Gender norms, Development and Feminicides: new evidence on the male backlash hypothesis

Alessandro Caggia, Anna Cristoforoni, Gabriele Mole'

Abstract

This paper examines the relationship between historical plough adoption and contemporary femicide rates. We employ an instrumental variable (IV) strategy based on the work of Alesina et al. (2013) to investigate this relationship, utilizing cross-country and regional data. We find that societies with a history of plough adoption, which are often considered more conservative, exhibit lower rates of femicide. Our analysis includes robustness checks and placebo tests to ensure the validity of our findings. In the mechanism section, we explore what theories may explain such results, and discuss the male backlash hypothesis.

1. Introduction

The relationship between historical agricultural practices and contemporary social outcomes has garnered significant attention in economic research. This paper investigates the unexpected inverse relationship between historical plough adoption and femicide rates. Societies with a history of plough adoption have been shown to be more often associated with more conservative gender norms and entrenched patriarchal structures. However, our findings reveal a counterintuitive empirical relationship: societies with greater historical plough adoption exhibit lower rates of femicide today. This paradox invites a deeper exploration of the theme and leads to a discussion of which factors may be the root factors determining male violence.

1.1. Motivation and Research Question

The motivation for this research stems from the need to understand the underlying causes of femicide, a critical issue affecting women's safety globally. Clearly, gender norms represent a crucial factor to be investigated as they embody the way in which women are perceived in society, and are expected to affect significantly the interaction between men and women. The issue with gender norms is that they are extremely difficult to measure and to proxy, and are endogenous to many other relevant factors. To explore the issue, we focus on the role of historical agricultural practices, particularly plough adoption, as an instrument for gender norms. Our primary research question is: How does historical plough adoption influence contemporary femicide rates, and what mechanisms underpin this relationship. By addressing this question, we aim to contribute to the broader discourse on gender norms, cultural persistence, and violence, offering insights that could inform policy interventions aimed at reducing femicide. Surely gender norms play a role, but this paper tries to focus on is that the role of gender norms and the interaction between gender norms with other societal and cultural factors is far more complex than normally though, with many considerations to be explored as there are many heterogenous and multilayered channels.

1.2. Contribution

This paper contributes to the literature in several ways. First, it provides empirical evidence for the paradoxical relationship between historical plough adoption and femicide rates. Second, we employ an instrumental variable (IV) strategy based on the work of Alesina et al. (2013) to establish a causal link between historical practices and contemporary outcomes. This approach allows us to control for confounding factors and enhance the robustness of our findings. Additionally, we integrate ethnographic data and conduct extensive robustness checks, including placebo tests and regional analyses, to validate our results. Our findings have significant implications for understanding the dynamics of gender norms and violence, suggesting that historical context plays a crucial role in shaping contemporary social issues. By employing a rigorous IV strategy, we offer a robust empirical framework that not only addresses potential endogeneity concerns but also enriches the discourse on gender dynamics and violence. Our findings encourage further exploration into the intricate interplay between historical practices and modern social issues, paving the way for future research in this domain.

1.3. Roadmap

The remainder of the paper is structured as follows. In Section 1, we review the relevant literature on gender norms, historical agricultural practices, and femicide, highlighting gaps that our research addresses. Section 2 outlines our theoretical framework and the male backlash hypothesis, detailing the mechanisms through which historical plough adoption may influence contemporary femicide rates. In Section 3, we present our empirical strategy, including data sources, the IV approach, and robustness checks. Section 4 discusses our findings, offering interpretations and implications for policy. Finally, Section 5 concludes the paper, summarizing our contributions and suggesting avenues for future research.

2. Literature Review

The intersection of historical agricultural practices and contemporary social issues has garnered significant attention in economic literature. This review synthesizes key theoretical frameworks and empirical findings surrounding the relationship between historical plough adoption and modern femicide rates, particularly focusing on which theories may explain such results.

2.1. Theoretical Perspective on Gender Norms

Gender norms are societal standards and expectations that dictate how women and men should be and behave

([EIGE(2023)]). These prescriptive norms are shaped by historical, social, and economic determinants, including fertility, labor force participation, education, marriage customs, and inheritance systems ([Giuliano(2017)]).

Today, every society holds specific assumptions about the desirable, proper roles of women and men, both in gender relations and within the labor market ([Pfau-Effinger(1998)]). Hence, gender norms are often embedded in modern societal structures and practices. These include, among others, institutional policiessuch as promotion requirements in government departmentsthe medias recurring imagery of women, television standards regarding what qualifies as news, and even the design of cities and architecture ([Cislaghi & Heise(2019)]).

Gender norms are internalized in the early stages of life, generating and reinforcing a "life cycle of gender socialization and stereotyping" ([EIGE(2023)]). Indeed, the structural and cultural networks in which individuals are embedded provide them with a system of norms that define what is socially acceptable and what is not ([Bicchieri(2006)]). For example, communities that uphold traditional gender roles may believe that men have a right to control or discipline women, thereby increasing the risk of domestic violence ([Flood & Pease(2009)]). Men may use violence as punishment when their female partners do not satisfy their needs, be them physical, sexual, or emotional. Moreover, violence is often legitimized as a means of restoring gender norms and contested masculinity ([Anderson & Umberson(2001)]). Consequently, when intimate partner violence (IPV) victims internalize this legitimization of violence and social prejudices, evidence suggests that they are more likely to blame themselves for the abuse and to be ashamed of it. ([Taccini & Mannarini(2023)]).

