistical matching framework and its current practical implementations. Additionally, I enhance the available statistical matching tools by integrating propensity score estimation and machine learning techniques, comparing the resulting synthetic datasets. To achieve this, I utilize the prepared synthetic data samples from D'Orazio (2017), derived from the European Union Statistics on Income and Living Conditions Survey (EU-SILC). The synthetic samples were originally created to evaluate hot deck matching techniques described in D'Orazio et al. (2006b), providing a baseline for my analysis.

The primary objective of this research is to assess whether the use of richer matching variables enhances the overall matching accuracy. Therefore, to achieve a comprehensive analysis, I introduce different matching variables. Initially, I combine propensity score matching (Rosenbaum and Rubin, 1983) with random forests and boosted models of classification and regression trees (Lee et al., 2010) within the non-parametric hot deck methodologies proposed by D'Orazio et al. (2006b). Next, I replicate the matching process by incorporating propensity scores that include relative weights obtained from the genetic matching algorithm (Diamond and Sekhon, 2013), as well as a new matching variable constructed from the sum of these resulting weights. Throughout the replication of D'Orazio (2017) paper, I ensure that all matching assumptions are satisfied, focusing solely on the performance of the matching variables.

The results reveal that the use of propensity scores as matching variables does not differ significantly from D'Orazio (2017)'s approach of using common variables directly. Either they perform worse when used directly, or when incorporating relative weights from the genetic matching algorithm. However, utilizing the direct sum of weights from the genetic matching algorithm emerges as the superior matching variable, effectively preserving the marginal and joint distributions of the imputed variables and, in some instances, outperforming the baseline from D'Orazio (2017).

This study contributes to the statistical matching literature by introducing novel techniques for creating matching variables with greater information content while controlling for relevance. The issue of dimensionality is a recurring challenge in the literature, often limiting authors to using a maximum of two matching

[1] Frequently unfeasible.

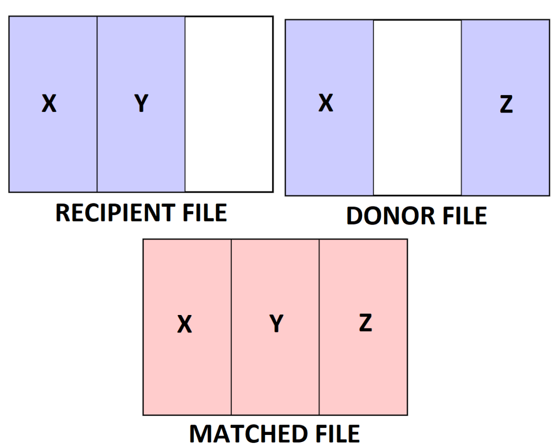
variables in their algorithms (D'Orazio, 2017; Donatiello et al., 2014; Leulescu and Agafitei, 2013). Thus, this paper makes a significant contribution by introducing a novel matching variable that incorporates more information without exacerbating the dimensionality problem in the matching algorithm. This new technique enhances the quality of synthetic datasets, thus facilitating more robust inference in subsequent analyses. Such improvements in the statistical matching domain hold considerable relevance for policymakers, especially in situations where budget constraints may hinder the collection of new surveys or the implementation of experiments.

The paper is structured as follows: Section 2.2 introduces the statistical matching procedure. Section 2.3 discusses the historical evolution of statistical matching and prior research findings. Section 2.4 provides a detailed review of all steps necessary for proper statistical matching. Section 2.5 introduces the new types of matching variables, outlines the matching framework, and presents the minimization problem. Section 2.6 describes the dataset and presents the results. Section 2.7 concludes.

2.2 Statistical Matching Introduction

Figure 2.1 illustrates the intuition of a simple matching in which the two linkage sources have incomplete information. Consider two representative datasets containing a set of variables, X, common to both data sources, but each dataset also contains other variables unique to that dataset (Y in one set, and Z in the other set). Using matching techniques on these common variables it is able to link records from one dataset (donor file) with another (recipient file) to form a single new dataset containing all variables X, Y and Z (matched file).

Figure 2.1: Scheme example of Survey Matching



On some fortunate occasions, surveys contain a code or a set of variables that uniquely identify the same individuals from one survey to another. Therefore, common variables X are those identification variables and the matching is named *exact matching* or record linkage. An example is the Mexican Labour Force Survey for Child Labour, which is a supplement to the Mexican Labour Force Survey. Within both surveys, there are specific code variables enabling the same individual in both surveys to be tracked. Therefore, using a unique identifier, it is possible to combine the information of both surveys for those individuals that are held in both datasets. Although this technique would be ideal, it is rare in National Surveys to avoid individual response fatigue emanating from length survey questionnaires. Moreover, data protection measures to protect individual confidentiality within survey data can lead to a less feasible exact match in other areas, making statistical matching more necessary.

Statistical matching (also known as data fusion, or synthetic matching) is the technique used in the literature that links records from different sources of information to create a database that allows for a complete analysis when exact matching is not feasible. Its main characteristic is that the population of the two samples to match does not overlap, and therefore, individuals are never jointly observed. Statistical matching uses the common variables between both sets to create a new synthetic dataset that enlarges the information scope from the existing data sources. For instance, if a research aims to obtain information from two different surveys such as the Mexican Time Use Survey and the Mexican National Survey on the Dynamics of Households Relationships, since both surveys sampled completely different individuals, by using the common variables it is possible to statistically match units with similar characteristics.[2]

2.3 Evolution of Statistical Matching

Statistical Matching arose from the need to analyze information that does not come from the same data source. Its origins in economics date around the mid-'60s in the US. In particular, Budd and Radner (1969) report the first documented case where this method is described as a plausible solution to a specific tax problem in the Bureau of Economic Analysis (BEA) of the US Department of Commerce. To improve the precision between the income amount and the addition of return tax income, the original base of the 1965 Income Supplement of the Current Population Survey (CPS) and the 1964 Tax Model were used, whose samples contained different individuals from the same population. By creating a microdata file, the tax return was estimated for those individuals who did not respond in the CPS. Simultaneously, Okner (1972) published another example of this methodology. Using the 1966 Tax File and the 1967 Survey of Economic Opportunity the objective of the matching was to create a comprehensive dataset that combined information on tax return and socio-demographic characteristics. Okner (1972) is considered the main origin of statistical matching since it segmented the two datasets using the different classifications of the common variables and selected randomly the missing record for each equivalent cluster.

[2] In the statistical matching framework, two observation units are considered similar if they respond the same answers for the same questions.

From there, other authors began to use matching methodologies, however, with a lack of consensus on their terminology and their respective definitions. It was not until Kadane (1978) that the matching problem was formalized and defined the concepts of Exact Matching (or record linkage) and Statistical Matching (or synthetic matching). Radner (1980) published an article from the US Federal Committee on Statistical Methodology in which they reported the basics of matching: definitions for both, exact matching and statistical matching; its general procedures, a summary of its applications at the moment and the type of errors that can be found once the matching has been carried out. This paper opened the doors to a new research path as Radner (1980) posed theoretical and practical questions whose answers were awaiting more advanced technologies.[3] These questions essentially covered the accuracy of statistical matching: 1) what are the important factors and conditions that result in a sufficiently accurate statistical match; 2) which are the most optimal methodologies and the relative precision between them; 3) how sensitive are the results obtained and 4) what is the best way to deal with problems that arise in practice. Today, there is still no clear consensus on many of the answers to these questions.

The theoretical answer to what conditions are necessary to proceed with the match was already discussed in the literature. Within the statistical theory, Sims (1972) and Kadane (1978) analyzed the need for independence between the variables aimed to be analyzed conditioned by the common variables used to make the match, the well-known: Conditional Independence Assumption (CIA). In fact, Sims (1972) was very critical of Okner (1972)'s methodology since the matching was proceeded assuming that the CIA was satisfied. Sims (1972) argued that when this condition does not hold, it is not possible to proceed with further inference of the synthetic database.

With new techniques developed during the subsequent years, Rodgers (1984) expanded the definitions proposed in Kadane (1978) and Radner (1980). In particular, Rodgers (1984) exposed the constrained and unconstrained matching methods, and the key differences between them. Moreover, Rodgers (1984) remarked on the usage of a proper distance function which determines the similarity between records, and pointed out that not all common variables need to be used. Key variables will require an exact matching (such as gender, geographical area, etc.) and other common variables may lack relevancy in the matching procedure. Rodgers (1984) evaluated previous empirical tests in the literature since they fell short of theoretical and empirical justification and defended the importance of complying with the CIA so that the results obtained have sufficient veracity. An important insight from Rodgers (1984) is that once the matching is done, it cannot be verified whether the CIA holds and, an *a priori* analysis is necessary: either empirically with a base that contains all the information, or through theoretical economic assumptions.

Following this reasoning, D'Orazio et al. (2006b) theoretically evaluates the results of statistical matching under the CIA assumption. D'Orazio et al. (2006b) expands the analysis elaborating on the correct use

[3] At that time, exact and statistical matching was carried out manually or with very primitive computers.

of auxiliary information (i.e. a dataset that contains all the variables X,Y,Z) in case CIA cannot be justified. Furthermore, D'Orazio et al. (2006b) compiles and amplifies the theoretical framework on statistical matching through the use of statistical tools. In line with the computational development of the time, it captures all the possible parametric and non-parametric methodologies to achieve a good matching, either to create a new synthetic data set (micro approach) or a contingency table (macro approach). Rodgers (1984)'s examples of unconstrained and constrained matching is nested within the non-parametric micro approach, named "Nearest Neighbour Distance Hot Deck", which is presented along two other non-parametric matching methods: "Random Hot Deck" and "Rank Hot Deck".

Rässler (2012)[4] complements D'Orazio et al. (2006b) theoretical framework by addressing the area of the quality of the matching results. Aware of the limits of matching since all the variables are not jointly observed, depending on the explanatory power of the common variables Rässler (2012) distinguishes 4 levels of validity for the matching result. The first requires that the imputed target variable maintains its marginal distribution and its joint distribution with the matching variables. The second level states that the correlation between the target variables from both datasets is maintained. The third guarantees that the true joint distribution (X, Y, Z) is represented in the synthetic dataset. And, finally, the fourth level claims that the true values of the imputed target variable are reproduced in the synthetic file.[5]

Although the early literature had considered the importance of including a weighted variable to set the matching framework (Rodgers, 1984; Rubin, 1986), Andridge and Little (2009) first introduced the use of sample weights to apply matching with actual survey data. This innovation led D'Orazio et al. (2010, 2015) to evaluate which approaches may be followed when the data used comes from complex survey designs and was applied in D'Orazio (2017) through the StatMatch package included in the R interface[6]. Among others, the StatMatch functions have been used for more comprehensive analysis of household living conditions (Donatiello et al., 2014)[7], quality of life (Leulescu and Agafitei, 2013) [8], informal economy (Fernández and Villar, 2017b)[9], education (Iztueta et al., 2017)[10], and energy policies (Douenne et al., 2020)[11], Zhang et al. (2018)[12].

However, the matching approach has been utilized for diverse data types beyond survey data. For instance, in the experimental literature, matching has been employed as the main method for causal inference. On the whole, if there are two groups in an experiment, treatment and control, matching consists of finding those in-

[4]First edition of the book was Rässler (2002)

[5]The fourth level is empirically impossible to know, and the second and third levels require previous assumptions to be justified, the first level can be estimated empirically.

[6]The StatMatch package was built in 2014, however, it was released in January 2017 with the inclusion of the uncertainty analysis with categorical variables

[7]Combines information of household income and consumption expenditure.

[8]Matches information objective and subjective well-being

[9]Uses two different waves to obtain employment characteristics of informal workers

[10]Combines teachers and student-level data

[11]Combines energy taxes and household income characteristics in France

[12]Studies household energy distribution at neighbourhood scale.

dividuals in the control group that are as similar as possible to those in the treatment group. Their objective is to calculate the average treatment effect, or, the average treatment effect on the treated otherwise. Despite this, there is still no consensus on how to classify this type of matching in the previous Statistical Matching literature, due to the nature of its methodology, it could be catalogued as unconstrained matching, according to the definitions of Radner (1980). This results from the fact that individuals in the control group (donor file) are selected with replacement and only require a smaller size of the treatment group (recipient file). To this end, propensity score matching has been widely used in the experimental literature.[13]

Propensity score matching is a matching method that leverages the probability of treatment assignment based on observed baseline characteristics. Rosenbaum and Rubin (1983) introduced propensity score matching as a way to identify individuals in a control group with similar characteristics to those in the treatment group. Following this line of study, propensity score matching has evolved to find optimal matches both in the quality of the result and in its computational cost. That is why the calculation of the propensity score using a logit regression has incorporated the use of machine learning techniques that omit parametric restrictions (Abadie and Imbens, 2011). Specifically, Lee et al. (2010) compare the quality of matches according to how the propensity score has been calculated. Lee et al. (2010) use artificial data and compares seven different scenarios to define the propensity score and considers a linear model, with quadratic terms and with interactions. Furthermore, the analysis estimates the propensity score in six different methods: logistic regression, classification and regression trees (CART), pruned CART, bagged CART, random forest, and boosted CART; in three different cohort sizes, for a sample of 500 observations, 1000 and 2000. Lee et al. (2010) conclude that the methods that perform better in most of the different scenarios are random forests and boosted CART.

On the statistical matching framework, Kum and Masterson (2010) introduced the propensity score estimation to proceed with a constrained nearest neighbour distance matching. Kum and Masterson (2010) classified some common variables as donation classes (or strata variables), that segment the data in different clusters. The approach estimates three different propensity scores with a logistic model. The first uses the overall sample, the second divides the sample into cohorts using strata variables and calculates the propensity score for each of the divided cohorts, and the third propensity score incorporates even more segmentation in the sample. The matching proceeds by first linking the third propensity score followed by the second and first one in case of unpaired matches. As a result, the use of propensity score reduces the dimensionality of the distance minimization problem and its computational cost. Kum and Masterson (2010) demonstrate their algorithm using national statistical surveys from the US: the 2001 Survey of Consumer Finances and the 2002 Annual Demographic Survey. Rios-Avila (2014, 2018) replicated the same algorithm for two different studies in which both of them used the American Time Use Survey. Moreover, using the Dutch Population Census of (2001), Waal (2015) compares among propensity score estimation, unconstrained nearest neighbour distance and random hot deck methods. It concludes that matching using propensity score estimation worsens the results in comparison to using distance hot deck.

[13] King (2018) reports that more than 93000 observational studies have used propensity score matching.

Another key dimension for matching is the distance function required to select the matches Rodgers (1984). For propensity score matching, the distance function is usually the absolute value between the two propensity score values as it is a single variable. But when matching is done considering multiple common covariates, the most followed distance function in the literature is the Mahalanobis distance. However, the Mahalanobis distance does not work optimally when all common covariates are categorical and not continuous. In that case, it is recommended to use Manhattan distance (D'Orazio, 2017)[14]. Regarding this, Diamond and Sekhon (2013) and Sekhon and Grieve (2012) introduce an improvement of the Mahalanobis distance through a genetic matching algorithm named GenMatch. This algorithm minimizes the Mahalanobis distance by using a weight matrix. Sekhon (2011) compares GenMatch results with propensity score matching results calculated with logit regression using LaLonde (1986)[15] data.[16] Sizemore and Alkurdi (2019) compares Lee et al. (2010) and Diamond and Sekhon (2013) machine learning techniques incorporated into the matching methods. Sizemore and Alkurdi (2019) reviews coarsened exact matching, propensity score matching estimated with logistic regression, random forest and boosted CART, and genetic matching algorithm to obtain the most accurate average treatment effect on the treated across six different datasets. (Sizemore and Alkurdi, 2019)'s results show that Genetic matching doesn't perform optimally in high dimensional data, despite its computational power. Besides, propensity score matching using random forest and boosted CART do not improve the results obtained from propensity score matching estimated with logistic regression.

This paper aims to bring together all the advances in the different branches of literature, including the different approaches for matching. It first compares the conventional non-parametric hot deck methods of D'Orazio et al. (2006b) and D'Orazio (2017) with the logit propensity score matching approach inspired by Kum and Masterson (2010), adding the estimation techniques of random forests and boosted CART (Lee et al., 2010; Sizemore and Alkurdi, 2019). Subsequently, it compares the hot deck methods in D'Orazio (2017) with the same propensity scores using the relative weights obtained from the genetic matching algorithm of Diamond and Sekhon (2013) and incorporates a novel matching variable with an alternative use of the GenMatch relative weights. By implementing the last developments from the experimental literature, more information could be included in the matching. Therefore, the matches between the two different surveys are expected to be more precise. As Sizemore and Alkurdi (2019) bring to light, results may change depending on the characteristics of the datasets, and thus, this work contributes to the literature by analyzing the novel techniques in a statistical matching framework.

[14]This study applies the Manhattan distance, as utilizing a single variable for matching is equivalent to employing the absolute distance. See Appendix 2.B.1 for detailed definitions of Mahalanobis and Manhattan distance metrics

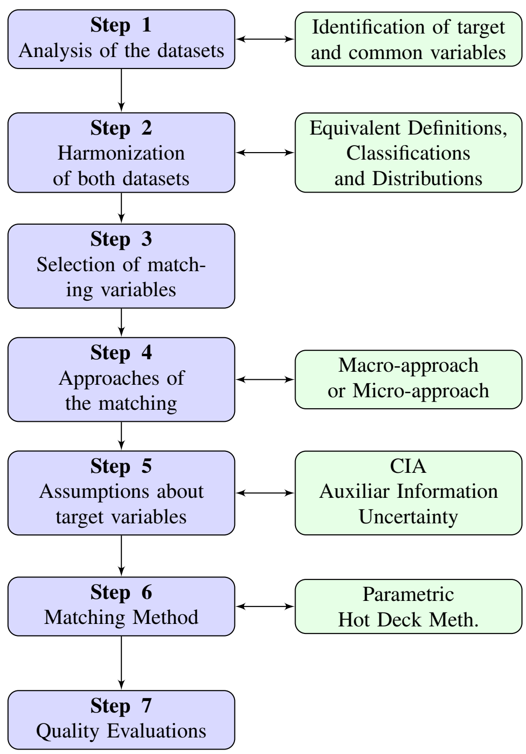
[15]LaLonde (1986) data set consists of a compiled dataset from the randomized job training experiment of the National Supported Work Demonstration (NSWD) and the Current Population Survey (CPS). Dehejia and Wahba (1999, 2002) created an experimental sample from the original LaLonde (1986) data, that has been widely employed in the literature for propensity score matching comparisons (Diamond and Sekhon, 2013; King et al., 2017; Ho et al., 2011; Schmidt and Martinello, 2019).

[16]Sekhon (2011) shows all the functions used that are included in the `Matching` package of the R interface

2.4 Practical Application of Statistical Matching

This section reviews all the required steps to correctly perform a statistical matching following the approaches of D'Orazio et al. (2006b) and Rässler (2012). As outlined in Figure 2.2 it includes, a preliminary analysis of both datasets, which implies the identification of target and common variables. Secondly, the harmonization of both datasets through definitions, classifications and distributions. Thirdly, from the set of common variables, selection of the matching variables. Therefore, depending on the explanatory power of the matching variables, the selection of the matching approach requires assumptions to consider the relationship of the target variables. Once the matching framework is determined, the most suitable matching method is applied, either using parametric methods or non-parametric methods (Hot Deck techniques). Finally, it includes an evaluation of the quality of the results. To obtain accurate results, every step of the matching procedure needs to be addressed carefully to avoid meaningless results.

Figure 2.2: Scheme of Statistical Matching application



2.4.1 Analysis of the datasets

Consider two different datasets A and B. Among all the variables covered by A and B, interest lies in the relationship between variables which are not jointly observed, the target variables. In other words, the main goal of the matching is to obtain more information about the relationship between Y and Z, being Y included exclusively in A and Z included exclusively in B. Therefore, Y is defined as the target variable in A and Z is the target variable in B. Common variables X, in contrast, are variables that appear in both sources A and B. These variables should ensure coherence between both datasets A and B, since, one of the key assumptions of statistical matching is that both datasets are independent samples of the same population. Therefore, apart from the study of the metadata (survey design, stratification, sample units), a harmonization process of the common variables should be followed in both surveys.

2.4.2 Harmonization and reconciliation

Common variables in both surveys should be akin in definitions, classifications and distributions so they are comparable (D'Orazio et al., 2006b). Harmonization and reconciliation actions are essential to correct any discrepancy between both datasets A and B. Firstly, common variables are re-coded to have homogeneous definitions and classifications in both datasets. Thus, some variables are aggregated from the individual level to the household level, or some continuous variables are categorized. Moreover, it requires an adjustment for missing data. For instance, if file A reports individual values from the age of 15 to 90 and file B reports the age from 12 to 85, the synthetic variable *age* will be defined for those individuals who are between 15 and 85 years old.

In the statistical content, each common variable should show similar marginal distributions across both datasets. This is a key condition to validate whether both samples are independently and randomly selected to represent the same population. The differences in the weighted frequency distributions of each common variable are computed to assess the degree of similarity or dissimilarity.[17] Those variables that cannot be harmonized are discarded from the matching.

When dealing with survey data, the randomness of both target (Y and Z) and common variables X is introduced by the sampling design. Thus, sample weights play a key role in the harmonization and matching process since they compensate for missing values and measurement errors. Consider W_{Ai} the sample weights for observation i in A, and W_{Bj} the sample weights of observation j in B. Two populations are considered equivalent if $\sum_{i=1}^{n} W_{Ai}$ is similar to $\sum_{j=1}^{m} W_{Bj}$, acknowledging that they are rarely equal for different national statistical surveys. In case there are substantial differences, it is necessary to take a calibration step of one of the sample weights. However, re-weighting techniques may be hard to implement and do not guarantee satisfactory results (Leulescu and Agafitei, 2013).

[17] Similarity/Dissimilarity indicators are detailed in the quality evaluation subsection

2.4.3 Selection of matching variables

Although there may be many variables X common in datasets A and B, not all of them have equal relevance for the matching procedure. The choice of the matching variables depends on the harmonization process and the predictive power over the target variables. Matching variables in A and B should be selected either by theoretic relevance or using statistical techniques (D'Orazio, 2017).[18] For instance, denote $X_Y \subseteq X$ as the common variables in A that are more relevant to explain Y and $X_Z \subseteq X$ as the common variables in B that are more relevant to explain Z. The group of common variables that are used in the matching procedure X_M should be such that: $X_Y \cap X_Z \subseteq X_M \subseteq X_Y \cup X_Z$. In other words, the selected matching variables X_M belong between the intersection and union of those variables with a significant predictive power over target variables Y and Z. This is a crucial step since directly influences the results of the matching, even more than the matching technique. Moreover, the number of variables selected is a compromise between choosing too few, which misses essential information and too many which increases the computational cost.

From the chosen set of matching variables X_M, some are selected to proceed with an exact matching. These variables segment the dataset into groups named matching classes, donation classes or strata, and the matching is performed in each group. For example, if variable "gender=(male, female)" generates a donation class, then, the matching is done separately for male observations and for female observations. Matching classes are determined by two main conditions: they should not create empty groups in either samples A or B, and the sample containing the imputed variable should have groups with more or equal observations.

In this paper, I focus on the optimal selection of matching variables by implementing novel methodologies that aim to optimize both the use of relative predicting power and the matching algorithm. Firstly, the introduction of the propensity scores as a unique matching variable by Kum and Masterson (2010) is applied. This propensity score is estimated using a logistic regression in which the regressors are the matching variables. Kum and Masterson (2010) prove that matching using propensity score estimation preserves the prediction power between the matching X_M and target variables Y and Z, and do not alter the assumptions about the relationship of the target variables Y and Z. Moreover, the use of the propensity score reduces the computational cost of the matching problem and the use of a higher list of matching variables does not affect it. Therefore, the first analysis compares the use of original matching variables and propensity score matching variables estimated in three different formulas: logistic regression, random forest and boosted CART. The incorporation of supervised machine learning techniques to calculate the propensity scores is used to examine whether the matching improvements obtained in the experimental literature are reflected in this framework.

Moreover, another aspect to consider is the relative importance of the matching variables. Although they all have a significant relevance among the set of common variables, some matching variables would have greater

[18] Depending on the nature of the common and target variables the following statistical techniques can be used: pairwise correlations, Spearman's rank correlations, Chi-square association measures, linear regressions, or CART techniques for non-linear relations between target and common variables.

predictive power than others. This analysis assesses this characteristic by the use of the importance weights from the genetic matching algorithm of Diamond and Sekhon (2013). The more explanatory the variable is, the higher value is obtained in the resulting weights from the genetic matching algorithm, and in case a variable is insignificant with respect to the others, the resulting weight will be near the zero value. Thus, this work addresses a second matching problem which examines whether there is an improvement from the use of estimated propensity scores using these relative importance weights. Moreover, the procedure allows me to examine, following the incorporation of importance weights calculated from the genetic matching algorithm, whether further analysis of the relevance of each of the common variables is no longer necessary. This would imply that all common variables can be used in the matching. Additionally, the procedure computes a new matching variable with the sum of the relative important weights.

2.4.4 Approaches

There are two levels in which the different data sources are integrated: macro and micro-level. When the objective of the matching is to obtain an estimation of the joint distribution of the target variables Y and Z, it is called the macro-approach. Therefore the matching estimates directly the parameters of interest, such as the correlation coefficient between Y and Z or the contingency table of $Y \times Z$.

The micro-approach, on the contrary, creates a *synthetic* dataset in which both target and common variables (X, Y, Z) are included. The resulting dataset is called synthetic since its records are not directly obtained from a population sample but from the appropriate use of other sources of information. This new dataset can be built by the imputation of one of the target variables from one dataset to the other. For instance, consider $A = (X,Y)$ and $B = (X,Z)$. Without loss of generality, the observations of the variable Z are filled in sample A which results as $\hat{A} = (X,Y,\hat{Z})$ where $\hat{Z}$ is the variable Z imputed in sample A and $\hat{A}$ is the resultant synthetic dataset. This way of imputations is termed the naive approach according to D'Orazio et al. (2006b) and is more common among non-parametric matching methods. The resultant sample $\hat{A}$ follows the same distribution as A, and thus, needs the assumption that the hypothetical distribution of the real (X, Y, Z) is equivalent to the distribution of $A = (X,Y)$.

Another way to create this new dataset is through file concatenation (Rubin, 1986), that is, considering the synthetic dataset is the general $A \cup B$ in which the variables Z of B are added in sample A and the variables Y of A are added in B. This approach is more common when matching with parametric methods and assumes that distributions of A and B are equivalent. A criticism of this method is that it creates the perception of doubling the original number of the total observations (Moriarity and Scheuren, 2003), which D'Orazio et al. (2006b) solves with the computation of new adapted weights (i.e. $W_{A\cup B}$). However, when dealing with data originating from surveys, to create the new weights $W_{A\cup B}$, it is necessary to calculate what the assigned weight of each unit in file A would be if it had been stratified within the same standards (or variables) as in file B. And, analogously, the weights of the units of B if they had been collected according to the stratification of A.

For this, it is necessary that the stratification variables of file A are equally defined in file B and vice versa. In practice obtaining this information can be complicated, although it is still not clear which approach (naive or concatenation) acquires better results (D'Orazio et al., 2006b).

2.4.5 Assumptions

Since the target variables Y and Z are not jointly observed, and the overarching aim is to learn about the relationship between these as well as the joint (X, Y, Z) distribution, it is useful to discuss *a priori* assumptions about the relationship between the target variables.

The first assumption to consider is the Conditional Independence Assumption (CIA) which claims the independence between Y and Z conditional on X. In essence, it states that all the information that relates Y to Z is contained in X. When this assumption holds, it is possible to estimate the joint distribution of (X, Y, Z). In other words, the synthetic dataset is equivalent to collecting a new survey of the same population where all (X, Y, Z) is contained. As a consequence, point parameters can be estimated such as the mean and covariance matrix of the joint distribution (X, Y, Z), or the correlation between Y and Z. However, it also constitutes a problem when it does. It makes no sense to proceed with a parametric analysis of these two variables combined with X, since as it is stated, all the information is contained in X. Although this condition rarely holds, it reinforces the validity of the choice of good predictive matching variables among the set of common variables to create an accurate synthetic dataset.

To obtain point parameters of the overall distribution of (X, Y, Z), D'Orazio et al. (2006b) claims an alternative of using external auxiliary information when the CIA does not hold. This additional information (denoted as C) should contain the joint distribution of (X, Y, Z) or at least (Y, Z) and it can be incorporated directly into the matching process. Therefore, when choosing the concatenation approach, target variables Z and Y are imputed respectively in A and B from sample C with the overall resulting synthetic data set $A \cup B \cup C$. In contrast, when the naive approach is performed the matching becomes a two-step process, in which firstly the target variable Z_C included in C is imputed in sample A (with respect to Y, or X and Y when available) and, thereafter, the values of Z_B in sample B are filled in A respect to X and the previous imputed Z_C.

However, these additional sources may not be compatible with the matching data sets A and B unless their sample units are generated from the same population. In that case, the auxiliary information can be used to evaluate the accuracy of the matching as in Ho et al. (2011). For instance, whether the correlation between Y and Z in sample C is preserved in the synthetic data set.

When CIA and Auxiliary information is not available the assumption of statistical matching consists in the uncertainty evaluation. The goal of statistical matching becomes the estimation of intervals of information about the relationship between Y and Z instead of point estimates. For instance, the macro-approach finds an interval of plausible distribution parameters, and the micro-approach constructs a compilation of synthetic data sets. Moriarity and Scheuren (2001, 2003) review Kadane (1978) and Rubin (1986) papers in which the notion of uncertainty was introduced in the statistical matching framework. Both coincide in repeating the matching for all the plausible values of the covariance of the target variables Y and Z such that the whole covariance matrix of X, Y and Z is positive definite. The continuous case is analysed by Moriarity and Scheuren (2001, 2003); Kadane (1978); Rubin (1986); Rässler (2012). Simplified to correlations and denoting the correlation between the target variables as ρ_{YZ}, all the plausible values that ρ_{YZ} can obtain are bounded by:

$$\rho_{XY}\rho_{XZ} - \sqrt{(1-\rho_{XY}^2)(1-\rho_{XZ}^2)} \leq \rho_{YZ} \leq \rho_{XY}\rho_{XZ} + \sqrt{(1-\rho_{XY}^2)(1-\rho_{XZ}^2)} \tag{2.1}$$

in which ρ_{XY} represents the known correlation between Y and X included in sample A, and ρ_{XZ} represents the known correlation between X and Z included in B. These bounds reaffirm that the higher the correlations between X and Y and X and Z, the shorter the interval between them and as a consequence the lower the uncertainty.

For categorical data, the uncertainty is reflected in the probability distribution of the contingency table of the target variables conditional on the common variables, i.e. YZ|X (D'Orazio et al., 2006a). Therefore, denoting $P_{YZ|X}$ as the probability of the target variables Y and Z conditional on X, the probability interval is presented with the Frechet-Hoeffding bounds (D'Orazio et al., 2017):

$$max(0; P_{Y|X} + P_{Z|X} - 1) \leq P_{YZ|X} \leq min(P_{Y|X}, P_{Z|X}) \tag{2.2}$$

where $P_{Y|X}$ is the probability of Y conditional on X included in sample A and $P_{Z|X}$ is the probability of Z conditional on X included in B. Akin to the continuous case, these intervals can be sharpened by using logical constraints (D'Orazio et al., 2006a). One type of constraint will be the structural zeros, i.e., with certainty, a particular event cannot happen due to existing laws. For example, in a country in which child labour is prohibited, it cannot be accepted to find surveyed individuals who are both nine years old and working. The other type is the inequality constraint, which means that the particular event is unlikely to happen, however, requires a subject matter expert to be validated. For example, the probability of being a blue-collar worker having a PhD is lower than being a white-collar worker.

These intervals are distinct from the widely known confidence intervals due to the nature of the uncertainty. While the uncertainty of the confidence intervals comes from the sample design, the uncertainty in the statistical matching framework comes from the data generation process. In fact, Rässler (2012) links the uncertainty intervals of the statistical matching with the ones of Manski (1995)'s *identification problem*. Manski (1995) states that having partial identification of the data, all models are plausible and describe the bounds of the interval with all the plausible estimates. Rässler (2012) applies the study of the upper and lower bounds of Manski (1995)'s intervals and simplifies the continuous interval bounds of the covariance matrix of the target variables Y and Z from Kadane (1978) and Rubin (1986) to the correlation bound of equation (2.1).

2.4.6 Matching methods

To build the synthetic file, there are three different matching frameworks to consider: parametric, non-parametric and mixed. Parametric matching performs similarly to an imputation method (Little and Rubin, 2002) and, usually, under a concatenation approach. This procedure fills the synthetic data $A \cup B$ by predicting the missing variable Z in A and, the missing variable Y in B. The parametric model is obtained either by estimating the conditional matching mean (the expectation of the missing variable given the observed variables) or the random imputation of missing values of a predictive distribution. The first method fills the missing values directly from the regression line, and the second obtains the mean and variance of the distribution and randomly selects the values to fill. However, neither of them is always considered the most accurate matching method since estimated values may not equal any observed value, or the distribution of target variables in the synthetic file may be completely different from the original distribution due to the absence of relative variability in the covariates.

In contrast, non-parametric methods do not require the specification of any family of distributions or their characteristics. Its procedure is denoted as hot deck imputation (Singh et al., 1993), whose methodology consists of filling in the missing values in one of the samples with the observed ones from the other sample. There are three main hot deck procedures described by D'Orazio et al. (2006b): nearest neighbour hot deck, random hot deck and rank hot deck. These are reviewed in detail below since they will be evaluated in the matching procedure used in this paper.

Finally, the mixed method combines both parametric and non-parametric procedures. It first proceeds with a regression step that estimates the missing values followed by a nonparametric matching step using any of the hot deck methods (Kadane, 1978; Moriarity and Scheuren, 2001, 2003). The estimation of the propensity scores using logistic regression can be considered, to a certain extent, a mixed approach.

As stated above, the three hot deck methods are reviewed in detail here with an illustrative example replicated from Rodgers (1984) and D'Orazio et al. (2006b) using two artificial and simplified datasets. Consider two different files A and B. Define $A = (X,Y)$ and $B = (X,Z)$ such that Y is the target variable included just in A, Z the target variable included just in B and X are the common variables available in both datasets. For this particular example, file A, Table 2.1, contains 8 observations and variable Y represents the logarithm of personal earnings - ln(per earn). File B, Table 2.2, contains 6 observations and the variable Z, the logarithm of property income - ln(prop inc). Moreover, both sets contain two common variables, X_1 - gender and X_2 - age, and sample weights W^A and W^B. The objective of the statistical matching is to combine these two files $A = (X,Y)$ and $B = (X,Z)$ using a naive approach with hot deck techniques and thus, to obtain one synthetic dataset containing all the variables (X_1, X_2, Y, Z, W). To do so, one of the files, *the donor file*, allocates its units of observations to the other file, *the recipient file* employing different matching methods. For this specific example, file A with 8 records, is the recipient file and B, with 6 observations, is the donor file.[19] Both common

[19]For explanatory reasons, I follow the example from Rodgers (1984), which does not affect the analysis since both population samples are the same (24 individuals each). However, in common practice, the smallest sample is chosen to be the recipient file since

variables X_1 and X_2 perform as matching variables and, in particular, variable X_1 segments the data in two matching classes of zeros and ones (i.e. individuals with $X_1 = 0$ in A will match only with individuals in B such that $X_1 = 0$).

Table 2.1: File A

Unit	X_1^A gender	X_2^A age	Y ln(per earn)	W^A weights
A1	1	42	9.156	3
A2	1	35	9.149	3
A3	0	63	9.287	3
A4	1	55	9.512	3
A5	0	28	8.484	3
A6	0	53	8.891	3
A7	0	22	8.425	3
A8	1	25	8.867	3
Mean	0.50	40.38	8.97	
SD	0.53	15.32	0.38	

Source Rodgers' (1984) Simplified Example of Statistical Matching

Table 2.2: File B

Unit	X_1^B gender	X_2^B age	Z ln(prop inc)	W^B weights
B1	0	33	6.932	4
B2	1	52	5.524	4
B3	1	28	4.224	4
B4	0	59	6.147	4
B5	1	41	7.243	4
B6	0	45	3.230	4
Mean	0.50	43.00	5.55	
SD	0.55	11.58	1.57	

Source Rodgers' (1984) Simplified Example of Statistical Matching

2.4.6.1 Nearest Neighbour Distance Hot Deck

This technique inputs for the most similar value given a distance function d which identifies the similarities between the observations. When matching variables are continuous, the literature recommends using the Mahalanobis distance function to find optimally the similarities, and when the matching variables are categorical,

the repetition of the donor records would modify the distribution of the imputed variable D'Orazio et al. (2006b). Tables B1 and B2 in the Appendix 2.B.2 show Tables 2.1 and 2.2 disaggregated from the weights, i.e. the 24 observations in each of them to reproduce the matching as Rodgers (1984).

it is best to use the Manhattan distance.[20] In case only one matching variable is used for finding similarities, then the absolute distance is valid. Depending on the imputation form, there are two different procedures that can classify the matching: Unconstrained statistical matching and Constrained statistical matching. Unconstrained statistical matching finds the nearest neighbour with no restrictions in the imputation of records from the donor file to the recipient file, i.e., elements from the donor file could be repeated or ignored depending on how near they are from other elements of the recipient file. In essence, unconstrained statistical matching solves the minimization problem of equation (2.3):

$$\min_{\substack{i=1,\dots,8\\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6} d_{ij} = \min_{\substack{i=1,\dots,8\\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6} |X_{2i}^{A} - X_{2j}^{B}| \tag{2.3}$$

where d_{ij} represents the distance function which finds the similarities between observation i in A and observation j in B. In this example, the minimization problem is equivalent to computing the absolute value (which is the distance function) between the observation i of the matching variable age in sample A (X_{2i}^{A}) and the observation j of the matching variable age in sample B (X_{2j}^{B}), for each of the categories of variable X_1. Table 3 depicts the matching result. The first column shows which are the links between each record in A and B and the second column reflects the exact matching of variable X_1. Column (3) and (4) presents the variable X_2 from sample A and B, with the absolute distance between these two in column (5)[21]. Column (6) is the target variable Y in the recipient sample A and column (7) is the imputed target variable Z from the donor file B. Finally, column (8) shows the weight values from sample A. Therefore, the first entry in the synthetic dataset is a match between observation A1 from the recipient file A (Table 2.1) and observation B5 from the donor file B (Table 2.2). This entry's matching variable X_1 (gender) is 1, which is the same in both original datasets. The value of age in the recipient file is 42 (X_2^A), and the value of age in the donor file is 41 (X_2^B), therefore the absolute distance (d_{ij}) is 1. The variable Y (Z) is simply the observed values from the recipient (donor) files for each matching observation 9.156 (7.243). The weights W^A retain the value from the recipient file only.

There are two elements to consider in these results. Firstly, the records are imputed from sample B of column (1). Some of the observations of sample B are repeated, such as record B5 which is linked with the records A1 and A2 in sample A, and others are ignored, such as record B6. This occurs due to the selection with replacement of the unconstrained matching. As a consequence, the distribution of the imputed variable Z (column (7)) differs from the original sample B. For instance, the mean and standard deviation of the variable Z in the original sample B are 5.55 and 1.57 (Table 2.2), while the mean in the synthetic dataset is 6.30 and the standard deviation 1.06 (Table 2.3). The difference in distribution values of the variable Z informs about the quality of the matching (Rässler, 2012).

In contrast, constrained statistical matching imposes that each record from the donor file is imputed only

[20] See Appendix 2.B.1 for more information about Mahalanobis and Manhattan distance.

[21] Variable X_2^B from the donor sample B is presented in the table for explanatory reasons. In practice, only variable X_2^A from the recipient file A would be included in the synthetic dataset

Table 2.3: Nearest Neighbour Distance Hot Deck: Unconstrained Match

Unit Match	$X_1^A = X_1^B$	X_2^A	X_2^B	d_{ij}	Y	Z	W^A
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A1 B5	1	42	41	1	9.156	7.243	3
A2 B5	1	35	41	6	9.149	7.243	3
A3 B4	0	63	59	4	9.287	6.147	3
A4 B2	1	55	52	3	9.512	5.524	3
A5 B1	0	28	33	5	8.484	6.932	3
A6 B4	0	53	59	6	8.891	6.147	3
A7 B1	0	22	33	11	8.425	6.932	3
A8 B3	1	25	28	3	8.867	4.223	3
Mean	0.50	40.38	43.25	4.88	8.97	6.30	
SD	0.53	15.32	12.40	3.00	0.38	1.06	

Source Rodgers' (1984) Simplified Example of Statistical Matching

once during the match, i.e. in the synthetic dataset, every record from the recipient file will be matched with one different record of the donor file. In case both recipient and donor files share the same size, then the match will use all the elements of both files. As a consequence, for this type of match, it is essential that the population size of the donor file should be greater or equal to the population size recipient file. Keeping on the same example as above with sample A (recipient) and sample B (donor), the constrained matching between these two files is represented as the following minimization problem:

$$\min_{\substack{i=1,\dots,8\\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6}(d_{ij} * w_{ij}) \tag{2.4}$$

subject to:

$$\sum_{j=1}^{6} w_{ij} = W_i^A = 3 \quad \text{for} \quad i = 1,\dots,8, \tag{2.5}$$

$$\sum_{i=1}^{8} w_{ij} \leq W_j^B = 4 \quad \text{for} \quad j = 1,\dots,6, \tag{2.6}$$

As it can be noticed, the sample weights of both samples are included in the minimization problem (Equation (2.4)) and in the constraints (Equations (2.5,2.6)). Being i an observation of the recipient file A and j an observation of the donor file B, w_{ij} defines the weight that the match between observation i and observation j obtains. This weight, multiplied by the distance function d_{ij}, needs to satisfy the condition of Equation (2.5) which states that every record from the recipient file A should be represented. For this reason, the sum is equal to three, which is the value of the weights in sample A. The second condition that needs to be satisfied (Equation (2.6)) is that the records of the donor file B should be as much represented as possible. Notice that the lower or equal condition precludes the repetition of donor records.[22]

[22] In particular, when both donor and recipient files have the same amount of observations, the inequality is transformed into equality.