This patriarchal discourse on the legitimization of violence is not confined to a single cultural setting; rather, it has been documented worldwide including Christian religious communities, Muslim-majority, Asian, and African societies ([Lomazzi(2023)]). For instance, in Christian communities in Tennessee, USA, women experiencing marital abuse are often shamed and silenced through religious beliefs that emphasize the sanctity of marriage and obedience to one's partner ([Knickmeyer et al.(2010)]). Similarly, in the Islamic Republic of Pakistan, Muslim women have come to accept both verbal and physical violence as a justified response to (perceived) disobedience ([Tarar & Pulla(2014)]). In Latin America, the term *Machismo* denotes male dominance and superiority. The latter, justified by patriarchal norms and strengthened through cultural values, are oftentimes encouraged by ones parents and by the society itself. Alongside, the Hispanic view of the females role, which is based on the Virgin Mary, underlies submissiveness and saintliness. This perspective imposes "limiting and stigmatizing effects" on Latin American women, while playing a key role in perpetuating patriarchal attitudes overtime. Norms similar to *Machismo* also appears in Arab, Asian, and Western societies in various forms. For instance, in Saudi Arabia, women are forbidden from driving, whereas among U.S. "tough" cowboys and urban ghetto males, respect is earned only from other men after proving to be a "tough man". Likewise, Japanese women were often forced to leave their job upon marriage ([Segrest & Domke-Damonte(2003)]). In 1996, nearly half of all women in each cohort left their premarital employment after getting married or having a child ([Yu(2005)]).

Hence, patriarchy manifests not only in individual attitudes but also in broader unequal social structures and sexist norms. Alarmingly, patriarchal systems, in addition to enabling violence against women, also contribute to the lack of willingness of governmental and non-governmental actors to intervene and protect survivors of gender-based violence ([Saunders et al.(2022)]). This lack of protection, coupled with traditional, masculinity-based social norms and widespread tolerance towards violence against women, may ultimately lead to femicide ([UN Women(2024)]).

2.2. Persistence of Gender Norms from Ancient to Present Times

Our empirical strategy relies on the assumption that gender norms are persistent overtime. Hence, this section reviews the literature on the intergenerational transmission of gender norms and their long-lasting cultural footprint. Research in economics and related disciplines shows that cultural values and beliefs often exhibit remarkable persistence across generations. For instance, Bisin and Verdier (2000) build a model of cultural transmission that formalizes the idea that parents willingly transmit their own beliefs to their children, with some cultural traits evolving more slowly than others and persisting over many generations. Giavazzi et al. (2014) empirically confirm that persistence is specific only to some beliefs. Indeed, deeply rooted values such as religion, family norms, and political orientation tend to remain roughly stable overtime, whereas attitudes strongly impacted by social interaction such as trust can change more rapidly. Furthermore, Bisin and Verdier (2001) find that the speed of the change in preferences mainly depends on whether the latter have been vertically transmitted, i.e., from parents to children, or horizontally or obliquely transmitted, i.e., trasmitted by peers, society, institutions. While vertical transmission tends to preserve cultural traits, horizontal and oblique transmission can reinforce these traits. At the same time, environmental stability plays a key role in explaining cultural persistence. Giuliano and Nunn (2021) show that when ancestors lived in a stable environment, traditions are more likely to be transmitted to generations as they contain valuable information to help new generations adapt to the environment and survive. Conversely, in societies where the ancestral environment drastically varied from one generation to another, traditions were less adaptive and cultural change occurred more rapidly. These insights suggest that cultural persistence mainly arises through intergenerational transmission within families, reinforced by social interactions, institutions, and role models, and is more pronounced for traits that are historically adaptive in stable environments. Scholars across disciplines have shown that gender norms often persist across generations. An influential line of research traces modern gender norms back to ancient agricultural practices. Ester Boserup (1970) brought forward the seminal hypothesis that pre-industrial farming techniques not only shaped the gendered division of labor in the past, but also had long-lasting effects on modern gender culture. In societies where shifting hoe agriculture was a traditional practice, women actively participated in farming; conversely, in societies that relied on plough agriculture which requires greater upper body

strength and is less compatible with childcare men specialized in agricultural work, whereas women focused on domestic activities. Overtime, this division of labor translated into the belief that a womans natural place is the home. Alesina et al. (2013) after documenting a strong negative correlation between the use of the plough and female participation in agricultural work during pre-industrial times show that historical farming technologies continue to shape modern gender norms: both countries and immigrant communities with plough-intensive tradition tend to exhibit more traditional gender attitudes. Giuliano (2014) finds similar patterns: in societies that practiced or continue to practice plough agriculture, fathers are granted greater parental authority, inheritance rules favor male heirs, and womens mobility outside the household is more restricted. Intergenerational cultural transmission is believed to underlie this persistence in gender norms. Cultural values and beliefs are typically transmitted across generations, displaying a fair stability overtime (/citeGiuliano2014). Alesina et al. (2013) show that gender culture is inherently sticky: by focusing on children of immigrants in the US and Europe, the researchers test whether the traditional practice of plough agriculture affects female labor force participation, holding constant the external environment. They find that even when individuals are exposed to the same labor markets, institutions, and policies, an ancestral tradition of plough agriculture remains associated with more unequal gender norms and lower female labor force participation. This provides direct evidence that the transmission of cultural beliefs from parents to children lies at the heart of the persistence in gender norms. This mechanism the intergenerational transmission of gender norms has also been investigated in other studies. For instance, Fernandez and Fogli (2009) investigated the work and fertility behavior of second-generation American women. They found that, all else equal, women whose country of ancestry is characterized by a higher female labor force participation tend to work more, whereas those whose ancestry is from countries with higher fertility rates have more children. Moreover, native US women also appear to be influenced in their labor supply decisions by the gender attitudes of their mother-in-laws country of origin ([Bredtmann et al.(2020)]). Another branch of research highlights how kinship and family structures exert a long-lasting impact on gender norms. Pre-industrial societal characteristics, such as matrilineal and patrilineal inheritance, play an important role in explaining present gender roles. Under patrilocality a social system in which women move to live with or near the husbands family after marriage mothers have stronger incentives to invest in their sons health and education, since daughters leave the household upon marriage ([Giuliano(2020)]). Instead, under matrilineality the determination of inheritance and descent through the maternal family line women are in control of the lineage and resources, thereby holding a higher status. For instance, Lowes (2017) shows that matrilineal kinship reduces a husbands authority over his wife. Indeed, women in matrilineal systems exhibit less altruism and cooperation towards their spouses due to a greater bargaining power within the household and a reduced fear of retaliation. Alongside, Gneezy et al. (2009) conducted an experiment focusing on the Maasai in Tanzania and the Khasi in India a patriarchal and matrilineal systems, respectively. The researchers found that matrilineal women tend to exhibit a more competitive behavior than matrilineal men, whereas in patriarchal societies the opposite holds true. Robinson and Gottlieb (2019) also explored the role of matrilineality in Africa, showing its robust association with narrowing the gender gap in political participation. They argue that this effect stems from the ability of matrilineal systems to perpetuate progressive gender norms by shaping expectations about womens role in a society and by transmitting them across generations. The persistence of gender norms is reinforced not only through cultural transmission within the family known as *vertical* transmission but also through peers (*horizontal* transmission) and role models (*oblique* transmission) ([Giuliano(2020)]). Research shows that males are subject to a stronger peer pressure to conform to gender norms compared to females, especially during adolescence. Boys who violate their social networks gender norms are at higher risk of harassment and maltreatment by their peers, while affiliating with a boy who violated gender norms comes at the greatest social cost of all (Masters et al., 2020). In Tanzania, peer gender norms are significantly associated with mens tendency to perpetrate intimate partner violence, even after controlling for their own gender attitudes, with effects strengthening as network cohesion increases ([Mulawa et al.(2017)]). Chen et al. (2025) further show that greater exposure to peers working mothers improves girls performance in mathematics in Chinese middle schools, while boys outcomes remain unaffected. Teachers can also contribute to the transmission of gender norms: Wolter et al. (2015) find that preschool teachers transmit gendered expectations about reading. Boys appear less motivated to read during preschool, and less able in primary school, when their teacher supports traditional gender attitudes compared to egalitarian ones. Mass media also play a central role, often perpetuating gender stereotypes by depicting women as caregivers and men as breadwinners and leaders ([Seluman et al.(2024)]). For example, men exposed to television clips that sexually objectify women report a greater tendency to harass women than those not exposed, an effect driven by heightened conformity to patriarchal gender norms ([**?**]). This body of evidence demonstrates that gender norms are remarkably persistent across time, transmitted through families, kinship systems, peer networks, role models, and institutions. This validates our core assumption that past gender norms, proxied by historical agricultural practices particularly, ancestral plough use can be used to predict modern gender norms.