In this example, Sample A has eight records with weights of three on each. Therefore Sample A has a total of 24 observations. Analogously, sample B has 6 records with weights of four per record. Thus, it also contains a total of 24 observations. Since both files have the same number of observations, all of them should participate in the minimization problem. Therefore, to solve the minimization problem of Equations (2.4, 2.5 and 2.6) is equivalent to replicate each of the records depending on their sample weights of both files[23] and compute the distance function for all of them, as shown in equation (2.7):

$$\min_{\substack{i=1,\dots,24 \\ j=1,\dots,24}} \sum_{i=1}^{24} \sum_{j=1}^{24} d_{ij} = \min_{\substack{i=1,\dots,24 \\ j=1,\dots,24}} \sum_{i=1}^{24} \sum_{j=1}^{24} |X_{2i}^{A} - X_{2j}^{B}| \tag{2.7}$$

In this case, i represents each of the twenty-four observations in the recipient sample A, j represents each of the twenty-four observations in the donor sample B and d_{ij} is the distance between observation i of the recipient file and observation j of the donor file. As before, the distance function is the absolute value for the variable age (X_2). Table 2.4 represents the result from this constrained matching procedure.

Table 2.4: Nearest Neighbour Distance Hot Deck: Constrained Match

Unit Match (1)	$X_1^A = X_1^B$ (2)	X_2 (3)	X_2^B (4)	d_{ij} (5)	Y (6)	Z (7)	W (8)
A1 B2	1	42	52	10	9.156	5.524	1
A1 B5	1	42	41	1	9.156	7.243	2
A2 B3	1	35	28	7	9.149	4.243	1
A2 B5	1	35	41	6	9.149	7.243	2
A3 B4	0	63	59	4	9.287	6.147	3
A4 B2	1	55	52	3	9.512	5.524	3
A5 B1	0	28	33	5	8.484	6.932	3
A6 B4	0	53	59	6	8.891	6.147	1
A6 B6	0	53	45	8	8.891	3.230	2
A7 B1	0	22	33	11	8.425	6.932	1
A7 B6	0	22	45	23	8.425	3.230	2
A8 B3	1	25	28	3	8.867	4.223	3
Mean	0.50	40.38	43.00	6.46	8.97	5.55	
SD	0.53	15.32	11.58	5.81	0.38	1.57	

Source Rodgers' (1984) Simplified Example of Statistical Matching

The first entry in the synthetic dataset is a match between observation A1 from the recipient file A (Table 2.1) and observation B2 from the donor file B (Table 2.2). As in the unconstrained case, the matching variable X_1 (gender) is 1, which is the same in both original datasets. The value of age in the recipient file is 42 (X_2^A), however, here the value of age in the donor file is 52 (X_2^B), therefore the absolute distance (d_{ij}) is 10. The variable Y (Z) is simply the observed values from the recipient (donor) files for each matching observation 9.156 (5.524). The weights W are not the same as the value from the recipient file only, since for this case is

[23]These tables are represented in Appendix section 2.B.2

equal to 1, meaning that the link between A1 and B2 has been generated only once. However, the sum of the weights of all times that A1 is involved in a match is equal to 3, the same as W^A. Notice that this is repeated for every record in A, satisfying condition (2.5). The equivalent happens for the donor records of B. The sum of the weights of each observation in B is equal to 4, as stated in condition (2.6). Therefore, every element of both files is represented in the synthetic dataset. As a result, it is worth noticing that the distribution of variable Z in sample B (mean and standard deviation) is maintained in the synthetic file, although, the distance between records is higher than the unconstrained case (6.46 compared to 4.88).

Comparing both approaches, the unconstrained matching problem is faster and easier to solve than the constrained matching problem, since it lacks restrictions on the data. Moreover, it guarantees higher similarities of the records. However, it does not preserve the distribution of the imputed variable as well as compared with the constrained approach, making this difference more notable when the sizes of both samples highly differ.

2.4.6.2 Random Hot Deck

In the Random Hot Deck approach, both files A and B are stratified according to different categories from the matching variables in X_A and X_B. Therefore, for each observation in a given group of A, an observation from the same group in B is randomly selected and imputed in A. As a consequence, the choice of matching variables to create the different donation classes becomes the main challenge (Waal, 2015). The random hot deck methodology applies categorical variables to create classes, such as gender, race, geographical area, etc. However, it is possible to apply one single continuous variable homogeneous in X_B and X_A to randomly select from a set of near observations. That is, fixing *a priori* a distance $\delta > 0$, the donor units are chosen in a subset of the closest records of the recipient. For instance, $d_{AB}(X_{g,A}, X_{g,B}) \leq \delta$.[24] Although this procedure is unconstrained when data from complex sample surveys are used, donors can be selected with probability proportional to their survey weights as a way of maintaining the same distribution from the original to the synthetic dataset (*weighted random hot deck*). This paper uses the weighted random hot deck technique for the analysis. However, for the sake of a better understanding, and since all sample weights in A have the same value, I illustrate here the original random hot deck. Moreover, for this example, only the variable gender X_1 has been used to create two different donation classes.

The results presented in Table 2.5 are compared with the unconstrained nearest neighbour from Table 2.3. Firstly, the random hot deck method allows replication of records from the donor file B as in the unconstrained nearest neighbour (such as B1 and B3), however, due to the randomness, all records are represented. Therefore, despite the matching distance result being larger (column (5)), the distribution of the variable Z is better preserved in the random hot deck case (column (7)).

[24] An example would be X_g being the variable age and δ the value 5, the donor chooses values whose distance from the recipient is no greater than five years.

Table 2.5: Random Hot Deck

Unit Match	$X_1^A = X_1^B$	X_2^A	X_2^B	d_{ij}	Y	Z	W^A
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A1 B2	1	42	52	10	9.156	6.932	3
A2 B3	1	35	28	7	9.149	4.223	3
A3 B4	0	63	59	4	9.287	6.147	3
A4 B5	1	55	41	14	9.512	7.243	3
A5 B1	0	28	33	5	8.484	6.932	3
A6 B6	0	53	45	8	8.891	3.230	3
A7 B1	0	22	33	11	8.425	6.932	3
A8 B3	1	25	28	3	8.867	4.223	3
Mean	0.50	40.38	39.88	7.75	8.97	5.73	
SD	0.53	15.32	11.44	3.77	0.38	1.59	

Realization of the Random Hot Deck with matching class Gender

2.4.6.3 Rank Hot Deck

The Rank Hot Deck technique follows the same logic as the Nearest Neighbour Distance: for each observation in the recipient file A the nearest neighbour from the donor file B is selected according to a distance function. It also distinguishes between unconstrained and constrained approaches depending on the imputation of the observations. In this case, however, from the selected matching variables, only one is used to calculate the distance.[25] Moreover, the order of variables is taken into account for the matching (regardless of whether they are continuous or categorical). All the observations in A and B are ranked separately according to the values of the ordinal matching variable. Equation (2.8) represents the ranking of the recipient file A and Equation (2.9) represents the ranking for the donor file B.

$$\hat{F}_X^A = \frac{\sum_{i=1}^{8} W_i^A I(X_{A,i} \leq x)}{\sum_{i=1}^{8} W_i^A} \tag{2.8}$$

$$\hat{F}_X^B = \frac{\sum_{j=1}^{6} W_j^B I(X_{B,j} \leq x)}{\sum_{j=1}^{6} W_j^B} \tag{2.9}$$

where $\hat{F}_{X_A}^A(x)$ is the cumulative distribution of the matching variable and W_i^A the weights of each observation in the recipient file A. Analogously, $\hat{F}_{X_B}^B(x)$ is the cumulative distribution of the matching variable and W_j^B the weights of each observation in the donor file B. Table 2.6 and 2.7 represent Tables 2.1 (file A) and 2.2 (file B) reordered following the distribution of the variable age (X_2).

Thus, Equation (2.10) states the unconstrained minimization problem in which each element from A is associated with an element in B depending on the cumulative distribution function. In other words, the distance between observations is calculated with the cumulative function and not the real values.

[25]The rest of the matching variables are used to create matching classes

Table 2.6: Rodgers' (1984) File A ranked by Age

Unit	X_1^A gender	X_2^A age	Y ln(per earn)	W^A weights	$\hat{F}_{X_2}^A(x)$
A7	0	22	8.425	3	3/24
A8	1	25	8.867	3	6/24
A5	0	28	8.484	3	9/24
A2	1	35	9.149	3	12/24
A1	1	42	9.156	3	15/24
A6	0	53	8.891	3	18/24
A4	1	55	9.512	3	21/24
A3	0	63	9.287	3	24/24
Mean	0.50	40.38	8.97		
SD	0.53	15.32	0.38		

Table 2.7: Rodgers' (1984) File B ranked by Age

Unit	X_1^B gender	X_2^B age	Z ln(prop inc)	W^B weights	$\hat{F}_{X_2}^B(x)$
B3	1	28	4.224	4	4/24
B1	0	33	6.932	4	8/24
B5	1	41	7.243	4	12/24
B6	0	45	3.230	4	16/24
B2	1	52	5.524	4	20/24
B4	0	59	6.147	4	24/24
Mean	0.50	43.00	5.55		
SD	0.55	11.58	1.57		

$$\min_{\substack{i=1,\dots,8\\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6} d_{ij} = \min_{\substack{i=1,\dots,8\\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6} \mid \hat{F}_X^A(x) - \hat{F}_X^B(x) \mid \tag{2.10}$$

Considering again gender as a class variable with exact matching, the unconstrained rank hot deck matching result is shown in Table 2.8. The first thing to notice is that the distance function is no longer comparable with the other hot deck techniques. However, the evaluation of the quality of the matching relies on terms of distribution conservation of the imputed target variable Z. In this example, the unconstrained rank hot deck performs better than the unconstrained nearest neighbour distance of Table 2.3, since the mean and standard deviation of the imputed Z are more similar to the original distribution in Table 2.2. However, these results are worse than the ones in Table 2.5 with the random hot deck method.

For the constrained case, the matching imposes the same restrictions as in the constrained nearest neighbour

Table 2.8: Rank Hot Deck: Unconstrained

Unit Match (1)	$X_1^A = X_1^B$ (2)	X_2^A (3)	X_2^B (4)	d_{ij} (5)	Y (6)	Z (7)	W^A (8)
A7 B1	0	22	33	0.208	8.425	6.932	3
A8 B3	1	25	28	0.083	8.867	4.223	3
A5 B1	0	28	33	0.041	8.484	6.932	3
A2 B5	1	35	41	0	9.149	7.243	3
A1 B5	1	42	41	0.125	9.156	7.243	3
A6 B6	0	53	45	0.083	8.891	3.230	3
A4 B2	1	55	52	0.041	9.512	5.524	3
A3 B4	0	63	59	0	9.287	6.147	3
Mean	0.50	40.38	41.5	0.07	8.97	5.93	
SD	0.53	15.32	9.72	0.04	0.38	1.41	

Realization of the Unconstrained Rank Hot Deck ranked by Age and classified by Gender

distance method, i.e., records from the donor file in B just can be imputed ones. As a consequence, the minimization problem reproduces the Equations (2.4), (2.5) and (2.6), but compares the distance with the cumulative function of the matching variable age (X_2) instead of their real values. Equations (2.11), (2.12) and (2.13) state the minimization problem for the rank hot deck.

$$\min_{\substack{i=1,\dots,8 \\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6}(d_{ij} * w_{ij}) = \min_{\substack{i=1,\dots,8 \\ j=1,\dots,6}} \sum_{i=1}^{8}\sum_{j=1}^{6} \mid \hat{F}_X^A(x) - \hat{F}_X^B(x) \mid * w_{ij} \tag{2.11}$$

subject to:

$$\sum_{j=1}^{6} w_{ij} = W_i^A = 3 \quad \text{for} \quad i = 1,\dots,8, \tag{2.12}$$

$$\sum_{i=1}^{8} w_{ij} \leq W_j^B = 4 \quad \text{for} \quad j = 1,\dots,6, \tag{2.13}$$

Results of the matching are presented in Table 2.9. Checking only the marginal distribution of the imputed variable Z, it seems that the constrained rank hot deck would obtain the same results as the constrained nearest neighbour hot deck (Table 2.4), although they present different linkages. Moreover, both constrained cases obtain exactly the same distribution of the variable Z as in the original file B (Table 2.2), since sample A and sample B have the same population size (24 observations each). In practice, it is very unlikely to find two different survey samples which represent exactly the same number of individuals. Therefore, despite using a constrained approach, rarely the marginal distribution of the imputed variable will be conserved as well as the ones in these examples. Moreover, not only the marginal distribution of the imputed target variable evaluates the matching. Joint distributions of the target variable with the matching variables are also analysed to evaluate it. The quality evaluation subsection reviews which are the most appropriate measures to validate a matching performance.

Table 2.9: Rank Hot Deck: Constrained

Unit Match (1)	$X_1^A = X_1^B$ (2)	X_2 (3)	X_2^B (4)	d_{ij} (5)	Y (6)	Z (7)	W (8)
A7 B1	0	22	33	0.208	8.425	6.932	3
A8 B3	1	25	28	0.083	8.867	4.223	3
A5 B1	0	28	33	0.041	8.484	6.932	1
A5 B6	0	28	41	0.291	8.484	7.243	2
A2 B5	1	35	41	0	9.149	7.243	3
A1 B6	1	42	45	0.041	9.156	3.230	3
A6 B6	0	53	45	0.083	8.891	3.230	1
A6 B2	0	53	52	0.083	8.891	5.524	2
A4 B2	1	55	52	0.041	9.512	5.524	2
A4 B4	1	55	59	0.125	9.512	6.147	1
A3 B4	0	63	59	0	9.287	6.147	3
Mean	0.50	40.38	43.00	0.09	8.97	5.55	
SD	0.53	15.32	11.58	0.05	0.38	1.57	

Realization of the Constrained Rank Hot Deck with records ranked by Age and classified by Gender

2.4.7 Quality evaluations of the Statistical Matching Results

Rässler (2012) defines 4 levels of validity for the matching result. The first requires the preservation of the marginal distribution of the imputed target variable and its joint distribution with the matching variables (empirically assessable). The second level states that the correlation between the target variables from both datasets is also preserved, and the third guarantees that the true joint distribution (X, Y, Z) is represented in the synthetic dataset (which can only be assured via previous assumptions: CIA and auxiliary information). The fourth level claims that the true values of the imputed target variable should be reproduced in the synthetic file (which never can be known). Therefore, here are presented the four indicators of the similarity between the distributions that measure the differences between original and synthetic distributions. These indicators are included in the StatMatch package of D'Orazio (2017) and are described as follows:

The first one that is included is the total variation distance function (tvd), which is a dissimilarity indicator. It compares the relative frequency of marginal and joint distributions.

$$\Delta_{or,syn} = \frac{1}{2}\sum_j |p_{or,j} - p_{syn,j}| \tag{2.14}$$

Where $p_{or,j}$ represents the relative frequencies for the *original* distribution ($0 \leq p_{or,j} \leq 1$) and $p_{syn,j}$ for the *synthetic* distribution. D'Orazio (2017) assumes that both distribution are considered equivalent when $\Delta_{1,2} \leq 0.03$ following the Agresti's rule of thumb (Agresti, 2012).

Following the same nomenclature, the second approach measures the overlap between both distributions using:

$$O_{or,syn} = \sum_j \min(p_{or,j}, p_{syn,j}) \tag{2.15}$$

The resulting values range from 0 to 1 (1 meaning that both distributions are equal). This indicator is related to the total variation distance since the overlap can be rewritten as $O_{or,syn} = 1 - \Delta_{or,syn}$.

The third measure, named the Bhattacharyya coefficient, is another measure for similarity within the values 0 and 1, with being 1 the equivalence of equal distributions (D'Orazio and D'Orazio, 2022). This measure is defined in Equation (2.16).

$$B_{or,syn} = \sum_j \sqrt{p_{or,j} \times p_{syn,j}} \tag{2.16}$$

The fourth and last indicator presented in D'Orazio and D'Orazio (2022) is Hellinger's distance. It measures the relative dissimilarity between both distributions in which the value 0 means that both distributions are equal.

$$d_{H,or,syn} = \sqrt{\frac{1}{2} \sum_j (\sqrt{p_{or,j}} - \sqrt{p_{syn,j}})^2} \tag{2.17}$$

Although there is not an official threshold which indicates the equivalence of two distributions, the literature considers two different distributions equivalent when $d_{H,or,syn} \leq 0.05$ or 5% (Donatiello et al., 2014; Leulescu and Agafitei, 2013). Therefore, the quality of results in this paper is presented using the Hellinger distance indicator.

2.5 Introduction of new types of matching variables

2.5.1 Propensity Score Matching

The propensity score estimation is defined as the conditional probability of being assigned to treatment T given a set of covariates or characteristics X:

$$e(X_i) = P(T = 1 | X = X_i) \tag{2.18}$$

The propensity score is used to identify a control group whose characteristics are similar to the treatment group. For each participant in the treatment group, they are matched with another individual in the control group sharing the same characteristics as measured by the propensity score. Usually, the propensity scores are estimated using parametric regressions such as logit, probit or linear regressions. Rosenbaum and Rubin (1983) state that if the propensity score is well specified then the distribution of characteristics, X, will be the same in the treatment and control groups. In fact, conditional on the true propensity score, the covariates and the treatment assignment are independent (i.e. CIA holds for true propensity scores).

In the statistical matching framework, this technique has been used when the number of covariates is large since it reduces the dimensionality problem by collapsing all the common variables into one unique variable: the propensity score. For instance, in an example where, after the pertinent analysis, eight common variables are selected as matching variables, in which the minimization problem finds the combined nearest distance among the eight variables. If the data used for the matching is large enough (as in the case of national surveys), the computational cost to create a synthetic dataset is immense. Therefore, when using the propensity score of those eight variables as the unique matching variable, the cost of solving the problem is automatically reduced (Kum and Masterson, 2010). In this paper, I evaluate three ways of estimating the propensity score: with logit regression, random forest and boosted Classification and Regression Trees (CART). The logit regression model predicts the assignment of belonging into the treatment or control group by the selected covariates:

$$logit(T = 1|X = x_i) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k \tag{2.19}$$

And thus, the estimated propensity score is:

$$\hat{e}(X = x_i) = \frac{exp(logit(T = 1|X = x_i))}{1 + exp(logit(T = 1|X = x_i))} \tag{2.20}$$

For the sake of comparison, logit is used instead of other parametric regressions, since it has been the standardized way of proceeding in the literature (Kum and Masterson, 2010; Lee et al., 2010). The other two are explained in the subsection below.

2.5.2 Machine Learning: Classification And Regression Trees

Machine learning aims to predict missing outcomes without the need for further inference. However, although it may seem that machine learning and econometrics are two disparate branches, their main components are comparable for greater understanding (Athey and Imbens, 2019). The terms such as training sample, features or weights in the machine learning literature, are equivalent to the terms estimation sample, covariates or regression parameters, respectively in the econometrics literature.[26] We can find two types of prediction problems: supervised learning problems, where both, regressors and the outcome, are observed; and unsupervised learning problems where covariates are grouped in clusters.

Classification And Regression Tree (CART) methods, included in the supervised learning problems, are the evolution of decision tree methods developed by Breiman et al. (1984). They consist of a hierarchically

[26]For better understanding, the term feature from the machine learning literature is referred to as covariate, regressor or independent variable.

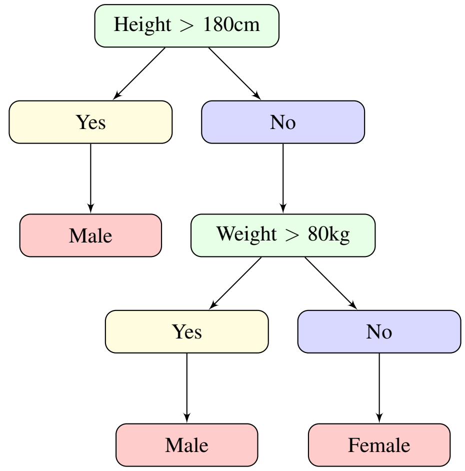


Figure 2.3: Scheme example of CART methods

organized structure which splits the dataset into disjoint subsets based on a particular value (or interval) of the independent variable. For all the observations that belong to the same subset, the same prediction is made: the mean, when the outcome is continuous (regression trees), and the mode for categorical outcomes (classification trees). Therefore, the different CART methodologies have as an objective to find the covariates that give information more efficiently after each split (i.e., the ones with the lowest variance within each subset), overcoming the overfitting issue that would result from partitioning the subsets too many times.

Figure 2.3 depicts a simplified representation of the procedure of a simple decision tree. For instance, given a population dataset, the algorithm classifies the gender of a particular individual (male or female) depending on the values of two main variables: height and weight. As it is shown, the overall data is split firstly by the covariate height which distinguishes between individuals taller and shorter of 1.80 meters. It results in two disjoint subsets: one subset, those who are taller than 1.80, is composed of only men. Therefore, the algorithm will finish here for this subset. However, the remaining subsample includes both genders and thus, requires a further partition. This second split uses the covariate weight that distinguishes between individuals with weights higher and lower than 80kg. As a result, there are two new disjoint subsets. The subset that classifies those individuals who weigh more than 80kg is composed of only men, and the remaining subset is composed of only women. Therefore, the algorithm stops here achieving a perfect split of the data using two covariates.

Among the different algorithms developed from CART methods, this study focuses on two main ones that are applied in propensity score matching literature and gave the best performance (Lee et al., 2010; Sizemore and Alkurdi, 2019): Random Forest and Boosted CART. Random forests (introduced in Breiman (2001)) is a reiteration of the decision tree process wherein each tree the splitting covariates are randomly sampled with replacement of the overall set of explanatory variables. This selection, by default, is the square root of the total number of covariates. For instance, if the sample that has been analysed is segmented by 9 different variables, the random forests algorithm would randomly take 3 of these covariates to relate them with the outcome. The performance of the procedure improves as the number of repetitions and covariates increases.

Boosted CART (Friedman, 2002) is similar to the random forest by using multiple tree iterations and employs an optimisation strategy which minimizes the error. The procedure starts with a prediction of a simple tree model where an initial subsample of covariates is chosen. After that, it calculates the residuals of the estimation sample (the difference between the real outcome and its prediction). Thus, on the new iteration, those observations that were wrongly classified in the previous iteration, have a greater priority in the selection of the new splitting covariates. Therefore, for each new tree, predictors are improved by reducing the error.

2.5.3 Genetic Matching

The propensity score fails on the specification of the model when the distribution of the common variables is significantly different after matching (Diamond and Sekhon, 2013), and thus, it can lead to bad matches in the synthetic dataset. Diamond and Sekhon (2013) presents the iterative Genetic Matching (GenMatch) algorithm that searches for the weight that minimizes the matching distance between covariates. The genetic match function uses the definition of the Mahalanobis distance but includes a positive definite weight matrix ω as follows:

$$\Delta(X_i, X_j) = \{(X_i - X_j)^T (S^{-1/2})^T \omega S^{-1/2} (X_i - X_j)\}^{\frac{1}{2}} \tag{2.21}$$

where $\Delta(X_i, X_j)$ is the Mahalanobis distance, X_i is the observation i in the matching variable X of one sample which searches for the most similar value X_j which is the observation j of the matching variable X on the other sample. $S^{-1/2}$ is the Cholesky descomposition of the variance of the covariates S ($S = S^{-1/2}(S^{-1/2})^T$).

This weight matrix is a positive diagonal matrix that sets a value for each of the covariates depending on the relevance of that variable for identification. For each iteration, the algorithm balances the covariates with statistics tests (univariate p-test or univariate Kolmogorov-Smirnov test) and the weight matrix is adjusted to prune those observations that are not significant from the tests. Using this algorithm, Diamond and Sekhon (2013) not only finds a dynamic algorithm to solve the matching problem but also, with this weight matrix, informs about which covariates are more meaningful. When the propensity score is added to the X matrix, the weight matrix would show whether it is representative of the covariates. Since the resulting elements of the diagonal of the weights ω matrix take into account the different measurement units of all the variables, they can be applied through a direct multiplication to each observation, without the need to previously standardize all the variables. Therefore, one way in which these weights are applied in this analysis is to calculate the propensity scores as shown below.

$$P(T = 1 | X = \omega X_i) = \beta_0 + \beta_1 \omega_{1,1} X_1 + \cdots + \beta_K \omega_{K,K} X_K \tag{2.22}$$

A second way these weights can be applied is by constructing a new value by adding each observation multiplied by these weights (2.23). The use of this sum to proceed with the matching is justified for the following reasons. Firstly, given that its origin comes from an iterative algorithm that minimizes matches between the different observations of two different samples, it guarantees that the weights relative to each variable are optimal. Furthermore, as explained before, since it harmonizes the different units of measurement among all the variables, this sum does not require prior standardizations for each of the variables, since the algorithm itself takes this factor into account. Finally, the sum for each observation results in a unique value, as in the propensity score case, which guarantees variability for matching.

$$SW = \sum_{k=1}^{K} \omega_{k,k} X_k \tag{2.23}$$

2.5.4 Transformation of the minimization problem

To further understand how the new matching variables are implemented, in this subsection, I outline the minimization problem with the constrained nearest neighbour distance hot deck.[27] Recall from the constrained Equations (2.4), (2.5) and (2.6) in which a distance function searches the similarities between observations taking into account the representation of each observation (using the sample weights) and makes sure that the records of the donor file were imputed without replacement (Equation (2.6)). Moreover, the minimization problem uses directly the matching variables with more predicting power. Here, the set of matching variables is used to compute a unique matching value to solve the same problem.

Consider $A = (X_A, Y_A, W_A)$, $B = (X_B, Z_B, W_B)$ where X_A is the k-vector of matching variables in A, Y_A the vector of variables included just in A and W_A the sample weights of A. Analogously, X_B is the k-vector of matching variables in B, Z_B the vector of variables included in just B and W_B the sample weights of B. Additionally, the number of observations in each dataset is defined as $|A| = n$ and $|B| = m$ and A and B represent approximately the same population, i.e. $\sum_{i=1}^{n} W_{Ai} \approx \sum_{j=1}^{m} W_{Bj}$. Let be $d_{ij} = (d_{ij1}, \ldots, d_{ijk})$ the vector of absolute distances of the common variables, such that $d_{ijp} = |X_{Aip} - X_{Bjp}|$ where X_{Aip} is the common variable p of individual i in file A and X_{Bjp} is the common variable p of individual j in file B.

Without loss of generality, denote A as the *recipient file* and B as the *donor file*. Therefore, following Rodgers (1984) framework the minimization problem will be:

$$\min_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} \sum_{i=1}^{n} \sum_{j=1}^{m} (d_{ij} * w_{ij}) = \min_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} \sum_{i=1}^{n} \sum_{j=1}^{m} (((|X_{Ai1} - X_{Bj1}| + \cdots + |X_{Aik} - X_{Bjk}|) * w_{ij}) \tag{2.24}$$

[27] Weighted random hot deck and rank hot deck methods are not explained here, however, the procedure is alike

subject to:

$$\sum_{j=1}^{m} w_{ij} = W_{Ai} \quad \text{for} \quad i = 1, \ldots, n, \tag{2.25}$$

$$\sum_{i=1}^{n} w_{ij} \leq W_{Bj} \quad \text{for} \quad j = 1, \ldots, m, \tag{2.26}$$

Where $w_{ij} \geq 0$ represents the weights between the match of individual $i \in A$ and $j \in B$.

Following Kum and Masterson (2010) the dimensionality of the minimization problem (2.24) can be reduced to one single dimension by the use of propensity scores or the sum of GenMatch weights. For the propensity score case, denote $T = \mathbf{1}_A$ the indicator of belonging to file A, therefore, $T_i = 1$ if $i \in A$ and $T_i = 0$ if $i \in B$. Therefore, $S_i = P(T = 1|X_{Ai})$ is the propensity score of individual i. In the case of the GenMatch weights sum, define $S_i = \sum_i \omega_{i,i} X_{Ai}$ as the sum of matching variables multiplied by their corresponding GenMatch weight of observation i in A. The minimization problem with propensity scores becomes:

$$\min_{\substack{i=1,\ldots,N \\ j=1,\ldots,N}} \sum_{i=1}^{N} \sum_{j=1}^{N} |S_i - S_j| * w_{ij} \tag{2.27}$$

subject to:

$$\sum_{j=1}^{m} w_{ij} = W_{Ai} \quad \text{for} \quad i = 1, \ldots, n, \tag{2.28}$$

$$\sum_{i=1}^{n} w_{ij} \leq W_{Bj} \quad \text{for} \quad j = 1, \ldots, m, \tag{2.29}$$

2.5.4.1 Propensity score computation

In practice, when working with data, to obtain the propensity scores Equation (2.27), it is necessary to add one step before applying the matching functions. For instance, to each of both files A and B the variable T is added as described in Equation (2.30). This variable identifies whether observation i corresponds to file A or B. This is the same reasoning as the observation i belongs to the treatment or control group in the propensity score matching literature. Once variable T is created, both files A and B are appended creating a unique $A \cup B$ set.

$$T_i = \begin{cases} 1 & if \quad i \in A \\ \\ 0 & if \quad i \in B \end{cases} \tag{2.30}$$

Continuing, the Genetic Matching algorithm from Diamond and Sekhon (2013) is applied to obtain the vector including the diagonal values of covariate balance matrix ω in Equation (2.21).[28] This matrix has two different roles. The first one is used to obtain three different weighted propensity scores: estimated using a

[28]The function of use is `GenMatch()` from the package `rgenoud` in R created by Sekhon (2011).

parametric logistic function, random forest and a boosted CART function derived from a general boosted modelling function (as in Equation (2.22)).[29] The second role is to add the factor between the matching variables and the GenMatch weight matrix (as in Equation (2.23)). Once the propensity score and the rest of the values are included in the set $A \cup B$, this union set is separated into their original files A and B using the variable T. Thereafter the three different hot deck techniques are applied to proceed with the matching.

2.6 Application of the matching techniques

2.6.1 Data

With the purpose of evaluating the three hot deck matching algorithms with the different matching variables, the analysis uses the artificial sample data included in the StatMatch R-package from D'Orazio (2017). This data set is inspired by the information usually covered in the European Union Statistics on Income and Living Conditions Survey (EU-SILC) and includes two different samples with the purpose of analysing all the Statistical Matching applications. Sample A has 3,009 observations which include the target variable (Y) income. Sample B has 6,686 observations with labour market status as the target variable (Z). A rich set of common variables, including age, gender and marital status is present in both datasets, as well as survey weights which assure the representation of two equivalent populations.

2.6.2 Results

The objective of the first scenario is to replicate and compare the results of D'Orazio (2017) with more recent matching techniques. As described in the data section, artificial data was created specifically for this type of analysis in the paper. Therefore, the selection of target and common variables, as well as the harmonization of the databases, are already determined. The definitions of all common variables are described in Table B3 in the Appendix accompanied by the corresponding harmonization categories for each variable. Additionally, Table B4 in the Appendix outlines the categorical distribution of each common variable, including the utilization of the Hellinger distance as a dissimilarity measure among distributions. Since all common variables present a Hellinger distance lower than 5%, it is verified that all the variables have equivalent distributions[30]. Furthermore, D'Orazio (2017) chooses sample A as the recipient file, so net income becomes variable Y and chooses sample B as the donor file, with labour status becoming the imputed variable Z. Also, when taking into account

[29]Remaining in R, the first propensity score is estimated using the parametric `glm()` logit function. The second one, with the `cforest()` computes the propensity score using random forest. Finally, the propensity scores estimation from boosted CART function is obtained by the `ps()` function which is an application from the general boosted modelling function `gmb()` that computes, particularly, propensity scores.

[30]As mentioned in Section 2.4.7 the rule of thumb agreed in the literature to consider two distributions equivalent is to present a Hellinger distance lower than 5% (Donatiello et al., 2014; Leulescu and Agafitei, 2013)

the sample weights, both samples represent a similar number of population[31], the calibration of the survey weights is not required for this analysis.

For the choice of common variables to be used in the matching, D'Orazio (2017) presents two different analyses for each of the target variables Y and Z. For net income (Y), which is a continuous variable, it is based on the adjusted-R^2 using Spearman's rank correlation. In this case, the common variables with the greatest explanatory power on the variability of net income are age and gender. For the case of labour status (Z), since it is a categorical variable, Cramer's V measure is used.[32] The common variables that have the greatest explanatory power on labour status are age and education level. Therefore, the variables with the overall best prediction power for the two target variables are age, gender and education level. Thus, in a first analysis, these three matching covariates are employed to calculate the three different propensity scores (with the logistic regression, random forest and boosted CART). In contrast D'Orazio (2017) uses age for the distance function in the minimization problem and gender and geographic area as donation classes, justifying that when the segmentation is given by gender and education level, there are two clusters with very few observations that can cause matching errors.

2.6.2.1 Use of the propensity score matching variables

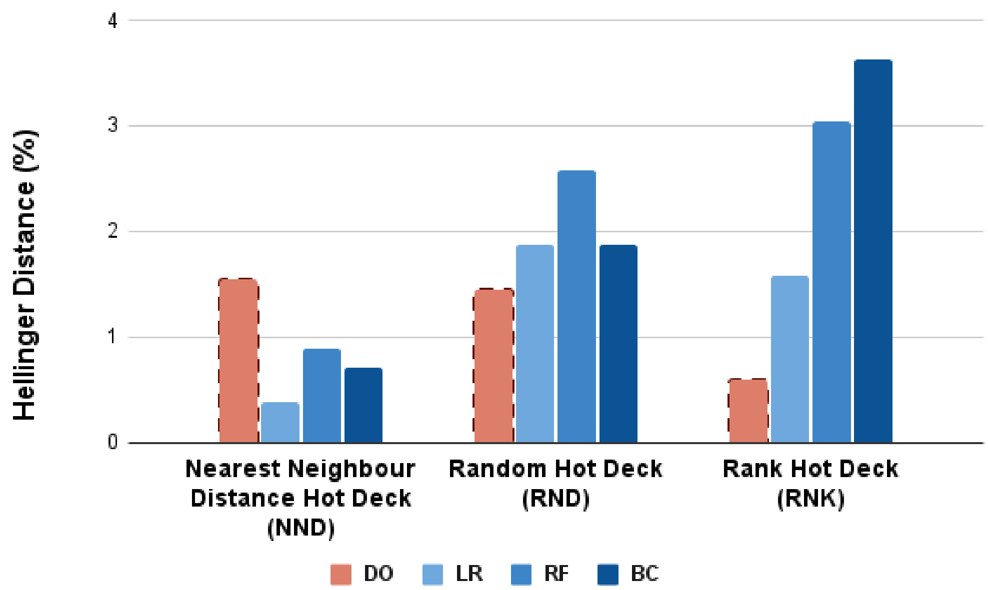
With the aim to create a synthetic dataset (micro-approach), D'Orazio uses three of the previously described methods: constrained nearest neighbour hot deck, weighted random hot deck and constrained rank hot deck. In the case of this paper, these three methods are used, keeping the donation classes' sex and geographic area, but changing the matching variable age for three different propensity scores: one calculated through logistic regression, another through the random forest and the last one using boosted CART. Therefore, a total of twelve different synthetic datasets are created: each of the three methods uses the four different matching variables.

To evaluate the quality of each of the twelve different procedures, the similarity of the marginal distributions of labour status and the joint distributions of labour status and sex, age, education and geographical area are compared. Figure 2.4 represents the results for the marginal distribution of labour status (Z). The graph depicts 12 different bars which indicate each of the methods. The first group of four bars represents the matching using the constrained nearest neighbour distance hot deck method (NND). The second group of four refers to the weighted random hot deck method (RND) and the third group is for the matchings using the constrained rank hot deck (RNK). Moreover, in each of the method groups, a different matching variable has been used. The first one (DO) indicates that the matching variable used is age as in D'Orazio (2017). (LR) indicates the use of the propensity score value estimated by a logit regression. (RF) indicates the propensity score computed

[31] The 3,009 observations of sample A represents a population of 5,094,952 individuals, and the 6,686 observations of sample B represents a population of 5,157,582 individuals

[32] Cramer's V measures how strongly two categorical fields are associated with the chi-square test of independence. In social sciences, a Cramer's V value above 0.35 is often considered a strong association, between 0.25 and 0.35 medium association and between 0 and 0.25 indicates a weak association, Waal (2015)

Figure 2.4: Marginal distributions



Note: Resulting Hellinger distances by measuring the dissimilarities of the marginal distributions of the target variable Z: labour status. Each subgroup corresponds to a distinct methodology employed: Nearest Neighbour Hot Deck, Random Hot Deck, and Rank Hot Deck. Within each method group, the first bar labelled as **DO** pertains to the resultant matching variable age used in D'Orazio (2017), serving as the baseline reference for this analysis. The subsequent bars are the ones calculated using propensity scores: **LR** - represents the matching variable calculated with Propensity Score estimated with Logistic Regression; **RF** - signifies the matching variable derived from Propensity Score estimated with Random Forest; and **BC** - pertains to the matching variable derived through Propensity Scores estimated using Boosted CART.

with random forest and (BC) boosted CART. These matching variables have been computed using solely the selected matching variables in D'Orazio (2017): age, gender, education level, and geographical area.

The vertical axis indicates the level of dissimilarity percentage between the distribution of labour status (Z) in the original sample and in the synthetic samples. The equivalence of the distributions is evaluated using Hellinger's distance, where a lower value indicates a higher degree of similarity, being 5% the highest value of the Hellinger distance which indicates that the two distributions are equivalent. Figure 2.4 shows that all 12 synthetic datasets preserved the marginal distribution of the imputed labour force variable since all of them have less than 3.5% of dissimilarity. In particular, it is noted that the most accurate method is the constrained nearest neighbour distance using propensity scores with logit regression (NND.LR), which is equivalent to the method used in Kum and Masterson (2010). Table B5, in the Appendix, shows the resulting Hellinger distance values for each matching variable and hot deck methodology. For the purpose of comprehensive comparison, I have additionally incorporated in Table B5 the resultant Hellinger distance when the propensity scores were calculated using all common variables: age, age squared, gender, education, geographical area, household size, marital status and urban area. The outcomes demonstrate similar results with improvements observed in the RNK method while deteriorations within the NND method.

Figure 2.5 depicts the joint distribution between the imputed variable labour force and the variables used for the matching: gender, age, education level and geographical area. Overall, the twelve synthetic datasets show equivalence in the joint distribution between labour status and both, gender and geographical area. When comparing the joint distribution of labour force (Z) and the matching variable gender, D'Orazio (2017) obtains better results by employing age in the minimization problem (in particular through the use of the rank hot deck method). The results presented between labour status and geographic area neither show an improvement due to the use of propensity scores. When the matching method is the random hot deck or rank hot deck, Figure 5 indicates that both the synthetic files that have used random forest or boosted CART propensity scores (RND.RF, RND.BC, RNK.RF and RNK.BC), do not produce equivalent joint distributions. For instance, their level of dissimilarity is higher than the horizontal blue line which indicates differences in more than 5%.[33]

In the case of the variables age and level of education, the results differ more compared to the other two previous variables. In fact, none of the joint distributions between education level and employment status could be considered equivalent, and only the nearest neighbour distance and rank hot deck using the age variable, (NND.DO and RNK.DO) have a value less than 5%. The horizontal black line shows the dissimilarity level of 10%, for the case of the joint distribution of labour force and age, the graph shows how divergent these distributions are from the original sample. This arises from the higher segmentation in both cases compared to the case of the variables sex and geographic area. Despite these results, Figure 2.5 shows that the nearest neighbour distance hot deck is the method that comes closest to better results.

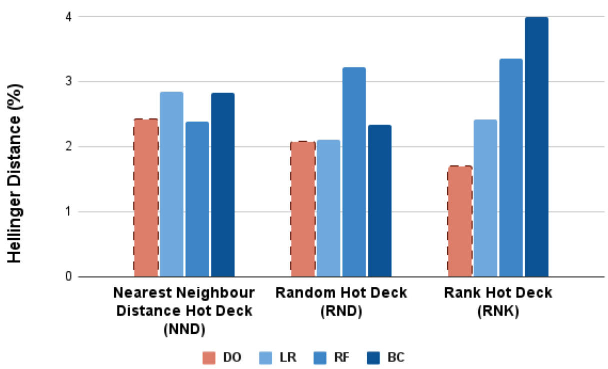
2.6.2.2 Use of genetic matching algorithm weights

In the second phase of the analysis, the weight matrix of the genetic matching algorithm was integrated into the matching procedure. These weights were applied in two different manners: firstly, for the computation of the three different weighted propensity scores, and secondly, for the aggregation of the matching variables. Following the Sekhon (2011) procedure, a balanced matrix using all the common variables was created: age, age squared, gender, education, geographical area, household size, marital status and urban area. This Balance matrix aims to evaluate whether the distribution discrepancies between common variables are minimized after the matching. Therefore, by default, the Balance matrix is equal to the matching variables but can contain additional variables that are transformed in various ways. After applying the genetic matching algorithm with the set of common variables and the Balance matrix the relative importance weight matrix is obtained. The diagonal elements for this matrix are shown in Table 2.10 ranked from the most relevant variable (gender) to the least (marital status).

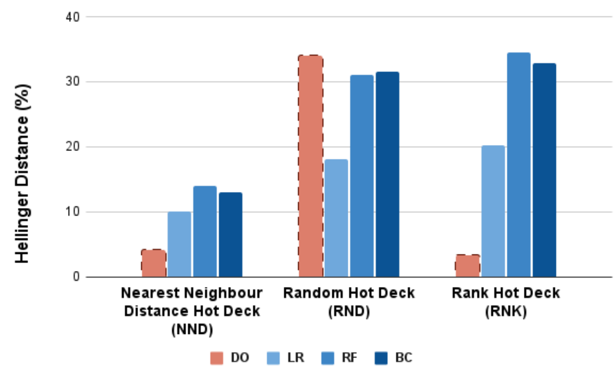
In alignment with the findings of D'Orazio (2017), the Weight matrix ω exhibits high importance to variables such as gender, age and geographical area as the most influential factors for the matching. In D'Orazio (2017), these variables are designated as donation classes (gender, and geographical area) and matching vari-

[33] Tables B6 to B9 presents the exact numbers of the resulting Hellinger distance of the joint distributions of labour status and gender (Table B6), age (Table B7), education level (Table B8), and geographical area (Table B9). These tables additionally include the results when all common variables have been used for the computation of the propensity scores.