2.3. Instrumenting Gender Norms: The Alesina et al. (2013) Plough IV and Its Applications

To address endogeneity issues concerning the relationship between historical plough use measured at the ethnicity level and gender norms, Alesina et al. (2013) employ an Instrumental Variable approach based on crops differential ancestral geo-climatic suitability to plough technology. This strategy is motivated by Pryor (1985), who argues that different types of cultivated crops benefit differently from plough use. Crops that require large tracts of land to be prepared quickly benefit more from plough technology, whereas crops that require less intensive soil preparation or could be grown in soils where ploughing is difficult benefit less from it. The former referred to as *plough-positive* crops include teff, wheat, rye, and wet rice; conversely, the latter defined by Prior (1985)

as *plough-negative* crops comprise maize, sorghum, millet, and various types of root and tree crops. For each ethnic group in the Ethnographic Atlas using the FAOs Global Agro-Ecological Zones data the authors compute the share of land within 200 km of the groups centroid that was suitable for plough-positive and plough-negative crops, respectively, normalized by overall arable land. Two ethnicity-level measures of suitability to plough use are then constructed: the ancestral suitability of an ethnic groups land for plough positive, and the ancestral suitability of an ethnic groups land for plough-negative crops. These variables provide exogenous indicators of how apt each ethnicitys ancestry to the adoption of the plough was, and are thus used by Alesina et al. (2013) to instrument historical plough use by ethnic group. Next, historical plough use (the treatment) and ancestral crop suitability (the instruments) are linked to present-day populations on a granular spatial scale by matching each of the Ethnologue language groups to one of the Ethnographic Atlass ethnic groups. Lastly, both the treatment and the instrumental variables are translated from the ethnicity level to modern district and country averages using population weights. In the first stage of the IV strategy, while suitability for plough-positive crops is shown to predict a higher probability of ancestral plough adoption, suitability for plough-negative crops predicts the opposite. Hence, the instruments successfully isolate exogenous variation in the treatment variable. In the second stage, Alesina, Giuliano, and Nunn (2013) find that individuals descending from plough-intensive societies tend to embrace more unequal gender attitudes. This plough-based IV has also been employed in numerous other empirical studies wishing to identify the causal effect of (modern) gender norms on various outcomes. For instance, Uberti and Douarin (2022) investigate how the trajectory of female labor force participation (FLFP) has been altered by traditional plough use during economic development. They show that a significant U-shaped pattern of female labor force participation only emerges in countries with a strong tradition of plough adoption, indicating a tendency of FLFP to decline during the early stages of development and to rise later. By contrast, the FLFPs path remains flat in countries with little to no legacy of plough agriculture. Similarly, Hazarika (2018) uses ancestral plough suitability as an instrument for womens participation in the economic and public spheres to study corruption. The author finds that once cultural endogeneity is accounted for, the correlation between higher womens participation and lower corruption disappears. The instrument has also been employed to investigate fertility behavior. Alesina et al. (2011) find that societies that traditionally adopted the plough today exhibit lower fertility rates, a pattern the authors attribute to the fact that in plough-based systems women and children were less useful as agricultural laborers, leading to a cultural preference for smaller families. Finally, research has examined outcomes related to gender violence. In ethnic groups with a legacy of plough usewhere women historically contributed less to productionwomen not only face higher levels of domestic violence but are also more likely to view such violence as acceptable ([Alesina et al.(2020)]).

2.4. Femicide and its causes

Definition and theoretical perspectives

Femicide is generally defined as intentional murder of women because they are women. As the gender component is essential in the motivation behind the killing, it is commonly committed by intimate partners or in settings where structural gender inequality prevails (Grzyb et al., 2018). The term first appeared in public in 1976 during the first International Tribunal on Crimes against Women thanks to the feminist sociologist Diana Russell (Russell, 1976).