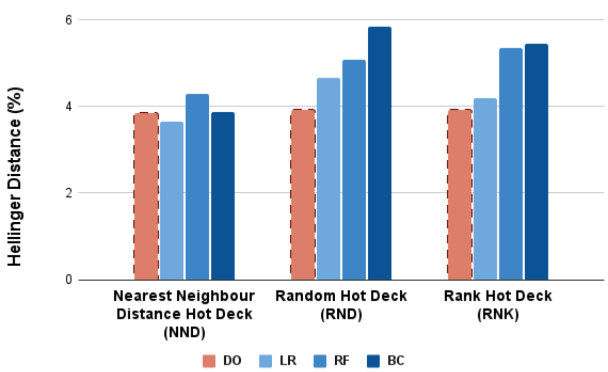
Figure 2.5: Joint distributions



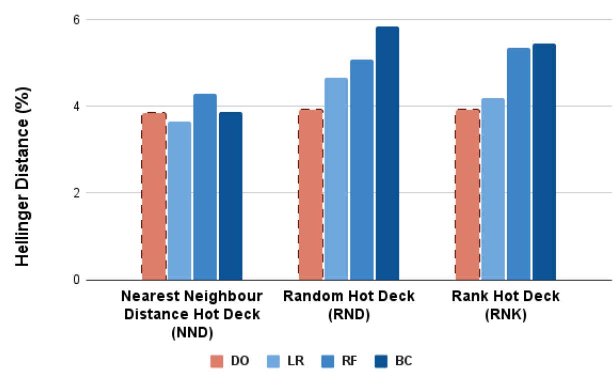
(a) Gender and Labour Status



(b) Age and Labour Status



(c) Education and Labour Status



(d) Area and Labour Status

Note: Resulting Hellinger distances by measuring the dissimilarities of the joint distributions of the target variable Z: labour status, and the relevant matching variables: gender, age, education level, and geographic area. Each subgroup corresponds to a distinct methodology employed: Nearest Neighbour Hot Deck, Random Hot Deck, and Rank Hot Deck. Within each method group, the first bar labelled as **DO** pertains to the resultant matching variable age used in D'Orazio (2017), serving as the baseline reference for this analysis. The subsequent bars are the ones calculated using propensity scores: **LR** - represents the matching variable calculated with Propensity Score estimated with Logistic Regression; **RF** - signifies the matching variable derived from Propensity Score estimated with Random Forest; and **BC** - pertains to the matching variable derived through Propensity Scores estimated using Boosted CART.

Table 2.10: Diagonal elements of the Weight matrix ω

Variable	ω
Gender	996.22
Age	909.39
Geographical Area	585.97
Household size	431.99
Education level	238.97
Urban area	215.54
Marital status	202.81

ables (age). Therefore, the GenMatch algorithm consistently displays its capability in identifying pivotal matching variables that are crucial for the comprehensive matching process. Consequently, this analysis encompasses additional results stemming from the application of the GenMatch weights using solely the selected matching variables.

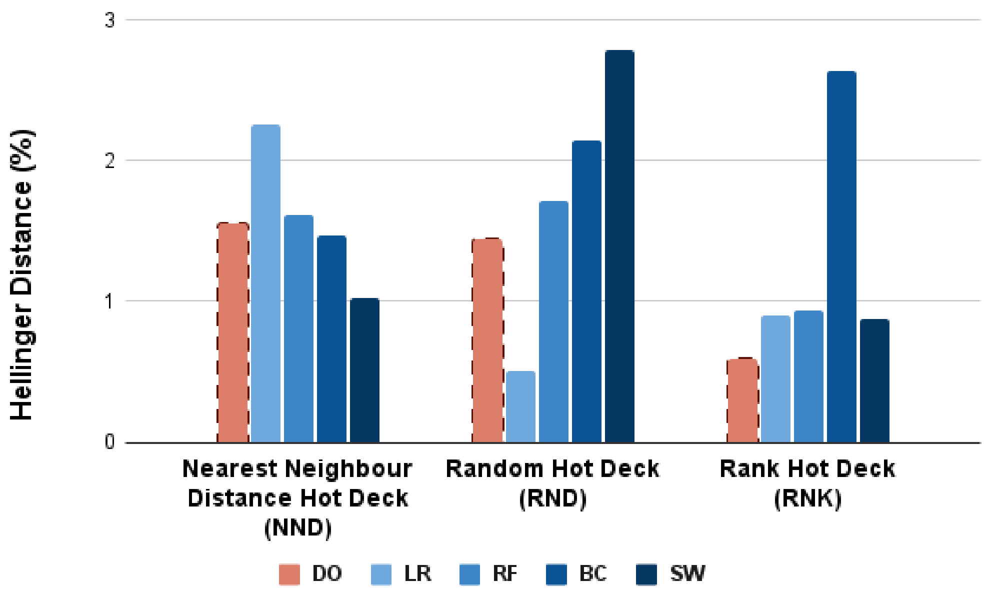
Taking into account the three different matching methods and all different matching variables, the accuracy of the fifteen[34] different synthetic datasets are evaluated. Firstly, the evaluation centres around the marginal distribution of the imputed variable labour status. Figure 2.6 depicts the Hellinger distance indicator of dissimilarity. The nomenclature for each bar aligns with the earlier results in Figure 2.4, therefore, NND, RND and RNK[35], indicate the three different hot deck methods and DO, LR, RF, BC, SW indicate which is the variable used for the matching. Notice that, the matching variable age used in D'Orazio (2017) remains the same (DO), but the propensity scores are computed using the weight matrix ω (Logistics Regression - LR, Random Forest - RF and Boosted CART- BC), and there is a new variable added, which is the sum of the common variables with the weights (SW). It is worth noting that the results illustrate a desirable preservation of the marginal distribution, as evidenced by Hellinger's dissimilarity measure being under 3%. In fact, the outcomes display relative improvement when compared to previous matchings conducted without the use of weights (Figure 2.4). Table B10 presents the numerical outcomes displayed in Figure 2.6. It includes the results derived from using exclusively the selected matching variables from D'Orazio (2017) (gender, age, education level, and geographical area) employed for the computation of the propensity score with the GenMatch weights and the sum of the GenMatch weights matching variables. While the disparities indicated by the Hellinger distance measures are not extensively pronounced, a discernible trend is observed. In general, the RNK method exhibits a marginal deterioration, while the NND and RND methods display an improvement by reducing the number of variables for the computation of the matching variable.

Figure 2.7 shows the Hellinger distance for the joint distribution between the imputed variable labour status and the common variables gender, age, education level and geographical area. Although in this phase of the analysis, all common variables have been used for the computation of the matching, we only evaluate these

[34]The 12 synthetic datasets considered above and the 3 synthetic datasets (denoted *SW*) including the sum of matching variables multiplied by the importance weights of GenMatch.

[35]Nearest neighbour distance, weighted random and rank hot deck methods

Figure 2.6: Marginal distributions

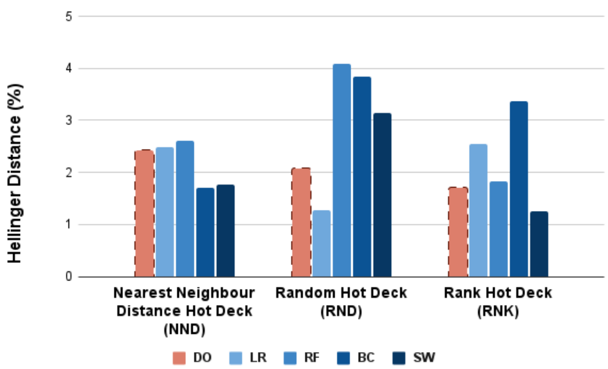


Note: Resulting Hellinger distances by measuring the dissimilarities of the marginal distributions of the target variable Z: labour status. Each subgroup corresponds to a distinct methodology employed: Nearest Neighbour Hot Deck, Random Hot Deck, and Rank Hot Deck. Within each method group, the first bar labelled as **DO** pertains to the resultant matching variable age used in D'Orazio (2017), serving as the baseline reference for this analysis. The subsequent bars are the ones calculated using propensity scores: **LR** - represents the matching variable calculated with Propensity Score estimated with Logistic Regression; **RF** - signifies the matching variable derived from Propensity Score estimated with Random Forest; and **BC** - pertains to the matching variable derived through Propensity Scores estimated using Boosted CART. Finally, the last bar of each method group represents the matching variable derived from the sum of the GenMatch Algorithm weights - **SW**.

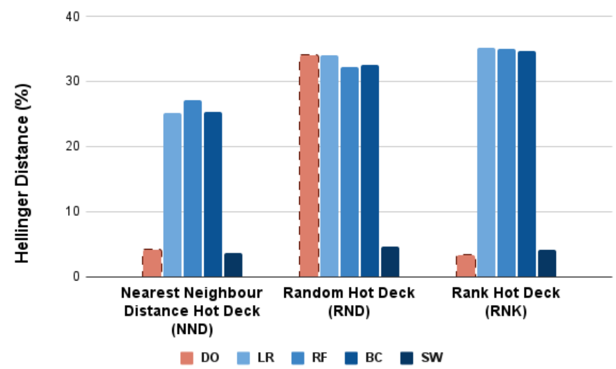
four since, as analyzed by D'Orazio (2017), they are the most explanatory of the target variables. When comparing the performance of these matches, in general, it is observed that the added use of relative weights ω in the propensity scores does not significantly improve the precision of the match. For instance, despite the similarity in the case of the variable gender and geographic area, (in which in both cases most of the methods present less than a 5% discrepancy) for both age and education level, the level of divergence from the original distribution using propensity scores are much worse. However, it should be noted that the use of the direct sum of the variables with their relative weights (SW) produces better outcomes. In the case of the joint distribution between education level and labour status, despite not being able to consider equivalence in the distributions, the results are obtained with less dissimilarity. The resulting Hellinger distance is lower than 10%, which is considered not excellent, but still valid (Donatiello et al., 2014). This is more clearly reflected in the case of the joint distribution between age and labour status, which, unlike the others, regardless of the hot deck method used, preserves the original distribution (Hellinger distance lower than 5%).

Lastly, for the purpose of comprehensive comparison, Tables B11 to B14 in the Appendix, present all the numerical outcomes resulting from calculating the matching variables using the selected ones from D'Orazio (2017): gender, age, educational level, and geographical area. The joint distributions of "labour status - gender", and "labour status - geographical" area remain similar to the outcomes resulting from the computation of the matching variables (LR, RF, BC, and SW) using all common variables. However, for the joint distributions of "labour status - age", and "labour status - educational level" there is a notable improvement in the results for the propensity score variables, although in some instances the Hellinger distance is still considered too high, more than 10%.

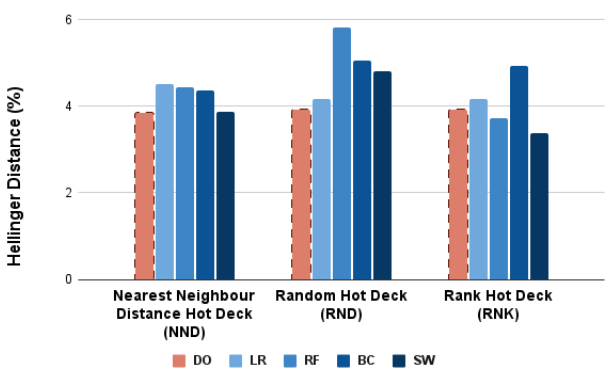
Figure 2.7: Joint distributions



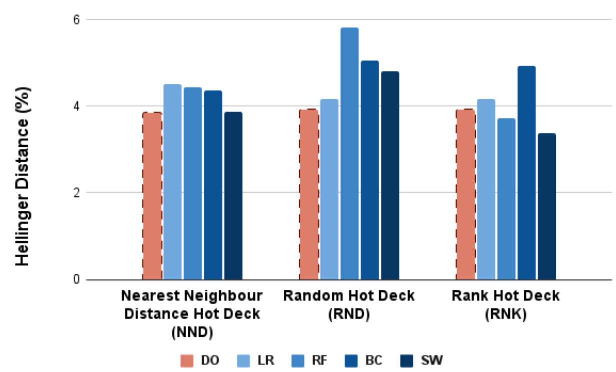
(a) Gender and Labour Status



(b) Age and Labour Status



(c) Education and Labour Status



(d) Area and Labour Status

Note: Resulting Hellinger distances by measuring the dissimilarities of the joint distributions of the target variable Z: labour status, and the relevant matching variables: gender, age, education level, and geographic area. Each subgroup corresponds to a distinct methodology employed: Nearest Neighbour Hot Deck, Random Hot Deck, and Rank Hot Deck. Within each method group, the first bar labelled as **DO** pertains to the resultant matching variable age used in D'Orazio (2017), serving as the baseline reference for this analysis. The subsequent bars are the ones calculated using propensity scores: **LR** - represents the matching variable calculated with Propensity Score estimated with Logistic Regression; **RF** - signifies the matching variable derived from Propensity Score estimated with Random Forest; and **BC** - pertains to the matching variable derived through Propensity Scores estimated using Boosted CART. Finally, the last bar of each method group represents the matching variable derived from the sum of the GenMatch Algorithm weights - **SW**.

2.7 Conclusion

Statistical matching serves as a useful tool in the academic literature when the need arises to combine information from multiple sources to comprehend a specific phenomenon. This technique enables the estimation of variables that are not observed together by utilizing a set of common variables. However, not all common variables can be utilized due to concerns such as their lack of relevance or computational complexity. Consequently, a subset of common variables is selected as the matching variables, with the aim of producing synthetic datasets that best preserve the original distributions. In this study, particular attention is devoted to the selection of matching variables and their impact on the quality of the matchings. To this end, new matching variables are introduced, incorporating propensity scores and machine learning techniques, and are compared with the direct use of variables as presented in D'Orazio (2017).

The results of the analysis indicate that, overall, the use of propensity scores to create matching variables does not lead to a significant improvement in the quality of the matchings. However, the incorporation of the sum of genetic matching weights demonstrates promising performance, outperforming the original matching variable from D'Orazio (2017) in some cases. Thus, the sum of the weights proves to be a viable matching variable for reproducing the statistical matching analysis using two real survey datasets. However, it is important to acknowledge that the use of the Genetic Matching (GenMatch) algorithm entails a computational cost drawback, particularly when handling large datasets such as those encountered in national representative surveys. In such contexts, the results underscore the validity of employing a selection of matching variables rather than the entirety of common variables. Yet, this approach leads to the compromise of one pivotal objective of the analysis presented - namely, the avoidance of evaluating which are the selected matching variables for the overall statistical matching.

In the next chapter, this technique is applied to create a synthetic dataset from the Mexican Time Use Survey and the Mexican National Survey on the Dynamics of Household Relationships. The primary objective is to investigate whether Intimate Partner Violence is associated with the individual allocation of time to work, housework, caregiving, or leisure activities.

2.B Appendix Chapter 2

2.B.1 Definitions

2.B.1.1 Mahalanobis Distance

Mahalanobis distance is a metric that measures the distance between a point and a distribution while accounting for the correlations between the different covariates or dimensions within the data. In mathematical terms, given a point denoted as x and a probability distribution Q characterized by a mean μ and positive-definite covariance matrix Σ, the Mahalanobis distance measures the separation between x and Q as in equation 2.31. The inclusion of the term Σ^{-1} incorporates the bespoke correlations between the different dimensions or covariates within the analytical dataset. Consequently, this metric demonstrates greater versatility than the Euclidean distance, especially when handling data that exhibits diverse scales of correlations. Notice that when the covariance matrix Σ reduces to the Identity Matrix, the Mahalanobis distance is equivalent to the Euclidean distance.

$$d_M(x,Q) = \sqrt{(x-\mu)^T \Sigma^{-1} (x-\mu)} \tag{2.31}$$

The Mahalanobis distance has been frequently applied within matching algorithms (D'Orazio et al., 2006b; Lee et al., 2010; Sekhon, 2011; Diamond and Sekhon, 2013; Sizemore and Alkurdi, 2019), primarily due to its versatile nature in accommodating correlations within covariates. However, this distant function exhibits diminished effectiveness when a substantial majority of the selected covariates are categorical rather than continuous. In such scenarios D'Orazio (2017) recommends the utilization of the Manhattan distance instead.

2.B.1.2 Manhattan Distance

The Manhattan distance, also referred to as the Taxicab distance, is a metric that computes the distance between two points or vectors by summing the absolute differences of their Cartesian coordinates. The inspiration behind its name is drawn by the orthogonal grid layout of streets of the Manhattan Island, and the taxi routes along avenues and streets. In other terms, and as the equation 2.32 shows, the distance d_T between two vectors $p = (p_1, p_2, ..., p_n)$ and $q = (q_1, q_2, ..., q_n)$ corresponds to the cumulative sum of the absolute difference between two points along the coordinate axes. The Manhattan distance is commonly used in machine learning algorithms due to the efficacy when dealing with high-dimensional datasets.

$$d_T(p,q) = \sum_{i=n}^{n} |p_i - q_i| \tag{2.32}$$

Therefore, in instances where $n = 1$, such as the scenario of the matching method using propensity scores (and subsequently, the summation of GenMatch weights), the distance d_T is equal to the absolute difference between the resultant matching variables.

2.B.2 Matching methods tables

Table B1: File A disaggregated

Unit	X_1^A gender	X_2^A age	Y ln(per earn)	W^A weights
A1	1	42	9.156	1
A1	1	42	9.156	1
A1	1	42	9.156	1
A2	1	35	9.149	1
A2	1	35	9.149	1
A2	1	35	9.149	1
A3	0	63	9.287	1
A3	0	63	9.287	1
A3	0	63	9.287	1
A4	1	55	9.512	1
A4	1	55	9.512	1
A4	1	55	9.512	1
A5	0	28	8.484	1
A5	0	28	8.484	1
A5	0	28	8.484	1
A6	0	53	8.891	1
A6	0	53	8.891	1
A6	0	53	8.891	1
A7	0	22	8.425	1
A7	0	22	8.425	1
A7	0	22	8.425	1
A8	1	25	8.867	1
A8	1	25	8.867	1
A8	1	25	8.867	1
Mean	0.50	40.38	8.97	
SD	0.53	15.32	0.38	

Source Rodgers' (1984) Simplified Example of Statistical Matching

Table B2: File B disaggregated

Unit	X_1^B gender	X_2^B age	Z ln(prop inc)	W^B weights
B1	0	33	6.932	1
B1	0	33	6.932	1
B1	0	33	6.932	1
B1	0	33	6.932	1
B2	1	52	5.524	1
B2	1	52	5.524	1
B2	1	52	5.524	1
B2	1	52	5.524	1
B3	1	28	4.224	1
B3	1	28	4.224	1
B3	1	28	4.224	1
B3	1	28	4.224	1
B4	0	59	6.147	1
B4	0	59	6.147	1
B4	0	59	6.147	1
B4	0	59	6.147	1
B5	1	41	7.243	1
B5	1	41	7.243	1
B5	1	41	7.243	1
B5	1	41	7.243	1
B6	0	45	3.230	1
B6	0	45	3.230	1
B6	0	45	3.230	1
B6	0	45	3.230	1
Mean	0.50	43.00	5.55	
SD	0.55	11.58	1.57	

Source Rodgers' (1984) Simplified Example of Statistical Matching

2.B.3 Summary statistics

Table B3: Coherence on the definition of variables

Variable	Definition in Sample A and Sample B
HH.P.id	unique unit identifier of the type aa.bb where aa identifies the Household while bb identifies the household member
area5	large geographic area, factor with 5 levels: 'NE'=North–East, 'NO'=North–West, 'C'=center, 'S'=South, 'I'=islands
urb	Degree of urbanization, factor with 3 levels: '1'=densely populated area, '2'=intermediate area, '3'=thinly populated area
hsize	integer, size of the household in which the person lives
hsize5	factor with 5 levels derived from hsize, where the 5th level '≥ 5' denotes 5 and more people in the household
age	integer, the person's age
c.age	factor, age categorized in 5 classes
sex	factor, the person's gender: '1'=male, '2'=female
marital	factor, the person's marital status: '1'=never married, '2'=married, '3'=other (separated, widowed, divorced)
edu7	factor, the person's highest education level attained, follows the ISCED-97 categories: '0'=pre–primary education, '1'=primary education, '2'=lower secondary education, '3'= (upper) secondary education, '4'= post–secondary non tertiary education, '5'=first stage of tertiary education (not leading directly to an advanced research qualification), '6'=second stage of tertiary education (leading to an advanced research qualification)
ww	numeric, the unit's weight

Table B4: Summary Distributions for Sample A and Sample B with the Hellinger Distance measurement

Variable	Sample A	Sample B	Hellinger Distance (%)
Area			2.54%
NE	23.85108	26.93917	
NO	20.61527	19.36315	
C	21.37817	20.77327	
S	23.64949	22.78076	
I	10.50598	10.14366	
Urban			0.82%
1	43.03618	43.89683	
2	40.44337	40.34773	
3	16.52045	15.75544	
Household size			1.65%
1	17.459225	16.566937	
2	41.092670	42.924071	
3	23.451481	22.924596	
4	14.253275	14.366363	
$>= 5$	3.743348	3.218033	
Age			1.29%
$[16,34]$	23.98236	23.46207	
$(34,44]$	19.37460	20.16357	
$(44,54]$	18.38099	17.26977	
$(54,64]$	14.28559	14.73367	
$(64,104]$	23.97646	24.37092	
Sex			0.38%
1	48.17474	48.71864	
2	51.82526	51.28136	
Marital Status			0.96%
1	29.94446	29.30632	
2	55.70804	55.41597	
3	14.34750	15.27771	
Education level			1.15%
0	2.6613821	2.9002045	
1	18.5263947	19.3360287	
2	31.1413802	31.1031554	
3	33.5457835	32.7168462	
4	2.8142158	2.7359128	
5	10.9567272	10.9447791	
6	0.3541165	0.2630734	

2.B.4 Table Results

Hellinger distance (%) differences comparison of matching variables derived from Propensity scores

Each column in the table corresponds to a distinct hot deck methodology: Nearest Neighbour Hot Deck (NND), Random Hot Deck (RND), and Rank Hot Deck (RNK). Each row corresponds to the computation of the variable utilized for the matching: Age variable from the study by D'Orazio (2017); Propensity Score derived through Logit Regression; Propensity Score derived from Random Forest; Propensity Score derived via Boosted CART. The cells categorized under the "Matching variables" columns signify that the chosen matching variables from D'Orazio (2017) have been employed for the computation of the propensity scores. The cells grouped in the "Common variables" columns denote that all common variables have been employed for the computation of the propensity scores.

Table B5: Results for preserving labour status **marginal distributions**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	1.55	1.44	0.59	1.55	1.44	0.59
PS Logit Regression	0.38	1.88	1.58	2.31	1.82	1.38
PS Random Forest	0.89	2.58	3.04	1.38	3.37	2.19
PS Boosted CART	0.7	1.87	3.63	2.12	0.51	1.69

Table B6: Results for preserving the joint distributions between **sex and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	2.42	2.07	1.7	2.42	2.07	1.7
PS Logit Regression	2.85	2.11	2.41	2.7	2.09	2.55
PS Random Forest	2.39	3.22	3.35	2.14	5.11	2.63
PS Boosted CART	2.83	2.34	4	2.45	0.88	2.16

Table B7: Results for preserving the joint distributions between **age and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	4.14	34	3.38	4.14	34	3.38
PS Logit Regression	10.03	18.14	20.28	26.96	32.29	32.24
PS Random Forest	14.05	31.1	34.53	26.67	34.39	35.15
PS Boosted CART	12.91	31.58	32.87	26.03	34.1	35.57

Table B8: Results for preserving the joint distributions between **education and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	7.9	15.81	6.77	7.9	15.81	6.77
PS Logit Regression	5.58	10.34	9.35	11.2	14.46	14.91
PS Random Forest	5.79	13.42	16.66	12.8	16.28	16.7
PS Boosted CART	6.49	14.44	15.54	11.29	15.48	16.75

Table B9: Results for preserving the joint distributions between **area and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	3.85	3.93	3.91	3.85	3.93	3.91
PS Logit Regression	3.65	4.65	4.2	5.12	4.89	4.43
PS Random Forest	4.29	5.07	5.34	4.27	7	4.31
PS Boosted CART	3.86	5.83	5.45	4.42	3.83	4.55

Hellinger distance (%) differences comparison of Matching variables derived from Propensity scores and GenMatch Weights Matching Variables

Each column in the table corresponds to a distinct hot deck methodology: Nearest Neighbour Hot Deck (NND), Random Hot Deck (RND), and Rank Hot Deck (RNK). Each row corresponds to the computation of the variable utilized for the matching: Age variable from the study by D'Orazio (2017); Propensity Score derived through Logit Regression and the Genetic Matching (GenMatch) algorithm; Propensity Score derived from Random Forest and the GenMatch algorithm; Propensity Score derived via Boosted CART and the GenMatch algorithm; Summation of the Weights obtained from the GenMatch algorithm. The cells categorized under the "Matching variables" columns signify that the chosen matching variables from D'Orazio (2017) have been employed for the computation of the final computed variables. The cells grouped in the "Common variables" columns denote that all common variables have been employed for the computation of the final computed variables.

Table B10: Results for preserving labour status **marginal distributions**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	1.55	1.44	0.59	1.55	1.44	0.59
PS Logit Regression	0.79	2.36	1.97	2.25	0.51	0.9
PS Random Forest	0.83	1.5	2.05	1.61	1.71	0.94
PS Boosted CART	0.37	1.47	3.63	1.47	2.14	2.64
Sum GenMatch weights	1.14	1.81	1.98	1.02	2.78	0.88

Table B11: Results for preserving the joint distributions between **sex and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	2.42	2.07	1.7	2.42	2.07	1.7
PS Logit Regression	2.75	2.93	2.28	2.49	1.27	2.54
PS Random Forest	2.56	3.5	3.08	2.6	4.09	1.82
PS Boosted CART	2.71	3.25	3.87	1.7	3.85	3.36
Sum GenMatch weights	2.79	2.57	3.1	1.76	3.15	1.26

Table B12: Results for preserving the joint distributions between **age and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	4.14	34	3.38	4.14	34	3.38
PS Logit Regression	4.64	7.51	17.7	25.17	34.06	35.21
PS Random Forest	13.59	32.47	34.36	27.11	32.23	34.97
PS Boosted CART	13.03	30.51	33.28	25.37	32.55	34.65
Sum GenMatch weights	3.49	3.49	4	3.7	4.54	4.04

Table B13: Results for preserving the joint distributions between **education and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	7.9	15.81	6.77	7.9	15.81	6.77
PS Logit Regression	4.58	6.41	7.42	10.14	15.59	16.82
PS Random Forest	6.01	15.36	16.31	11.77	15.49	16.74
PS Boosted CART	6.53	13.58	15.46	10.45	15.88	16.83
Sum GenMatch weights	5.05	7.04	7.24	6.03	8.44	6.58

Table B14: Results for preserving the joint distributions between **area and labour status**

	Matching variables			Common variables		
Procedure	NND	RND	RNK	NND	RND	RNK
D'Orazio (2017)	3.85	3.93	3.91	3.85	3.93	3.91
PS Logit Regression	3.39	5.12	4.05	4.51	4.16	4.17
PS Random Forest	3.57	5.09	4.29	4.43	5.81	3.72
PS Boosted CART	4.08	4	5.43	4.36	5.05	4.92
Sum GenMatch weights	3.68	4.48	4.3	3.87	4.8	3.38

This page is intentionally left blank.

Chapter 3

Intimate Partner Violence and Time Use in Mexico

Abstract

This research investigates whether individuals exhibit differences in time allocation in their routine activities according to their experience of intimate partner violence. Given the absence of a singular survey encompassing all necessary information to address this research question, a statistical matching methodology is employed, merging data from two national representative surveys conducted in Mexico. Leveraging a sophisticated matching variable derived from the Genetic Matching algorithm, two separate synthetic datasets are generated for women and men, comprising all pertinent variables for the analysis. While the overall examination does not reveal specific discernible patterns, noteworthy findings do emerge. For women, differences in time allocation are observed concerning their working hours, on average, as well as the time devoted to childcare and family gatherings. As for men, disparities in intimate partner violence perpetration are detected in the allocation of time across various activities, particularly among those working less than 40 hours per week. These insights shed light on the intricate dynamics between intimate partner violence and individuals' routine activities, may provide valuable implications for understanding and addressing this multifaceted issue.

3.1 Introduction

Intimate Partner Violence (IPV) has emerged as a pressing social and public concern in Mexico and other regions worldwide. According to the 2016 Mexican National Survey on the Dynamics of Household Relationships, nearly 44% of women reported experiencing at least one incident of IPV from a current or former partner in their lifetime. Existing literature indicates gender-based differences in time allocation influenced by social norms (Ahn et al., 2004; Aguiar and Hurst, 2007; Connelly and Kimmel, 2009; Sevilla-Sanz et al., 2010;

Sevilla-Sanz and Fernandez, 2006), with deviations from these norms potentially leading to violence (Jewkes, 2002; Anwary, 2015; Aguiar and Hurst, 2007; Simister, 2013; Yount et al., 2014; Li and Wang, 2022; Svec and Andic, 2018). In Mexican culture, masculine stereotypes depict men as responsible for productive work, head of households, and a degree of control over both public and private spheres. Conversely, women's work is often relegated to domestic chores, remaining invisible and undervalued economically and socially (Inmujeres, 2004; Lamas, 2007; Terrazas-Carrillo and McWhirter, 2015). Consequently, Mexican women face heightened vulnerability to violence when challenging gender norms, such as by engaging in the labour market (Angelucci, 2008; Canedo and Morse, 2021; Gupta et al., 2018) or resisting their husbands' demands (Agoff et al., 2006, 2007).

The analysis presented here aims to address whether routine activities defined by gender roles are related to indicators of intimate partner violence and whether individuals exhibit distinct time allocation behaviours in the presence of violence. The findings presented in Chapter 1 of this thesis established that married Mexican women who engaged in the labour market are more susceptible to experiencing IPV compared to their not-employed counterparts. However, it remains unclear whether this association is solely linked to employment itself, the potential reduction of time dedicated to traditional female activities such as housework or childcare, or a potential absence of leisure time. Furthermore, the allocation of time among perpetrators of violence is a relatively unexplored area. While Chapter 1 found no disparities in employment status as a determinant of violence perpetration among married Mexican men, this paper aims to uncover other unidentified distinctions in the weekly routine activities of perpetrators.

As no single dataset contains information on both time use and the incidents of IPV, this study employs statistical matching tools from D'Orazio et al. (2006b). The data sources for fusion are the 2016 Mexican National Survey on the Dynamics of Household Relationships (ENDIREH 2016) and the 2014 Mexican Time Use Survey (ENUT 2014). Building upon the research contribution in Chapter 2, a leading-edge matching variable derived from the Genetic Matching algorithm by Diamond and Sekhon (2013) is utilized to construct two synthetic datasets: one for women and another for men. Each synthetic dataset incorporates relevant variables associating IPV indicators with various weekly activities, aggregated as work, housework, caregiving, and leisure. The matching performance does not fulfil the Conditional Independence Assumption (CIA), indicating that the independence between IPV and time use variables, given certain common variables, cannot be guaranteed. To complement the validity of the matching, an exploration of uncertainty is presented.

The primary objective of this work is to identify behavioural patterns related to routine activities. Mele (2009) associates lifestyle and daily activities with individuals' vulnerability to victimization, particularly IPV, which the author argues is often repetitive in nature. The empirical strategy of this study involves comparing time allocation for the four main weekly activities in the presence or absence of violence. Moreover, the literature establishes employment status as a robust indicator associated with IPV in Mexico (Angelucci, 2008; Villarreal, 2007; Terrazas-Carrillo and McWhirter, 2015; Canedo and Morse, 2021). Hence, this research analyzes whether individuals exhibit different behaviours depending on the number of hours they work. The main hypothesis posits that increased work hours lead to reduced time allocated to other activities like housework, caregiving, or leisure. However, the presence of IPV may disrupt this time allocation pattern, as seen in Simister (2013).

The results, obtained after accounting for uncertainty, cannot be definitively conclusive for both synthetic samples. Overall, there is no clear indication of specific patterns in individuals' time allocation in the presence of violence, regardless of gender. Nor is there evidence to support the importance of housework time in explaining the existence of violence. However, the findings suggest that female victims of IPV tend to work more hours, on average. Moreover, upon comparing the time allocated to work with the time dedicated to the rest of their activities, the evidence suggests that as victims' working hours increase, they exhibit an inclination to allocate more time to caregiving activities, particularly childcare. On the other hand, women working less than full-time jobs (less than 40 hours a week) and experiencing IPV in the previous year, spend less time in family meetings. These results align with previous findings that consider children (Mele, 2009) and family networks (Lanier and Maume, 2009; Agoff et al., 2007; Tur-Prats, 2021) as key factors influencing the risk or protection against IPV.

As for the men's sample, the results provide varying and unclear insights into identifying perpetrators. The only conclusive interpretation is that men who perpetrated violence in the previous year tend to spend more time in caregiving activities, specifically childcare for children aged 0 to 14 years. The presence of children in the household could potentially escalate intimate conflicts (Mele, 2009), either due to an increased burden of domestic chores or the knowledge that victims may be less likely to leave. The rest of the results lack sufficient clarity to discern the transmission of violence from job stress, as argued in Jewkes (2002) and Fox et al. (2004). Likewise, there is inadequate information to fully understand the effects of alcohol and sports events, which some studies associated with the perpetuation of IPV (Angelucci, 2008; Jewkes, 2002; Card and Dahl, 2011; Dickson et al., 2016). Nonetheless, it is noteworthy that men who work less than 40 hours a week appear to be more sensitive to changes in IPV.

A key contribution of this paper is through exploring novel technologies to address previously unanswerable questions. Existing research already employs statistical matching techniques from D'Orazio et al. (2006b). For example, some studies examine household living conditions by comparing household income and consumption expenditure (Donatiello et al., 2014; Tran and Osier, 2023; Serafino and Tonkin, 2017), quality of life by measuring poverty estimates (Leulescu et al., 2011; Webber and Tonkin, 2013), working conditions (Rezaei Ghahroodi, 2023), and objective and subjective well-being (Leulescu and Agafitei, 2013). Others investigate the informal economy by linking employment characteristics of informal workers (Fernández and Villar, 2017a), or education indicators by combining teachers and student-level data (Iztueta et al., 2017). Most of these studies acknowledge the challenges of finding extensive surveys that include sensitive data, such as the one examined in this research. Thus, this study offers novel approaches to gain deeper insights into the phenomenon of intimate partner violence, potentially leading to more effective policies to address and combat it.

The subsequent sections of the analysis are organized as follows. Section 3.2 reviews existing literature linking gender roles with IPV. Section 3.3 presents the two data sources that will be statistically matched. Section 3.4 provides an overview of the matching procedure, while Section 3.5 describes in detail the matching performance. Section 3.6 presents the results obtained from the synthetic dataset. Section 3.7 concludes by discussing the results and noting the limitations and opportunities for future research.

3.2 Literature Review

3.2.1 Gender Roles and Intimate Partner Violence

A substantial body of literature examines the gender disparities in time use between Europe and the US, with studies by Ahn et al. (2004); Aguiar and Hurst (2007); Connelly and Kimmel (2009); Gimenez-Nadal and Sevilla (2012); Gimenez-Nadal et al. (2012); Sevilla-Sanz and Fernandez (2006); Sevilla-Sanz et al. (2010) shedding light on this topic. In general, men tend to allocate more time to work or leisure activities, while women tend to spend a greater amount of time on housework and childcare responsibilities. These patterns of time allocation are also influenced by marital status and the number of children. Similar findings emerge from research conducted in Mexico, highlighting the influence of gender stereotypes on time distribution between men and women (Romo, 2020; Lamas, 2007; Inmujeres, 2004). Men generally allocate more time to paid work outside the home, often involving longer commutes. Women, regardless of their labour market participation, dedicate significantly more time to unpaid work such as housework and caregiving compared to men. This pattern further intensifies after marriage.

Consequently, it is not surprising that the breaking of these gender roles can lead to conflict and rejection (Anwary, 2015). Resource theories and the male backlash theory interpret this violent rejection as an attempt to restore the husband's control over the empowerment of women (Guarnieri and Rainer, 2021). For instance, existing literature explores how women's employment destabilizes their role within the family and may provoke violence (Gage and Thomas, 2017; Anderberg and Rainer, 2013; Canedo and Morse, 2021). Conversely, other studies suggest that when men work less or remain unemployed, they may resort to violence to reestablish dominance over their partners (Macmillan and Gartner, 1999; Dalal, 2011; Bhalotra et al., 2021b,a; Costa et al., 2016; Tur-Prats, 2021). Chapter 1 of this thesis provides new evidence indicating that employed Mexican women are more susceptible to experiencing IPV, especially when their partners are not employed.

Moreover, existing literature suggests that women may alter their individual preferences in time allocation to avoid further violence. The gender-deviance neutralization hypothesis posits that both men and women adjust their contribution to household tasks based on their adherence to social norms. Bittman et al. (2003) found that working married women reduce their housework time until their earnings equal those of their husbands. Once they begin earning more, they increase their time spent on house chores to compensate for the additional salary. Similarly, Simister (2013) examine data from various countries to investigate whether the gender-deviance neutralization hypothesis holds true under the threat of IPV. Their findings indicate that when women earn significantly more than their husbands, there is an increase in the husbands' alcohol consumption, leading to a higher risk of gender-based violence. In response, women may assume greater responsibility for housework, as decreasing bargaining power to diminish their husbands' violent tendencies.

In a similar line of research,Yount et al. (2014) evaluates the association between market work, subsistence work, and non-economic work (house production and care work) with the exposure to lifetime, recent, and chronic physical and sexual IPV in Egypt. Among their results, they showed that women exposed to IPV in the past spend more time in domestic work. Market and subsistent work increased the chances of recent IPV, while care work didn't bring any significant changes. In addition, Li and Wang (2022) studies the relationship between remunerated and non-remunerated work in rural China finding that similar or higher relative wages increased wives' psychological and physical violence that reduced when there was an equal distribution of the housework. In fact, when husbands contributed more to housework, the odds of psychological violence decreased.

Interestingly, the association between leisure time and IPV has not been extensively studied. Fox et al. (2004) and Jewkes (2002) suggest that an excess of work and lack of time for oneself can elevate the risk of partner conflict. Consequently, it can be assumed that the more leisure time a couple enjoys, the less violence may occur in the household. However, certain activities during leisure time have been linked to an increased likelihood of IPV, such as alcohol consumption and sports mass media events. Drinking alcohol is a prevalent social activity in Mexico, often associated with the reinforcement of specific hegemonic masculinities (Ibarraran-Bigalondo, 2020). Numerous studies have linked alcohol consumption to a rise in gender-based violence (Foran and O'Leary, 2008; Jewkes, 2002; Simister, 2013; Angelucci, 2008; Orozco et al., 2012). Angelucci (2008) provides evidence of the connection between increased income transfers, reduced alcohol consumption, and a subsequent decline in physical abuse in Mexico.

Furthermore, Card and Dahl (2011) analyze police reports of violent incidents and establish a link between domestic violence and the emotional cues associated with unexpected losses of football teams. Similarly, Dickson et al. (2016) find that football derby matches of Celtics against Rangers in Glasgow increase the chances of home violence reports, irrespective of the expected outcome of the match. In Mexico, apart from football, other popular sports activities like boxing and wrestling involve a significant amount of violence (Allen, 2017; Glenday, 2013). Although there is no direct evidence of how these activities relate to gender-based violence, Melzer (2002) reveals that men engaged in activities with high levels of physical violence are also more prone to being physically violent at home.

3.3 Data

This paper utilizes two survey datasets: the 2016 Mexican National Survey on the Dynamics of Household Relationships (Encuesta Nacional Sobre la Dinámica de las Relaciones en los Hogares, ENIDREH) and the 2014 Mexican Time Use Survey (Encuesta Nacional sobre Uso del Tiempo, ENUT). The ENDIREH 2016 survey specializes in collecting information about women's experiences of violence in various spheres, including school, work, community, family, and intimate relationships. Interviewers select one woman aged 15 years or older from each household to conduct face-to-face interviews, ensuring a safe and secure environment for the respondents. This survey also includes data on household and demographic characteristics for each member of the household, enabling the identification of demographic characteristics of husbands and the occurrence of intimate partner violence among married women.

The ENUT 2014 survey offers valuable insights into the time use of men and women aged 12 and over during a week. Its primary objective is to comprehensively measure all forms of work individuals engage in, both paid and unpaid, with a focus on highlighting the significance of domestic production and its contribution to the economy. For this purpose, every individual in each household reports the number of hours and minutes they dedicate to specific activities. The survey captures a wide range of activities, including paid work, household chores, caregiving, leisure, and other engagements. Since individuals may simultaneously participate in activities such as housework and family care, the reported variables do not sum to a total of 168 hours per week (24 hours per day for seven days).

Both ENDIREH 2016 and ENUT 2014 surveys are nationally representative and are administered by the National Institute of Statistics and Geography in Mexico (INEGI). The harmonization and comparison of both surveys are facilitated by INEGI which provides the same standards and methodologies. The sampling unit in both surveys is the household, using the same sampling frame: the INEGI National Housing Framework 2012, derived from cartographic and demographic information collected during the 2010 Population and Housing Census. Both surveys employ a two-stage simple random sampling design, where the primary sampling unit (PSU) is the geographic location, and the second is the households.

The selected individuals from both surveys are men and women aged between 15 and 64 years old, married or living in a couple, who have provided responses to all key variables relevant to the research. The total sample consists of 46,033 married couples of men and women from the ENDIREH survey and 8,370 couples from the ENUT. Table 3.1 presents the sample sizes and the population represented. Although there is a two-year gap between the two surveys, the populations can be considered similar, as argued by Masterson (2011, 2010); Rios-Avila (2014, 2016, 2018), who advocate using different survey years for matching as long as both surveys maintain some level of similarities.

Table 3.1: Comparison between ENDIREH 2016 and ENUT 2014 by sample size and population size

	Household Level		Individual Level		Sample Level	
	ENDIREH	ENUT	ENDIREH	ENUT	ENDIREH	ENUT
Sample Size	121,999	15,059	451,548	56,274	46,033	8,370
Population Size	33,002,969	32,130,805	122,746,883	120,205,418	24,853,462	35,772,654

Note: This table presents the sample and population size at the Household level, Individual level, and Selected Matching sample level for ENDIREH 2106 and ENUT 2014 respectively. Despite the Population size differences, they still preserve some level some level of similarity (Masterson, 2010, 2011; Rios-Avila, 2014, 2016, 2018).

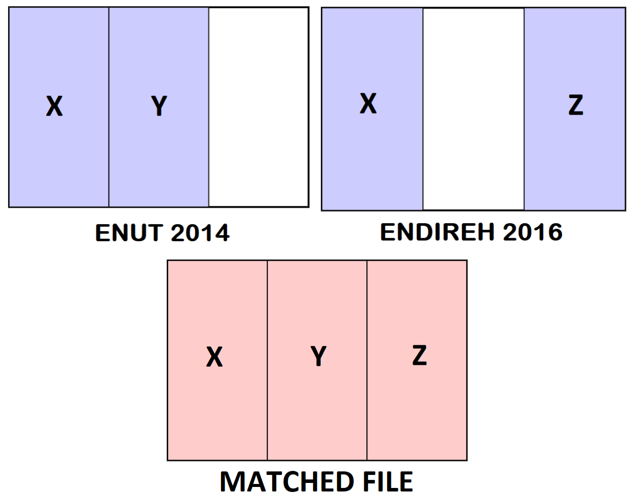
In this research, the survey samples are divided into two subsets, one for women and the other for men. This division enables separate analyses of the characteristics related to women's likelihood of becoming victims based on their activities and men's likelihood of becoming perpetrators based on their engagement in specific activities.