The phenomenon is usually theoretically understood in the context of patriarchal domination (Cameron and Frazer, 1987; Caputi, 1987; Russell and Harmes, 2001; Taylor and Jasinski, 2011), while empirical studies tend to focus on the features and contexts that identify femicide as a social phenomenon (Campbell and Runyan, 1998; Collins, 2008). Further theoretical traditions adopt a criminological approach, framing femicide within the broader field of homicide studies (Liem and Pridemore, 2013; Bonanni et al., 2014; Dixon et al., 2008; Campbell et al., 2007; Dobash and Dobash, 2011) or focusing on honor crimes and the impact of past colonial domination on femicides (Shalhoub-Kevorkian (2002; Kulczycki and Windle (2011). Alternatively, the human rights perspective conceives femicides broadly, encompassing every form of extreme violence against women (Domazetoska et al., 2014; Filip and Platzer 2015; Laurent et al., 2013).

Stylized facts

According to UN WOMEN (2023), 85.000 women were intentionally killed in 2023, 51,000 of whom by intimate partners or family members, a common statistical proxy for femicide. Although men represent the majority of homicide victims (80% of total homicide victims), they are less likely to be killed in a domestic environment. Indeed, only 12% of male homicide victims are killed by an intimate partner or family member, compared to 60% of female homicide victims, suggesting the private component is core to fatal violence towards women. The connotation of this domestic dimension of femicide varies across continents. In Europe and in the Americas, intentional female homicides are mostly committed by intimate partners - 64% and 58%, respectively - whereas in the rest of the world the majority of reported cases involve family members (59%). Hence, it is important to account for both measures when assessing femicide at the global level.

Femicide statistics - proxied by female homicide by intimate partner or family member - display significant geographical heterogeneity. Africa presents the higher figures both in absolute number (21,700) and relative to the population (2.9 victims per 100,000). The Americas (1.6 victims per 100,000) and Oceania (1.5 victims per 100,000) show slightly lower rates of femicides, while Asia and Europe show the mildest ones (0.8 and 0.6 per 100,000, respectively). Due to limited data availability, the analysis of time trends and the assessment of COVID-19's impact are currently restricted to the Americas and Europe only. For instance, in the Americas, femicide rates remained quite stable between 2010 and 2023, whereas Europe displayed a drop (-20%) over the same period. Only some sub-regions (Northern America, Southern Europe, and Western Europe) saw notable increases in femicide rates in 2020, suggesting heterogeneity in Covid-19's effects.

Explanations Through a systematic review, Garcia-Vergara et al. (2022) summarizes the factors that are correlated with Intimate Partner Femicide (IPF) both at the individual level - i.e., the characteristics of the victims and the perpetrators - and at the aggregate level.

At the individual level, victims who live distant from their family, are socially isolated, and unemployed are more likely to experience femicide. Perpetrators are more likely to commit femicide if they are older than their spouse, have low levels of education, face precarious or no employment, lack adequate welfare support, have a history of migration, display severe alcohol consumption, or have a previous criminal record or poor mental health. The relevance of mental health is corroborated by evidence in criminology. For instance, several studies have highlighted how early exposure to trauma - such as violence and socio-economic hardship - and its persistence overtime increase the likelihood of committing violence and femicide (Narvey et al, 2024). Alongside, low-economic social status also increases the likelihood of committing crime by limiting social connections and access to jobs. This, in turn, lowers a mans self-esteem, thereby increasing his fear of losing control over his female partner, further incrementing his chances of exerting violence and committing femicide (Anderson, 1997).

At the aggregate level, femicides tend to be more common in rural areas (Capaldi et al., 2019) and where accessing weapons is easier. According to UNODC (2019), wealth inequality and total unemployment are predictors of femicides. However, culture plays a role too: the honor-based violence in South Asia (Ali et al., 2021) and the Latin American machismo (Corradi & Stangherlin, 2019) are notable examples of beliefs and traditions fostering violence and killings against women. Interestingly, Heise and Kostman (2015) find that the effect of economic measures on partner violence is no longer significant after controlling for gender norms. They also find that more educated girls are less likely to experience violence, and that women are more likely to experience femicide when they are employed in countries with a low share of working women. Consequently, UNODC (2022) underscores the importance of changing gender norms in the fight against femicides.

Institutions are also relevant. Research has found that a stronger rule of law is associated with both lower gender inequality (Barajas-Sandoval et al., 2022) and lower femicides (Testa et al., 2017), whereas corruption is related to higher incidents of violence against women (Velasco, 2020). Moreover, weak institutions often lead to impunity. Several studies on Latin America highlight that a lack of rule of law leaves femicides unpunished (Walsh & Menjívar, 2016a; Walsh & Menjívar, 2016b; Walsh & Menjívar, 2016c; Beck (2024; Beck & Stephen, 2021), thus undermining actions aimed at their reduction.

Nonetheless, existing evidence does not suggest that femicides' criminalization leads to their reduction. Saccomanno (2017) finds that making femicide legally punishable does not predict femicide rates, while low rule of law and limited female representation in decision-making bodies - such as national parliaments - emerge as the most significant factors to explain variation in femicide trends. Similarly, Devries et al. (2023) shows that there is no consistent correlation between femicide-specific legislation and lower incidence rates, a finding that may be attributed to the lack of effective implementation of such laws.

Policies aimed at reducing femicides exhibit mixed results. Relying on 66 countries, Whittington et al. (2023) find no correlation between the strength of implemented anti-femicide policies and the level of femicides at the end of the study period, only reporting a correlation with structural factors such as income and inequality. Global and local prevention programs range from awareness campaigns to emergency shelters. However, their effectiveness varies: NGO-led interventions such as community-based advocacy - e.g., SASA! in Uganda - show promise, whereas many government initiatives lack appropriate funding or strategic planning (Bott et al., 2019). Integrated approaches that combine legal reform, economic empowerment, education, and healthcare access tend to be the most effective (García-Moreno et al., 2020). Meanwhile, Moscoso provides evidence of policies' potential negative consequences: in Ecuador, municipalities that implemented a stricter enforcement coupled with an increased level of women's empowerment experienced rising rates of gender-based violence, suggesting a backlash effect.

2.5. Measuring Gender Norms and Femicides

Femicides Both femicides and gender norms are challenging to measure quantitatively. UNODC (2022) statistically classifies femicide as an intentional homicide with a gender component. Intentional homicide is defined as an unlawful death inflicted upon a person with the intent to cause death or serious injury and is hence characterized by three distinct elements:

- Objective criterion: the killing of a person by another person;
- Subjective criterion: the intent of the perpetrator to kill or injure the victim;
- Legal criterion: the unlawfulness of the killing.