3.4 Statistical Matching: Overview of the Methodology

In this analysis, the primary objective is to conduct a data matching process between the ENDIREH and ENUT surveys, as depicted in Figure 3.1. Embracing a micro-approach, the aim is to construct a novel synthetic dataset that incorporates both target and common variables (X, Y, Z). Each survey will encompass specific target variables, denoted as those variables solely observed within their respective surveys. Specifically, the ENDIREH survey's target variable pertains to intimate partner violence (IPV), while the ENUT survey focuses on the use of time activities. Moreover, these surveys share a subset of variables that are common across both datasets. To facilitate the matching procedure effectively, we designate the ENUT survey as the recipient dataset due to its relatively lower number of observations, while the ENDIREH survey will serve as the donor dataset.

To achieve the matching, I follow the steps described in Chapter 2.1. First, I harmonize and redefine the common variables present in both surveys to ensure consistency and compatibility. Next, I select the matching variables, which must meet two key criteria: they should exhibit good similarities between the surveys and be predictive of at least one of the target variables. Once the matching variables are identified, I proceed with a two-step procedure for the matching process. I combine matching methods presented in D'Orazio et al. (2006b) with the sum of GenMatch weights from Diamond and Sekhon (2013), as introduced in the previous chapter. Non-parametric linking methods such as the random hot deck, nearest neighbour distance hot deck, and rank hot deck will be utilized, along with the matching variable derived from the sum of the GenMatch weights.

Figure 3.1: Statistical Matching



Note: Graphic representation of the Statistical Matching procedure

Some matching variables can be selected to become strata variables, dividing the samples into subsets for more targeted matching. This ensures that, for example, individuals living in rural areas will not be matched with individuals from urban areas, enhancing the quality of the matching. After performing the matching, I will assess the quality of the synthetic dataset and select the optimal one based on a dissimilarity measuring. However, due to the unavailability of evidence to support the Conditional Independence Assumption (CIA), I will incorporate an evaluation of uncertainty using Frechet bounds.

3.5 Statistical Matching: Application

3.5.1 Identification of the Target Variables

Intimate Partner Violence. The ENDIREH 2016 survey comprises 36 questions that address any experiences of violence occurring in the previous year. The questions cover the main categories of Intimate Partner Violence (IPV), encompassing physical, emotional, sexual, and economic violence. Respondents are asked about incidents such as being pushed, slapped, humiliated, forced into sexual intercourse, or having money spent without consent, representing each category of violence. The respondents indicate the frequency of each violent tactic as "one time, few times, several times, or never," enabling the measurement of violent occurrences. For this analysis, we define the Intimate Partner Violence variable as a binary dummy variable, assigning a value of 1 for respondents who have experienced any type of violence at least once, and 0 for those who have not. Consequently, women who respond affirmatively to any of the 36 questions are identified as victims of IPV, and their partners as perpetrators. Statistics from the survey reveal that 29.16% of households experienced some form of violence during the previous year.

Time Use Variables. The ENUT 2014 survey records the duration, in hours and minutes, that individuals in the household dedicate to various activities, which can be categorized as Work, Housework, Care, and Leisure. The ***Work*** variable includes time spent on work-related activities and commuting. ***Housework*** accounts for the time allocated to tasks involved in household consumption, food preparation, housecleaning, laundry, maintenance, shopping, payments, and household management. ***Care*** encompasses activities related to the care of dependent relatives, children of different age groups (from 0 to 5 years old, and from 0 to 14 years old) and other family members (from 15 to 59 years old, and relatives 60 years old or more). Lastly, the ***Leisure*** variable encompasses time spent on recreational activities, sports, games, social gatherings, and media usage.

It is important to note that as these variables are based on self-reported beliefs about time allocation, simultaneous performance of different activities, such as cleaning while caring for a child, cannot be captured. Moreover, frequently, the sum of total hours is decompensated and does not sum up to 168. To address this, a readjustment of the time allocation is performed to ensure that the total hours do not exceed 168 hours per week, following the approach of Pacheco and Florez (2014) and Romo (2020). Equation (3.1) shows the formula used in which $T'\lambda_i$ is the adjusted allocation of time, $T\lambda_i$ is the recorded one, and $\sum_l T\lambda_l$ is the total sum of activities.

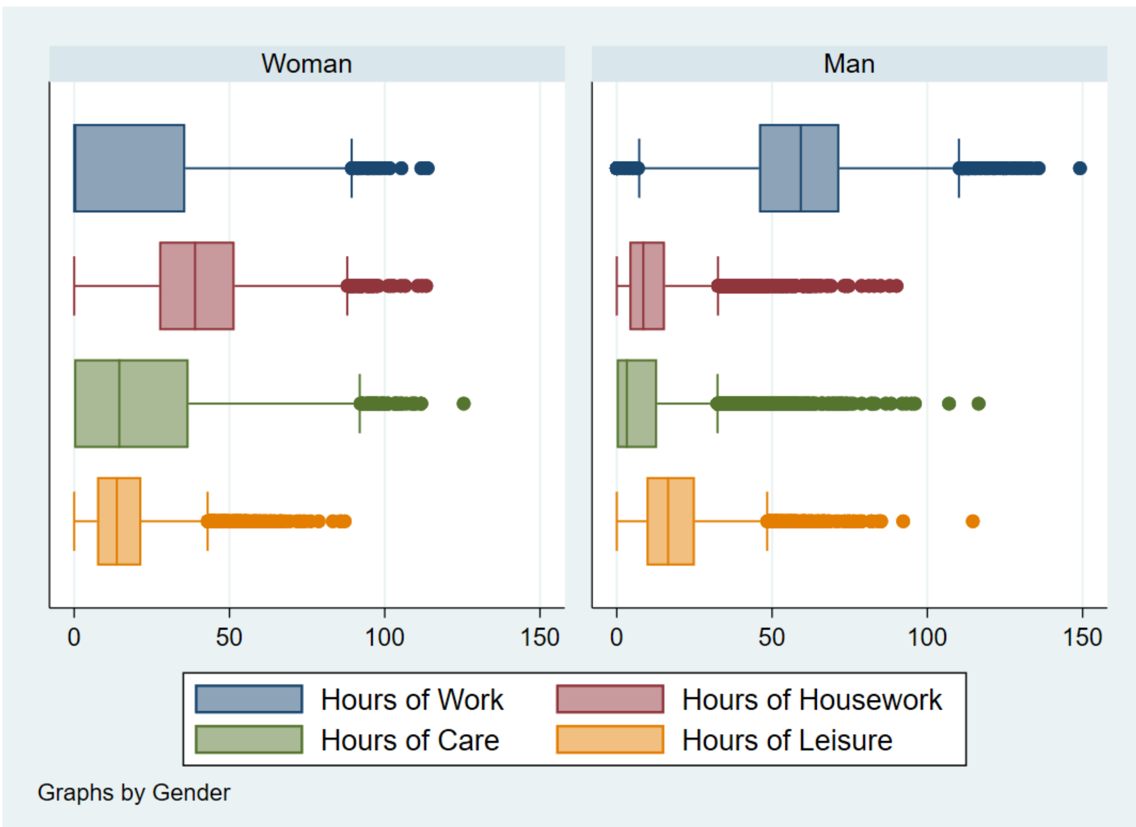
$$T'\lambda_i = T\lambda_i \times \frac{168}{\sum_l T\lambda_l} \tag{3.1}$$

Figure 3.2 displays the distribution of the four main activities by gender, revealing clear gender disparities in time allocation among the Mexican population. Women exhibit a median of 0 working hours, while men display a median of more than 50 working hours per week. In contrast, women allocate more time, on average, to housework and caregiving activities compared to men. Notably, no significant gender differences are observed in leisure time allocation. These findings underscore the existence of distinct task specifications influenced by social norms, as highlighted in previous reports by Romo (2020).

3.5.2 Common variables: Harmonization and reconciliation

To ensure the effectiveness of the statistical matching process, careful selection of common variables is essential. These variables must be comparable in their definitions and distributions. Fortunately, the Mexican Institute of Statistics (INEGI) maintains consistent criteria for many common questions across its surveys over time, providing an advantage for the matching process. Table 3.2 presents the selected common variables for this matching, organized into categories such as individual characteristics, household structure, housing condition, and geographic characteristics. The selection of these variables involved a rigorous comparison of their definitions (as shown in Table C1) and a harmonization process to ensure homogeneity (D'Orazio et al., 2006b). Importantly, the same set of common variables was used for both the sample of women and the sample

Figure 3.2: Gender differences in the allocation of time



Note: Boxplot distribution of the four main activities (Work, Housework, Care, and Leisure) by gender. The box spans the interquartile range (25^{th} percentile to 75^{th} percentile), with the hinge of the box representing the median value. The whiskers extend to 1.5 times the interquartile range, leaving the rest of the observations plotted in individual dots. Source: ENUT 2014.

of men in this analysis. This uniform approach facilitates a robust and consistent matching procedure.

Table 3.2: Classification of Common of Variables

Common Variables	
Individual Characteristics	Indigenous, Head of the household, Age (grouped), Marital Status, Level of education (grouped), Employment status, Student, Retired, Homemaker, Partner employment, Partner indigenous, Age gap (grouped), Education gap (grouped)
Household Structure	Children under 18, Children under 16, Children under 12, Children under 5
Housing Condition	Number of Bedrooms (grouped)
Geographic Characteristics	State, Rural

Note: Identification of the common variables in both surveys ENDIREH 2016 and ENUT 2014

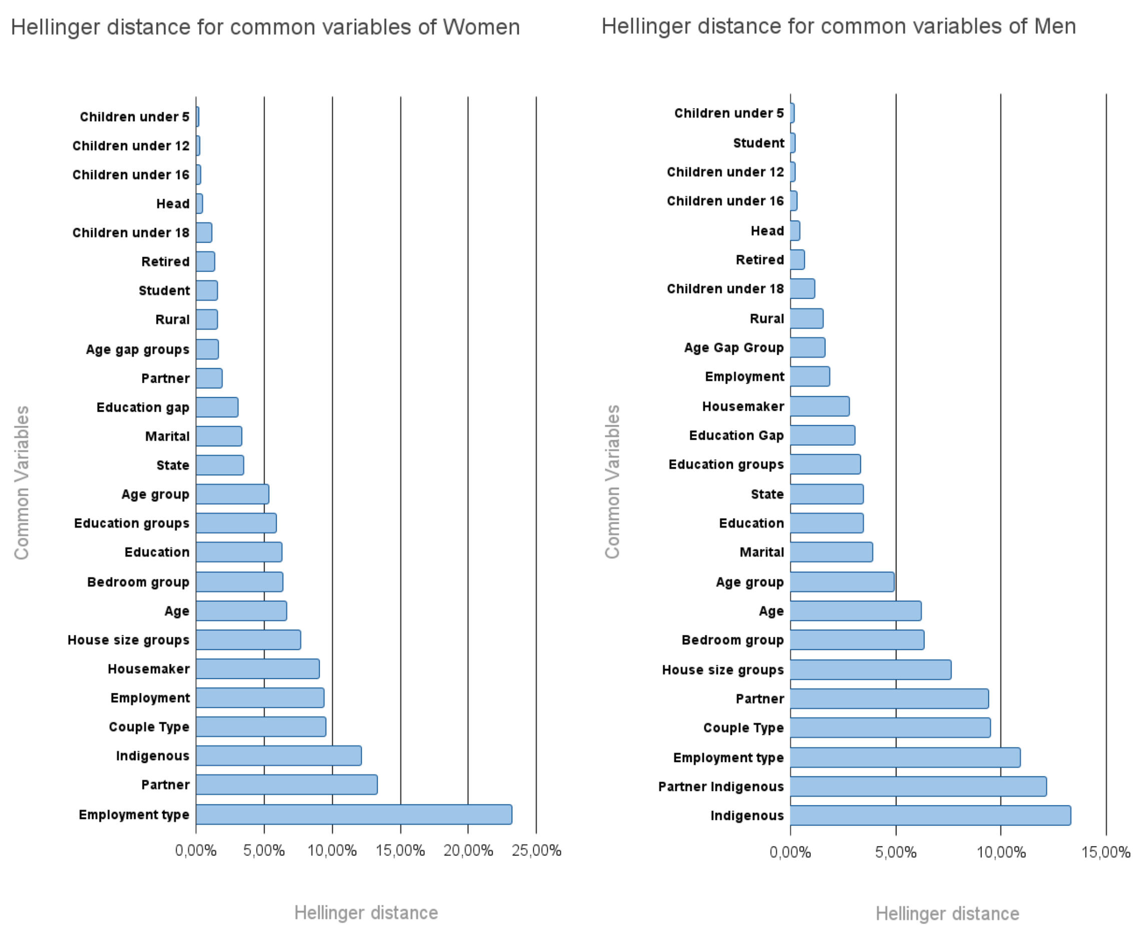
3.5.3 Selection of matching variables

In the matching process, not all common variables are selected as matching variables (X_M). Instead, the matching variables are derived from the combination of the best predictors for the target variables Y and Z, represented as X_Y and X_Z respectively. This selection follows the principle that X_M lies between the intersection and the union of X_Y and X_Z, denoted as $X_Y \cap X_Z \subseteq X_M \subseteq X_Y \cup X_Z$, with the additional requirement of similarities among their distributions. Therefore, a common variable is considered a matching variable if it fulfils two main conditions: firstly, the distribution of the variable in both datasets should exhibit similarity, which can be quantified using the Hellinger distance. Secondly, the variable must possess a certain level of explanatory power for the target variables, which varies depending on the nature of these target variables. This approach ensures that only relevant and informative common variables are used in the matching process, enhancing the accuracy and validity of the synthetic dataset construction.

3.5.3.1 Hellinger Distance

The Hellinger distance is a widely accepted measure in the literature to quantify the similarities between distributions (D'Orazio et al., 2006b; D'Orazio, 2017; Donatiello et al., 2014, 2016; Leulescu et al., 2011; Leulescu and Agafitei, 2013; Webber and Tonkin, 2013; Serafino and Tonkin, 2017; Iztueta et al., 2017; Rezaei Ghahroodi, 2023). According to the established rule of thumb, distributions with a Hellinger distance lower than 5% are considered similar, those between 5% and 10% may be ambiguous, and distances larger than 10% are regarded as different. In this analysis, Figure 3.3 illustrates the Hellinger distances for the two separate datasets of men and women. For the women's data sample, variables such as *Children under 5, 12, 16 and 18 years old, whether the head of the household is female if they live in rural areas, the state where they live*, and *the education gap* exhibit Hellinger distances below the 5% threshold, indicating similarity. Variables *Age group* and *level of education* are slightly above the threshold but can still be considered for analysis if they prove to be good predictors of the target variables. However, variables *whether they considered themselves or their partner indigenous or not* and *type of employment* should not be included as matching variables due to their significant discrepancy. Similar variables are excluded from the sample of men.

Figure 3.3: Hellinger distance of the common variables



Note: Graphic representation for men and women of the Hellinger distance of the common variables between ENDIREH 2016 and ENUT 2014. The variables are ordered from lower distance to higher.

3.5.3.2 Best Predictors of the Target Variables

The second criterion for selecting the matching variables involves determining which common variables serve as good predictors for at least one of the target variables. Figure C2 and C3 in Appendix Section 3.C.3 present the prediction results, summarized as follows:

Predictors of Intimate Partner Violence (IPV) variable: As IPV is a dichotomous variable, the selection of the best predictors is based on Chi-square-based association measures (Cramer's V) or proportional reduction of the variance measures. Figure C2 displays the best predictors for the various IPV variables in the case of women. Common variables such as *state, household size, rural status, age group, presence of children under 5, 12, 16, and 18, education groups, employment status,* and *homemaker status* emerge as the most significant predictors for IPV. Similarly, for men, the best predictor variables include *state, household size, rural status, age group, presence of children under 5, 12, 16, and 18, education groups,* and *partner's employment status*.

Predictors of Time Use variables: As the variables of Work, Housework, Housecare, Leisure, Other, and Sleeping are continuous, the adjusted R^2 is computed for each couple response predictor. Figure C3 presents the results of the best predictors. For women, the variables *household size, age group, presence of children under the age of 5, 12, 16, and 18, employment status,* and *homemaker status* are found to be the most influential predictors for the allocation of hours in working, housework, care, and leisure activities. In the case of men, the best predictor variables are *household size, age group, presence of children under 5, 12, 16, and 18, education groups,* and *employment status*.

3.5.4 Matching

3.5.4.1 Genetic Matching Algorithm

In this study, I extend the matching procedure proposed by D'Orazio et al. (2006b) by incorporating an additional step using the sum of the matching variables' weights obtained from the Genetic Matching (GenMatch) algorithm. The GenMatch algorithm is an iterative process that searches for a weight matrix of the common variables to minimize the matching distance between the two datasets. Once the weight matrix ($\omega_{k,k}$) is derived from GenMatch, I construct a new value, denoted as SW_i, by summing the products of the weight value ($\omega_{k,k}$) with the value of each common variable (X_{ki}), as shown in Equation (3.2)[1]. The results derived in Chapter 2 demonstrated the efficacy of this technique in preserving both the marginal distribution of the target variable and the joint distribution with the matching variables.

$$SW_i = \sum_{k=1}^{K} \omega_{k,k} X_{ki} \tag{3.2}$$

The benefit of using the Sum of the Weights values lies in its ability to consolidate multiple pieces of information into a single value, enhancing the efficiency of the matching procedure compared to the approach by D'Orazio et al. (2006b). However, it is important to consider that the GenMatch algorithm may face computational limitations when dealing with large datasets (Sizemore and Alkurdi, 2019). To mitigate this, it is crucial to use an optimal number of variables, and the selected matching variables that are either continuous or should not have an excessive number of categories. For instance, among the variables indicating whether children under 5, 12, 16, and 18 are present in the household (assumed to provide similar information), I select the one with the highest explanatory power on both Intimate Partner Violence (IPV)[2] and time allocation[3], which is

[1] For more detailed information about the Genetic Matching algorithm and the construction of the sum of weights, consult Chapter 2 of this thesis.

[2] See Figure C2 (a) and (b) in Appendix 3.C.3.1

[3] See Figure C3 (a) and (b) in Appendix 3.C.3.2

children under 12. Moreover, although the household size and state demonstrate good predictive ability for IPV, their 4 and 32 categories, respectively, could significantly increase the computational cost. Consequently, for practicality, I restrict the final set of matching variables used in the GenMatch algorithm to a total of 6 variables for each implementation. In particular, within this selection, only one variable is continuous (the variable age).

Table 3.3: Matching Variables

Women		Men	
Variable	Weight (ω)	Variable	Weight (ω)
Homemaker	696.77	Partner Employment	847.45
Age	578.24	Employment	740.22
Children under 12	576.06	Children under 12	560.56
Education level	537.98	Rural	408.98
Rural	358.65	Education level	275.77
Employment	253.77	Age	9.93

Note: Resulting weights from the Genetic Matching (GenMatch) Algorithm in descendent order of relevance.

Table 3.3 provides an overview of the selected matching variables along with their corresponding weights obtained from the GenMatch algorithm. The algorithm's implementation for the women's sample assigns greater relevance to the homemaker activity, followed by the age and having children under 12 years old. Additionally, within the men's sample, the algorithm prioritizes the importance of their spouse's employment, their own employment and the presence of children. This indicates that both the employment status and the presence of children are commonly influential determinant variables across both the men's and women's samples. However, it should be underlined that for men, the variable age obtained a considerably low weight (contrary to women), indicating its relatively diminished significance in comparison to other variables. This suggests that if an attempt were made to conduct the matching as performed in D'Orazio (2017) using the unique continuous matching variable available age (baseline in Chapter 2), the preservation of the distributions in the resulting men's synthetic dataset could potentially deteriorate. Hence, the adoption of GenMatch weights stands as an advancement compared to the initial Statistical Matching framework (D'Orazio et al., 2006b), despite its inherent limitations. It is noteworthy that the GenMatch algorithm took a total of 11 days to compute the weights for both samples.

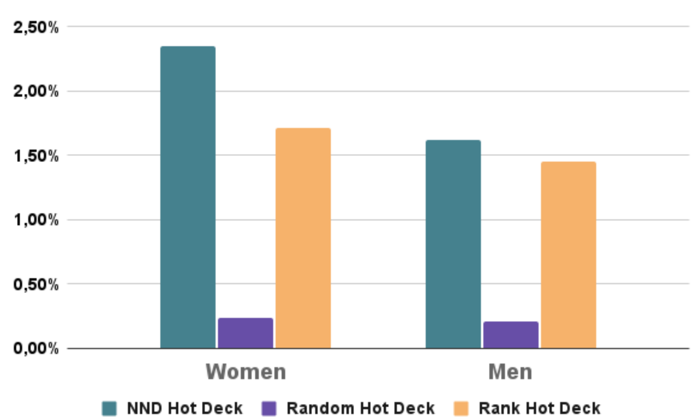
3.5.4.2 Hot Deck Methods

Upon constructing the matching variable, the statistical matching is finalized using three non-parametric procedures proposed by D'Orazio et al. (2006b): Nearest neighbour distance hot deck, Random hot deck, and Rank hot deck. The hot deck methods aim to impute values in the recipient file (ENUT 2014 in this study) using information from the donor file (ENDIREH 2016) based on the similarities observed in the matching variables. The choice of the specific hot deck method is determined by the metric used to measure the similarities among the matching variables, which distinguishes the three procedures[4]. In this research, I perform the matching using all three hot deck methods and subsequently select the one that yields the most favourable qualitative results for both the sample of women and the sample of men. Additionally, to enhance the matching process for each sample, I designate two strata variables[5]: *children under 12 years old* and the indicator of *individuals living in rural or urban areas*. These variables are chosen based on their lower Hellinger distance and their explanatory power with respect to the target variables in both cases (as depicted in Figure 3.3 and Figures C2, C3).

3.5.5 Quality Evaluation and Selection of the Synthetic Datasets

To assess the quality of the matching, several validity checks are conducted. The first check involves comparing the similarity of the distributions between the target variable IPV from the donor dataset (ENDIREH 2016) and the synthetic dataset. Following established practices in the literature, the Hellinger distance measurement is utilized, as illustrated in Figure 3.4. The results indicate that all three hot deck methodologies performed well, as the Hellinger distances for both samples are below 5%. Particularly, the random hot deck method yielded the most favourable results, with a Hellinger distance of less than 0.5%.

Figure 3.4: Hellinger distance of marginal distribution of IPV



Note: Hellinger distance of the marginal distribution of target variable IPV for each hot deck methodology for the sample of women and men, separately.

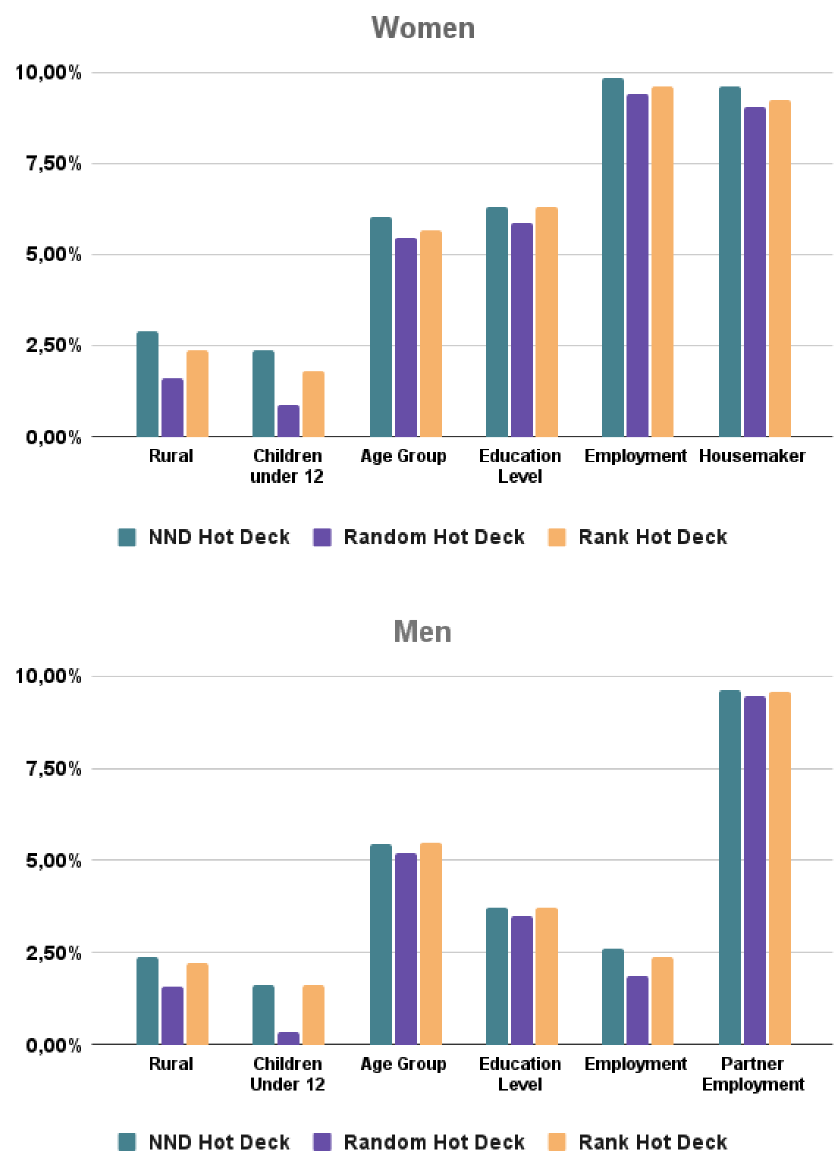
To further evaluate the matching, the dissimilarities among the joint distributions of the target variable IPV with the matching variables are examined, as shown in Figure 3.5. For the sample of women, the joint distribution with the strata variables *rural* and *children under 12* demonstrates good preservation, while the remaining matching variables display Hellinger distances exceeding 5% but remaining below 10%, which is considered tolerable. In the case of the sample of men, the matching process overall better preserves the joint distributions of the variables *rural, children under 12, education level*, and *employment*. However, *age group* and *partner's*

[4]For a comprehensive understanding of the hot deck methods and their distinctions, refer to Chapter 2 Part 2.4.6.

[5]These variables partition the dataset into distinct clusters.

employment status exhibit dissimilarities ranging from 5 to 10%. Therefore, based on the results of the validity checks, the random hot deck method is determined to be the best performer overall for both women's and men's matchings. Consequently, this method is selected for generating the synthetic datasets.

Figure 3.5: Hellinger distance of the joint distribution between IPV and matching variables



Note: Hellinger distance for the joint distribution for each sample of women and men separately.

3.5.5.1 Uncertainty Assumption

The theoretical proposition states that the Conditional Independence Assumption (CIA) holds if the target variables - Intimate Partner Violence (IPV) and Time Use variables - are independent when conditioned on the chosen matching variables. In a practical application, upholding the CIA is feasible if at least one of the matching variables closely resembles either of the target variables. However, in this context, none of the selected matching variables can be considered a proxy for either IPV or Time Use variables. Therefore, given the impossibility of verifying whether the CIA holds, two solutions proposed by D'Orazio et al. (2006b) are employed to gain insights into the joint distribution of the target variables (IPV and Time Use) and the matching variables. The first solution involves creating an Auxiliary dataset that includes all variables, especially the target variables, to obtain information about their joint distribution. However, in this specific analysis, no alternative survey containing both IPV and the allocation of hours to daily activities is available, rendering the use of auxiliary information infeasible for evaluating the matching.

The second solution to address the limitation of the CIA involves deriving intervals of uncertainty based on the available information. This approach involves constructing the probability distribution of the contingency table YZX using the Fréchet bounds. Equation (3.3) defines the Fréchet bounds, where $P_{Y|X}$ denotes the probability of Y (Time Use variables) conditional on X (Common variables), and $P_{Z|X}$ represents the probability of Z (IPV) conditional on X (Common variables).

$$max(0; P_{Y|X} + P_{Z|X} - 1) \leq P_{YZ|X} \leq min(P_{Y|X}, P_{Z|X}) \tag{3.3}$$

To implement the Fréchet bounds method, the first step is to harmonize the weights of the two samples using the survey weights of each dataset (Renssen, 1998; D'Orazio et al., 2006b; Donatiello et al., 2014). The weights are harmonized with the joint distribution of all matching variables, including *state, household size, rural, age, children under 12, education level, employment status* (for women and men), *homemaker* (only for women), and *partner employment status* (only for men). In the second step, the time use variables for women and men are categorized following economic criteria and aiming to achieve uniform distributions as much as possible. Table C3 show the distribution of the different categories[6].

The resulting contingency tables with the lower and upper Fréchet bounds are presented in Appendix Tables C4 for women and C5 for men. These tables provide interval values that associate IPV with specific Time Use activities. Therefore, the column of the CIA represents the contingency table of IPV and the specific Time Use activity, and the Lower and Upper bound are the extreme values. For example, in Table C4, the frequency of the cell corresponding to women who work 0 hours per week ($Work = 0$) and do not experience IPV ($IPV = 0$) should lie between 38.19% and 38.31%, with the CIA estimate at 38.24%. Tables C4 and C5 also reveal the average width of the uncertainty bounds, which, except for Women's working time, appears wider than 5% (7%, on average).

[6]The categorization follows the economic criteria and tries to achieve uniform distributions as much as possible, thus avoiding cells or categories with little or none observations.

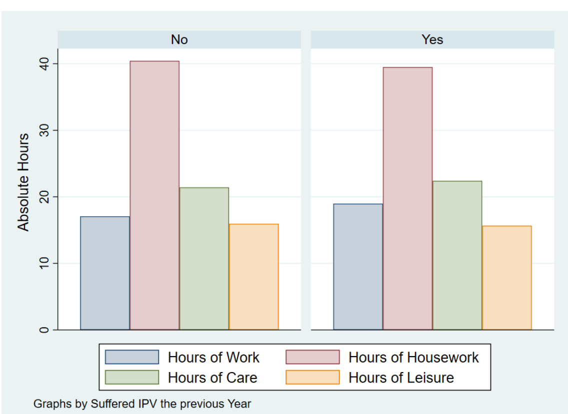
3.6 Results from the Synthetic Dataset

In a general context, when the Conditional Independence Assumption (CIA) holds, it ensures the robustness and reliability of the outcomes within the synthetic dataset, encompassing the associated confidence intervals (D'Orazio et al., 2006b). However, the subsequent results do not incorporate the customary confidence intervals, nor do they attempt to allude to the significance of mean disparities. This is attributed to the prevailing uncertainty linked with most of the Fréchet bounds presented in Appendix Tables C4 and C5.

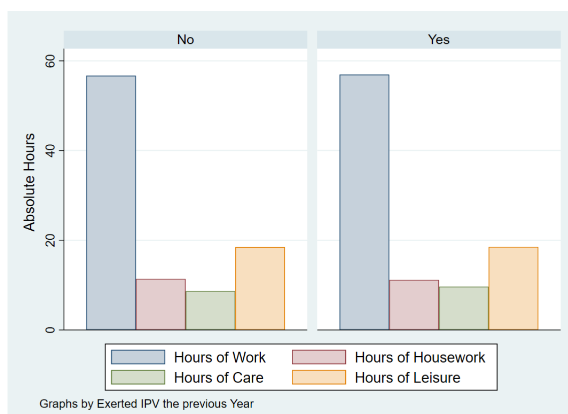
3.6.1 Basic Means

Figure 3.6 presents a comparison of the average hours dedicated to work, housework, care, and leisure for women and men in cases where there is intimate partner violence (IPV) and when there is none. Upon examination, there are minimal observable differences in activity allocation for both women and men. Among women who experience IPV, there is a slightly higher average time spent on work and caregiving activities, while slightly less time is dedicated to housework. For men, the only discernible distinction is that those who perpetrate violence allocate slightly more time to caregiving tasks. However, overall, the variations in time allocation between those experiencing IPV and those who do not, are not particularly pronounced for either gender.

Figure 3.6: Mean of Hour by IPV



(a) Women



(b) Men

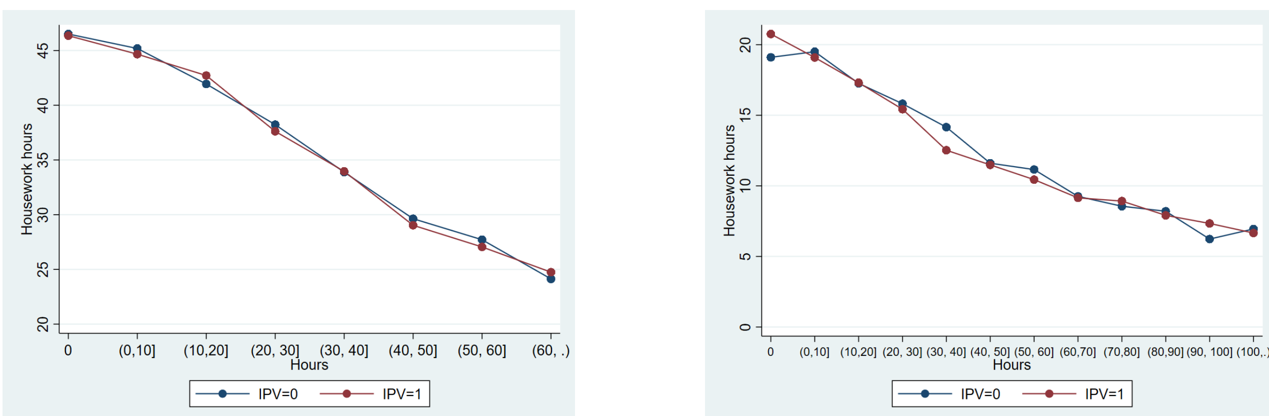
Note: Mean of hours of work, housework, care, and leisure for women and men, comparing if there was IPV the previous year or not. Source: (a) Synthetic Dataset for Women and (b) Synthetic Dataset for Men.

3.6.2 Differences in Time Use by Working Hours

The subsequent step in the analysis involves examining whether differences exist in the relationship between the number of hours spent in paid employment and the allocation of time for other activities, specifically, whether such differences are associated with the presence of IPV. Previous literature, as discussed in Chapter 1 of this thesis, has indicated that the employment status of Mexican women is linked to IPV. Thus, this analysis aims to explore whether the number of hours spent in paid work affects women and men differently in terms of time allocation for other weekly activities and whether these differences also relate to the occurrence of IPV.

To investigate this, I calculate the mean hours spent on housework, care activities, and leisure for every 10 hours spent working, separately for women and men. Figure 3.7 illustrates the association between hours spent on housework and hours spent on work for both genders. The findings show that as the number of hours spent working increases, the time dedicated to housework decreases for both women and men. However, it is noteworthy that women who work an average of 60 hours or more per week still spend more time on housework than men who do not work. Despite these observations, no significant differences in the occurrence of IPV can be identified for either women or men based on the presented data.

Figure 3.7: Mean of Housework hours by Amount of working hours

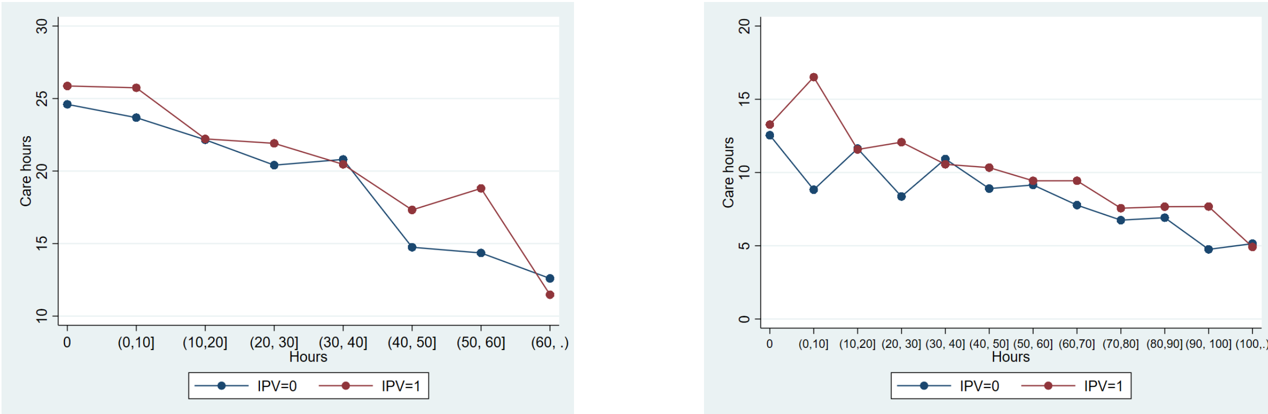


(a) Women (b) Men

Note: The horizontal axis measures the hours of work in intervals of 10 hours. The vertical axis measures the mean of housework hours dedicated to each interval of working hours. Source: (a) Synthetic Dataset for Women and (b) Synthetic Dataset for Men.

In Figure 3.8, the correlations between the hours spent working and the hours dedicated to caring for another relative in the household are depicted. In this analysis, clear differences are observed in the time allocation between situations with and without IPV. Generally, despite some fluctuations, women who are victims of IPV tend to dedicate more time to caring activities compared to those who are not victims. Similarly, men who perpetrate IPV also allocate more hours to care activities than those who do not engage in such behaviour. These findings suggest that the presence of IPV is associated with differences in the time spent on caring activities for both women and men.

Figure 3.8: Mean of Care hours by working hours

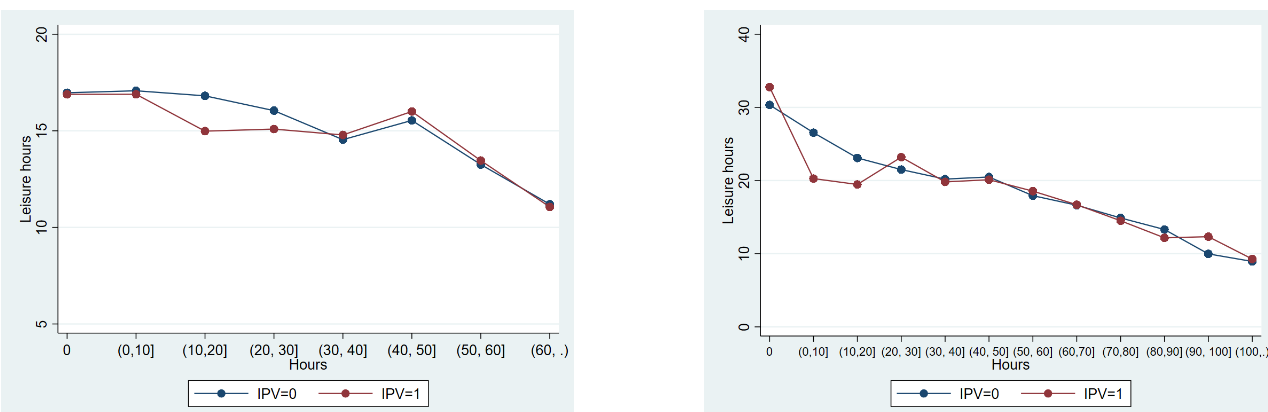


(a) Women (b) Men

Note: The horizontal axis measures the hours of work in intervals of 10 hours. The vertical axis measures the mean of care hours dedicated to each interval of working hours. Source: (a) Synthetic Dataset for Women and (b) Synthetic Dataset for Men.

Figure 3.9 presents a comparison of the time spent in leisure activities with the time spent working, focusing on the differences in IPV status. For women victims of IPV who work between 10 and 30 hours per week, the data shows that they spend less time in leisure activities compared to women who are not victims. A similar pattern is observed in the case of men who perpetrate IPV. Specifically, among men who are actively employed but work less than 20 hours per week, those who exert violence tend to have less leisure time compared to men who do not engage in violent behaviours.

Figure 3.9: Mean of Leisure hours by working hours



(a) Women (b) Men

Note: The horizontal axis measures the hours of work in intervals of 10 hours. The vertical axis measures the mean of leisure hours dedicated to each interval of working hours. Source: (a) Synthetic Dataset for Women and (b) Synthetic Dataset for Men.

Figures 3.10 and 3.11 provide a detailed analysis of the differences in correlations between the most relevant activities within each category and the hours of work, focusing on IPV status for each gender. Among the selected activities were hours dedicated to food production, cleaning the house, childcare for individuals aged

0 to 14 years old, and time spent in family gatherings, mass media consumption, or physical exercise. These activities have been selected as being the most representative of each aggregated activity as shown in Figure 3.2.

For women, there is a similar pattern in the hours spent on food production and cleaning as observed in the overall time spent on housework. In the case of men, there is greater fluctuation in the time spent cleaning for those who perpetrate violence. Specifically, men who work 20 to 40 hours a week and engage in more violent behaviours spend less time cleaning compared to those who are less violent, while those who work 10 hours or less per week and exert violence spend almost an hour more on cleaning than less violent counterparts. No significant differences in IPV were observed for men working more than 40 hours.

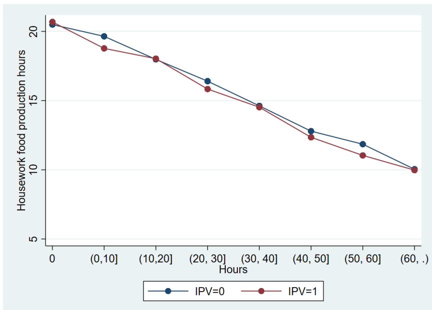
Regarding childcare for children under 14 years old, both men and women exhibit similar behaviour as they do in the overall caregiving time (Figure 3.8). Overall, despite the fluctuations, those experiencing IPV spend more time in childcare compared to those who do not.

Finally, further analysis compares the time spent on work with common leisure activities: family gatherings, mass media consumption, and physical exercise. For women victims of IPV who work 30 hours or less per week, fewer hours were dedicated to activities involving reuniting with family and friends (family gatherings). Similarly, men who perpetrate violence and work 20 hours or fewer per week also spent less time participating in family gatherings. Regarding mass media consumption, women working between 10 and 40 hours per week tend to consume less media if they are victims of IPV. However, once they work more than 40 hours, they tend to consume more media compared to those who did not experience IPV. Men, on the other hand, showed more consistent behaviour in mass media consumption regardless of IPV status. Finally, the analysis of the time spent on physical exercise revealed that men who either barely work 20 hours a week or work more than 80 hours a week tend to spend more time on physical activities if they have harmed their spouses. However, no clear pattern of differences was evident for women in relation to physical exercise.