Femicides are gender-motivated intentional homicides, i.e., violence that is directed against a woman precisely because she is a woman or that affects women disproportionately. This gender-related component is usually the hardest to capture statistically. States usually adopt different classifications and frameworks to record femicides, and heterogeneity in legal definitions leads to similar cases being handled differently depending on the legislation, undermining comparability. Variations in enforcement capabilities and legal culture further contribute to underreporting. To address this, UNODC (2022) has introduced a statistical framework for measuring femicides, aimed at improving reporting and harmonization.

Such a statistical framework disentangles intentional female homicides according to the relationship between the victim and its perpetrator. These homicides can be committed by intimate partners, other family members, or by other perpetrators, be them known or unknown. The first two categories account for the vast majority of all reported femicides and require objective and commonly recorded information to be identified. The third category requires additional statistical information - such as the perpetrator's previous criminal record or involvement in the sex industry - which is rarely available in official databases and represents only a small share of total recorded femicides. Hence, the UNODC's framework proposes *intentional female homicide by an intimate partner or family member* as the best available proxy for femicide, as it combines validity, relevance, and statistical feasibility. Furthermore, UNODC (2019) has highlighted that such homicides are overwhelmingly related to gender roles, strengthening their suitability as a proxy. Nonetheless, Sciarrino and Todesco (2025) argue that relying solely

on the victim-perpetrator relationship may obscure the motives underlying such killing. Drawing on Italian data, they show that most female homicides committed by family members other than intimate partners were not driven by gender-related motives, while misclassification of intimate partner cases amounted to only one among five cases. Thus, although concerns regarding the intimate partner proxy appear negligible, the family-member category may serve as a valuable supplementary proxy in developing countries, where honor killings are still relatively common and usually committed by the male partner's relatives.

Other common proxies are either difficult to collect uniformly or even more prone to underreporting e.g., data on womens disappearances, the presence of stalking or harassment or they are subject to emotional and visibility biases, such as media coverage. The most valid alternative is the total number of intentional female homicides. While it risks not to capture entirely the gender motivation behind the killing, it is nevertheless a comparable and commonly recorded measure that might overcome the issue of a narrow statistical definition of femicides, and it is relatively less prone to underreporting. Because four in ten homicides remain unclassified as they lack proper supporting data (UNODC, 2019), capturing the overall level of violence through a wider statistical indicator can be a relevant substitute of more precise categorizations in absence of supporting information (Sciarrino and Todesco, 2025).

We employ both the intentional female homicide by an intimate partner or family member and intentional female homicide as proxies for femicides. However, employing intentional homicides as a proxy still presents limitations. For instance, intentional homicides require a judicial authority to ascertain the presence of the subjective will to kill or to harm a woman, as well as the illegality of the act. It is likely that the intentional component is downplayed by judges in context where femicide is more prevalent. Additionally, in many Middle East countries *de facto* honor killings are still allowed, hence failing the unlawfulness criteria of the definition. We address this concern in the robustness section.

Gender Norms Gender norms are usually measured through the means of surveys, indexes, or objective and historical proxies. Common survey data include the World Values Survey or the General Social Survey, where gender norms are quantified through respondents' agreement to statements such as:

- When jobs are scarce, men should have more right to a job than women;
- "Men make better political leaders than women;"
- "University education is more important for a man than for a woman.

However, such surveys are typically prone to measurement error and social desirability bias. Furthermore, people tend to interpret questions from the lens of their culture (cultural bias) limiting comparability across heterogeneous contexts. Another concern is that individual preferences often differ from social norms. People tend to behave differently in public while potentially privately disagreeing with the norm, limiting the validity of aggregating individual preferences as a proxy of a group norm (Bursztyn et al., 2020). Gender norms indexes tend to overcome this issue by aggregating in a single measure both subjective and objective indicators such as formal laws, informal practices, and prevailing attitudes that disadvantage women spanning domains like the family code, physical integrity (including violence norms), son preference, access to resources, and civil liberties. A notable example is the Social Institutions and Gender Index (SIGI) developed by the OECD. One downside is that interpretability is limited and complementarity among the components is assumed. Other common critiques include the risk of oversimplification (Barnat et al., 2019) and capturing formal over substantial enforcement. Objective proxies include demographic patterns, namely male to female birth ratio (due to sex-selective practices) or fertility stopping rules, and observable gender gaps, such as female labor force participation rate, the gender gap in education, or womens share in parliament. Some scholars have argued that such proxies risk to be correlated with other economic or social variables such as urbanization or economic need, hence failing to capture the latent gender norm (Schmid and Elliot, 2023). They also argue that such proxies fail to represent the private component of preferences, being limited to public outcomes. Reverse causality and endogeneity remain a threat in identifying the role of gender norms for economic and social outcomes, as economic and social outcomes usually do have an influence on those objective proxies of norms.

A common solution to identification problems is relying on historical variables. Alesina et al (2013) adopt the plough use in pre-industrial context as a proxy for gender norms, finding significant persistence on female labor force participation and political representation. Alesina et al (2016) show that ethnic groups with historically patriarchal practices (e.g. high bride-price, patrilocal marriage, or a limited productive role for women) tend to have more domestic violence today. Other historical proxies include the kinship systems and marriage customs (Lowes, 2020). Historical proxies suffer some limitations. Because we are dealing with data retrieved from the past, measurement error remains a potential concern. As the presence of past gender norms is usually coded as a dichotomic variable, such approach has been accused of relying on an oversimplified conceptual framework and coding. In Alesina et al. (2013) the authors acknowledge they cannot test for timing of plough adoption. Dynamic changes are hard to test while it remains nontrivial to disentangle the effect of other historical forces. This might harm the assumption of cultural persistence that justifies the role of such proxies. Endogeneity might still arise if past gender norms are related to other confounding variables that are still related to todays outcomes (past economic development). Finally, such proxies are usually taken at the aggregate level, possibly failing to capture the role of individual preferences.