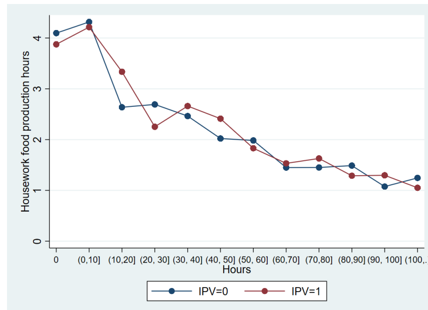
Overall, these findings highlight the complex relationships between working hours, leisure activities, and IPV status for both women and men. The results suggest that employment and time allocation may play a role in the dynamics of IPV, but additional research is needed to fully understand the underlying factors driving these potential patterns.

Figure 3.10: Mean of hours allocated in different activities by working hours a week

(a) Hours of Food Production

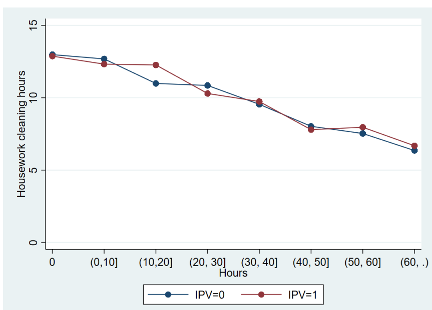


Women

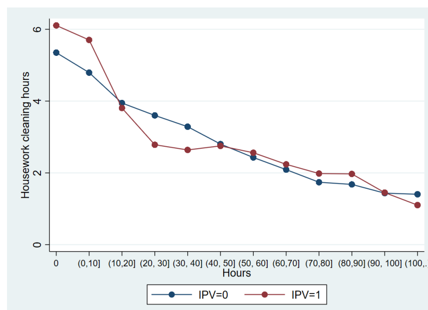


Men

(b) Hours of Cleaning

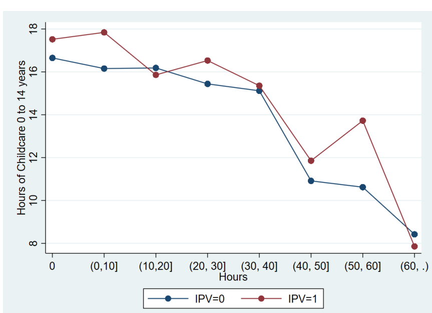


Women

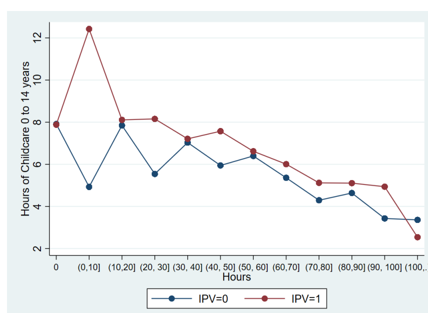


Men

(c) Hours of Childcare 0 to 14 years



Women

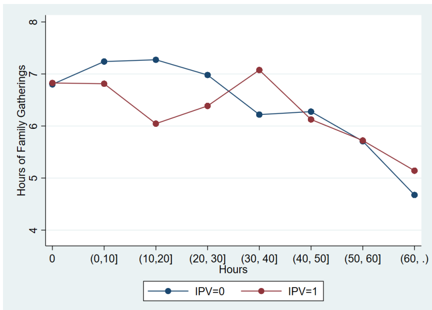


Men

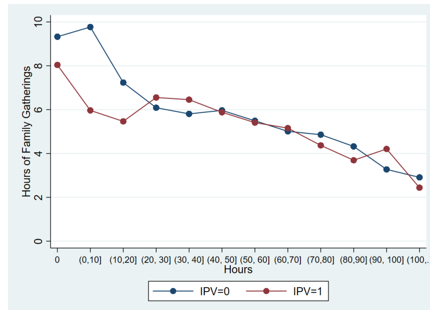
Note: The horizontal axis measures the hours of work in intervals of 10 hours. The vertical axis measures the mean of hours for the different activities dedicated to each interval of working hours. Source: First column, Synthetic Dataset for Women and Second column, Synthetic Dataset for Men.

Figure 3.11: Mean of hours allocated in different activities by working hours a week

(a) Hours of Family Gathering

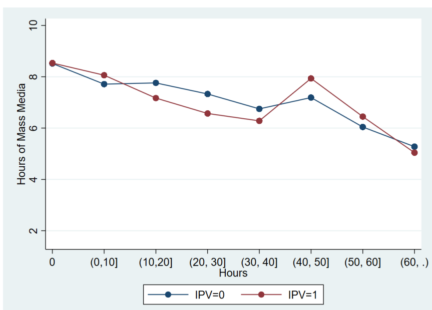


Women

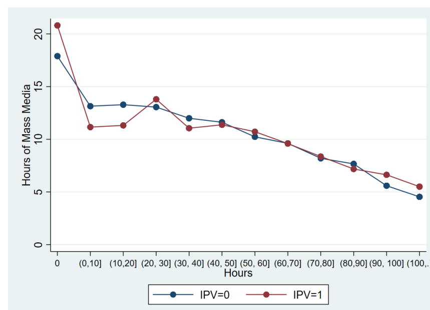


Men

(b) Mass Media Consumption

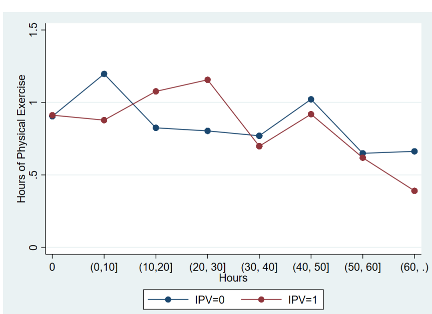


Women

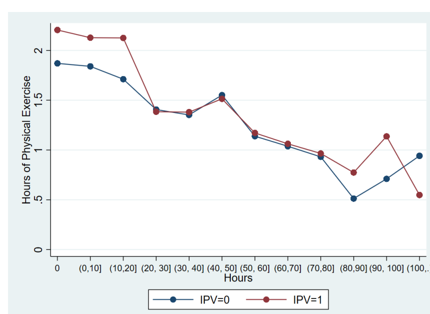


Men

(c) Physical Exercise



Women



Men

Note: The horizontal axis measures the hours of work in intervals of 10 hours. The vertical axis measures the mean of hours for the different activities dedicated to each interval of working hours. Source: First columns, Synthetic Dataset for Women and Second columns, Synthetic Dataset for Men.

3.7 Conclusion: Discussion, Limitations and Future Research

The analysis conducted in this study aimed to shed light on whether Intimate Partner Violence could be identified during routine activities and whether individuals behave differently in their time allocation in the presence of violence. Due to the unavailability of a single survey containing all necessary information, I employed a statistical matching methodology based on D'Orazio et al. (2006b), using a sophisticated matching variable derived from the Genetic Matching algorithm to create separate synthetic datasets for women and men. Although no specific patterns were identified overall, some handy findings emerged from the analysis.

Women who experienced IPV the previous year, tend to spend slightly more time in paid work, on average. Moreover, as the amount of time women victims of IPV spend on paid work increases, they allocate more time to care activities, especially childcare, compared to women who are not victims of IPV. Additionally, women victims of IPV who work fewer than 40 hours per week also spend less leisure time, particularly in family gatherings. These findings suggest that differences in IPV among women are closely related to the time they spend with their families. These results are consistent with previous studies by Mele (2009) and Lanier and Maume (2009), which indicated that childcare not only serves as a source of conflict but also acts as a barrier to leaving a harmful relationship. Furthermore, women who work less than 40 hours per week and become more economically dependent on their husbands are at a higher risk of experiencing IPV when they allocate less time to family gatherings, highlighting the importance of social support in case of victimization. Interestingly, contrary to expectations, there are no significant differences in how women allocate their time to housework in the presence of IPV compared to when there is no violence. This suggests that compensating for housework as a strategy to avoid violence, as proposed by Simister (2013), may not be a prevalent response among women in this context.

Similar patterns are observed for men regarding the time they dedicate to work. The differences in time allocation are most evident among men who work less than 40 hours a week. Notably, violent men in this category spend more time in caring activities, and this difference becomes even more pronounced when they care for children between the ages of 0 and 14 years. Additionally, there are notable differences in overall leisure time, particularly in activities such as family gatherings and mass media consumption, as well as in cleaning activities. However, it is important to note that these differences are highly fluctuating. One plausible explanation for these results could be the limited representation of men who work less than 40 hours in the sample (approximately 15% of the total sample of men). Due to this limited representation, it is essential to explore other characteristics, apart from the hours spent working, that might contribute to male perpetration of violence against their partners. This evidence highlights the need for further research to identify other factors that may play a significant role in driving violent behaviours among men in intimate relationships.

This research has encountered several limitations that warrant acknowledgement. The first limitation arises from the assumptions made in the statistical matching process and the validity of the synthetic datasets. Due to the absence of auxiliary information and the inability to verify the Conditional Independence Assumption (CIA), an uncertainty analysis was performed. The average width of the contingency tables was 7%, with the exception of the work categories of women (Table C4), which showed a narrower width, suggesting a better match. However, despite these efforts, the results obtained from the synthetic datasets cannot be definitively claimed as either statistically significant or indicative of causal relationships. The uncertainty surrounding the matching process necessitates a cautious interpretation of the findings. Further research with alternative matching methodologies and more robust uncertainty analysis would be required to strengthen the validity of the results and draw stronger conclusions.

Another significant challenge encountered in this analysis pertains to the computational cost associated with using large datasets. While the matching variable derived from the Genetic Matching algorithm provided valuable information and improved matching outcomes, certain constraints had to be addressed to ensure a feasible analysis. Notably, the common variables *state* and *household size* were excluded from the GenMatch algorithm, despite their demonstrated explanatory power with the target variables and their similar distributions in both the ENUT 2014 and ENDIREH 2016 surveys. The decision to exclude these variables was driven by the need to limit the number of matching variables in the GenMatch, as *state* had 32 categories, and *household size* had 4 categories, contributing to a considerable computational burden when dealing with large datasets. Consequently, a careful and comprehensive preliminary analysis needs to be conducted to select the most relevant matching variables, as the GenMatch algorithm becomes less flexible when subsequent changes are introduced. It is essential to strike a balance between ensuring rigorous analysis and managing computational complexity, especially when working with extensive datasets.

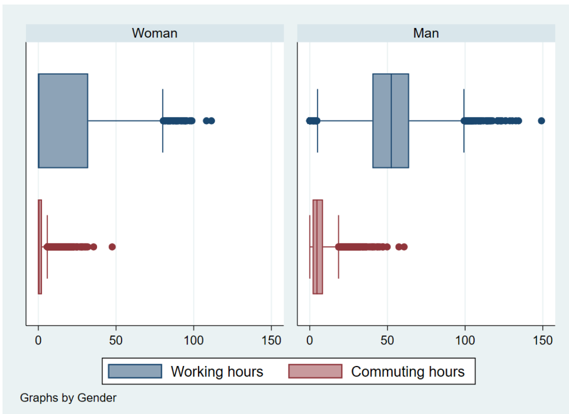
Furthermore, the research itself encountered certain limitations related to the data source, particularly with the ENUT 2014 survey. One significant issue pertains to the time recording system used, which resulted in overtime reports. Unlike the diary system, where individuals record their activities in a 24-hour diary with 30-minute intervals, the ENUT survey requires respondents to provide the total number of hours and minutes spent on each activity. This difference in data collection methods introduces challenges in accurately capturing and representing individuals' daily routines. Moreover, the lack of distinction between primary and secondary activities in the ENUT survey poses another limitation. As a result, it becomes difficult to identify which activities were performed simultaneously or to account for overlapping activities. This lack of precision becomes evident when summing up the reported hours for each activity, as the total does not always align with the expected 168 hours in a week. These data limitations can introduce potential biases and inaccuracies in the analysis and interpretation of results, warranting caution in drawing definitive conclusions.

In future research aiming to detect intimate partner violence from daily routines, it is crucial to explore surveys that incorporate key day-to-day indicators to enhance the understanding of this complex issue. One essential aspect to consider is measuring the time spent on specific activities that might be associated with intimate partner violence, such as time spent drinking alcohol or the type of mass media consumption. Furthermore, understanding the day-to-day social network of individuals is crucial. Future research should include questions about the extent to which spouses spend time together during the week and who the individuals are that they regularly interact with, such as relatives, friends, and co-workers. Examining patterns of social interactions can shed light on potential sources of support, stress, or influence in the context of intimate partner relationships. Overall, by incorporating these additional indicators and exploring day-to-day patterns in intimate partner relationships, future research can provide a more comprehensive understanding of intimate partner violence and potentially identify early warning signs or risk factors. This knowledge can be instrumental in developing effective interventions and support systems to address and prevent intimate partner violence in Mexico and beyond.

3.C Appendix Chapter 3

3.C.1 Hours dedicated to each activity by gender

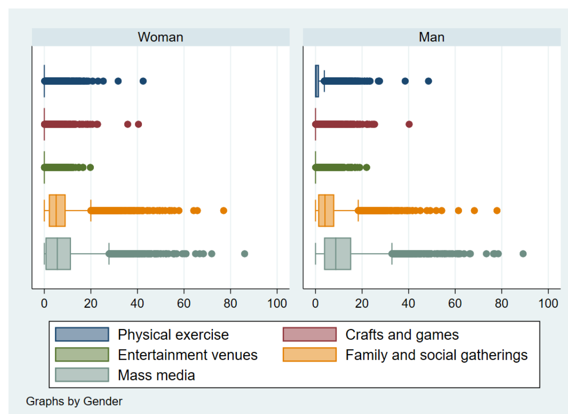
Figure C1: Mean of Hours By Gender



(a) Hours of Work

(b) Hours of Housework

(c) Hours of Care



(d) Hours of Leisure

The box spans the interquartile range (25^{th} percentile to 75^{th} percentile), with the hinge of the box representing the median value. The whiskers extend to 1.5 times the interquartile range, leaving the rest of the observations plotted in individual dots. Source: ENUT 2014.

3.C.2 Description of ENDIREH and ENUT common variables

Table C1: Definition of Variables

	ENIREH 2016	ENUT 2014
Variable	Definition	Definition
State	Federative Entity Key	Entity
Bedroom Groups	How many rooms are they used to sleep in? (not counting the hallways)	How many rooms are used for sleeping, not counting corridors?
Indigenous	According to your culture, does (NAME) consider yourself indigenous?	According to your culture, do you consider yourself indigenous?
House Size Group	How many people normally live in this house, counting young children and the elderly (also count domestic workers who sleep here)?	How many people normally live in this house, counting small children and the elderly? (also count domestic workers who sleep here)
Rural	Domain: Rural	Locality
Head of the Household	How is (NAME) related to the head of household?: Head	What is (NAME) related to the head of the household?: Head
Age group	How old is (NAME)?	How old is (NAME)?
Under 18	How old is (NAME)?	How old is (NAME)?
Under 16	How old is (NAME)?	How old is (NAME)?
Under 12	How old is (NAME)?	How old is (NAME)?
Under 5	How old is (NAME)?	How old is (NAME)?
Marital Status	Currently: Do you live with your partner in a free union?	Currently: Do you live with your partner in a free union?
Education groups	Up to what year or grade did (NAME) pass in school? (Level)	Up to what year or grade did you pass in school?
Employment	Did you work last week?	During the past week, did you work at least one hour?
Student	Last week did you: Are you a student?	Therefore, last week did you spent time studying?
Retired	Last week did you: Are you retired?	Therefore, last week were you retired?
Homemaker	Last week did you: Do you do housework?	Therefore, last week did you do housework and childcare?
Partner Employment	Did you work last week?	During the past week, did you work at least one hour?
Partner Indigenous	According to your culture, does (NAME) consider yourself indigenous?	According to your culture, do you consider yourself indigenous?
Age gap group	How old is (NAME)?	How old is (NAME)?
Education gap	Up to what year or grade did (NAME) pass in school? (Level)	Up to what year or grade did you pass in school?

Definition of the common variables in both surveys: ENDIREH 2016 and ENUT 2014.

Table C2: Categories of Common Variables

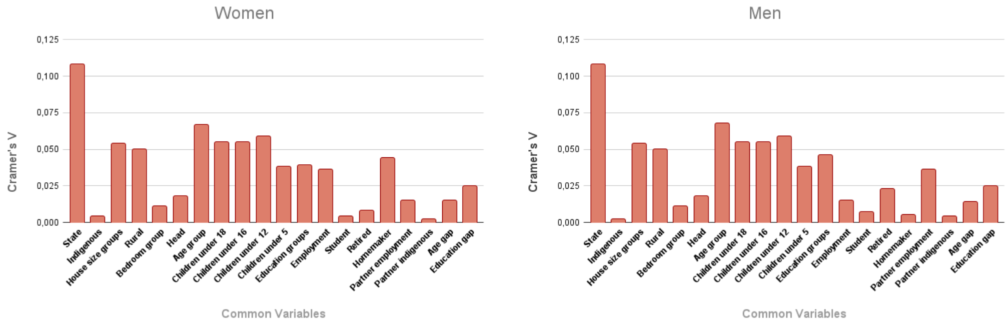
Variable	Categories	Label
State	32	1 = Aguascalientes; 2 = Baja California; 3 = Baja California Sur; 4 = Campeche; 5 = Coahuila; 6 = Colima; 7 = Chiapas; 8 = Chihuahua; 9 = Distrito Federal; 10 = Durango; 11 = Guanajuato; 12 = Guerrero; 13 = Hidalgo; 14 = Jalisco; 15 = Edo. de Mexico; 16 = Michoacan; 17 = Morelos; 18 = Nayarit; 19 = Nuevo Leon; 20 = Oaxaca; 21 = Puebla; 22 = Queretaro; 23 = Quintana Roo; 24 = San Luis Potosi; 25 = Sinaloa; 26 = Sonora; 27 = Tabasco; 28 = Tamaulipas; 29 = Tlaxcala; 30 = Veracruz; 31 = Yucatan; 32 = Zacatecas;
Bedroom Groups	4	1 = 1 Bedroom; 2 = 2 Bedrooms; 3 = 3 Bedrooms; 4 = 4 or more Bedrooms
Indigenous	2	1 = Yes; 2 = No
House Size Group	4	1 = 2 individuals; 2 = 3 individuals, 3 = 4 individuals; 4 = 5 or more individuals
Rural	2	1 = Rural; 2 = Urban
Head of the Household	2	1 = Yes; 2 = No
Age group	4	1 = 15-24; 2 = 25-44; 3 = 45-64; 4 = 65-99
Under 18	2	1 = Yes; 2 = No
Under 16	2	1 = Yes; 2 = No
Under 12	2	1 = Yes; 2 = No
Under 5	2	1 = Yes; 2 = No
Marital Status	2	1 = Married; 2 = Living together
Education Groups	3	1 = Primary; 2 = Secondary; 3 = Tertiary
Employment	2	1 = Employed; 2 = Not employed
Student	2	1 = Yes; 2 = No
Retired	2	1 = Yes; 2 = No
Homemaker	2	1 = Yes; 2 = No
Partner Employment	2	1 = Yes; 2 = No
Partner Indigenous	2	1 = Yes; 2 = No
Education gap	5	1 = -2; 2 = -1; 3 = 0; 4 = 1; 5 = 2

Label of the common variables' categories after the harmonization process. Sources: ENDIREH 2016, and ENUT 2014.

3.C.3 Predictors of Target Variables

3.C.3.1 Predictors of IPV

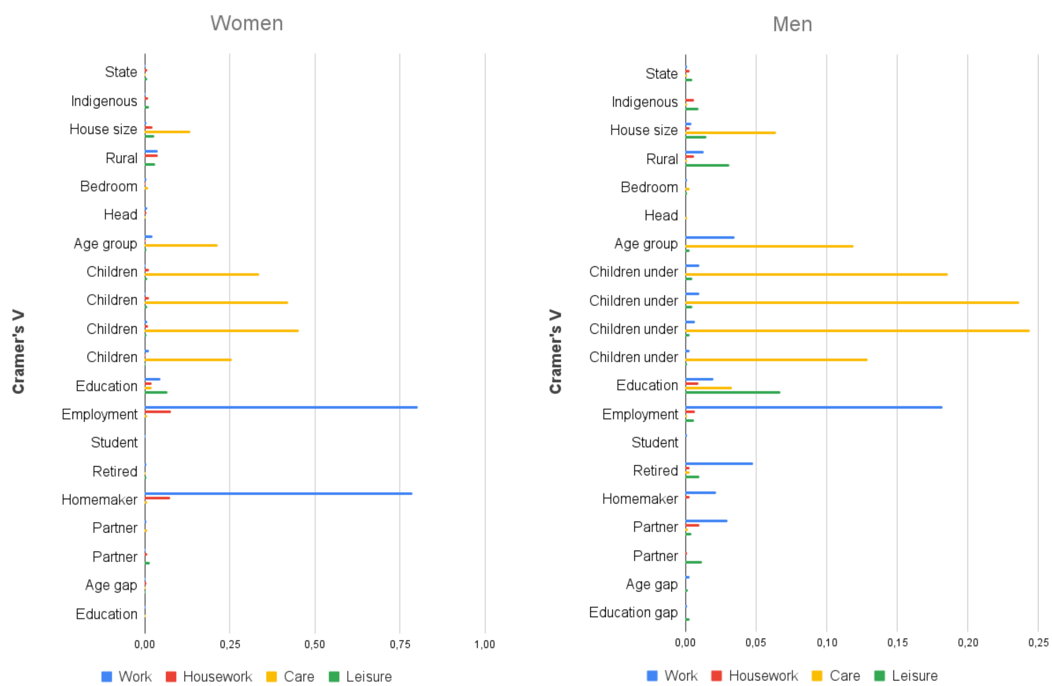
Figure C2: Cramer's V on IPV



Cramer's V value to analyse the most powerful determinants of IPV among all common variables for the sample of Women and Men separately.

3.C.3.2 Predictors of Time Use Variables

Figure C3: Adjusted-R^2



Adjusted-R^2 value to analyse the most powerful determinants of Time Use variables (work, housework, care, and leisure) among all common variables for the sample of Women and Men separately.

3.C.4 Uncertainty Assumption

Table C3: Distribution of Categorical Time Use Variables

Activity	Women		Men	
Work	Hours	Distribution	Hours	Distribution
	0	52.76%	(0, 40)	14.90%
	(0, 20)	10.49%	(40, 55)	31.01%
	(20, 40)	11.70%	(55, 70)	30.63%
	(40, 60)	17.61%	(70, .)	23.46%
	(60, .)	7.44%		
Total		100%		100%
Housework	Hours	Distribution	Hours	Distribution
	(0, 20)	11.96%	(0, 3.5)	24.92%
	(20, 40)	41.02%	(3.5, 8)	26.26%
	(40, 50)	19.64%	(8, 20)	24.21%
	(50, .)	27.38%	(20, .)	24.61%
Total		100%		100%
Care	Hours	Distribution	Hours	Distribution
	0	27.20%	0	33.09%
	(0, 14)	23.64%	(0, 3)	18.33%
	(14, 42)	24.43%	(3, 8)	16.30%
	(42, .)	24.72%	(8, 20)	15.52%
			(20, .)	16.76%
Total		100%		100%
Leisure	Hours	Distribution	Hours	Distribution
	(0, 7)	22.83%	(0, 8.5)	23.23%
	(7, 14)	27.38%	(8.5, 15)	25.20%
	(14, 40)	44.48%	(15, 24)	26.19%
	(40, .)	5.31%	(24, .)	25.38%
Total		100%		100%

Categorization of the Time Use Variables - work, housework, care, and leisure - depending on their economic definitions and distribution that avoid empty joint distributions with IPV.

Table C4: Frechet Bounds: Women

Work	IPV	Lower Bound	CIA	Upper Bound	Average Width
0	0	0.3819	0.3824	0.3831	0.0227
(0, 20)	0	0.0235	0.0348	0.0441	
(20, 40)	0	0.0283	0.0443	0.0574	
(40, 60)	0	0.0590	0.0790	0.0987	
(60, .)	0	0.0155	0.0288	0.0387	
0	1	0.1572	0.1578	0.1584	
(0, 20)	1	0.0108	0.0201	0.0314	
(20, 40)	1	0.0099	0.0230	0.0390	
(40, 60)	1	0.0183	0.0380	0.0581	
(60, .)	1	0.0054	0.0153	0.0285	
Housework	**IPV**	**Lower Bound**	**CIA**	**Upper Bound**	**Average Width**
(0, 20)	0	0.0366	0.0590	0.0741	0.0840
(20, 40)	0	0.1666	0.2270	0.2800	
(40, 50)	0	0.0662	0.1151	0.1493	
(50, .)	0	0.1131	0.1683	0.2151	
(0, 20)	1	0.0105	0.0256	0.0479	
(20, 40)	1	0.0472	0.1002	0.1606	
(40, 50)	1	0.0177	0.0520	0.1009	
(50, .)	1	0.0297	0.0765	0.1317	
Care	**IPV**	**Lower Bound**	**CIA**	**Upper Bound**	**Average Width**
0	0	0.1367	0.1516	0.1678	0.0733
(0, 14)	0	0.0801	0.1269	0.1632	
(14, 42)	0	0.0906	0.1405	0.1831	
(42,)	0	0.1047	0.1503	0.1912	
0	1	0.0383	0.0544	0.0694	
(0, 14)	1	0.0194	0.0557	0.1025	
(14, 42)	1	0.0287	0.0714	0.1212	
(42,)	1	0.0320	0.0728	0.1184	
Leisure	**IPV**	**Lower Bound**	**CIA**	**Upper Bound**	**Average Width**
(0, 7)	0	0.0788	0.1227	0.1558	0.0811
(7, 14)	0	0.0931	0.1480	0.1910	
(14, 40)	0	0.1953	0.2582	0.3183	
(40,)	0	0.0191	0.0353	0.0455	
(0, 7)	1	0.0206	0.0538	0.0976	
(7, 14)	1	0.0242	0.0672	0.1222	
(14, 40)	1	0.0577	0.1178	0.1807	
(40,)	1	0.0035	0.0138	0.0299	

Frechet bounds for the Women's synthetic dataset. The CIA column corresponds to the contingency table of the two joint variables - IPV and time use variables - work, housework, care, and leisure, respectively. In fact, the cumulative sum of all CIA values of each routine activity amounts to 1 approximately.

Table C5: Frechet Bounds: Men

Work	IPV	Lower Bound	CIA	Upper Bound	Average Width
(0, 40)	0	0.0498	0.0736	0.0906	0.0836
(40, 55)	0	0.1090	0.1626	0.2061	
(55, 70)	0	0.1116	0.1689	0.2164	
(70,)	0	0.0731	0.1239	0.1645	
(0, 40)	1	0.0119	0.0290	0.0528	
(40, 55)	1	0.0284	0.0718	0.1254	
(55, 70)	1	0.0272	0.0747	0.1320	
(70,)	1	0.0218	0.0623	0.1132	
Housework	**IPV**	**Lower Bound**	**CIA**	**Upper Bound**	**Average Width**
(0, 3.5)	0	0.0897	0.1417	0.1829	0.0866
(3.5, 8)	0	0.0933	0.1450	0.1863	
(8, 20)	0	0.0840	0.1320	0.1678	
(20,)	0	0.0866	0.1301	0.1629	
(0, 3.5)	1	0.0235	0.0646	0.1167	
(3.5, 8)	1	0.0247	0.0661	0.1178	
(8, 20)	1	0.0240	0.0597	0.1077	
(20,)	1	0.0228	0.0556	0.0991	
Care	**IPV**	**Lower Bound**	**CIA**	**Upper Bound**	**Average Width**
0	0	0.1442	0.1829	0.2158	0.0678
(0, 3)	0	0.0648	0.1044	0.1342	
(3, 8)	0	0.0485	0.0877	0.1164	
(8, 20)	0	0.0443	0.0827	0.1108	
(20,)	0	0.0550	0.0911	0.1186	
0	1	0.0400	0.0729	0.1116	
(0, 3)	1	0.0164	0.0463	0.0859	
(3, 8)	1	0.0146	0.0432	0.0824	
(8, 20)	1	0.0122	0.0403	0.0787	
(20,)	1	0.0159	0.0434	0.0795	
Leisure	**IPV**	**Lower Bound**	**CIA**	**Upper Bound**	**Average Width**
(0, 8.5)	0	0.0827	0.1268	0.1621	0.0848
(8.5, 15)	0	0.0831	0.1333	0.1721	
(15, 24)	0	0.0915	0.1433	0.1834	
(24,)	0	0.0974	0.1413	0.1763	
(0, 8.5)	1	0.0222	0.0575	0.1016	
(8.5, 15)	1	0.0232	0.0620	0.1122	
(15, 24)	1	0.0236	0.0637	0.1156	
(24,)	1	0.0264	0.0614	0.1053	

Frechet bounds for the Men's synthetic dataset. The CIA column corresponds to the contingency table of the two joint variables - IPV and time use variables - work, housework, care, and leisure, respectively. In fact, the cumulative sum of all CIA values of each routine activity amounts to 1 approximately.

This page is intentionally left blank.

Chapter 4

Severity of COVID-19 and Intimate Partner Violence in Mexico

Abstract

This chapter investigates the association between the incidence of Intimate Partner Violence (IPV) and the severity of the COVID-19 pandemic in Mexico. Using data from the 2021 Mexican National Survey of Households Relationships, I conducted regressions to examine the likelihood of experiencing IPV in relation to the Age-Standardized Mortality Rate (ASMR) of each Mexican municipality. This research provides a comprehensive analysis of the association between mortality rates and the probability of the four primary IPV forms (physical, emotional, economic and sexual), as well as their frequency. The results reveal a significant positive correlation between the severity of the virus and emotional, economic, and sexual IPV. In contrast, there is a negative association between ASMR and the occurrence of physical aggression. These findings suggest that as the COVID-19 death rate increased, more households experienced heightened tensions and conflicts related to financial and employment uncertainty. However, due to the fear of infection, perpetrators exhibited less physically harmful violence.

4.1 Introduction

The COVID-19 pandemic, caused by the SARS-CoV-2 virus, has brought about unprecedented disruptions to global health, economic and social systems. After the first cases were detected in China in December 2019, the World Health Organization (WHO) declared it a Public Health Emergency of International Concern due to the rapid spread of the virus. Although vaccines have helped mitigate the virus lethality, the COVID-19 pandemic persisted in various parts of the world until the 5th of May of 2023, when WHO declared an end to the COVID-19 emergency. Over the course of three years, the pandemic resulted in hundreds of millions of confirmed cases and approximately 7 million related deaths (WHO[1]). This global health crisis extended beyond public health concerns, profoundly inducing an economic and social disruption with a long-term effect on the livelihood of millions. Among the far-reaching repercussions, it encompassed the closure of schools and businesses, the loss of jobs, the rise of mental health issues and intra-household violence.

This research delves into the impact of Mexico's experience with the COVID-19 pandemic on a less visible, yet persistent pandemic: Intimate Partner Violence (IPV). Internationally, the emergence of the COVID-19 pandemic has sparked considerable speculation about its effect on women's victimization (Peterman et al., 2020; Gresham et al., 2021; Van Gelder et al., 2020; Pietromonaco and Overall, 2021). Lockdowns and changes in the economic routines implemented to curb the virus's spread have been linked to the confinement of victims with their abusers, leading to heightened tension exacerbated by financial instability. In the context of Mexico, the pandemic had multifaceted consequences. Pre-existing socio-economic disparities among the country's municipalities played a significant role in shaping the overall impact of COVID-19, resulting in a heterogeneous distribution of the virus's effects. Consequently, Mexico ranked among the countries with the highest recorded death toll worldwide, with nearly 334,000 deaths by the end of the pandemic (Hong et al., 2021). Furthermore, while the pandemic raged on, violence against women did not subside. Emergency calls related to domestic violence surged by 12% from 2020 to 2021, and the country witnessed nearly 4,000 cases of women being murdered, with over a third classified as femicides (UN Women, 2017).

While a substantial body of research has examined the pandemic's impact on various aspects of Mexican society, including employment dynamics and time allocation (Hoehn-Velasco et al., 2022), women's mental health and fertility (Silverio-Murillo et al., 2021), as well as domestic violence (Hoehn-Velasco et al., 2021a; Valdez-Santiago et al., 2021; Silverio-Murillo et al., 2023), much of this work has focused on the early stages of the pandemic, particularly the initial two months of lockdown. This study extends the analysis beyond this confinement timeframe, to fill the notable gap in evaluating the influence of COVID-19 severity, which extended well beyond the confines of strict quarantine measures, particularly throughout 2021. Furthermore, a dearth of literature exists examining the nexus between the pandemic and Intimate Partner Violence (IPV), partly due to the limited data available - primarily comprising helpline calls and police reports, which are susceptible to underreporting. This research uses the 2021 National Survey on the Dynamics of Household Relationships (ENDIREH), offering comprehensive insights into IPV experiences, types, and frequencies.

This analysis leverages death data from the Mexican Secretariat of Health to calculate the Aged-Standardized Mortality Rates (ASMR) from the onset of the pandemic until the end of March 2023. The ASMR serves as a reliable indicator for assessing the pandemic's severity across various regions (WHO, 2023), facilitating cross-municipality comparisons regarding virus impact. To address the emerging concerns regarding the accuracy of data provided by the Mexican Secretariat of Health, this study incorporates administrative death data from the National Institute of Statistics (INEGI). This dataset not only allows for the assessment of direct COVID-19

[1] Up to date data from https://covid19.who.int/

deaths but also enables control for excess mortality (Knaul et al., 2021; Calixto-Calderón et al., 2021; Pamplona, 2020; Singer, 2020). Consequently, by integrating all available data sources, this paper enhances the robustness of the examination of the association between mortality rates and IPV and its primary forms, including physical, emotional, economic and sexual violence.

Through Ordinary Least Squares (OLS) regression models, the results reveal a positive correlation between COVID-19 mortality rates and the likelihood of IPV, particularly concerning emotional, economic and sexual violence. Financial and economic insecurity is the primary explanatory mechanism for these findings. In regions marked by higher COVID-19 fatalities, health emergencies were declared, imposing corresponding confinement measures. Consequently, this led to a temporary economic standstill, thereby intensifying tensions within couples. However, the study also observes a negative correlation between the pandemic's severity and the frequency of physical aggressions. This finding may be attributed to the fear harboured by both victims and perpetrators of contracting the virus and being hospitalized. Consequently, individuals may have altered their behaviour to avoid this risk, which in turn resulted in a decrease in physical violence.

The chapter is structured as follows. Section 4.2 provides an overview of the COVID-19 pandemic's timeline in Mexico and how it relates to IPV victimization, examining changes in IPV rates before and after the pandemic. Section 4.3 reviews various virus-related mechanisms that may connect the severity of COVID-19 with IPV. It also summarizes the existent evidence on the relationship between COVID-19 severity and violence. Section 4.4 details all the data sources employed in the analysis. Section 4.5 defines the relevant variables of interest and presents primary descriptive statistics to provide an initial understanding of the data. Section 4.6 presents the main findings of the analysis. Section 4.7 discusses the results and exposes the main limitations of the study. Finally, section 4.8 concludes.

4.2 Institutional Setting

4.2.1 The COVID-19 pandemic in Mexico

The first case of COVID-19 infection in Mexico was confirmed on February 28^{th}, 2020, nearly two months after the World Health Organization's confirmation of the first case in Wuhan, China. Initially, Mexico's President Andrés Manuel López Obrador understated the seriousness of the virus. However, on March 30^{th}, 2020, the Mexican Federal Government proclaimed a nationwide health emergency, over a month subsequent to the emergence of the first case. By the end of May 2020, the COVID-19 outbreak was triggered by collapsing hospitals, struggling to provide treatment to all patients while imperiling healthcare workers who faced the risk of getting infected (Singer, 2020; Knaul et al., 2021).

As outlined by Felbab-Brown (2020), the pandemic's impact on Mexico was exacerbated by an already weakened healthcare system resulting from the austerity policies implemented under López Obrador administration. These policies led to a 3.3% reduction in healthcare expenditures, culminating in the dismissal of medical professionals and shortage of essential medications. Among the critical medical resources that faced cuts, the most substantial were ventilators and medical beds, with an uneven allocation between private and public hospitals. Adding to these challenges, in January 2020, the Government discontinued Seguro Popular, the prevailing health insurance option for informal sector workers lacking access to public social security.[2] Instead, the Government established the Institute of Health for Wellbeing (INSABI), with the aim of providing centralized cost-free social protection on a national scale. However, this initiative encountered operational inefficiencies that were exacerbated by the onslaught of the pandemic. In addition, in 2019, over half of the Mexican population workforce (56.2% of total workers) was engaged in informal employment, and this proportion grew following the overture of lockdown measures (Ibarra-Olivo et al., 2021; Hoehn-Velasco et al., 2022). As a consequence, a substantial portion of the population found themselves less shielded against the virus's transmission.

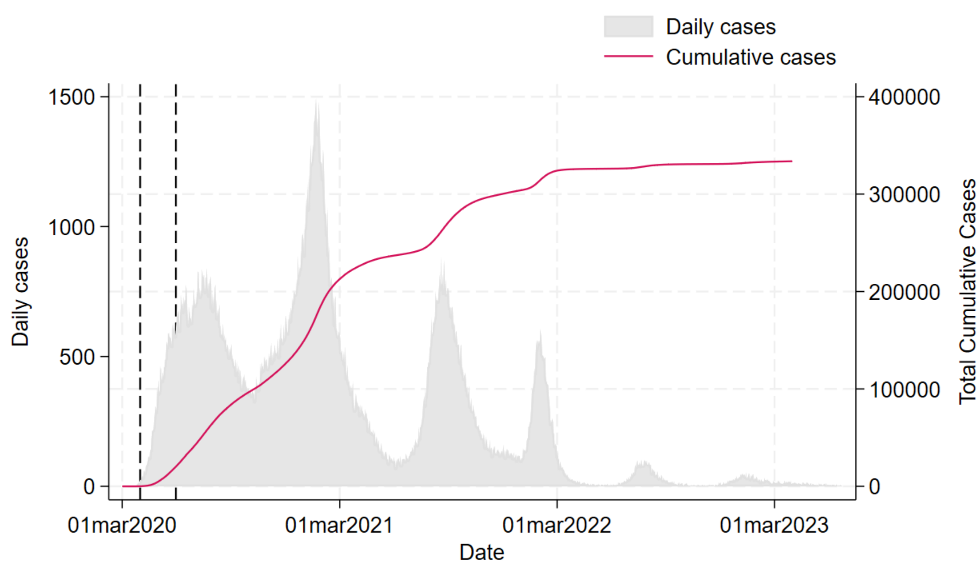
The sole nationwide measure employed was the implementation of a national lockdown, spanning from March 31^{st} to May 30^{th}, 2020 (Atondo and Atondo, 2021). Despite its short phase, this lockdown increased rates of poverty, extreme poverty, and inequality, based on the Gini index (Salas et al., 2020; Huesca et al., 2021). For instance, the country faced a recession with a decline of almost 8% in GDP in 2020 with little aid from the Mexican government to most affected groups (Hoehn-Velasco et al., 2022). As a result, Mexican individuals faced changes in employment, income, and time use during the recession.

Moreover, although complete control over the virus transmission was lacking, the national confinement strategy transitioned into a region-specific "traffic light" system for reopening, controlled by the federal government of each of the 32 states. As detailed by Hoehn-Velasco et al. (2021b), every 14 days each state revised its epidemiological risk of transmission to inform state-level responses. Each federal government considered ten indicators encompassing the territorial spread of the virus, medical response capacity, and epidemic severity. Each indicator was measured in points ranging from 0 to 4. Consequently, a state fell within the "green traffic line" category when they reported from 0 to 8 points. This category enabled the state to reinstate regular economic and labour activities, including the reopening of schools and social activities. When a state accumulated between 9 and 15 points, the state fell within the "yellow traffic line" designation. This permitted economic activity to proceed, albeit with a reduction in social activities to 60% of their normal capacity and the closure of schools. For states reporting points between 16 to 31, the state fell within an "orange traffic line" phase. This category involved allowing essential economic activities, or, for non-essential occupations, the workplace reduced its own intensity levels. This phase also enhanced extreme protection to vulnerable individuals, school closure and social activities reduced to 30% of its capacity. Lastly, states with the highest alert level

[2] Mexico's public health insurance is provided through institutions like the Mexican Institute for Social Security (IMSS) and the Institute for Social Security of State Employees (ISSSTE), which cater to salaried public and private sector employees.

were assigned the "red traffic light" with 32 to 40 points. These states were confined to the lockdown with the only function of essential economic activities. The traffic light system remained in effect until April 30^{th}, 2022.

Figure 4.1: Daily and cumulative COVID-19 death cases



The shaded grey bars in the chart represent the daily crude death counts, while the red line depicts the cumulative crude deaths over time. The two vertical shaded lines indicate the period of lockdown. Data source: Mexican Health of Secretariat.

According to Knaul et al. (2021), the arrival of COVID-19 vaccines in Mexico did not occur until May 2021, and even then, the distribution was marked by deficiencies, insufficiencies, and glaring inequitable access. This situation exacerbated existing economic disparities among the states, with wealthier regions receiving more extensive vaccine coverage than their economically disadvantaged counterparts.[3] Moreover, the transparency of data during the vaccine campaign reported by the Mexican Health Secretariat was questionable, characterized by contradictory information and frequent alterations to the campaign timeline. The most recent available data from the Mexican Health Secretariat, as of January 25^{th}, 2022, indicates that over 83 million Mexicans had received at least one vaccine dose.[4]

Therefore, the inefficient coping strategies employed by both the national and federal governments in combating the coronavirus resulted in an uncontrollable escalation of COVID-19 casualties. Mexico found itself among the most devastatingly affected countries globally, with an excess of 330,000 cumulative deaths at the beginning of 2023. Furthermore, in the face of the potential economic repercussions, some decentralized federal governments faced allegations of lacking transparency in reporting positive COVID-19 cases and deaths, presumably to avoid shifting the traffic light status.[5]

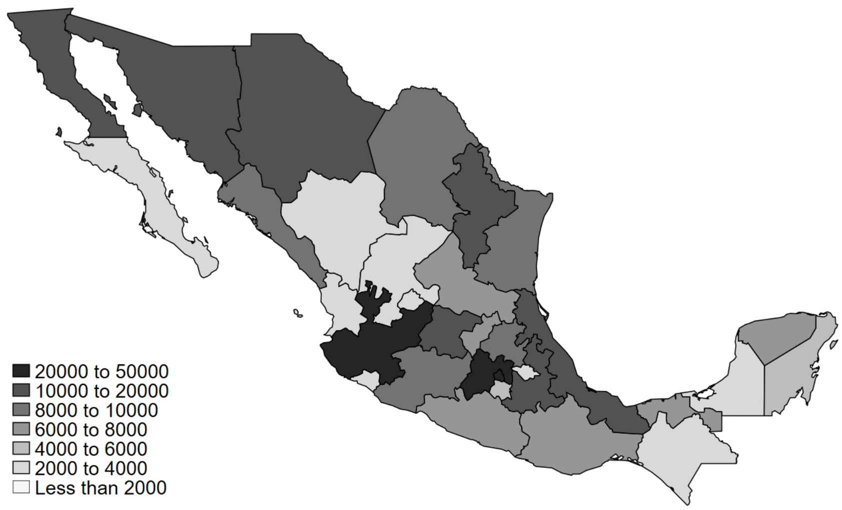
[3] To illustrate this situation, whereas almost 80% of Baja California's population had received at least one vaccine dose, only 21% of Chiapas's population had similar access.

[4] Source: https://www.gob.mx/salud/prensa/034-en-mexico-83-3-millones-de-personas-vacunadas-contra-covid-19?idiom=es

[5] One of the states embroiled in significant controversy over its handling of the virus and resistance to changing the green traf-

As illustrated in Figure 4.1, the months covered by the lockdown imposition - represented as the vertical shaded lines - primarily coincided with the initial outbreak of the pandemic, during which case numbers were at their lowest in comparison to the subsequent periods. Following the implementation of quarantine measures, the virus has given rise to a minimum of five distinct waves, with the first three proving to be the most severe and consequential covering the 2020 and 2021 years.

Figure 4.2: Covid daily confirmed cases and deaths differences in each State



Total crude death counts in each Mexican state, spanning from the onset of the pandemic until March 31^{st}, 2023. Source: Mexican Health of Secretariat.

Figure 4.2 displays the total deaths related to COVID-19 in each of the Mexican states. One immediately noticeable observation is that the impact of the virus did not manifest uniformly across all Mexican states. Clear disparities are evident, with entities such as Ciudad de Mexico, Baja California, and Veracruz experiencing a notably higher rate of spread in contrast to others such as Campeche, Durango, and Zacatecas. Consequently, it is plausible that the pre-pandemic socio-economic distinctions among these various states have significantly contributed to the overall impact of COVID-19.

4.2.2 Intimate Partner Violence in Mexico Before and After the COVID-19 Pandemic

Mexico has garnered notoriety for its persistently elevated incidence of violence against women, a phenomenon predating the COVID-19 pandemic (UNODC, 2018).[6] The official initiation of the war on drugs on December 11^{th}, 2006, marked a pivotal juncture in the country, initiating the adoption of a militarized approach to pub-

fic light status was Chiapas. https://www.eleconomista.com.mx/estados/Chiapas-y-el-Covid-19-un-tratamiento-fuera-de-la-realidad-20210211-0110.html
https://vlex.com.mx/vid/chiapas-covid-19-tratamiento-857475281

[6] https://politica.expansion.mx/mexico/2022/03/07/datos-sobre-la-violencia-contra-las-mujeres-mexico

lic security - a model that persists nearly two decades later. Subsequently, from 2007, instances of violence and homicides involving firearms in Mexico escalated to a point where public spaces ceased to be safe for both men and women (Data Cívica, 2019).[7] However, neither the confines of homes were safe for women. According to Equis (2020), between 2000 and 2018, 57% of women's murders occurred within the family context.

Notwithstanding the concerning statistics, it is imperative to recognize that homicides merely represent the tragic conclusion of a broader issue. Pre-pandemic data drawn from the 2016 Mexican National Survey of the Dynamics of Households Relationships (ENDIREH) unveils that 44% of women aged 15 and older have encountered at least one incident of Intimate Partner Violence (IPV) during their lifetimes. The reported breakdown of this violence indicates that 40.1% is characterized as emotional violence, 20.9% as economic violence, 17.9% as physical violence, and 6.5% as sexual violence (Castro, 2019; Equis, 2020).

These outcomes did not ameliorate with the onset of the pandemic. As reported by the United Nations (UN Women, 2017), Mexico has been grappling with an alarming prevalence of femicides[8], and witnessed an estimated average of approximately 10.5 to 11 feminicides per day. The Executive Secretariat of Public Security of the National System reported that in 2021 alone, there were 1006 registered femicides - an unprecedented high in the nation's history - alongside the 2747 female deaths that were categorized as "intentional homicides". Concurrently, there was a 12% increase in emergency calls (from 260,067 in 2020 to 291,333 calls in 2021) and the National Shelter Network provided assistance to 45,490 women and children who were survivors of male violence (UN Women, 2017).

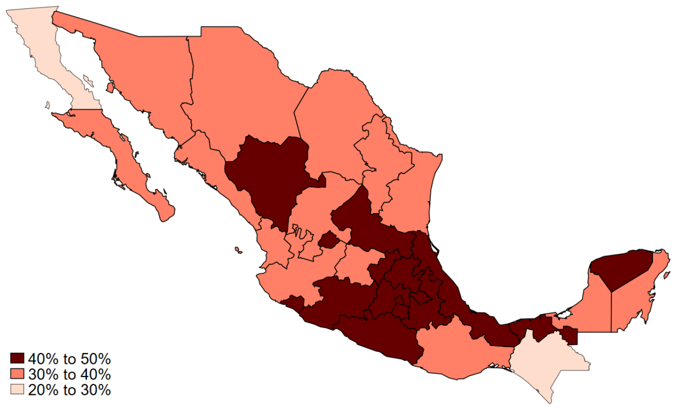
Additionally, data derived from the 2021 Mexican National Survey of the Dynamics of Households Relationships (ENDIREH) reveals that, while the overall incidence of Intimate Partner Violence (IPV) has decreased since 2016, a persistent 40% of women aged over 15 still report experiencing some form of violence - whether physical, emotional, sexual, or economic - from a partner or ex-partner. Moreover, as illustrated in the upper map of Figure 4.3 (a), the prevalence of intimate partner violence across a woman's lifetime does not exhibit a uniform distribution among the 32 states. The southern region of the country emerges as more perilous for women, with the exception of the state of Chiapas. Figure 4.3 (b) presents the average proportions of women who endured any of the four violence types during the 12 months preceding October 2021. The distribution of intimate partner violence (IPV) showcases a more consistent pattern across states, with rates ranging between 20% and 25%. Notable exceptions include Baja California and Chiapas, reporting percentages between 10% and 15%, and Aguascalientes, Colima, Michoacan, Querétaro, and Guerrero, registering percentages between 25% and 30%.

[7] Notably, in 2007, the recorded count of female homicides was 1089, remarkably lower than the 3824 homicides in 2019 (Equis, 2020).

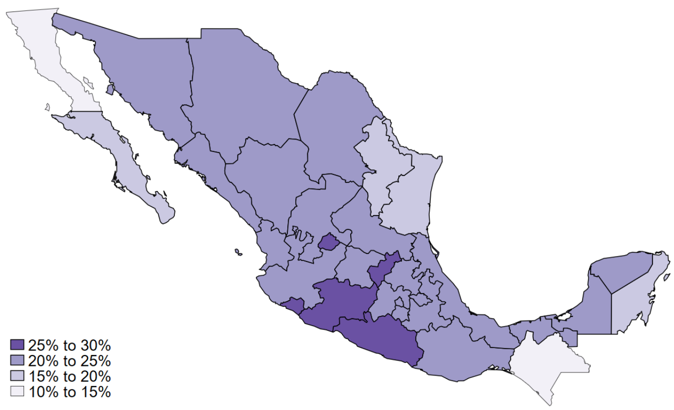
[8] As defined by Mexican law, encompasses a range of circumstances, including killings of women subjected to sexual violence, mutilation, acts of necrophilia, family violence context, or by a current or former intimate partner (UNODC, 2018).

Figure 4.3: Intimate Partner Violence Prevalence from current or previous relationship

(a) Percentage of women who ever experienced IPV in 2021



(b) Percentage of women who experienced IPV in the previous 12 months



The first map (a) shows the lifetime experience of Intimate Partner Violence (IPV) from a current or former relationship. The second map (b) shows the IPV experience suffered during the previous 12 months. Source: ENDIREH 2021.

4.3 Literature Review

4.3.1 Mechanisms that relate COVID-19 stressors with Intimate Partner Violence

Beyond the direct effects of the imposed lockdown, there are other determinants that either directly or indirectly could have affected the existing IPV. Current literature has studied the mechanisms by which the COVID-19 pandemic could affect the occurrence of Intimate Partner Violence (Peterman et al., 2020; Gresham et al., 2021; Van Gelder et al., 2020; Pietromonaco and Overall, 2021; Opanasenko et al., 2021). Using a sample of U.S. National Institutes of Health, Gresham et al. (2021) found that IPV victimization is positively related to COVID-19 stressors and vulnerabilities. Those stressors and vulnerabilities include economic insecurity, social isolation and virus-related shocks:

Economic hardship and financial stress. The economic shock from the COVID-19 pandemic was shaped differently depending on past economic vulnerabilities and socioeconomic conditions of each region or state (Peterman et al., 2020; Masferrer et al., 2022). Economic insecure populations are associated with weaker health services which may facilitate the prevalence of the COVID-19 disease. The unavoidable increase in unemployment or decrease in income may have a huge impact on poorer households. A reduction in wealth is translated to a reduction in essential goods and fewer resources to reduce stress which may lead to couple conflicts (Jewkes, 2002). Even for those who chose to expose themselves to getting infected to keep up with their economic activity, there is increased stress due to the uncertainty of economic insecurity and general well-being. Fox et al. (2004) find that in the US the men's perceptions of financial well-being, relative income, and the number of debts are negatively associated with the likelihood of exerting violence against their wives.

Moreover, these economic and unemployment shocks can lead to shifts in who is the main provider between partners. There is an extended literature that evaluates the link between IPV with the relative gain/loss of women's economic power in the household. Some evidence shows that the lack of women's economic independence may increase their victimization (Anderberg et al., 2015; Villarreal, 2007). Conversely, other papers show that male partners increase the use of violence when their female partners are active in the labour force Canedo and Morse (2021). Finally, regarding the employment status of men, Bhalotra et al. (2021b) find that unemployed men are more likely to become aggressive against their partners when they are working, i.e., women become the main provider of the household.

In this context, Mexico is an interesting scenario to analyse. Firstly, there is already mixed evidence of how women's employment status affects the likelihood of IPV. While Villarreal (2007) finds that, overall employed women suffer less physical and sexual violence than non-employed women, Canedo and Morse (2021), targeting women with lower financial capabilities, are more likely to suffer from IPV when they work. Therefore, women's unemployment shock from the pandemic could lead to an unclear violent scenario.

Secondly, using the Mexican Labour Force Survey (ENOE), Hoehn-Velasco et al. (2022) compare the effect of the nationally imposed lockdown on men's and women's employment. During the months of quarantine, both genders suffered from job loss in the formal and informal sectors. After the opening of isolation restrictions, men became employed faster than women, however, the job recovery happened mostly in the informal sector and without a full recovery of income levels before the pandemic. Therefore, from one side, the relative contribution of women's income to the household is very similar before and after the pandemic, which may have zero or negative effects on their risk of victimization. However, on the other side, the overall financial insecurity from the swift between formal to informal employment with lower earnings may increase couples' conflict.

Finally, Hoehn-Velasco et al. (2021b) evaluates the deep effects of the lockdown on the Mexican labour market and geographic differences. States with higher contact restrictions (known as red zone) produced 1% more employment losses (or reductions) than the rest of the orange, yellow and green zones. However, they do not find geographically significant differences in employment for those states that had higher COVID-19 cases or deaths during the first months of the pandemic. As is shown in Figure 4.1 the most severe pick of the pandemic happened between the end of 2020 (November) and the beginning of 2021 (February). Therefore, there may be significant differences across the different zones towards the Mexican employment in the long run that were not captured in Hoehn-Velasco et al. (2021b).

Lockdown and social distance measures. Although imposed quarantines reduced the speed of virus infection, many authors presaged the increase of domestic violence by naming it *the shadow pandemic* (Peterman et al., 2020; Pietromonaco and Overall, 2021; Van Gelder et al., 2020; Gresham et al., 2021; Agüero, 2021; Bettinger-Lopez and Bro, 2020; Lersch and Hart, 2022). All the literature agrees with the hypothesis that “Stay-at-home" restriction policies confined victims to their abusers, increasing the likelihood of suffering more violent incidents. In a recent systematic review of 32 studies, Kourti et al. (2023) shows that COVID-19 actually increased intimate partner violence cases in every analysed country, especially during the first weeks of the COVID-19 lockdown. Among those research studies, Piquero et al. (2020) finds a significant change in the trends of domestic violence during the lockdown period in Dallas (US). Incidence in domestic violence increased the two first weeks after the confinement order was implemented, and decreased thereafter. Similar results were found in Nix and Richards (2021).

However, other evidence shows either no significant association between the lockdown period and the reports of IPV (Agüero et al., 2022; Reingle Gonzalez et al., 2020; Piquero et al., 2021; Jetelina et al., 2021; Silverio-Murillo et al., 2023; Peitzmeier et al., 2022) or a negative association (Lersch and Hart, 2022; Hoehn-Velasco et al., 2021a; Valencia-Londoño and Nateras-González, 2020). Other papers find mixed evidence depending on the measurement of IPV (Bullinger et al., 2021; Peitzmeier et al., 2022; Silverio-Murillo et al., 2023). For instance, in the city of Chicago (US), Bullinger et al. (2021) compares the 911 calls of domestic violence, the crimes reported to the police, and the arrests made. They find an increase in 911 calls, but a decrease in reports and arrests. Finally, using an online survey in the State of Michigan (US), Peitzmeier et al. (2022) finds no increase in IPV occurrence between the months of March and June 2020, however, finds an increase in the severity of the violence among those who were already victims. This finding is worth noticing since it suggests that the COVID-19 pandemic anxiety did not produce abusive behaviours in those who otherwise would have never become abusive, however, it may increase the harshness to those who already were abusive. This research addresses this question in the case of Mexico.

During the lockdown in Mexico, Hoehn-Velasco et al. (2021a) and Valencia-Londoño and Nateras-González (2020), found a decrease in reported cases of gender-based violence. Although in the first three months of 2020, there was an increase in emergency call reports compared to the same time period in 2019, during the quarantine months, there was a decrease in calls. The authors argue this phenomenon happened because of two main reasons: Firstly, there was a decreased exposure of women to high-risk environments like public spaces, where there are high rates of violence against women. Secondly, limited operation of the public identities in charge of collecting violence reports. Moreover, Silverio-Murillo et al. (2023) using data from Mexico City, finds a 17% increase of hotline calls regarding emotional violence, while during the same period, the police reports of domestic violence decreased a 22%. Despite the surge in IPV-related calls during the confinement period, this increase was not statistically significant.

Silverio-Murillo et al. (2023) confirms the methodological limitations in the existing literature coming from the nature of the data. To capture either Intimate Partner Violence or Domestic Violence, researchers mainly used police and helpline reports, which already were susceptible to being underreported before the pandemic (Anderberg et al., 2015). During the months of confinement, victims of violence may be even less willing to report since they are sheltered with their perpetrators who can closely monitor their activities and communications (Piquero et al., 2020; Peterman et al., 2020; Hoehn-Velasco et al., 2021a; Kourti et al., 2023; Jetelina et al., 2021). My paper contributes to the literature by using National survey data specialized on Intimate Partner Violence which may capture better the violence than police-reported cases or arrests. Therefore, this survey data may bring a better understanding of what really happened and in the case of violence, which type of violence was the most predominant in Mexico.

Virus-specific types of violence. Not only does the imposed lockdown have a direct impact on victims of IPV, but also has an indirect effect still persistent after the post-quarantine months (Peterman et al., 2020). Mental health detriment, sleep difficulties, substance abuse, and behavioural problems among others can increase the risk of IPV (Gresham et al., 2021). Moreover, perpetrators may use unused coercive tactics directly related to the virus contagion: depriving sanitary material to prevent the spread of the virus (masks, tests, etc.), or blaming the victims for transmitting the virus. Finally, the lack of access to health services and support organizations inhibited many people to ask for help (Van Gelder et al., 2020). For instance, victims may be hesitant to seek help for fear of contracting the virus (Peterman et al., 2020) or because of the little response from the authorities (Van Gelder et al., 2020).

Furthermore, access to legal systems of safety support becomes limited, compounding the pre-existing challenges that victims face when searching for help. Stanley and Granick (2020)'s report explains how police officers may be less likely to assess cases of IPV since they become hesitant to enter private homes out of risk of disease exposure. This assumption is illustrated in Bullinger et al. (2021); Silverio-Murillo et al. (2023) with the difference between help-line calls and police arrests. As a consequence, violent offenders have convenient access to a victim, with the absence of a guardian making them unpunished for their abusive behaviour (Piquero et al., 2020). Therefore, in areas where the severity of the COVID-19 virus is higher we may expect worse functioning systems discouraging victims from reporting, and, as a consequence, shielding their perpetrators.

4.3.2 Severity of COVID-19 on IPV

During pandemic times the health system breaks from its normal functionalities to focus on the cure of the major virus. Front-line healthcare workers are often the first point of formal contact of seeking help from the victims (García-Moreno et al., 2015; Spangaro, 2017). However, in regions in which hospitals are saturated with COVID-19 patients this form of detection of Intimate Partner Violence (IPV) gets lost. Moreover, preexisting inequalities among the different regions highlight the imbalanced resources from the health institutions to battle the virus, and consequently, the level of response when a victim of violence is seeking help. Therefore, the prevalence of COVID-19 may indicate different behaviours in IPV.

For instance, Lersch and Hart (2022) assumes that COVID-19 positive cases and COVID-19 deaths may increase reported domestic violence from the opportunity theory (i.e. perpetrators have more opportunity scenarios to exert violence). In order to analyse that, they use domestic violence reports from the Annual Uniform Crime Reports of the Florida Department of Law Enforcement (FDLE), and the percentage of confirmed COVID-19 that died during the calendar year 2020 from the population of 2020 estimated by the FDLE crime data. However, despite their initial supposition, their evidence shows no significant association between COVID-19-positive cases and domestic violence and a significant negative association between COVID-19 death rates with domestic violence. The paper neither finds any significant association between the COVID-19 positive cases and deaths with the variation of domestic violence from 2019. This paper, however, uses police-reported incidents to measure domestic violence, which may not represent the actual number of domestic violence incidents.

In the context of Mexico, Masferrer et al. (2022) link the severity of COVID-19 with previous insecurity in Urban Mexico municipalities to understand whether the pre-existing vulnerabilities had an impact on the severity of the virus. Among their findings, they show that pre-pandemic intentional homicide rates were negatively associated with age-standardized crude deaths, whereas robbery was positively associated. Although Jewkes (2002) relates that street crime transfers to the intimate sphere, the literature lacks research which aims to link the severity of COVID-19 in Mexico with surveyed experiences of intimate partner violence, which can highly differ from the police-reported incidents.

4.4 Data

To investigate the correlation between COVID-19 severity with Intimate Partner Violence (IPV) I use various sources provided by the National Institute of Statistics and Geography of Mexico (INEGI) and the Mexican Secretariat of Health. Data concerning IPV was extracted from the Mexican National Survey on the Dynamics of Household Relationships 2021 (Encuesta Nacional Sobre la Dinamica de las Relaciones en los Hogares, ENDIREH 2021). ENDIREH is a nationally representative survey conducted by INEGI that since 2006 has gathered data on violence against women every five years. This survey contains information regarding household and individual characteristics for each household member and features a comprehensive questionnaire regarding women's experience of violence. Within each household, a woman aged 15 or older is selected to respond, indicating whether they have ever encountered abuse in various settings, including their school, workplace, community, home or intimate settings. The selected sample comprises women over 15 years old who can be either married, co-living with their partner, separated, divorced, or single with a boyfriend or ex-boyfriend. Women who are married to husbands residing abroad, widowed individuals and single women who never had a partner were excluded from the analysis. This exclusion was based on the lack of a clear mechanism of how they might become victims of IPV during the pandemic. Consequently, out of a total of 110,127 surveyed women, 77,704 were selected for the analysis.

The COVID-19 severity variable was assessed using the data of confirmed COVID-19 deaths, which was obtained from The Mexican Secretariat of Health. This authority regularly releases data encompassing positive and negative tests, suspected cases and COVID-19-related deaths[9]. This dataset allows for disaggregation by age (in 5-year intervals), municipality, state and gender. The data collection was finalized on March 31st, 2023, with comprehensive information available by municipality and age. It is noteworthy that, although COVID-19-related deaths continued to occur in 2022 and 2023, as depicted in Figure 4.1, the numbers have notably decreased in comparison to the levels observed in 2020 and 2021.

The data merging process between ENDIREH 2021 and the Mexican Secretariat of Health followed these steps. Out of the total 2457 Mexican municipalities, ENDIREH 2021 sampled data from individuals residing in 1262 different municipalities. This means that roughly half of the total municipalities are represented in the analysis, accounting for 51% coverage. Figure D1 in the Appendix shows a visual representation of the sampled municipalities from ENDIREH 2021. Despite it might seem like many municipalities are missing from the analysis, Figure D1 provides confidence that essential information is not overlooked. Additionally, out of the total 2457 municipalities, 31 didn't report any COVID-19-related deaths, which amounts to 390 observations in ENDIREH 2021. Consequently, the analysis of the association between mortality rates and Intimate Partner Violence was focused on individuals from municipalities where at least one COVID-19 death was registered.[10]

Finally, besides the challenge of data availability until the end of 2021, there have been concerns raised regarding the accuracy of COVID-19-related data provided by the Mexican Secretariat of Health. The primary critique revolves around unreported cases. To be classified as a COVID-19 death, a prior positive COVID-19 test is required. Consequently, there may be instances where deaths caused by the COVID-19 virus are not included in the official dataset because the individuals were not previously tested. This potential gap in the

[9] https://datos.covid-19.conacyt.mx/

[10] The Appendix Section 4.D.4 includes results obtained by assigning a value of 1 to municipalities without any recorded COVID-19 deaths. While the outcomes show similar findings, the model's goodness of fit (measured by the R^2) is lowered compared to the results presented in the main analysis.

data could introduce bias and potentially lead to inaccurate estimations of the relationship between COVID-19 severity and IPV incidents, as highlighted by Knaul et al. (2021). To address this potential issue, I have employed a sensitivity analysis approach by incorporating individual and regional controls into the model. Additionally, to bolster the robustness of the analysis, I have replicated it using official deceased data for the year 2021 from the Mexican Institute of Statistics (INEGI). This administrative dataset not only encompasses demographic information of the deceased, including gender, age, and location but also records the date and causes of death. However, it's important to note that this administrative dataset contains all deaths registered during 2021, which doesn't necessarily mean that all of them occurred in 2021. It includes deaths that occurred before 2021 and might miss other deaths that occurred in 2021. Therefore, while the INEGI dataset provides valuable information, the data obtained from the Mexican Secretariat of Health has been retained as the leading source of official COVID-19-related deaths.

4.5 Measurements

4.5.1 Intimate Partner Violence

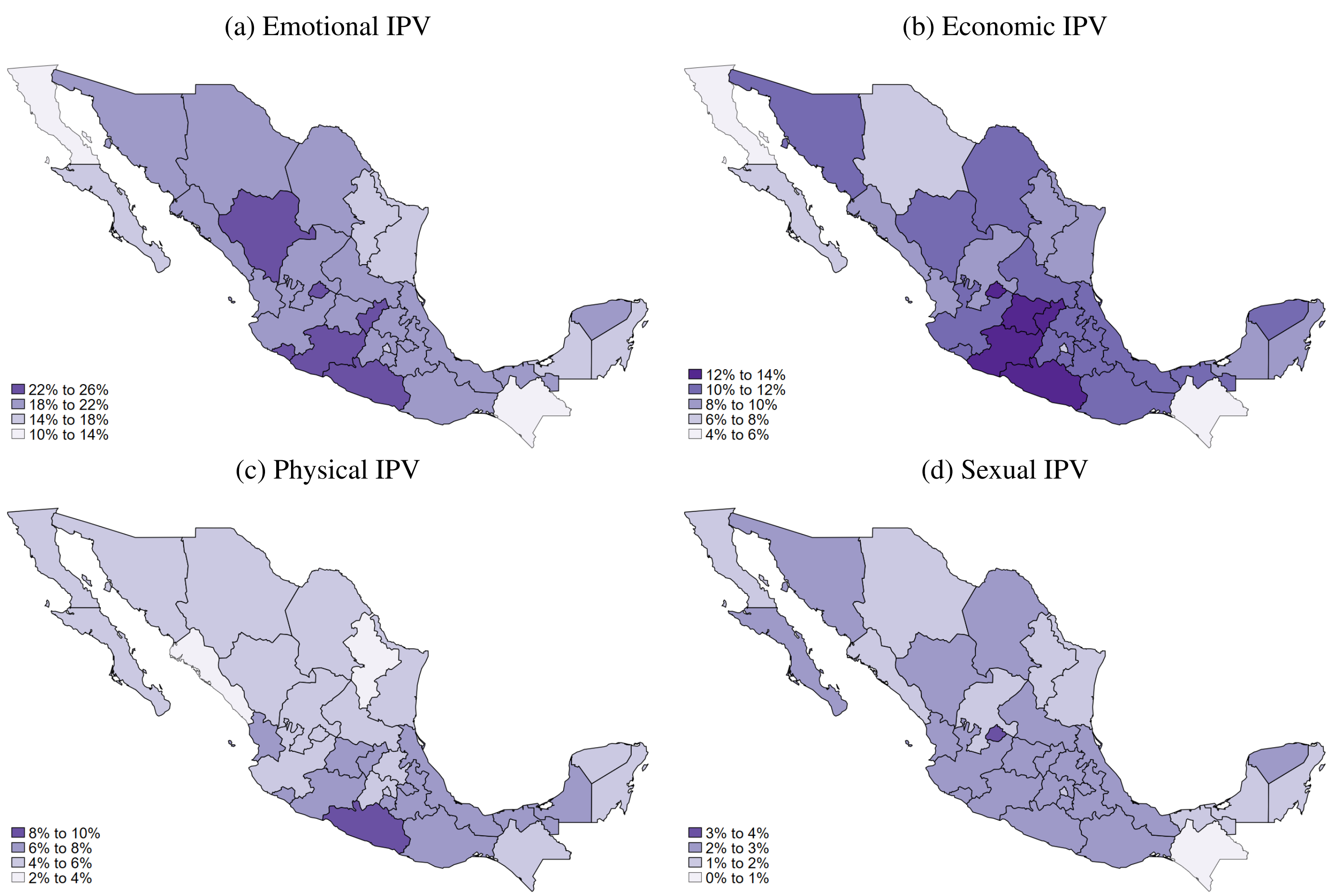
The Mexican National Survey of Households Relationships 2021 (ENDIREH, 2021) incorporates a collection of 38 questions pertaining to experiences of violence involving any partner or ex-partner, which occurred from October 2020 onward.[11] These questions are categorized into the four primary domains of Intimate Partner Violence (IPV): physical, emotional, sexual, and economic violence. The Appendix section 4.D.1 provides a detailed list of all 38 questions included in the survey, organized by each category. As illustrative examples of *physical violence*, the survey includes questions such as "Has your partner/ex-partner pushed you?" or "Has your partner/ex-partner slapped you?". For *emotional violence* the survey encompasses questions such as "Has your partner/ex-partner shamed you or humiliated you?". In terms of *sexual violence*, the ENDIREH 2021 survey features, among 6 other questions, "Has your partner/ex-partner threatened or blackmailed you to have sexual intercourse?". Lastly, questions related to *economic violence* tactics include "Has your partner/ex-partner forbidden you to work or study?" or "Has your partner/ex-partner spent your money without your consent?".

For each of these questions, respondents can provide answers indicating whether they experienced the specific violent tactic "one time, few times, several times, or never". For the first part of the analysis, the variable Intimate Partner Violence (IPV) is represented as a binary variable: $IPV = 1$ if the respondent answers "one time", "a few times" or "several times" to any of the 38 questions that embed the four forms of IPV, and $IPV = 0$ if the respondents answer "never" to every question. In aggregate, 22.5% of the sample has experienced at least one form of IPV from October 2020 until November 2021. In particular, emotional violence emerges as the

[11] The survey was conducted during October and November of 2021.

most prevalent form of violence, affecting 20.13% of the respondents, encompassing nearly 90% of the overall IPV. This form of violence is far followed by economic violence (8.95%), physical violence (5.64%), and sexual violence (2.24%).

Figure 4.4: Forms of Intimate Partner Violence



The chart displays the distribution of percentages for each form of Intimate Partner Violence (IPV), which includes physical, emotional, sexual, and economic violence in each Mexican state. These experiences of violence occurred during the period from October 2020 to November 2021.

Figure 4.4 provides an overview of the geographic disparities among the different types of violence. In general, Guerrero and Querétaro stand out as the states with the highest rates of IPV, while Chiapas and Baja California exhibit the lowest rates. Regarding emotional violence (Figure 4.4 (a)), the majority of the states fall within a range of percentages between 18% and 22%. However, the states of Guerrero, Michoacan, Colima, Querétaro, Aguascalientes and Durango report percentages exceeding 22%, indicating higher prevalence in these regions. These same states also exhibit elevated percentages concerning economic violence (Figure 4.4 (b)), and physical violence (Figure 4.4 (c)) with Guerrero notably exceeding in both categories.[12] Finally, in terms of sexual violence (Figure 4.4 (d)), although most states report percentages between 2% and 3%, Aguas-

[12]The state of Guerrero reports 13% of economic violence and 8.2% of physical violence experiences in the 12 months preceding October 2021.

calientes stands out with a higher rate of 3.04%.

In the second part of the analysis, I created an index variable from the 38 questions about IPV included in ENDIREH 2021. This variable assigns a score of 1 unit for each *one time* response, two points for *few times*, three points for *several times*, and 0 points for *never* answers to each question. This index variable reflects the frequency of partner/ex-partner abuse, with higher values indicating more frequent abuse. It has a Cronbach's alpha coefficient of 0.92.[13] I applied this scoring method separately for the four forms of IPV. Table 4.1 summarizes the frequencies for all forms of IPV. Overall, among the 17,487 women in the ENDIREH 2021 sample who had experienced at least one tactic of violence, the average frequency of violence was 8.5. This could result from various different tactics or from a few tactics being used repeatedly.[14] Consistent with the existence of violence, emotional violence is the most frequently reported by victims, while sexual violence is the least frequent. However, the gap in frequency between economic and physical violence is much narrower than found in previous studies.

Table 4.1: Intensity of IPV

	Mean	Std.Dev.	Min	Max	Obs
Overall IPV	8.45	11.09	1	113	17,487
Physical IPV	1.12	2.66	0	27	17,487
Emotional IPV	5.27	6.18	0	44	17,487
Sexual IPV	0.42	1.65	0	21	17,487
Economic IPV	1.65	2.97	0	21	17,487

The table presents the distribution of violence frequency for each form of Intimate Partner Violence (IPV), including physical, emotional, sexual, and economic violence, as well as the overall IPV. The data is based on women who have experienced at least one form of violence. Source: ENDIREH 2021.

4.5.2 Age-Standardized Mortality Rate

While studies like (Hong et al., 2021; Lersch and Hart, 2022) incorporate the incidence (positive COVID-19 cases) or the case fatality rates (ratio of COVID-19 deaths to the COVID-19 incidence) measures as well, this analysis specifically focuses on mortality rates (i.e., the COVID-19 related deaths). This decision is driven by several key reasons. Firstly, the mortality rate metric offers a more nuanced perspective on epidemic surveillance, shedding light on pre-existing inequalities in health infrastructures and their capacity to mitigate fatal outcomes of the virus. Secondly, the count of COVID-19 deaths played a pivotal role in the implementation of

[13]The Cronbach's alpha coefficient measures the internal consistency of a set of survey questions, indicating whether these questions measure the same characteristic. High Cronbach's alpha values, typically above 0.7, suggest that the respondents' answers are consistent(Frost, 2023)

[14]For example, if a victim shows a frequency of IPV with a value of 6, it could mean that the perpetrator used 6 different violent tactics *one time*, or used 2 violent tactics *several times*, and so on.

traffic light policies, which, in turn, led to lockdown measures with economic and social consequences.

Therefore, the Age-standardized Mortality Rate (ASMR) serves as the key indicator to measure the severity of the COVID-19 virus's impact across various regions (WHO, 2023). As defined by the World Health Organization, the ASMR is a weighted average of mortality rate for each age group, with the weights determined by the proportion of individuals corresponding to each age group. This measure is utilized to compare mortality rates across different entities, be it countries (Hong et al., 2021; Mattiuzzi et al., 2023; Moradzadeh et al., 2021), or municipalities (Masferrer et al., 2022) by adjusting for age-related variations in the population age distribution. The calculation of the ASMR is represented as follows:

$$ASMR_m = 10,000 \times \sum_{k=1}^{14} w_k \frac{d_{k,m}}{p_{k,m}} \tag{4.1}$$

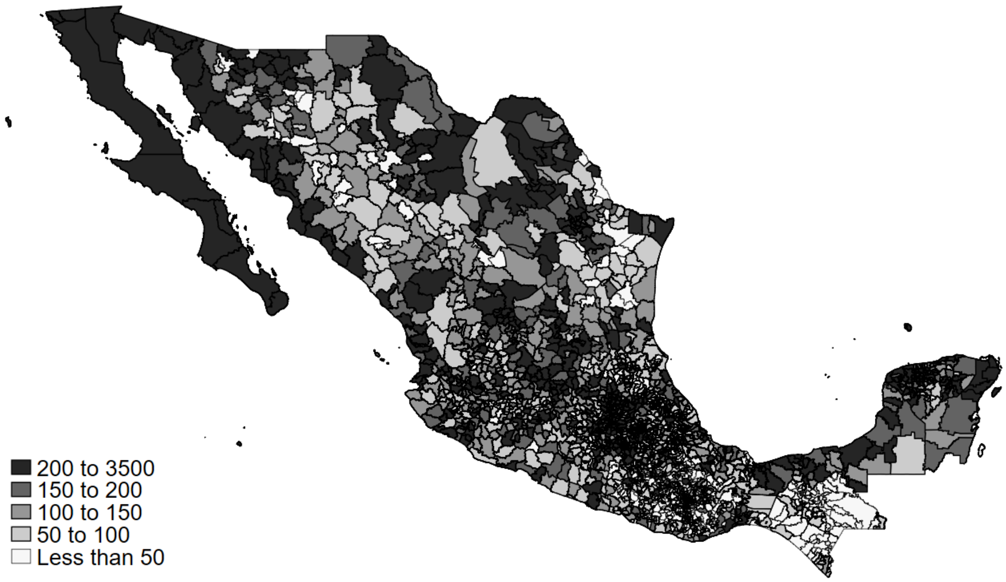
where $ASMR_m$ represents the age-standardized mortality rate, for municipality m, and k corresponds to the 14 age groups.[15] w_k signifies the proportion of the standard population of Mexico within the k^{th} age group, $d_{k,m}$ denotes the number of deaths in the k^{th} group within municipality m, and $p_{k,m}$ reflects the population in the k^{th} age group of municipality m. I estimate the age standard population using the Mexican official population projections of 2021 (Conapo, 2018). Using the age standard population of Mexico as the population reference allows for comparisons of COVID-19 severity - mortality rates - across various regions of the country, either at the municipal or state level (Hong et al., 2021). Figure 4.5 illustrates the regional differences in ASMR from the date of the first recorded COVID-19-related death (March 18, 2020) through the end of March 2023. It is important to note the heterogeneity among the Mexican municipalities. Given the concerns regarding the potential under-reporting of COVID-19 deaths by the Mexican Secretariat of Health (Knaul et al., 2021), the map presented in Figure 4.5 enhances the robustness of the analysis. Despite the visible heterogeneous ASMR distribution, it is notable that the northwest and central regions of the country were the most affected areas, in stark contrast to the relatively unaffected region around the state of Chiapas.

4.5.2.1 Excess of Mortality

As mentioned in the Data section, concerns have been raised about the transparency of COVID-19 death data reported by the Mexican Secretariat of Health (Knaul et al., 2021). Therefore, I obtained official death data for the year 2021 from the Mexican Institute of Statistics (INEGI). This dataset covers various recorded causes of death, including COVID-19 and the COVID-19 vaccine-related deaths. Additionally, it provides information on comorbidities, which are underlying conditions highly correlated with COVID-19. In fact, another branch of the literature highlights the strong correlation between the risk of COVID-19 death with pre-existing comorbidities among the Mexican population (Calixto-Calderón et al., 2021; Pamplona, 2020; Singer, 2020).

[15] These age groups consist of 13 intervals from 0 to 65 years, in 5-year increments, and a final group encompassing ages from 65 to 99 years.

Figure 4.5: COVID-19 Death rates



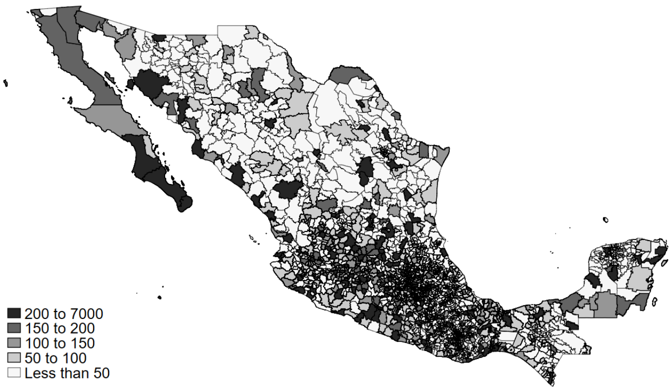
The map depicts the Age-Standardized Mortality Rates in each Mexican municipality, covering the period from the onset of the pandemic until March 31st, 2023. The values in the legend should be interpreted as X cases per 100,000 population (for example, 200 deaths per 100,000). Source: Mexican Health of Secretariat (CONACYT).

Therefore, from the INEGI 2021 death dataset, I created the following three variables. The first variable identifies individuals whose cause of death was registered as COVID-19 or COVID-19 vaccine-related. This variable accounts for directly related deaths resulting from the virus. The second variable includes individuals whose cause of death was registered as either COVID-19 or any of the following comorbidities: asthma, cardiovascular disease, Chronic Kidney disease, Diabetes mellitus, Chronic Obstructive Pulmonary Disease, Immunosuppression, Hypertension, or obesity (Calixto-Calderón et al., 2021). This variable controls for the possibility of a death caused by COVID-19 that may not have been registered as such or deaths strongly attributed to COVID-19. Finally, the third variable encompasses all deaths, regardless of the cause, to account for those indirect deaths that the pandemic circumstances may have caused. For instance, these indirect deaths could result from factors such as the shortage of facilities and overfull of patients, the delayed detection of other diseases, or postponed surgeries, leading to irreversible harm to patients. Since data on the municipality and age of the deceased are available, all variables are aged-standardized using the Equation (4.1).

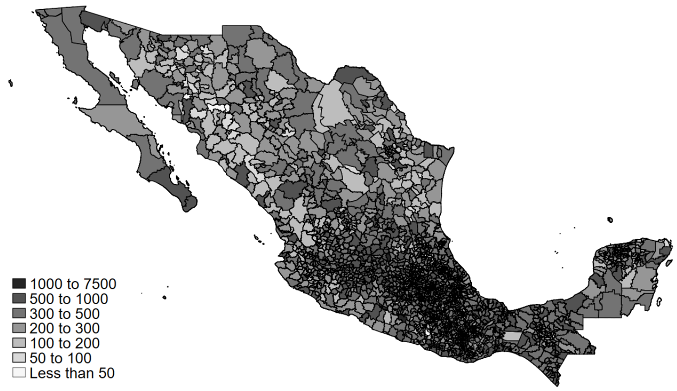
Figure 4.6 displays the Mexican maps depicting the different categories of registered deaths from INEGI 2021 - (a) COVID-19-related deaths, (b) Comorbidities-related deaths, and (c) all deaths. To facilitate a comparison with Figures 4.5 and 4.6 (a), the scales for COVID-19 deaths have been kept consistent. Figure 4.5 shows higher overall mortality rates across the entire country compared to those depicted in Figure 4.6 (a). It's worth noting that the ASMR data from the Mexican Secretariat of Health covers all COVID-19-related deaths from 2020 to March 2023, whereas the INEGI death data covers the registered deaths in 2021. This time discrepancy may explain the differences in the data. However, the affected areas remain consistent between the

Figure 4.6: Deaths Registries per Municipality

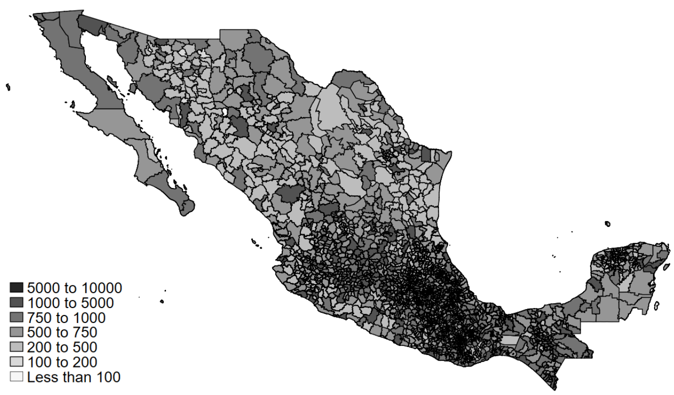
(a) COVID-19



(b) COVID-19 and Cormobidility



(c) All Deaths



The map illustrates the Age-Standardized Mortality Rates for COVID-19, COVID-19 and Comorbidity, and all deaths in each Mexican municipality during the registered year 2021. The values in the legend should be understood as X cases per 100,000 population (for instance, 200 deaths per 100,000). Source: Administrative Death data from INEGI.

two datasets.[16]

The other two maps in Figures 4.6 (b) and (c) illustrate that excess mortality is also associated with COVID-19 deaths. Comorbidity-related deaths in 2021 exhibit a similar pattern to COVID-19-related deaths, with higher prevalence in states where fewer COVID-19 deaths are registered. Lastly, Figure 4.6 (c) representing all registered deaths in 2021, reveals that the southern region of the country generally has higher mortality rates than the northern region. These variations across municipalities depicted in Figure 4.6 (a), (b) and (c) provide a robust basis for the analysis.

[16]Indeed, the consistency between the two datasets provides robust evidence that 2021 experienced higher mortality rates during the pandemic, indicating it was one of the most severe periods of the outbreak.