We employ as preferred operational measure for gender norms past plough adoption as in Alesina et al., (2013). We argue that, despite all limitations, this consolidated methodology has proven robust to a plethora of controls that reduce significantly potential concerns with endogeneity or confounding roles of other historical forces. Furthermore, the paper itself has shown that plough adoption is related to todays gender norms, proxied by objective gender gaps, making cultural persistence more credible. This reduces concerns of endogeneity compared to cross-sectional regression of femicides on gender norms.

2.6. Hypotheses and Empirical Evidence

See https://www.tandfonline.com/doi/full/10.1080/07418825.2024.2368135

- Amelioration (Anna)
- Backlash (Anna)

A third theoretical perspective posits a U-shaped association between gender norms and femicides (Vieratis et al., 2008): as society moves towards more equal gender norms, femicides may spike as a transitory backlash to perceived threats to patriarchal power (Brownmiller, 1975; Russell, 1975). With subsequent internalization of egalitarian norms and the implementation of policies that expand womens power, employment, and civic participation, victimization rates are expected to decline (Inglehart and Norris, 2005). Multiple studies corroborate this hypothesis. Bailey (1999) finds evidence of cross-sectional backlash effects of gender equality on gender-based violence (rapes) but long-term ameliorative trends through panel data. Whaley et al. (2013) finds evidence of a curvilinear relationship between gender equality and rates of inter-and intrasexual lethal violence using OLS after controlling for socio-demographic and individual variables.

3. Data

3.1. Key Variables

3.2. Data Sources

3.3. Data on soil type

3.4. Data on femicides

For our analysis of femicides, we use data from the United Nations Office on Drugs and Crime (UNODC), which provides information on intentional homicides disaggregated by the sex of the victim and the relationship of the perpetrator with the victim. Following the *Statistical Framework for Measuring the Gender-Related Killing of Women and Girls (also referred to as femicide/feminicide)* ([UN Women(2022)]), we employ as a proxy for femicides the intentional homicides of women, committed by an intimate male partner (IMP) or a family member.

This framework distinguishes between femicides committed within the domestic environment measured as intentional killings of women by an intimate male partner or family member, which we rely on and femicides occurring outside the domestic context.

The former measure has been widely adopted in the literature as a proxy for femicide ([Campbell et al.(2017)]; [Loinaz et al.(2018)]; [Sorrentino et al.(2022)]), as it enables cross-national comparability, involves minimal subjective judgment, and captures the majority of femicide cases ([UN Women(2022)]).

In contrast, the latter relies on supplementary data namely, sex/gender-related motives and indicators (SGRMIs) which are difficult to measure, often unavailable in existing datasets, and generally not comparable across national contexts ([Dawson & Carrigan(2021)]).

Utilizing homicide data helps mitigate concerns about under-reporting, a persistent issue in crime analysis. While gender-related violence metrics are vulnerable to reporting biases, people tend to report more in regions with a lower actual incidence ([Cullen(2023)]) homicides are less susceptible to concealment ([Fajnzylber et al.(2002)]; [Soares(2004)]).

To reduce measurement error and control for year-specific idiosyncratic shocks, we use average femicide rates per 100,000 population over the period 20052015. Our final merged sample comprises 140 countries from all continents.

3.5. Data for IV (AlesinaGiulianoNunn, 2013)

Overview. Our IV dataset replicates exactly the inputs and controls described, preserving sources, construction rules, and sample definitions. The endogenous regressor is *traditional plough use* (ancestral exposure), instrumented by exogenous, pre-industrial *agro-climatic suitability* for crops that are complementary to the plough ("plough-positive") versus substitutes for female-intensive hand cultivation ("plough-negative"). We adopt all historical, contemporary, and geographic controls as in the original specification, changing only the second-stage outcome (femicide rates) in our application.

Endogenous regressor: Traditional plough use. Traditional plough use is the estimated proportion of a countrys current population whose *ancestors* used the plough in pre-industrial agriculture. The mapping links ethnic groups in the Ethnographic Atlas to present-day language/ethnic distributions (Ethnologue/GREG), with documented imputation procedures when language data are missing; country-level averages are constructed accordingly. Alternative versions distinguish *indigenous* plough use from adoption post-European contact. We retain all variants and imputation schemes used for robustness.

Instruments: Plough-positive vs. plough-negative environments. The first stage uses two exogenous, pre-industrial agro-climatic measures computed from FAO GAEZ v3.0 at a 5-arc-minute grid: (i) average suitability for *plough-positive* cereals (wheat, barley, rye); and (ii) average suitability for *plough-negative* cereals (sorghum, pearl millet, foxtail millet). Suitability rasters for these six crops are extracted and aggregated as described below; these variables are strong predictors of traditional plough use in the first stage.

Geographic reference units and aggregation. All agro-climatic variables are constructed from GAEZ v3.0 cell-level suitability and then aggregated to historically relevant geography. For ethnographic controls and related historical geography, the Atlas centroid of each ethnic group is identified and a 200 km radius buffer is used to compute fractions suitable for specific crops (including the six instrument crops), which are then averaged to the country level by ancestral composition. We follow this procedure to ensure that instrument variation reflects *ancestral environments* rather than contemporary cultivation.

First-stage strength and specification. Instruments enter linearly as "plough-positive environment" and "plough-negative environment." First-stage F-statistics for the joint significance of the two instruments exceed conventional thresholds across specifications, and the coefficients on plough-positive suitability are large and precisely estimated; we replicate their construction and include the same historical controls and continent fixed effects in the first stage.

Historical Controls (Ethnographic Atlas)

We include the full set of pre-industrial covariates exactly as defined:

- **Historical economic development** (settlement complexity; Atlas v30; 18).
- **Political hierarchies** (jurisdictional layers beyond local community; Atlas v33; 15).
- **Presence of large domesticated animals** (indicator from Atlas v40).
- **Intensity of agriculture** (indicator for intensive/ irrigated; Atlas v28).
- **Absence of private property (land inheritance)** (Atlas v75).
- **Post-marital residence rules** (matrilocal, patrilocal; Atlas v12).
- **Family structure** (nuclear vs. extended; Atlas v8; indicators).
- **Subsistence composition** (shares of hunting and herding; Atlas v2, v4; category medians).
- **Atlas sampling year** (average observation year of ancestors; Atlas v102).