4.5.3 Individual and Regional Characteristics

Finally, the analysis incorporates control variables at both the individual and regional levels. From the same ENDIREH 2021 survey, I collect demographic characteristics such as marital status, age bracket, educational attainment, ethnicity, residence in a rural area, and the presence of children in the household are collected. Table 4.2 provides an overview of the distribution of each variable. A significant proportion of women in the sample are aged between 25 and 44 years old (45%) and 45 to 64 years old (30%). More than half of the respondents have educational levels up to secondary education, including those who completed secondary studies or preparatory education, while 21% pursued graduate studies. Moreover, a substantial majority of the surveyed sample is married (almost 70%), has at least one child alive (over 80%) and resides in urban areas, accounting for more than three-quarters of the sample.

Table 4.2: Individual Characteristics

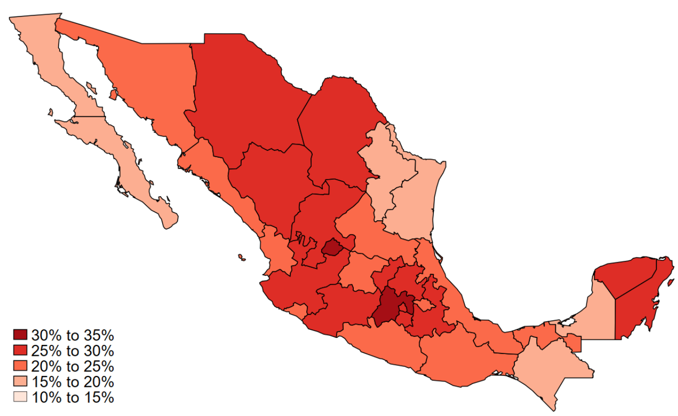
	Percentage
Age Group	
15-24	15.30%
25-44	44.91%
45-64	30.41%
65-97	9.38%
Education Level	
Primary	24.77%
Secondary	54.23%
University	21.01%
Marital Status	
Married	69.92%
Separated or Divorced	13.39%
Single	16.69%
Number of Children	
0 Children	18.53%
1 Child	15.49 %
2 Children	25.83%
3 Children	20.92%
More than 4 Children	19.23%
Rural Area	24.43%
Observations	77,704

Distribution of individual characteristics within the sample. All variables are categorized. Source: ENDIREH 2021.

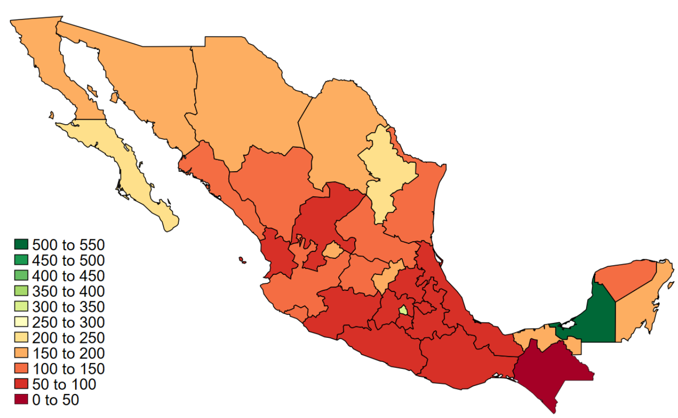
This analysis includes regional characteristics to provide valuable insights into the potential relationships between regional socioeconomic conditions and prior Intimate Partner Violence (IPV) rates. Therefore, the analysis incorporates the average percentages of IPV in 2016 and the GDP per capita in 2019 for each Mexican state. Both of these indicators were sourced from the Institute of Statistics of Mexico (INEGI). Although the data is collected at the state level, Figure 4.7 illustrates that there are notable regional differences within the country. Firstly, regarding the percentage of IPV in 2016 (Figure 4.7 (a)), the north-central region of the country exhibits higher percentages of violence, particularly in states like Aguascalientes and Estado de Mexico. Additionally, in terms of 2019's GDP per capita (Figure 4.7 (b)), the northern states tend to have higher GDP figures compared to the southern states, with the exception of Campeche (coloured in Figure 4.7 (b) in green), which ranks as the richest state in the country. For instance, the correlation between lower GDP per capita and higher registered deaths in 2021 (Figure 4.6 (c)) underscores the pre-existing socioeconomic inequalities that affected the health system's capacity to mitigate excess mortality during the pandemic. Moreover, by comparing both maps (a) and (b), it can be inferred that prior to the pandemic, the wealthiest and poorest states of Mexico generally had lower rates of IPV. However, states with GDP per capita values closer to the median are the ones with higher percentages of IPV. This suggests an inverted U-shaped relationship between state GDP and IPV.

Figure 4.7: Regional Characteristics

(a) Percentage of IPV in 2016



(b) GDP per capita in 2019



Regional Characteristics of the 32 States of Mexico. Map (a) shows the percentage of Intimate Partner Violence (IPV) experienced in each state in the year 2016. Map (b) illustrates the disparities in GDP per capita among the Mexican states. Source: INEGI.

4.5.3.1 Individual and Regional Characteristics and IPV

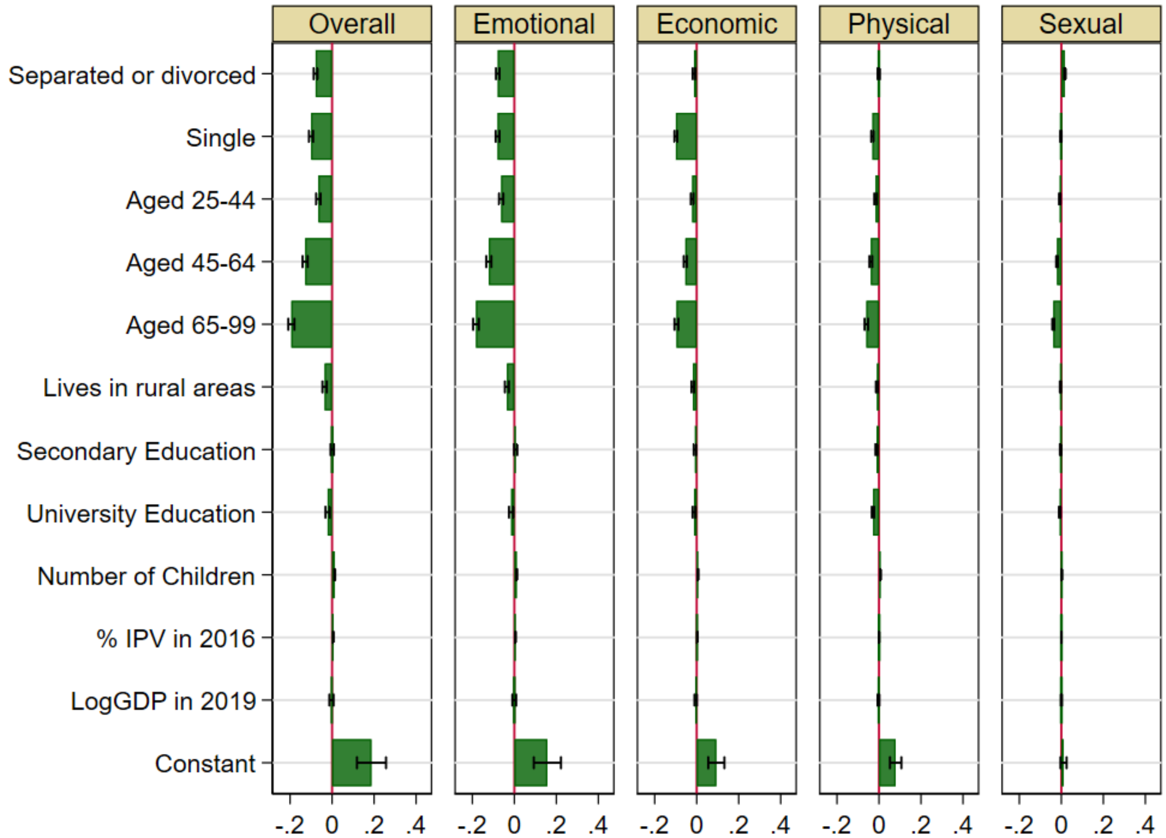
Conducting an OLS regression analysis to assess the association between various forms of Intimate Partner Violence (IPV) and individual and state-level characteristics is a valuable preliminary step. This approach allows for a comprehensive understanding of the key determinants that correlate with different forms of IPV and their relationships with individual and regional characteristics in 2021. By controlling for specific population segments that may be more vulnerable to experiencing IPV, regardless of how their vulnerability evolved during the pandemic, this analysis provides a more nuanced perspective on the complex factors contributing to the prevalence of IPV. Following Equation 4.2, IPV_{is} represents the type of IPV (overall, physical, emotional, sexual or economic) for each individual i of the sample who lives in state s. X_i are the individual characteristics of the individual i, and R_s are the regional characteristics of each state s.

$$IPV_{is} = \alpha_0 + \alpha_1 X_i + \alpha_2 R_s + u_{is} \tag{4.2}$$

The results presented in Figure 4.8 shed light on the initial findings (Table D1 offers a detailed view of each coefficient's magnitude). Focusing on the column representing Overall IPV, it is noticeable that separated and single women are significantly less likely to experience IPV compared to their married counterparts. Furthermore, there is an inverse relationship between age and the likelihood of experiencing any type of violence. Women with a university degree, in comparison to those with only primary education, exhibit a slightly lower susceptibility to becoming victims of IPV. Similar results are found for those women living in rural areas. Additionally, an increase in the number of children is associated with a higher likelihood of experiencing IPV. Interestingly, the association between IPV in 2016 and IPV in 2021 presents a positive and significant coefficient, however, with a very low magnitude which is not visually discernible in Figure 4.8

When examining the four distinct forms of IPV, Figure 4.8 illustrates that women who are single, over 45 years old, and possess graduate-level education are significantly less likely to be victims of physical IPV. Moreover, the patterns for emotional and economic IPV closely resemble those of overall IPV, with the key distinction being that emotional IPV tends to have larger coefficients with respect to economic IPV. Notably, GDP per capita only exhibits a negative significant association with economic IPV. Finally, in stark contrast to the other types of violence, separated women are significantly more likely to experience sexual violence compared to their married counterparts.

Figure 4.8: IPV experience on Individual and Regional Characteristics



OLS regression of all forms of Intimate Partner Violence (IPV), including physical, emotional, sexual, and economic violence, as well as overall IPV, against individual and regional characteristics. The baseline categories include married women, those aged 15 to 24 years old, residents of urban areas, and those with only primary education.

4.6 Results

4.6.1 Severity of COVID-19 on the prevalence of IPV

The initial part of this analysis studies the association between the severity of COVID-19 across the different municipalities with the incidence of Intimate Partner Violence (IPV) during the year 2021. In order to achieve a normally distributed variable, the Age-Standardized Mortality Rate (ASMR) is log-transformed before being integrated into the regression analysis. The OLS regression models, as outlined in Equation (4.3), follow a progressive approach. Initially, the IPV experienced by individual i within municipality m is regressed against the logarithm of the ASMR for the same municipality with no additional controls. Subsequently, the second model incorporates individual-level controls denoted as X_i. Finally, the third and last model extends the analysis by including state-level regional controls, labelled R_s, for each state within the country.

$$IPV_{im} = \beta_0 + \beta_1 log(ASMR_m) + \beta_2 X_i + \beta_3 R_s + u_{im} \tag{4.3}$$

Table 4.3 provides an overview of the results obtained from the three models as previously described. Each column corresponds to one of the models described above. In Column (1), the results of Equation 4.3 are presented, reflecting the simple regression model that examines the association between the probability of experiencing IPV and the Log-ASMR, without accounting for any control variables. Column (2) presents the results of Equation 4.3, where the analysis includes controls for individual characteristics. Notably, the results of Column (2) consistently demonstrate a significant positive association between IPV and the logarithm of ASMR, albeit with a relatively higher coefficient magnitude compared to the results in Column (1). Finally, Column (3) showcases the outcomes when both individual and regional controls are included. The resulting coefficient magnitude is slightly smaller than the one in Column (2). Therefore, it is important to consider the potential influence of omitted variable bias when introducing individual-level controls. Despite this, the results in Column (3) indicate that individuals residing in municipalities at the 25^{th} percentile of the logarithm of ASMR would have encountered a 0.015 increase in the likelihood of experiencing IPV if they resided in municipalities with a 75^{th} percentile of the logarithm of ASMR.

Table 4.3: Experience of Overall IPV

	(1)	(2)	(3)
Log Age-Standardized Mortality Rate	0.0154***	0.0228***	0.0219***
	(0.00428)	(0.00464)	(0.00465)
Individual Characteristics	NO	YES	YES
Regional Characteristics	NO	NO	YES
Constant	0.141***	0.196***	0.134***
	(0.0218)	(0.0241)	(0.0337)
Observations	77314	77314	77314
Adjusted R^2	0.001	0.021	0.024

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OLS regressions with Intimate Partner Violence (IPV) as the dependent variable. Each column represents a different regression model with the inclusion of controls (i.e., in column no controls are included; in column 2 individual characteristics are included as controls; in column 3 both individual and regional characteristics are included as controls). Individual characteristics include marital status, age group, rural settlements, education level, and number of children. Regional characteristics include state percentage of IPV in 2016, and state GDP per capita in 2019.

The natural progression in this research involves a more detailed examination of the distinct forms of Intimate Partner Violence (IPV) and their correlations with COVID-19 mortality rates. The 2021 ENDIREH survey classifies the different IPV questions into four different forms: physical, emotional, sexual and economic IPV. The aim of this analysis is to identify which form of violence is the most prevalent and explore its association with the severity of COVID-19. To achieve this, I employ an OLS model, as expressed in Equation (4.4). In this model, I regress $IPVform_{im}$, which represents each form of IPV (physical, emotional, sexual, and economic)for individual i in municipality m, against the logarithm of Age-Standardized Mortality Rate (ASMR) specific to each municipality. Furthermore, I account for individual characteristics (X_i) and state-level (R_s) characteristics as control variables.

$$IPV form_{im} = \beta_0 + \beta_1 log(ASMR_m) + \beta_2 X_i + \beta_3 R_s + u_{im} \tag{4.4}$$

Table 4.4 presents the results obtained from the regressions outlined in Equation 4.4. Each column corresponds to each form of violence. Emotional and economic violence exhibit significant positive associations with ASMR, both achieving p-values below 0.01. Intriguingly, sexual violence also demonstrates a significant positive correlation, albeit at a 10% significance level, while physical violence shows a positive association but lacks statistical significance. Therefore, these findings suggest that the mortality rates of the COVID-19 pandemic may have contributed to a slight increase in the prevalence of the most common forms of IPV, specifically emotional and economic violence, along with sexual violence. However, there is no significant evidence that relates the pandemic to the physical form of violence. This implies that perpetrators may have resorted to tactics that do not carry the risk of hospitalisation and viral infection.

Table 4.4: Experience of the distinct types of IPV

	Physical	Emotional	Sexual	Economic
Log Age-Standardized Mortality Rate	0.00189	0.0209***	0.00166*	0.00893***
	(0.00185)	(0.00431)	(0.000931)	(0.00286)
Individual Characteristics	YES	YES	YES	YES
Regional Characteristics	YES	YES	YES	YES
Constant	0.0739***	0.107***	0.00732	0.0736***
	(0.0145)	(0.0315)	(0.00752)	(0.0196)
Observations	77314	77314	77314	77314
Adjusted R^2	0.010	0.022	0.005	0.022

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OLS regressions with Intimate Partner Violence (IPV) as the dependent variable. Each column represents a distinct form of violence: physical, emotional, sexual and economic violence. Individual and regional characteristics are included in every regression. Individual characteristics include marital status, age group, rural settlements, education level, and number of children. Regional characteristics include state percentage of IPV in 2016, and state GDP per capita in 2019.

4.6.2 Frequency of Intimate Partner Violence

This section of the analysis aims to assess whether the severity of COVID-19 is linked to changes in the frequency of IPV among those who were already experiencing IPV. Evidence in Peitzmeier et al. (2022) indicates that, while the lockdown period didn't lead to new IPV conflicts, families with pre-existing IPV experienced an escalation in both, frequency and severity. The preceding analysis demonstrated a positive correlation between IPV and COVID-19 mortality rates, significant with all forms except physical violence. Consequently, this section of the analysis aims to explore whether, in addition to the observed positive correlation between IPV and log-ASMR, victims of IPV saw an uptick in the frequency of these violent incidents. Furthermore, this investigation seeks to identify which specific form of IPV becomes more prevalent in light of COVID-19-related deaths.

$$IPVintensity_{im} = \beta_0 + \beta_1 log(ASMR_m) + \beta_2 X_i + \beta_3 R_s + u_{im} \tag{4.5}$$

Therefore, within the sample of women who experienced at least one act of IPV during the previous year, I conduct an OLS regression model as outlined in Equation (4.5). In this model $IPVintensity_{im}$ represents the frequency of general IPV and its four specific types - overall, physical, emotional, sexual, economic - for individual i in municipality m, regressed agaisnt the log-ASMR and individual and regional control variables. The results presented in Table 4.5 reveal a contrasting picture compared to the existence of IPV. The association between most forms of IPV is negative. However, emotional and economic IPV appears to have a positive association, although not statistically significant. Moreover, physical violence is significantly negative associated with log-ASMR, suggesting that in municipalities with a more severe impact of the COVID-19 virus, victims of violence experienced fewer physical aggressions. These results align with the previous findings, supporting the assumption that perpetrators may have refrained from using physical aggression to avoid possible hospitalization, and consequently, the risk of virus infection. Table D4 in the Appendix provides these results, including all control coefficients. It is noteworthy to mention, that in contrast to the IPV existence results, separated women experienced a higher frequency of all forms of IPV compared to their married counterparts.

4.6.3 Excess Mortality

The final stage of this research evaluates the excess mortality resulting from COVID-19. Apart from the uncertainties surrounding the reported cases by the Mexican Secretariat of Health, there are other indirect factors that may have contributed to the fatal outcomes during the pandemic. These factors include hospital closures, which prioritized COVID-19 patients, and lockdown measures, which limited mobility, exercise, and access to fresh air, among others. These secondary consequences of COVID-19 could, in turn, have affected the likelihood of the existence of Intimate Partner Violence (IPV). These types of deaths are classified into three broad groups, each group being a subset of the next: direct COVID-19 deaths, deaths resulting from COVID-19 or comorbidity-related health issues, and all registered deaths in general. The OLS regression model employed for this analysis is described as follows:

$$IPV_{im} = \beta_0 + \beta_1 log(DeathType_m) + \beta_2 X_i + \beta_3 R_s + u_{im} \tag{4.6}$$

Table 4.5: Frequency of the distinct types of IPV

	Overall	Physical	Emotional	Sexual	Economic
Log Age-Standardized Mortality Rate	-0.0689	-0.148***	0.0822	-0.0101	0.00746
	(0.181)	(0.0461)	(0.0971)	(0.0268)	(0.0501)
Individual Characteristics	YES	YES	YES	YES	YES
Regional Characteristics	YES	YES	YES	YES	YES
Constant	6.148***	1.943***	2.640***	0.151	1.415***
	(1.167)	(0.326)	(0.687)	(0.184)	(0.335)
Observations	17404	17404	17404	17404	17404
Adjusted R^2	0.045	0.036	0.035	0.018	0.070

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OLS regressions with the frequency of Intimate Partner Violence (IPV) as the dependent variable. Each column represents a distinct form of violence: overall, physical, emotional, sexual and economic violence. Individual and regional characteristics are included in every regression. Individual characteristics include marital status, age group, rural settlements, education level, and number of children. Regional characteristics include state percentage of IPV in 2016, and state GDP per capita in 2019.

where IPV_{im} indicates whether a woman i residing in municipality m has experienced Intimate Partner Violence (IPV) in the past year, and $log(DeathType_m)$ denotes the natural logarithm of the type of Age-Standardized Death Rate in municipality m (comprising deaths from COVID-19, deaths from COVID-19 or comorbidities, or all death). Table 4.6 presents the outcomes for each regression specified in Equation (4.6). Despite the difference in magnitudes observed, the association between the incidence of IPV and the different recorded deaths consistently maintains its positive and statistically significant relationship. The dissimilarity in magnitude might be attributed to differences in the timing of data recording between the Mexican Secretariat of Health database and the National Institute of Statistics database, as mentioned in the measurement section. Nevertheless, it is crucial to highlight that these findings affirm the robustness of the influence of the COVID-19 outbreak on the prevalence of IPV in Mexico.

Table 4.6: Experience of IPV: Robustness Checks on the Type of Death

	ASMR	COVID-19	COVID-19 and Comorbidity	All deaths
Log Death Type	0.0219***	0.00656***	0.0158***	0.0137**
	(0.00465)	(0.00249)	(0.00598)	(0.00669)
Individual Characteristics	YES	YES	YES	YES
Regional Characteristics	YES	YES	YES	YES
Constant	0.134***	0.166***	0.0785	0.0831
	(0.0337)	(0.0365)	(0.0524)	(0.0581)
Observations	77314	74448	75387	75387
Adjusted R^2	0.024	0.024	0.024	0.024

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

OLS regressions with Intimate Partner Violence (IPV) as the dependent variable. Each column represents a distinct death measurement: Column 1 represents the Log Age-Standardized Mortality Rate (ASMR) calculated by the Mexican Secretariat of Health. Column 2 represents the Log Age-Standardized COVID-19 deaths registered in the death administrative 2021 data (INEGI). Column 3 represents the Log Age-Standardized COVID-19 and Comorbidity deaths, and column 4 represents the Log Age-Standardized All death data. Individual and regional characteristics are included in every regression. Individual characteristics include marital status, age group, rural settlements, education level, and number of children. Regional characteristics include state percentage of IPV in 2016, and state GDP per capita in 2019. The variations in sample sizes are due to the different municipalities that account for registered deaths in each case.

4.7 Discussion

The objective of this study is to explore the relationship between COVID-19 severity and Intimate Partner Violence (IPV), with a primary contribution being the utilization of specialized survey data on gender-based violence. In contrast to the findings of previous studies like (Lersch and Hart, 2022; Hoehn-Velasco et al., 2021a; Valencia-Londoño and Nateras-González, 2020), this research's findings demonstrate that in municipalities where the COVID-19 virus had a higher mortality rate, there were more incidences of IPV. As pointed out by Silverio-Murillo et al. (2023), the nature of the IPV data plays a crucial role in determining the results. Previous studies relied on IPV data from police reports or police arrests, which may not fully capture the reality experienced by victims but rather reflect their willingness to take legal action against their perpetrators. In this context, it is not surprising that in a more protected environment, possibly created in the setting of a specialized victimization survey, women who are victims of violence become more inclined to acknowledge and report their experiences with IPV.

The positive and significant correlation between IPV and mortality rates can be explained by various mechanisms, with economic hardship and financial stress stemming from the COVID-19 outbreak playing a crucial role. Research by Masferrer et al. (2022) highlights the positive association between pre-existing economic vulnerabilities and the severity of the virus's impact on different Mexican municipalities. In areas where the virus's mortality rates were particularly high, emergency alerts were activated, leading to widespread safety measures such as the suspension of non-essential jobs, school closures, and increased care of vulnerable individuals, as noted by Hoehn-Velasco et al. (2021b). Consequently, areas with a higher number of COVID-19-related deaths faced a greater risk of employment loss, as indicated by a 1% increase compared to areas with fewer deaths. This collective economic burden and uncertainty in households increased the likelihood of conflict among couples (Fox et al., 2004), providing an explanation for the observed positive correlation between IPV incidence and COVID-19 mortality rates.

Furthermore, this research offers insights into the specific types of violence that increased during the pandemic. The results indicate significant correlations with emotional, economic, and sexual violence, but not physical violence. This trend remains consistent when considering the frequency of IPV. In areas with higher COVID-19 mortality rates, there was a decrease in the frequency and use of physically abusive tactics. Several plausible explanations for these findings have been suggested (Jetelina et al., 2021). First, the reduced physical victimization may result from perpetrators being less inclined to exert violence which may lead victims to seek medical attention during the pandemic. Second, it's possible that perpetrators, satisfied with the level of control they impose, may not feel the need to employ as many physically abusive tactics. This theory aligns with the evidence that separated women experienced a higher frequency of all forms of IPV compared to their married counterparts. Additionally, as proposed by Piquero et al. (2020) snd Nix and Richards (2021), victims may have initially experienced more aggressive IPV during the early weeks of lockdown but adjusted their behaviour to maintain a more pleasant environment, knowing they had limited escape options. However, separated women may suffer escalated aggressions when perpetrators perceive a loss of control over them (Ornstein and Rickne, 2013; Rezey, 2020). Further research could delve into the distinct mechanisms through which the pandemic affected IPV victims based on their marital status. This could help identify policies and interventions that protect women regardless of their relationship status.

4.7.1 Limitations

This research has some limitations that should be considered. Firstly, there are concerns regarding the accuracy and access to COVID-19 reports from the Mexican Secretariat of Health. Criticisms have been raised regarding potential data manipulation to avoid changes in the emergency status, which might have led to unreported COVID-19-related fatalities. Moreover, the official website does not permit the downloading of the COVID-19 public data for researchers outside Mexico. Consequently, the only way to collect the COVID-19 death data was by manually copying the numbers from the interactive CONACYT webpage. The dataset available only included data updated to the day of the research, preventing the clustering of deaths around specific dates relevant to the study. Nonetheless, Figure 4.5 illustrates significant data heterogeneity, and when comparing results with the administrative death database from the National Institute of Statistics (INEGI),[17] the substantial positive correlation persists, although potential omitted variable bias should be acknowledged.

The second limitation of this research has been the absence of 2019 or 2020 data on Intimate Partner Violence (IPV). While this research employs survey data that provides a more nuanced understanding of violence against women, there has been a lack of information about the IPV situation since 2016, resulting in a 5-year gap between the past two ENDIREH surveys. Although the aim is to identify possible mechanisms to comprehend the increase in IPV stemming from the COVID-19 outbreak, having more recent IPV data could assist in isolating the pandemic's causal impact on this form of violence and mitigating potential confounding factors.

4.8 Conclusion

The inadequate management of the SARS-CoV-2 pandemic placed Mexico among the countries with one of the highest recorded death tolls worldwide, with nearly 334,000 deaths by the end of the pandemic. The pre-pandemic socioeconomic disparities among the 32 states of the country played a significant role in the overall impact of COVID-19, resulting in a heterogeneous distribution of the virus's severity. Previous studies have explored the effects of these disparities on employment, mental health, domestic violence, and violence in public spaces. However, most of them focused on the initial two-month period of the national lockdown. This research makes a unique contribution to the literature by analyzing a longer time frame, encompassing

[17] This administrative dataset contains all deaths registered during 2021, which does not necessarily mean that all of them occurred in 2021. For instance, it includes deaths that occurred before 2021 and might miss other deaths that occurred in 2021.

the deadliest phase of the pandemic. Its objective is to examine whether there is an association between the severity of COVID-19 and the likelihood of experiencing intimate partner violence (IPV). Additionally, this study utilizes survey data, which provides a more accurate reflection of victims' experiences compared to police records or arrests. The analysis reveals a positive correlation between the severity of COVID-19 and the incidence of IPV, particularly emotional, economic, and sexual IPV. This finding is likely explained by the fact that areas with higher COVID-19 death counts experienced more extensive government measures, such as economic stagnation and job losses, which might heightened tensions within couples. Furthermore, the analysis shows that in areas most affected by the virus, victims experienced less frequent physical aggression. Some researchers interpret this result as perpetrators being deterred by the fear of sending their victims to hospitals and exposing them to COVID-19 infection.

4.D Appendix Chapter 4

4.D.1 Definition of variables

4.D.1.1 Suffered IPV the previous year

From October 2020, has your partner...

Physical violence:

- Pushed you or pulled your hair?
- Slapped you?
- Tied you up?
- Kicked you?
- Thrown an object at you?
- Beaten you with his hands or an object?
- Tried to choke you or suffocate you?
- Hurt you with a knife?
- Fired a weapon on you?

Emotional violence:

- Shamed on you, or humiliated you?
- Ignored you, didn't take you into account, or didn't give affection?
- Told you that you are cheating on him?
- Made you feel fearful?
- Threatened to leave you, hurt you, take your children?
- Kept you at home and prohibited you to leave or receive visits?

- Stalked you, spies on you?
- Constant calls and messages to know where are you?
- Threatened you with a weapon?
- Threatened to kill you, hurt you or your children?
- Destroyed, thrown away, or hidden your things away?
- Stopped talking to you?
- Demands your passwords and checks your mail and phone?
- Turned your children and relatives against you?
- Have been mad because the house chores are not perfect?

Sexual violence:

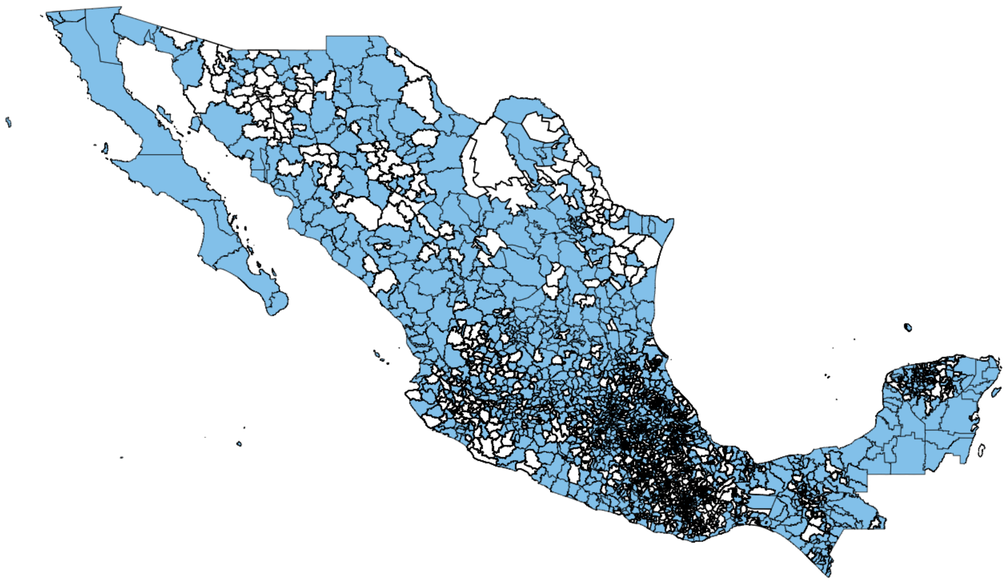
- Threatened or blackmailed to have sexual intercourse?
- Made you do things you do not like when you have sexual intercourse?
- Used physical force to make you have sexual intercourse?
- Forced you to watch porn?
- Forced sexual intercourse without protection?
- Sent messages with sexual content or insults with the aim of hurting
- Published information of you with the aim of hurting

Economic violence:

- Forbidden you to work or study?
- Taken over your money or spent it without your consent?
- Taken over your goods or properties?
- Spent the money you need to pay bills?
- Not given you money to pay bills or has threatened not to give you money for bills?
- Not given you money for home expenses, even if he has money?
- Demanded about how you spend your money?

4.D.2 Additional Graphs

Figure D1: Municipalities Represented in ENDIREH 2021



Source: ENDIREH 2021.

4.D.3 Complete Regressions

Table D1: IPV experience on Individual and Regional Characteristics

	Overall	Physical	Emotional	Sexual	Economic
Married or co-living					
Separated or divorced	-0.0782***	-0.00122	-0.0791***	0.0169***	-0.0129***
	(0.00445)	(0.00248)	(0.00428)	(0.00159)	(0.00304)
Single	-0.0999***	-0.0316***	-0.0798***	-0.00206	-0.0991***
	(0.00478)	(0.00266)	(0.00459)	(0.00171)	(0.00327)
Aged 15-24					
Aged 25-44	-0.0652***	-0.0167***	-0.0627***	-0.00741***	-0.0216***
	(0.00493)	(0.00274)	(0.00473)	(0.00176)	(0.00337)
Aged 45-64	-0.128***	-0.0390***	-0.122***	-0.0206***	-0.0540***
	(0.00549)	(0.00306)	(0.00528)	(0.00196)	(0.00376)
Aged 65-99	-0.194***	-0.0603***	-0.183***	-0.0384***	-0.0962***
	(0.00746)	(0.00415)	(0.00717)	(0.00267)	(0.00510)
Lives in rural areas	-0.0354***	-0.00998***	-0.0355***	-0.00401***	-0.0186***
	(0.00366)	(0.00203)	(0.00351)	(0.00131)	(0.00250)
Primary Education					
Secondary Education	0.000842	-0.0120***	0.00615	-0.00338**	-0.00832***
	(0.00407)	(0.00226)	(0.00391)	(0.00146)	(0.00278)
University Education	-0.0209***	-0.0289***	-0.0154***	-0.00753***	-0.0125***
	(0.00512)	(0.00285)	(0.00492)	(0.00183)	(0.00350)
Number of children	0.0128***	0.00752***	0.0115***	0.00293***	0.00677***
	(0.000996)	(0.000554)	(0.000957)	(0.000356)	(0.000681)
% IPV in 2016	0.00597***	0.000701***	0.00548***	0.000760***	0.00280***
	(0.000417)	(0.000232)	(0.000401)	(0.000149)	(0.000285)
logGDP	-0.00273	-0.00251	-0.000560	0.000391	-0.00487**
	(0.00312)	(0.00174)	(0.00300)	(0.00112)	(0.00214)
Constant	0.188***	0.0793***	0.157***	0.0106	0.0935***
	(0.0201)	(0.0112)	(0.0193)	(0.00718)	(0.0137)
Observations	77704	77704	77704	77704	77704
R^2	0.023	0.010	0.021	0.005	0.022

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D2: Experience of Overall IPV

	(1))	(2))	(3))
Log Aged-Standardized Mortality Rate	0.0154***	0.0228***	0.0219***
	(0.00428)	(0.00464)	(0.00465)
Married			
Separated or divorced		-0.0788***	-0.0790***
		(0.00475)	(0.00474)
Single		-0.0974***	-0.100***
		(0.00527)	(0.00505)
Aged 15-24			
Aged 25-44		-0.0665***	-0.0666***
		(0.00547)	(0.00542)
Aged 45-64		-0.131***	-0.131***
		(0.00653)	(0.00641)
Aged 65-99		-0.197***	-0.198***
		(0.00769)	(0.00761)
Lives in rural areas		-0.0235***	-0.0245***
		(0.00545)	(0.00518)
Primary Education			
Secondary Education		-0.00130	-0.00130
		(0.00410)	(0.00401)
University Education		-0.0251***	-0.0240***
		(0.00525)	(0.00523)
Number of children		0.0136***	0.0133***
		(0.00105)	(0.00104)
% IPV in 2016			0.00529***
			(0.000854)
logGDP in 2019			-0.0127**
			(0.00559)
Constant	0.141***	0.196***	0.134***
	(0.0218)	(0.0241)	(0.0337)
Observations	77314	77314	77314
Adjusted R^2	0.001	0.021	0.024

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D3: Experience of the distinct types of IPV

	Physical	Emotional	Sexual	Economic
Log Aged-Standardized Mortality Rate	0.00189	0.0209***	0.00166*	0.00893***
	(0.00185)	(0.00431)	(0.000931)	(0.00286)
Married				
Separated or divorced	-0.00104	-0.0798***	0.0169***	-0.0131***
	(0.00262)	(0.00432)	(0.00188)	(0.00322)
Single	-0.0315***	-0.0800***	-0.00206	-0.0992***
	(0.00276)	(0.00469)	(0.00160)	(0.00342)
Aged 15-24				
Aged 25-44	-0.0171***	-0.0640***	-0.00757***	-0.0225***
	(0.00292)	(0.00522)	(0.00172)	(0.00318)
Aged 45-64	-0.0397***	-0.124***	-0.0211***	-0.0554***
	(0.00340)	(0.00630)	(0.00197)	(0.00388)
Aged 65-99	-0.0609***	-0.187***	-0.0388***	-0.0982***
	(0.00423)	(0.00715)	(0.00230)	(0.00489)
Lives in rural areas	-0.00911***	-0.0253***	-0.00324**	-0.0140***
	(0.00246)	(0.00485)	(0.00139)	(0.00315)
Primary Education				
Secondary Education	-0.0123***	0.00399	-0.00360**	-0.00900***
	(0.00248)	(0.00393)	(0.00165)	(0.00288)
University Education	-0.0291***	-0.0185***	-0.00791***	-0.0136***
	(0.00280)	(0.00526)	(0.00197)	(0.00362)
Number of children	0.00758***	0.0120***	0.00298***	0.00706***
	(0.000609)	(0.00101)	(0.000310)	(0.000692)
% IPV in 2016	0.000654**	0.00480***	0.000698***	0.00248***
	(0.000299)	(0.000800)	(0.000168)	(0.000479)
logGDP in 2019	-0.00323	-0.0102**	-0.000430	-0.00918***
	(0.00240)	(0.00501)	(0.00129)	(0.00311)
Constant	0.0739***	0.107***	0.00732	0.0736***
	(0.0145)	(0.0315)	(0.00752)	(0.0196)
Observations	77314	77314	77314	77314
Adjusted R^2	0.010	0.022	0.005	0.022

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D4: Frequency of the distinct types of IPV

	Overall	Physical	Emotional	Sexual	Economic
Log Aged-Standardized Mortality Rate	-0.0689	-0.148***	0.0822	-0.0101	0.00746
	(0.181)	(0.0461)	(0.0971)	(0.0268)	(0.0501)
Married					
Separated or divorced	6.664***	1.066***	3.179***	0.658***	1.761***
	(0.440)	(0.109)	(0.236)	(0.0640)	(0.121)
Single	-0.926***	-0.162**	0.357**	0.186***	-1.308***
	(0.256)	(0.0685)	(0.156)	(0.0431)	(0.0474)
Aged 15-24					
Aged 25-44	0.0746	-0.0241	0.0570	0.00239	0.0394
	(0.281)	(0.0702)	(0.158)	(0.0371)	(0.0632)
Aged 45-64	-1.260***	-0.198**	-0.862***	-0.102**	-0.0988
	(0.315)	(0.0836)	(0.171)	(0.0458)	(0.0776)
Aged 65-99	-4.047***	-0.367***	-2.476***	-0.408***	-0.795***
	(0.419)	(0.118)	(0.225)	(0.0648)	(0.119)
Lives in rural areas	-0.477**	-0.103*	-0.306**	-0.0162	-0.0518
	(0.238)	(0.0588)	(0.130)	(0.0354)	(0.0626)
Primary Education					
Secondary Education	-0.532*	-0.384***	0.131	-0.124***	-0.155**
	(0.274)	(0.0678)	(0.146)	(0.0422)	(0.0673)
University Education	-1.623***	-0.747***	-0.403**	-0.233***	-0.240***
	(0.327)	(0.0714)	(0.190)	(0.0496)	(0.0748)
Number of children	0.429***	0.122***	0.173***	0.0440***	0.0895***
	(0.0602)	(0.0159)	(0.0316)	(0.00923)	(0.0155)
% IPV in 2016	0.0487**	-0.0113*	0.0368***	0.00560	0.0176**
	(0.0221)	(0.00625)	(0.0131)	(0.00343)	(0.00697)
logGDP in 2019	0.183	0.0256	0.190*	0.0252	-0.0577
	(0.191)	(0.0542)	(0.111)	(0.0314)	(0.0518)
Constant	6.486***	2.183***	2.675***	0.189	1.440***
	(1.261)	(0.352)	(0.724)	(0.195)	(0.350)
Observations	17404	17404	17404	17404	17404
Adjusted R^2	0.045	0.036	0.035	0.018	0.070

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D5: Experience of IPV: Robustness Checks on the Type of Death

	ASMR	COVID-19	COVID-19 and Comorbidility	All deaths
log_Death	0.0219***	0.00656***	0.0158***	0.0137**
	(0.00465)	(0.00249)	(0.00598)	(0.00669)
Married				
Separated or divorced	-0.0790***	-0.0804***	-0.0804***	-0.0804***
	(0.00474)	(0.00480)	(0.00475)	(0.00475)
Single	-0.100***	-0.102***	-0.101***	-0.101***
	(0.00505)	(0.00516)	(0.00512)	(0.00513)
Aged 15-24				
Aged 25-44	-0.0666***	-0.0637***	-0.0635***	-0.0634***
	(0.00542)	(0.00558)	(0.00553)	(0.00554)
Aged 45-64	-0.131***	-0.128***	-0.127***	-0.127***
	(0.00641)	(0.00653)	(0.00644)	(0.00646)
Aged 65-99	-0.198***	-0.194***	-0.193***	-0.193***
	(0.00761)	(0.00772)	(0.00761)	(0.00766)
Lives in rural areas	-0.0245***	-0.0304***	-0.0309***	-0.0317***
	(0.00518)	(0.00539)	(0.00533)	(0.00530)
Primary Education				
Secondary Education	-0.00130	0.00175	0.00206	0.00210
	(0.00401)	(0.00408)	(0.00406)	(0.00406)
University Education	-0.0240***	-0.0215***	-0.0208***	-0.0208***
	(0.00523)	(0.00533)	(0.00530)	(0.00529)
Number of children	0.0133***	0.0132***	0.0131***	0.0130***
	(0.00104)	(0.00105)	(0.00104)	(0.00104)
% IPV in 2016	0.00529***	0.00579***	0.00599***	0.00600***
	(0.000854)	(0.000812)	(0.000793)	(0.000804)
logGDP in 2019	-0.0127**	-0.00433	-0.000934	-0.000965
	(0.00559)	(0.00520)	(0.00515)	(0.00515)
Constant	0.134***	0.166***	0.0785	0.0831
	(0.0337)	(0.0365)	(0.0524)	(0.0581)
Observations	77314	74448	75387	75387
Adjusted R^2	0.024	0.024	0.024	0.024

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.D.4 Robustness Checks

Table D6: Experience of IPV: Robustness Checks

	Original	Including 0 COVID-19 deaths
Log Aged-Standardized Mortality Rate	0.0219***	0.0163***
	(0.00465)	(0.00381)
Married		
Separated or divorced	-0.0790***	-0.0790***
	(0.00474)	(0.00473)
Single	-0.100***	-0.100***
	(0.00505)	(0.00504)
Aged 15-24		
Aged 25-44	-0.0666***	-0.0659***
	(0.00542)	(0.00540)
Aged 45-64	-0.131***	-0.130***
	(0.00641)	(0.00636)
Aged 65-99	-0.198***	-0.196***
	(0.00761)	(0.00757)
Lives in rural areas	-0.0245***	-0.0266***
	(0.00518)	(0.00518)
Primary Education		
Secondary Education	-0.00130	-0.00153
	(0.00401)	(0.00400)
University Education	-0.0240***	-0.0240***
	(0.00523)	(0.00522)
Number of kids	0.0133***	0.0131***
	(0.00104)	(0.00103)
% IPV in 2016	0.00529***	0.00542***
	(0.000854)	(0.000841)
logGDP in 2019	-0.0127**	-0.0111**
	(0.00559)	(0.00550)
Constant	0.134***	0.154***
	(0.0337)	(0.0335)
Observations	77314	77704
Adjusted R^2	0.024	0.024

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Bibliography

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1):1–11.