Contemporary Controls

All modern covariates are included (measured circa 2000 unless noted):

- **log real GDP per capita** (WDI), **Communism indicator**, **Democracy** (Polity2 $>$ 0).
- **Religious composition** (shares Catholic, Protestant, other Christian, Muslim, Hindu).
- **Per capita oil production** (BP Statistical Review; per person per day).
- **Sectoral value-added shares** (agriculture, manufacturing, services; WDI).
- **Conflict exposure** (years of civil and interstate conflict since 1816).
- **European-descent population share**.

Geographic Controls

To isolate the instrument from other ancestral geography, we include:

- **Agricultural suitability** (fraction of ancestor land suitable for barley, wheat, rye, sorghum, foxtail millet, pearl millet within 200 km of Atlas centroid; from GAEZ).
- **Tropical/subtropical share** (GAEZ 2002 classification, same 200 km procedure).
- **Terrain ruggedness index** (grid-based variance in elevation).
- **Soil depth constraints** and **terrain slope** (GAEZ 2002).
- **Baseline climate** (mean temperature and precipitation, 19501959).

Instrument Construction Details

1. Extract GAEZ v3.0 raster layers for: wheat, barley, rye (plough-positive) and sorghum, pearl millet, foxtail millet (plough-negative); align to a common 5-arc-minute grid.
2. Compute cell-level suitability indices and aggregate to the relevant ancestral geography (as above) to form country-level averages for *plough-positive environment* and *plough-negative environment*.
3. Use these two measures jointly in the first stage to instrument traditional plough use; retain the full control set and continent fixed effects. Reported F-statistics and coefficient patterns in the original first stage are strong; our construction follows the same pipeline.

Sample, Imputation, and Alternative Definitions

Country coverage and variable availability match the original: typical country-level samples lie in the 160180 range depending on outcome and control availability. Missing language information is imputed either via the national official language or GREG ethnic groups, with parallel robustness tables. We also retain the "indigenous-only" plough-use definition for sensitivity analysis.

What We Reuse for Our Femicide Application

We *unchanged* reuse: (i) traditional plough use; (ii) the two GAEZ-based instrument components; and (iii) the complete historical, contemporary, and geographic control set above. Only the second-stage dependent variable differs (femicide rates), preserving the original identification logic and data environment.

3.6. Descriptive Statistics

Descriptive statistics reveal intriguing patterns in our key variables. The average femicide rate across the countries in our sample is approximately 3.5 per 100,000 women, with significant variation observed between regions. Notably, countries with a history of plough adoption exhibit lower average femicide rates (2.1 per 100,000) compared to those without plough adoption (5.8 per 100,000). This initial observation supports our hypothesis regarding the paradoxical relationship between conservative gender norms and femicide rates. Furthermore, the average GDP per capita in plough-adopting countries is significantly higher, suggesting that economic development may also play a role in shaping gender norms and violence against women. These descriptive statistics set the stage for our subsequent econometric analysis, where we employ an instrumental variable strategy to further explore the causal relationship between historical plough adoption and contemporary femicide rates.

4. Empirical Strategy

To investigate the relationship between historical plough adoption and contemporary femicide rates, we employ a robust empirical strategy that integrates an instrumental variable (IV) approach. This strategy is designed to address potential endogeneity issues arising from omitted variable bias and reverse causality.

4.1. Identification and Instrumental Variable Strategy

Our identification strategy follows alesina2013, using the **historical suitability for plough agriculture** as an instrument for contemporary gender norms. To address endogeneity issues concerning the relationship between historical plough use measured at the ethnicity level and gender norms,

Alesina et al. (2013) employ an Instrumental Variable approach based on crops differential ancestral geo-climatic suitability to plough technology. This strategy is motivated by Pryor (1985), who argues that different types of cultivated crops benefit differently from plough use. Crops that require large tracts of land to be prepared quickly benefit more from plough technology, whereas crops that require less intensive soil preparation or could be grown in soils where ploughing is difficult benefit less from it. The former referred to as *plough-positive* crops include teff, wheat, rye, and wet rice; conversely, the latter defined by Prior (1985) as *plough-negative* crops comprise maize, sorghum, millet, and various types of root and tree crops. For each ethnic group in the Ethnographic Atlas using the FAOs Global Agro-Ecological Zones data the authors compute the share of land within 200 km of the groups centroid that was suitable for plough-positive and plough-negative crops, respectively, normalized by overall arable land. Two ethnicity-level measures of suitability to plough use are then constructed: the ancestral suitability of an ethnic groups land for plough positive, and the ancestral suitability of an ethnic groups land for plough-negative crops. These variables provide exogenous indicators of how apt each ethnicitys ancestry to the adoption of the plough was, and are thus used by Alesina et al. (2013) to instrument historical plough use by ethnic group. Next, historical plough use (the treatment) and ancestral crop suitability (the instruments) are linked to present-day populations on a granular spatial scale by matching each of the Ethnologue language groups to one of the Ethnographic Atlass ethnic groups. Lastly, both the treatment and the instrumental variables are translated from the ethnicity level to modern district and country averages using population weights. In the first stage of the IV strategy, while suitability for plough-positive crops is shown to predict a higher probability of ancestral plough adoption, suitability for plough-negative crops predicts the opposite. Hence, the instruments successfully isolate exogenous variation in the treatment variable. In the second stage, Alesina, Giuliano, and Nunn (2013) find that individuals descending from plough-intensive societies tend to embrace more unequal gender attitudes. This plough-based IV has also been employed in numerous other empirical studies wishing to identify the causal effect of (modern) gender norms on various outcomes. For instance, Uberti and Douarin (2022) investigate how the trajectory of female labor force participation (FLFP) has been altered by traditional plough use during economic development. They show that a significant U-shaped pattern of female labor force participation only emerges in countries with a strong tradition of plough adoption, indicating a tendency of FLFP to decline during the early stages of development and to rise later. By contrast, the FLFPs path remains flat in countries with little to no legacy of plough agriculture. Similarly, Hazarika (2018) uses ancestral plough suitability as an instrument for womens participation in the economic and public spheres to study corruption. The author finds that once cultural endogeneity is accounted for, the correlation between higher womens participation and lower corruption disappears. The instrument has also been employed to investigate fertility behavior. Alesina et al. (2011) find that societies that traditionally adopted the plough today exhibit lower fertility rates, a pattern the authors attribute to the fact that in plough-based systems women and children were less useful as agricultural laborers, leading to a cultural preference for smaller families. Finally, research has examined outcomes related to gender violence. In ethnic groups with a legacy of plough usewhere women historically contributed less to productionwomen not only face higher levels of domestic violence but are also more likely to view such violence as acceptable ([Alesina et al.(2020)]).