Agoff, C., Herrera, C., and Castro, R. (2007). The weakness of family ties and their perpetuating effects on gender violence: A qualitative study in Mexico. *Violence against women*, 13(11):1206–1220.

Agoff, C., Rajsbaum, A., and Herrera, C. (2006). Perspectivas de las mujeres maltratadas sobre la violencia de pareja en México. *Salud pública de México*, 48:s307–s314.

Agresti, A. (2012). *Categorical data analysis*, volume 792. John Wiley & Sons.

Agüero, J., Field, E., Rodriguez Hurtado, I., and Romero, J. (2022). Covid-19, job loss, and intimate partner violence in Peru. *Job Loss, and Intimate Partner Violence in Peru*.

Agüero, J. M. (2021). Covid-19 and the rise of intimate partner violence. *World development*, 137:105217.

Aguiar, M. and Hurst, E. (2007). Measuring trends in leisure: The allocation of time over five decades. *The Quarterly Journal of Economics*, 122(3):969–1006.

Ahn, N., Jimeno, J. F., and Ugidos, A. (2004). 'Mondays in the Sun:'unemployment, time use, and consumption patterns in Spain. *Contributions to Economic Analysis*, 271:237–259.

Aizer, A. (2010). The Gender Wage Gap and Domestic Violence. *American Economic Review*, 100:1847–1859.

Allen, S. D. (2017). *A history of boxing in Mexico: Masculinity, modernity, and nationalism*. University of New Mexico Press.

Almanza-Avendaño, A. M., Romero-Mendoza, M. P., Luis, G.-S., and Hortensia, A. (2018). Feelings of insecurity regarding organized crime in Tamaulipas, Mexico. *Salud Pública de México*, 60:442–450.

Alvarado-Zaldivar, G., Salvador-Moysen, J., Estrada-Martinez, S., and Terrones-Gonzalez, A. (1998). Prevalence of domestic violence in the city of Durango. *Salud Pública de Mexico*, 40(6):481–486.

Anderberg, D. and Rainer, H. (2013). Economic abuse: A theory of intrahousehold sabotage. *Journal of Public Economics*, 97:282–295.

Anderberg, D., Rainer, H., Wadsworth, J., and Wilson, T. (2015). Unemployment and Domestic Violence: Theory and Evidence. *The Economic Journal*, 126:1947–1979.

Anderson, K. L. (1997). Gender, Status, and Domestic Violence: An Integration of Feminist and Family Violence Approaches. *Journal of Marriage and Family*, 59(3):655–669.

Andridge, R. R. and Little, R. J. (2009). The use of sample weights in hot deck imputation. *Journal of Official Statistics*, 25(1):21.

Angelucci, M. (2008). Love on the Rocks: Alcohol Abuse and Domestic Violence in Rural Mexico. *B.E. Journal of Economic Analysis and Policy: Contributions to Economic Analysis and Policy*, 8(1):43.

Anwary, A. (2015). Construction of hegemonic masculinity: Violence against wives in Bangladesh. In *Women's Studies International Forum*, volume 50, pages 37–46. Elsevier.

Athey, S. and Imbens, G. W. (2019). Machine Learning Methods Economists Should Know About. *American Review of Economics*, 11:685–725.

Atondo, J. M. O. and Atondo, L. A. O. (2021). COVID-19 y su incidencia en el empleo en México. *Sapienza: International Journal of Interdisciplinary Studies*, 2(3):17–31.

Au Yong Lyn, A. (2021). Vocational Training and Employment Outcomes of Domestic Violence Victims: Evidence from Preliminary Data on Chihuahua City. *CES Working Papers*.

Avila-Burgos, L., Valdez-Santiago, R., Híjar, M., del Rio-Zolezzi, A., Rojas-Martínez, R., and Medina-Solís, C. E. (2009). Factors associated with severity of intimate partner abuse in Mexico: Results of the first national survey of violence against women. *Canadian Journal of Public Health*, 100:436–441.

Ayón, C., Marsiglia, F. F., and Bermudez-Parsai, M. (2010). Latino family mental health: Exploring the role of discrimination and familismo. *Journal of community psychology*, 38(6):742–756.

Babcock, J. C., Waltz, J., Jacobson, N. S., and Gottman, J. M. (1993). Power and violence: the relation between communication patterns, power discrepancies, and domestic violence. *Journal of consulting and clinical psychology*, 61(1):40.

Bagley, B. M. (2012). *Drug trafficking and organized crime in the Americas: major trends in the twenty-first century*. Woodrow Wilson International Center for Scholars, Latin American Program.

Bates, L., Shuler, S., Islam, F., and Islam, K. (2004). Socioeconomic factors associated with domestic violence. *International Family Planning Perspectives*, 30(4):190–199.

Bergman, M., Cafferata, F. G., Ambrogi, J., and Pernigotti, F. (2021). Estudio sobre Homicidios en América Latina. *Centro de Estudios Latinoamericanos sobre Inseguridad y Violencia*.

Bettinger-Lopez, C. and Bro, A. (2020). A double pandemic: Domestic violence in the age of COVID-19. *Council on Foreign Relations*, 13:1–7.

Bhalotra, S., GC Britto, D., Pinotti, P., and Sampaio, B. (2021a). Job displacement, unemployment benefits and domestic violence. *CEPR Discussion Paper No. DP16350*.

Bhalotra, S., Kambhampati, U., Rawlings, S., and Siddique, Z. (2021b). Intimate partner violence: The influence of job opportunities for men and women. *The World Bank Economic Review*, 35(2):461–479.

Bhattacharya, M., Bedi, A., and Chhachhi, A. (2009). Marital violence and women's employment and property status: evidence from North Indian villages: IZA discussion papers 4361. *Institute for the Study of Labor*.

Bittman, M., England, P., Sayer, L., Folbre, N., and Matheson, G. (2003). When does gender trump money? Bargaining and time in household work. *American Journal of Sociology*, 109(1):186–214.

Bobonis, G. J., González-Brenes, M., and Castro, R. (2013). Public transfers and domestic violence: The roles of private information and spousal control. *American Economic Journal: Economic Policy*, 5(1):179–205.

Boyle, M. H., Georgiades, K., Cullen, J., and Racine, Y. (2009). Community influences on intimate partner violence in India: Women's education, attitudes towards mistreatment and standards of living. *Social science & medicine*, 69(5):691–697.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). Classification and Regression Trees. *CRC press*.

Budd, E. C. and Radner, D. B. (1969). The OBE size distribution series: methods and tentative results for 1964. *The American Economic Review*, 59(2):435–449.

Bullinger, L. R., Carr, J. B., and Packham, A. (2021). COVID-19 and crime: Effects of stay-at-home orders on domestic violence. *American Journal of Health Economics*, 7(3):249–280.

Calixto-Calderón, B., Vázquez-González, M. F., Martínez-Peláez, R., Bermeo-Escalona, J. R., García, V., Mena, L. J., Maestre, G., Parra-Michel, J. R., Ceja Bravo, L. A., and López-de Alba, P. L. (2021). Comorbilidad preexistente, el mayor factor de riesgo de mal pronóstico de COVID-19 en la población Mexicana. *Nova Scientia*, 13.

Canedo, A. P. and Morse, S. M. (2021). An estimation of the effect of women's employment on the prevalence of intimate partner violence in Mexico. *Journal of Interpersonal Violence*, 36(19-20):NP10594–NP10618.

Capaldi, D. M., Knoble, N. B., Shortt, J. W., and Kim, H. K. (2012). A systematic review of risk factors for intimate partner violence. *Partner abuse*, 3(2):231–280.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Carpenter, T. G. (2014). *Bad neighbor policy: Washington's futile war on drugs in Latin America*. St. Martin's Press.

Castañeda, J. (2010). Mexico's Failed Drug War, Economic Development Bulletin.

Castro, R. (2019). De parejas, hogares, instituciones y espacios comunitarios: violencias contra las mujeres en México (Endireh 2016). *Centro Regional de Investigaciones Multidisciplinarias, UNAM*.

Conapo (2018). Proyecciones de la población en México y sus entidades federativas 2016–2050. *CONAPO-UNFPA*.

CONDUSEF (2018). Las finanzas en pareja. https://revista.condusef.gob.mx/2018/02/las-finanzas-en-pareja/.

Connelly, R. and Kimmel, J. (2009). Spousal Influences on parents non-market time choices. *Review of Economics of the Household*, 7:361–394.

Costa, D., Hatzidimitriadou, E., Ioannidi-Kapolou, E., Lindert, J., Soares, J. J., Sundin, Ö., Toth, O., and Barros, H. (2016). Male and female physical intimate partner violence and socio-economic position: a cross-sectional international multicentre study in Europe. *Public Health*, 139:44–52.

Cutrona, S. A., Rosen, J. D., and Lindquist, K. A. (2022). Not just money. How organised crime, violence, and insecurity are shaping emigration in Mexico, El Salvador, and Guatemala. *International Journal of Comparative and Applied Criminal Justice*, pages 1–24.

Dabbura, I. (2018). K-means Clustering: Algorithm, Applications, Evaluation Methods, and Drawbacks. https://towardsdatascience.com/k-means-clustering-algorithm-applications-evaluation-methods-and-drawbacks-aa03e644b48a.

Dalal, K. (2011). Does economic empowerment protect women from intimate partner violence? *Journal of Injury and Violence Research*, 3(1):35.

Data Cívica, C. d. I. y. D. E. (2019). Claves para enternder y prevenir los asesinatos de mujeres en México. https://tinyurl.com/u599tjxf.

Dehejia, R. H. and Wahba, S. (1999). Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs. *Journal of the American Statistical Association*, 94(448):1053–1062.

Dehejia, R. H. and Wahba, S. (2002). Propensity Score Matching Methods for Non-experimental Causal Studies. *Review of Economics and Statistics*, 84(1):151–161.

Diamond, A. and Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics*, 95(3):932–945.

Dickson, A., Jennings, C., and Koop, G. (2016). Domestic violence and football in Glasgow: are reference points relevant? *Oxford Bulletin of Economics and Statistics*, 78(1):1–21.

Donatiello, G., D'Orazio, M., Frattarola, D., Rizzi, A., Scanu, M., and Spaziani, M. (2014). Statistical matching of income and consumption expenditures. *International Journal of Economic Sciences*, 3(3):50.

Donatiello, G., D'Orazio, M., Frattarola, D., Rizzi, A., Scanu, M., and Spaziani, M. (2016). The statistical matching of EU-SILC and HBS at ISTAT: where do we stand for the production of official statistics. In *DGINS–Conference of the Directors General of the National Statistical Institutes*, pages 26–27.

D'Orazio, M., Di Zio, M., and Scanu, M. (2006a). Statistical matching for categorical data: Displaying uncertainty and using logical constraints. *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM-*, 22(1):137.

D'Orazio, M., Di Zio, M., and Scanu, M. (2006b). *Statistical matching: Theory and practice*. John Wiley & Sons.

D'Orazio, M., Di Zio, M., and Scanu, M. (2017). The use of uncertainty to choose matching variables in statistical matching. *International Journal of Approximate Reasoning*, 90:433–440.

D'Orazio, M. and D'Orazio, M. M. (2022). Package 'StatMatch'. *Available Online at One of the Mirror Sites*.

Douenne, T. et al. (2020). The vertical and horizontal distributive effects of energy taxes. *The Energy Journal*, 41(3):231–254.

Dude, A. (2007). Intimate partner violence and increased lifetime risk of sexually transmitted infection among women in Ukraine. *Studies in Family Planning*, 38(2):89–100.

D'Orazio, M. (2017). Statistical matching and imputation of survey data with StatMatch. *Italian National Institute of Statistics: Rome, Italy*.

D'Orazio, M. et al. (2015). Integration and imputation of survey data in R: the StatMatch package. *Romanian Statistical Review*, 63(2):57–68.

D'Orazio, M., Zio, M. D., and Scanu, M. (2010). Old and new approaches in statistical matching when samples are drawn with complex survey designs. *Proceedings of the 45th "Riunione Scientifica della Societa'Italiana di Statistica", Padova*, pages 16–18.

ENIGH (2020). Encuesta Nacional de Ingresos y Gastos de los Hogares. `https://www.inegi.org.mx/temas/ingresoshog/`.

Equis, J. (2020). Las dos pandemias: Violencia contra las mujeres en México en el context de COVID-19.

Felbab-Brown, V. (2020). AMLO's feeble response to COVID-19 in Mexico. http://brookings.edu/blog/order-from-chaos/2020/03/30/amlos-feeble-response-to-covid-19-in-mexico.

Fernández, C. and Villar, L. (2017a). Taxonomía de la informalidad en América Latina. *Fesdesarrollo*.

Fernández, C. and Villar, L. (2017b). The impact of lowering the payroll tax on informality in Colombia. *Economía*, 18(1):125–155.

Flake, D. F. and Forste, R. (2006). Fighting families: Family characteristics associated with domestic violence in five Latin American countries. *Journal of Family Violence*, 21(1):19–29.

Foran, H. M. and O'Leary, K. D. (2008). Alcohol and intimate partner violence: A meta-analytic review. *Clinical Psychology Review*, 28(7):1222–1234.

Fox, G. L., Benson, M. L., DeMaris, A. A., and Wyk, J. V. (2004). Economic Distress and Intimate Violence: Testing Family Stress and Resources Theories. *Journal of Marriage and Family*, 64:793–807.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367–378.

Frost, J. (2023). Cronbach's Alpha: Definition, Calculations & Example. https://statisticsbyjim.com/basics/cronbachs-alpha/.

Frye, V., Manganello, J., Campbell, J. C., Walton-Moss, B., and Wilt, S. (2006). The distribution of and factors associated with intimate terrorism and situational couple violence among a population-based sample of urban women in the United States. *Journal of Interpersonal Violence*, 21(10):1286–1313.

Fuchsel, C. L. M., Murphy, S. B., and Dufresne, R. (2012). Domestic violence, culture, and relationship dynamics among immigrant Mexican women. *Affilia*, 27(3):263–274.

Gage, A. J. (2005). Women's experience of intimate partner violence in Haiti. *Social Science & Medicine*, 61(2):343–364.

Gage, A. J. and Thomas, N. J. (2017). Women's work, gender roles, and intimate partner violence in Nigeria. *Archives of Sexual Behavior*, 46(7):1923–1938.

Galvez, G., Mankowski, E. S., McGlade, M. S., Ruiz, M. E., and Glass, N. (2011). Work-related intimate partner violence among employed immigrants from Mexico. *Psychology of Men & Masculinity*, 12(3):230.

Garay, A. D. and Valle-Díaz-Muñoz, G. D. (2011). Una mirada a la presencia de las mujeres en la educación superior en México. *Revista Iberoamericana de Educación Superior (RIES)*.

García, M. L. S. and Adame, M. E. C. (2020). Las brechas de género y el empoderamiento femenino en México. *GénEroos. Revista de Investigación y Divulgación sobre los estudios de Género*, 27(28):219–246.

García-Moreno, C., Zimmerman, C., Morris-Gehring, A., Heise, L., Amin, A., Abrahams, N., Montoya, O., Bhate-Deosthali, P., Kilonzo, N., and Watts, C. (2015). Addressing violence against women: a call to action. *The Lancet*, 385(9978):1685–1695.

García-Ramos, A. (2021). Divorce laws and intimate partner violence: Evidence from Mexico. *Journal of Development Economics*, 150:102623.

Gimenez-Nadal, J. I., Molina, J. A., and Sevilla-Sanz, A. (2012). Social norms, partnerships and children. *Review of Economics of the Household*, 10:215–236.

Gimenez-Nadal, J. I. and Sevilla, A. (2012). Trends in time allocation: a crosscountry analysis. *European Economic Review*, 56(6):1338–1359.

Giraldo, O. (1972). El machismo como fenómeno psicocultural. *Revista Latinoamericana de Psicología*, 4(3):295–309.

Glenday, D. (2013). Professional wrestling as culturally embedded spectacles in five core countries: the USA, Canada, Great Britain, Mexico and Japan. *Revue de Recherche en Civilisation Américaine*, 4.

Government, M. (2020). Finanzas para nuevas parejas. Cuentas claras, relaciones largas. https://www.gob.mx/profeco/articulos/finanzas-para-nuevas-parejas-cuentas-claras-relaciones-largas?state=published.

Graham-Kevan, N. and Archer, J. (2003). Intimate terrorism and common couple violence: A test of Johnson's predictions in four British samples. *Journal of Interpersonal Violence*, 18(11):1247–1270.

Gresham, A. M., Peters, B. J., Karantzas, G., Cameron, L. D., and Simpson, J. A. (2021). Examining associations between COVID-19 stressors, intimate partner violence, health, and health behaviors. *Journal of Social and Personal Relationships*, 38(8):2291–2307.

Guarnieri, E. and Rainer, H. (2021). Colonialism and female empowerment: A two-sided legacy. *Journal of Development Economics*, 151:102666.

Gupta, J., Willie, T. C., Harris, C., Campos, P. A., Falb, K. L., Moreno, C. G., Olavarrieta, C. D., and Okechukwu, C. A. (2018). Intimate partner violence against low-income women in Mexico City and associations with work-related disruptions: a latent class analysis using cross-sectional data. *J Epidemiol Community Health*, 72(7):605–610.

Hadi, A. (2005). Women's productive role and marital violence in Bangladesh. *Journal of Family Violence*, 20(3):181–189.

Ho, D. E., Imai, K., King, G., and Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 15:199–236.

Hoehn-Velasco, L., Silverio-Murillo, A., Balmori de la Miyar, J. R., and Penglase, J. (2022). The impact of the COVID-19 recession on Mexican households: evidence from employment and time use for men, women, and children. *Review of Economics of the Household*, 20(3):763–797.

Hoehn-Velasco, L., Silverio-Murillo, A., and de la Miyar, J. R. B. (2021a). The great crime recovery: Crimes against women during, and after, the COVID-19 lockdown in Mexico. *Economics & Human Biology*, 41:100991.

Hoehn-Velasco, L., Silverio-Murillo, A., and de la Miyar, J. R. B. (2021b). The long downturn: The impact of the great lockdown on formal employment. *Journal of Economics and Business*, 115:105983.

Hong, D., Lee, S., Choi, Y.-J., Moon, S., Jang, Y., Cho, Y.-M., Lee, H., Min, S., Park, H., Hahn, S., et al. (2021). The age-standardized incidence, mortality, and case fatality rates of COVID-19 in 79 countries: a cross-sectional comparison and their correlations with associated factors. *Epidemiology and Health*, 43.

Howard-Bostic, C. D. (2013). Is mutual violent combat (mvc) a gender neutral conceptualization of intimate partner violence? *International Journal of Arts & Sciences*, 6(2):361.

Huesca, L., Llamas, L., Jara, H. X., Vargas Téllez, C. O., and Rodríguez, D. (2021). The impact of the COVID-19 pandemic on poverty and inequality in Mexico. *Revista mexicana de economía y finanzas*, 16(3).

Ibarra-Olivo, E., Acuña, J., and Espejo, A. (2021). Estimación de la informalidad en México a nivel subnacional. *CEPAL*.

Ibarraran-Bigalondo, A. (2020). Corridos, Tequila (and others), and Mexican/Chicano Masculinity. *Servicio de Publicaciones. Universidad de La Laguna*.

Ingoldsby, B. B. (1991). The Latin American family: familism vs. machismo. *Journal of Comparative Family Studies*, 22(1):57–62.

Inmujeres, E. (2004). ABC del género en la Administración Pública. *Instituto Nacional de las Mujeres, México*.

Iztueta, I. L., Navarro, I. G., and Murgiondo, J. E. (2017). Statistical matching in practice–An application to the evaluation of the education system from PISA and TALIS. *Revista de Investigación Educativa*, 35(2):371–388.

Jetelina, K. K., Knell, G., and Molsberry, R. J. (2021). Changes in intimate partner violence during the early stages of the COVID-19 pandemic in the USA. *Injury Prevention*, 27(1):93–97.

Jewkes, R. (2002). Intimate partner violence: causes and prevention. *The Lancet*, 359:1423–1429.

Johnson, M. P. (1995). Patriarchal Terrorism and Common Couple Violence: Two Forms of Violence against Women. *Journal of Marriage and Family*, 57(2):283–294.

Johnson, M. P. (2006). Conflict and Control: Gender Symmetry and Asymmetry in Domestic Violence. *Violence Against Women*, 12(11):1003–1018.

Johnson, M. P. and Ferraro, K. J. (2000). Research on domestic violence in the 1990s: Making distinctions. *Journal of Marriage and Family*, 62(4):948–963.

Johnson, M. P. and Leone, J. M. (2005). The differential effects of intimate terrorism and situational couple violence: Findings from the National Violence Against Women Survey. *Journal of Family Issues*, 26(3):322–349.

Kadane, J. B. (1978). Some Statistical Problems in Merging Data Files. *1978 Compendium of Tax Research*, page 159.

Kalmuss, D. and Straus, M. (1990). Physical violence in American families: Risk factors and adaptations to violence in 8,145 families. *Physical Violence in American Families: __. New Brunswick, NJ: Transaction Publishers*.

Kessler, R. C. and McRae Jr, J. A. (1982). The effect of wives' employment on the mental health of married men and women. *American Sociological Review*, pages 216–227.

Kim, J. C., Watts, C. H., Hargreaves, J. R., Ndhlovu, L. X., Phetla, G., Morison, L. A., Busza, J., Porter, J. D., and Pronyk, P. (2007). Understanding the impact of a microfinance-based intervention on women's empowerment and the reduction of intimate partner violence in South Africa. *American Journal of Public Health*, 97(10):1794–1802.

King, G. (2018). Matching Methods for Causal Inference. *Institute for Quantitative Social Science, Harvard University*.

King, G., Lucas, C., and Nielsen, R. (2017). The Balance-Sample Size Frontier in Matching Methods for Causal Inference. *American Journal of Political Science*, 61(2):473–489.

Kishor, S. and Johnson, K. (2004). Profiling Domestic Violence; A Multi-Country Study, 2004. *Macro International: Calverton, Maryland*.

Knaul, F. M., Touchton, M., Arreola-Ornelas, H., Atun, R., Anyosa, R. J. C., Frenk, J., Martínez-Valle, A., McDonald, T., Porteny, T., Sánchez-Talanquer, M., et al. (2021). Punt politics as failure of health system stewardship: evidence from the COVID-19 pandemic response in Brazil and Mexico. *The Lancet Regional Health-Americas*, 4:100086.

Koenig, M. A., Ahmed, S., Hossain, M. B., and Mozumder, A. (2003). Women's status and domestic violence in rural Bangladesh: individual-and community-level effects. *Demography*, 40(2):269–288.

Kourti, A., Stavridou, A., Panagouli, E., Psaltopoulou, T., Spiliopoulou, C., Tsolia, M., Sergentanis, T. N., and Tsitsika, A. (2023). Domestic violence during the COVID-19 pandemic: a systematic review. *Trauma, Violence, & Abuse*, 24(2):719–745.

Kreager, D. A., Felson, R. B., Warner, C., and Wenger, M. R. (2013). Women's education, marital violence, and divorce: A social exchange perspective. *Journal of Marriage and Family*, 75(3):565–581.

Kum, H. and Masterson, T. (2010). Statistical matching using propensity scores: Theory and application to the analysis of the distribution of income and wealth. *Journal of Economic and Social Measurement*, 35(3):177–196.

Kumar, D. (2023). A Complete Understanding of LASSO Regression. `https://www.mygreatlearning.com/blog/understanding-of-lasso-regression/`.

LaLonde, R. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review*, 76(4):604–620.

Lamas, M. (2007). El impacto de los estereotipos y los roles de género en México. *Instituto Nacional de las Mujeres, México*.

Lanier, C. and Maume, M. O. (2009). Intimate Partner Violence and Social Isolation Across the Rural/Urban Divide. *Violence Against Women*, 15(11):1311–1330.

Lawoko, S. (2008). Predictors of attitudes toward intimate partner violence: A comparative study of men in Zambia and Kenya. *Journal of Interpersonal violence*, 23(8):1056–1074.

Lee, B. K., Lessler, J., and Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, 29(3):337–346.

Lersch, K. and Hart, T. C. (2022). COVID-19, Violent Crime, and Domestic Violence: An Exploratory Analysis. *International Criminology*, 2(1):5–18.

Leulescu, A. and Agafitei, M. (2013). Statistical matching: a model based approach for data integration. *Eurostat-Methodologies and Working papers*, pages 10–2.

Leulescu, A., Agafitei, M., and Mercy, J.-L. (2011). Statistical matching: a case study on EU-SILC and LFS. In *ESSnet Data Integration Workshop. Madrid*.

Li, X. and Wang, J. (2022). Continuity and change: Violations of private patriarchal practices and domestic violence against rural wives in China. *Health Care for Women International*, 43(7-8):898–913.

Little, R. J. and Rubin, D. B. (2002). Bayes and multiple imputation. *Statistical Analysis with Missing Data*, pages 200–220.

Macmillan, R. and Gartner, R. (1999). When She Brings Home the Bacon: Labor-Force Participation and the Risk of Spousal Violence Against Women. *Journal of Marriage and Family*, 61(4):947–958.

Maldonado, I., Nájera, M., and Segovia, A. (2005). Effects of the Oportunidades Program in couple and Family relationships. *Contribution to the Gender Project on Oportunidades at the Colegio de Mexico. Mexico City: COLMEX*.

Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard University Press.

Marin, G. and Marín, B. (1991). Applied social research methods series, Vol. 23. Research with Hispanic populations. *Sage Publications, Inc Thousand Oaks*.

Masferrer, C., Chávez, O. R., et al. (2022). Are homicides and robberies associated with mortality due to COVID-19? Lessons from Urban Mexico. *Vienna Yearbook of Population Research*, 20(1):TBA–OLF.

Masterson, T. (2010). Quality of match for statistical matches used in the 1992 and 2007 LIMEW estimates for the United States. *Levy Economics Institute Working Paper*.

Masterson, T. (2011). Quality of match for statistical matches used in the development of the Levy Institute Measure of Time and Income Poverty (LIMTIP) for Argentina, Chile, and Mexico. *Levy Economics Institute of Bard College Working Paper*, 692.

Mattiuzzi, C., Henry, B. M., and Lippi, G. (2023). Correlation between relative age-standardized mortality rates and COVID-19 mortality over time in Italy. *Acta Bio Medica: Atenei Parmensis*, 94(2).

Mele, M. (2009). The time course of repeat intimate partner violence. *Journal of Family Violence*, 24:619–624.

Melzer, S. A. (2002). Gender, work, and intimate violence: Men's occupational violence spillover and compensatory violence. *Journal of Marriage and Family*, 64(4):820–832.

Miller, S. L. and Meloy, M. L. (2006). Women's use of force: Voices of women arrested for domestic violence. *Violence Against Women*, 12(1):89–115.

Moradzadeh, R., Jamalian, S. M., Nazari, J., Kamali, A., Sadeghi, B., Hosseinkhani, Z., Sofian, M., and Zamanian, M. (2021). Age-standardized mortality rate and predictors of mortality among COVID-19 patients in Iran. *Journal of Education and Health Promotion*, 10(1).

Moriarity, C. and Scheuren, F. (2001). Statistical matching: a paradigm for assessing the uncertainty in the procedure. *Journal of Official Statistics*, 17(3):407.

Moriarity, C. and Scheuren, F. (2003). A note on Rubin's statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics*, 21(1):65–73.

Naved, R. T. and Persson, L. Å. (2005). Factors associated with spousal physical violence against women in Bangladesh. *Studies in Family Planning*, 36(4):289–300.

Nix, J. and Richards, T. N. (2021). The immediate and long-term effects of COVID-19 stay-at-home orders on domestic violence calls for service across six US jurisdictions. *Police Practice and Research*, 22(4):1443–1451.

OECD (2017). *Building an inclusive Mexico: Policies and good governance for gender equality*. Organisation for Economic Co-operation and Development.

Okner, B. (1972). Constructing a new data base from existing microdata sets: the 1966 merge file. In *Annals of Economic and Social Measurement, Volume 1, number 3*, pages 325–362. NBER.

Opanasenko, A., Lugova, H., Mon, A. A., and Ivanko, O. (2021). Mental health impact of gender-based violence amid COVID-19 pandemic: a review. *Bangladesh Journal of Medical Science*, pages 17–25.

Ornstein, P. and Rickne, J. (2013). When does intimate partner violence continue after separation? *Violence Against Women*, 19(5):617–633.

Orozco, A. E., Nievar, M. A., and Middlemiss, W. (2012). Domestic violence in Mexico: Perspectives of Mexican counselors. *Journal of Comparative Family Studies*, 43(5):751–772.

Orozco, C. E. (1995). Beyond Machismo, La Familia, and Ladies Auxiliaries: A Histo riography of Mexican-Origin Women's Participation in Voluntary Associations and Politics in the United States, 1870 1990. *Mexican American Studies & Research Center, The University of Arizona*.

Pacheco, E. and Florez, N. (2014). Entre lo rural y lo urbano. Tiempo y desigualdades de género. *B. García y E. Pacheco (coords.), Uso del tiempo y trabajo no remunerado en México*, pages 263–323.

Pamplona, F. (2020). La pandemia de COVID-19 en México y la otra epidemia. *Espiral (Guadalajara)*, 27(78-79):265–302.

Parrado, E. A. and Zenteno, R. M. (2002). Gender differences in union formation in Mexico: Evidence from marital search models. *Journal of Marriage and Family*, 64(3):756–773.

Peitzmeier, S. M., Fedina, L., Ashwell, L., Herrenkohl, T. I., and Tolman, R. (2022). Increases in intimate partner violence during COVID-19: Prevalence and correlates. *Journal of Interpersonal Violence*, 37(21-22):NP20482–NP20512.

Pence, E., Paymar, M., and Ritmeester, T. (1993). *Education groups for men who batter: The Duluth model*. Springer Publishing Company.

Peterman, A., Palermo, T., and Bredenkamp, C. (2011). Estimates and determinants of sexual violence against women in the Democratic Republic of Congo. *American Journal of Public Health*, 101(6):1060–1067.

Peterman, A., Potts, A., O'Donnell, M., Thompson, K., Shah, N., Oertelt-Prigione, S., Van Gelder, N., et al. (2020). *Pandemics and violence against women and children*, volume 528. Center for Global Development Washington, DC.

Peterman, A., Schwab, B., Roy, S., Hidrobo, M., and Gilligan, D. O. (2015). Measuring women's decision-making: indicator choice and survey design experiments from cash and food transfer evaluations in Ecuador, Uganda, and Yemen. *IFPRI Discussion Paper 1453*.

Pietromonaco, P. R. and Overall, N. C. (2021). Applying relationship science to evaluate how the COVID-19 pandemic may impact couples' relationships. *American Psychologist*, 76(3):438.

Piquero, A. R., Piquero, N. L., and Kurland, J. (2021). A case study of domestic violence arrests during the COVID-19 pandemic in Miami-Dade County. *Victims & Offenders*, 16(8):1077–1088.

Piquero, A. R., Riddell, J. R., Bishopp, S. A., Narvey, C., Reid, J. A., and Piquero, N. L. (2020). Staying home, staying safe? A short-term analysis of COVID-19 on Dallas domestic violence. *American Journal of Criminal Justice*, 45:601–635.

Pyke, K. D. (1996). Class-based masculinities: The interdependence of gender, class, and interpersonal power. *Gender & Society*, 10(5):527–549.

Radner, D. (1980). Report on exact and statistical matching techniques. *For sale by the Supt. of Docs., USGPO*.

Rahman, M., Hoque, M., Makinoda, S., et al. (2011). Intimate partner violence against women: Is women empowerment a reducing factor? A study from a national Bangladeshi sample. *Journal of Family Violence*, 26(5):411–420.

Ramirez, H. (2011). Masculinity in the workplace: The case of Mexican immigrant gardeners. *Men and Masculinities*, 14(1):97–116.

Rässler, S. (2002). *Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches*, volume 168. Springer, Verlag New York.

Rässler, S. (2012). *Statistical matching: A frequentist theory, practical applications, and alternative Bayesian approaches*, volume 168. Springer Science & Business Media.

Reingle Gonzalez, J. M., Molsberry, R., Maskaly, J., and Jetelina, K. K. (2020). Trends in family violence are not causally associated with COVID-19 stay-at-home orders: A commentary on Piquero et al. *American Journal of Criminal Justice*, 45:1100–1110.

Rennison, C. M. and Welchans, S. (2002). Intimate Partner Violence. *Bureau of Justice Statistics: Special Report*.

Renssen, R. H. (1998). Use of statistical matching techniques in calibration estimation. *Survey Methodology*, 24:171–184.

Rezaei Ghahroodi, Z. (2023). Statistical matching of sample survey data: application to integrate Iranian time use and labour force surveys. *Statistical Methods & Applications*, pages 1–29.

Rezey, M. L. (2020). Separated women's risk for intimate partner violence: a multiyear analysis using the National Crime Victimization Survey. *Journal of Interpersonal Violence*, 35(5-6):1055–1080.

Riding, A. (2011). *Distant neighbors: A portrait of the Mexicans*. Vintage.

Rios-Avila, F. (2014). Quality of Match for Statistical Matches Using the American Time Use Survey 2010, the Survey of Consumer Finances 2010, and the Annual Social and Economic Supplement 2011. *Levy Economics Institute of Bard College*, 798.

Rios-Avila, F. (2016). Quality of match for statistical matches used in the development of the Levy Institute Measure of Time and Consumption Poverty (LIMTCP) for Ghana and Tanzania. *Levy Economics Institute, Working Paper*, 873.

Rios-Avila, F. (2018). Quality of Match for Statistical Matches Using the American Time Use Survey 2013, the Survey of Consumer Finances 2013, and the Annual Social and Economic Supplement 2014. *Levy Economics Institute of Bard College*, 914.

Rivera-Rivera, L., Lazcano-Ponce, E., Salmerón-Castro, J., Salazar-Martínez, E., Castro, R., and Hernández-Avila, M. (2004). Prevalence and determinants of male partner violence against Mexican women: A population-based study. *Salud Pública de México*, 46(2):113–121.

Rodgers, W. L. (1984). An Evaluation of Statistical Matching. *Journal of Business and Economic Statistics*, 2(1):91–102.

Romo, G. A. C. (2020). Trabajo, tiempo libre y movilidad cotidiana en México, 2014. *Revista CIFE: Lecturas de Economía Social*, 22(36):155–181.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rubin, D. B. (1986). Statistical Matching Using File Concatenation with Adjusted Weights and Multiple Imputations. *Journal of Business and Economic Statistics*, 4(1):87–94.

Salas, C., Quintana, L., Mendoza, M. Á., and Valdivia, M. (2020). Distribución del ingreso laboral y la pobreza en México durante la pandemia de la Covid-19. Escenarios e impactos potenciales. *El Trimestre Económico*, 87(348):929–962.

Schmidt, S. and Martinello, A. (2019). *Machine Learning as a Data Driven Approach to Automate Multivariate Matching Methods*. PhD thesis, Lund University.

Schuler, S. R., Hashemi, S. M., Riley, A. P., and Akhter, S. (1996). Credit programs, patriarchy and men's violence against women in rural Bangladesh. *Social Science & Medicine*, 43(12):1729–1742.

Sebba, L. (1984). Crime seriousness and criminal intent. *Crime & Delinquency*, 30(2):227–244.

Sekhon, J. S. (2011). Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software*, 42(7).

Sekhon, J. S. and Grieve, R. D. (2012). A matching method for improving covariate balance in cost-effectiveness analyses. *Health Economics*, 21(6):695–714.

Sellin, T. and Wolfgang, M. E. (1964). The measurement of delinquency. *Wiley*.

Serafino, P. and Tonkin, R. (2017). Statistical matching of European Union statistics on income and living conditions (EU-SILC) and the household budget survey. *Eurostat: Statistical Working Papers. Luxembourg: Publications Office of the European Union.*, 10:933460.

Sevilla-Sanz, A. and Fernandez, C. (2006). Social norms and household time allocation. *University of Oxford*.

Sevilla-Sanz, A., Gimenez-Nadal, J. I., and Fernández, C. (2010). Gender roles and the division of unpaid work in Spanish households. *Feminist Economics*, 16(4):137–184.

Shirk, D. A. (2011). The Drug War in Mexico: Confronting a Shared Threat. Technical report, Council on Foreign Relations.

Silverio-Murillo, A., de la Miyar, J. B., and Hoehn-Velasco, L. (2023). Families under confinement: Covid-19 and domestic violence. In *Crime and Social Control in Pandemic Times*, volume 28, pages 23–41. Emerald Publishing Limited.

Silverio-Murillo, A., Hoehn-Velasco, L., de la Miyar, J. R. B., and Rodríguez, A. (2021). COVID-19 and women's health: examining changes in mental health and fertility. *Economics Letters*, 199:109729.

Simister, J. (2013). Is Men's Share of Housework Reduced by 'Gender Deviance Neutralization'? Evidence From Seven Countries. *Journal of Comparative Family Studies*, 44(3):311–325.

Sims, C. (1972). Comments on: "Constructing a new data base from existing microdata sets: the 1966 merge file", by BA Okner. *Annals of Economic and Social Measurements*, 1:343–345.

Singer, M. (2020). Deadly companions: COVID-19 and diabetes in Mexico. *Medical Anthropology*, 39(8):660–665.

Singh, A., Mantel, H., Kinack, M., and Rowe, G. (1993). Statistical matching: use of auxiliary information as an alternative to the conditional independence assumption. *Survey Methodology*, 19(1):59–79.

Sizemore, S. and Alkurdi, R. (2019). Matching methods for causal inference: a machine learning update. *Seminar Applied Predictive Modeling*, August 18, 2019.

Spangaro, J. (2017). What is the role of health systems in responding to domestic violence? An evidence review. *Australian Health Review*, 41(6):639–645.

Stanley, J. and Granick, J. S. (2020). The limits of location tracking in an epidemic. *American Civil Liberties Union*, pages 1–9.

Svec, J. and Andic, T. (2018). Cooperative decision-making and intimate partner violence in Peru. *Population and Development Review*, 44(1):63.

Swiss, S., Jennings, P. J., Aryee, G. V., Brown, G. H., Jappah-Samukai, R. M., Kamara, M. S., Schaack, R. D., and Turay-Kanneh, R. S. (1998). Violence against women during the Liberian civil conflict. *Jama*, 279(8):625–629.

Terrazas-Carrillo, E. C. and McWhirter, P. T. (2015). Employment status and intimate partner violence among Mexican women. *Journal of Interpersonal Violence*, 30(7):1128–1152.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Tran, T. T. H. and Osier, G. (2023). Identifying the disadvantaged in Luxembourg–Measuring multidimensional poverty by statistical matching. *Statec*, 133.

Tur-Prats, A. (2021). Unemployment and intimate partner violence: A cultural approach. *Journal of Economic Behavior & Organization*, 185:27–49.

UN Women, Segob, I. (2017). La violencia feminicida en México, aproximaciones y tendencias 1985-2016. https://mexico.unwomen.org/es/digiteca/publicaciones/2017/12/violencia-feminicida.

UNODC (2018). Global Study on Homicide: Gender-related killing of women and girls.

UNODC (2019). Global Study on Homicide.

Valdez-Santiago, R., Híjar, M., Martínez, R. R., Burgos, L. Á., and Monreal, M. d. l. L. A. (2013). Prevalence and severity of intimate partner violence in women living in eight indigenous regions of Mexico. *Social Science & Medicine*, 82:51–57.

Valdez-Santiago, R., Villalobos-Hernández, A., Arenas-Monreal, L., Flores, K., and Ramos-Lira, L. (2021). Violencia en el hogar contra mujeres adultas durante el confinamiento por la pandemia de Covid-19 en México. *salud Pública de México*, 63(6):782–788.

Valencia-Londoño, P. A. and Nateras-González, M. E. (2020). Violencia en contra de las mujeres como discriminación en contextos de violencia criminal: el caso del feminicidio en Medellín y el Estado de México. *Revista Criminalidad*, 62(1):59–85.

Van Gelder, N., Peterman, A., Potts, A., O'Donnell, M., Thompson, K., Shah, N., and Oertelt-Prigione, S. (2020). COVID-19: Reducing the risk of infection might increase the risk of intimate partner violence. *E Clinical Medicine*, 21.

Villarreal, A. (2007). Women's employment status, coercive control, and intimate partner violence in Mexico. *Journal of Marriage and Family*, 69(2):418–434.

Vyas, S. and Watts, C. (2009). How does economic empowerment affect women's risk of intimate partner violence in low and middle income countries? A systematic review of published evidence. *Journal of International Development: The Journal of the Development Studies Association*, 21(5):577–602.

Waal, T. D. (2015). Statistical matching: Experimental results and future research questions. *CBS | Discussion Paper 2015*, 19.

Webber, D. and Tonkin, R. (2013). Statistical matching of EU-SILC and the Household Budget Survey to compare poverty estimates using income, expenditures and material deprivation. *Eurostat-Methodologies and Working Papers, Luxembourg: Publications Office*.

WHO (2012). Understanding and addressing violence against women: Intimate partner violence. Technical report, World Health Organization.

WHO (2021). Violence against women prevalence estimates, 2018: global, regional and national prevalence estimates for intimate partner violence against women and global and regional prevalence estimates for non-partner sexual violence against women. Technical report, World Health Organization.

WHO (2023). Age-standardized mortality rate (per 100 000 population). Technical report, World Health Organization.

Wolf, E. (2002). Lower wage rates for fewer hours? A simultaneous wage-hours model for Germany. *Labour Economics*, 9(5):643–663.

Yount, K. M. (2005). Resources, family organization, and domestic violence against married women in Minya, Egypt. *Journal of Marriage and Family*, 67(3):579–596.

Yount, K. M. and Carrera, J. S. (2006). Domestic violence against married women in Cambodia. *Social Forces*, 85(1):355–387.

Yount, K. M., Zureick-Brown, S., and Salem, R. (2014). Intimate partner violence and women's economic and non-economic activities in Minya, Egypt. *Demography*, 51(3):1069–1099.

Zamudio Sánchez, F. J., Ayala Carrillo, M. d. R., and Arana Ovalle, R. I. (2014). Mujeres y hombres: desigualdades de género en el contexto mexicano. *Estudios Sociales (Hermosillo, Son.)*, 22(44):251–279.

Zhang, W., Robinson, C., Guhathakurta, S., Garikapati, V. M., Dilkina, B., Brown, M. A., and Pendyala, R. M. (2018). Estimating residential energy consumption in metropolitan areas: A microsimulation approach. *Energy*, 155:162–173.