5. Results: Cross-Sectional Evidence

5.1. Baseline OLS Estimates

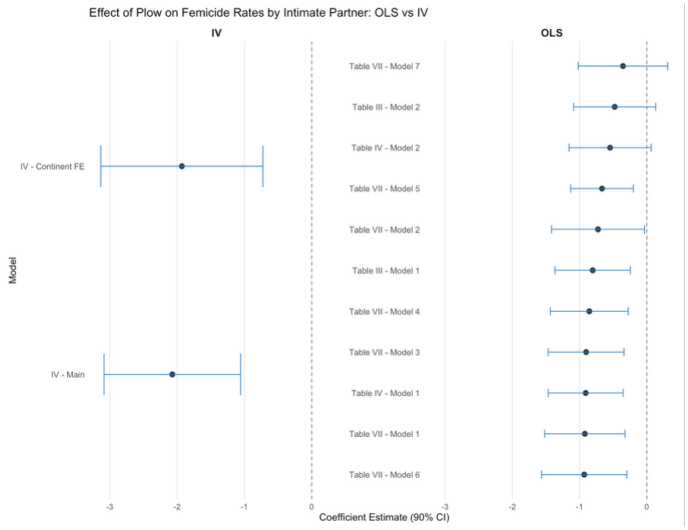
We begin by estimating the relationship between historical plough adoption and contemporary femicide rates using cross-sectional OLS specifications. Across all model variants, the coefficient on *plough adoption* is consistently negative and statistically significant. This pattern is robust to the inclusion of extensive control variables that account for geographic conditions, colonial history, institutional development, religious composition, and conflict exposure.

Importantly, in every specification we include the *average male homicide rate* as a control. This serves as a proxy for the general level of violence in a country and ensures that our results capture femicide-specific dynamics rather than broader cross-country variation in violent crime. After conditioning on these covariates, the OLS estimates continue to indicate that greater historical plough adoption is associated with *lower contemporary femicide rates.*

5.2. Instrumental Variable Estimates

To address concerns of omitted variable bias and reverse causality, we employ the instrumental variable (IV) strategy proposed by alesina2013. Specifically, we use the historical agro-climatic suitability for plough-positive versus plough-negative crops as instruments for traditional plough use. The first-stage results are strong: the F-statistics for instrument relevance exceed conventional thresholds.

The IV coefficient on plough adoption is negative and statistically significant, even after includign continent FE. This implies that countries whose ancestors were more likely to adopt the plough exhibit significantly lower femicide rates today. The persistence of this effect after instrumenting suggests a causal interpretation: historical agricultural practices shaped gender norms in ways that continue to affect contemporary violence against women.



5.3. Robustness and Placebo Tests

We conduct a series of robustness checks to ensure that our results are not driven by specification choices or influential observations.

- **Alternative controls.** Including additional covariates-such as economic development, education, and religion-does not materially affect the estimated coefficient.
- **Placebo outcome.** When we replace femicide rates with male homicide rates as the dependent variable, we find no systematic relationship with plough adoption, confirming that the effect is specific to gender-based violence.
- **Influential observations.** DFBETA statistics reveal that no single country exerts undue influence on the estimated coefficients.
- **Regional heterogeneity.** The negative association holds across continents, although magnitudes differ somewhat across regions (see Appendix for details).
- **Machine Learning Specification** We follow Baiardi and Naghi (2024) who adopt different machine learning techniques to test the robustness of the original results of Alesina et al. (2013). We estimated again both OLS and IV specifications (full details in the Appendix) by employing different machine learning techniques, namely lasso, trees, neural net, random forest, boosting, ensemble and best. The advantage of this approach is the flexibility in selecting controls and their functional form. This is particularly convenient in a setting with many controls but relatively few observations. Results are overall comparable to baseline specifications.

5.4. Interpretation and Mechanism

Taken together, the cross-sectional evidence suggests a counterintuitive pattern: although plough adoption historically entrenched conservative gender norms, these same norms are associated with *lower femicide rates* in the present. One potential explanation is a *male backlash mechanism*, whereby stricter gender roles reduce the perceived threat of female empowerment and thus dampen violence against women. Another complementary explanation involves *endogenous reporting*: in societies with stronger traditional norms, reporting of intimate partner violence may differ systematically. We explore these mechanisms further in the regional analysis section.

To check: role of education attainment and labor force participation.

6. Looking ahead: Potential extensions

6.1. Regional Analysis

A promising extension of our work is to exploit **regional-level data** within countries. Such data would allow us to refine identification and address several concerns:

1. **Endogenous reporting.** By comparing regions within the same country (e.g., U.S. counties or European NUTS units), we can mitigate cross-country differences in reporting standards of femicides.
2. **Role of institutions.** Regional variation enables us to hold national-level institutions constant, thus isolating the role of cultural persistence from institutional quality.
3. **Regional fixed effects.** Subnational analysis allows the inclusion of country or region fixed effects, potentially increasing statistical power by adding more observations.

Implementation strategies. We propose two complementary approaches.

- **Region-level aggregation.** Assign to each region the plough instrument and the average femicide rate, then estimate specifications analogous to the national-level regressions.
- **Extension to individual-level datasets.** When microdata contain ethnicity identifiers (e.g., the World Values Survey), we can assign ancestral plough use at the respondent level. Using census data, we may then compute each regions average plough intensity based on its ethnic composition.

Limitations. Several challenges remain. First, detailed regional data on femicides are scarce and often inconsistently reported. Second, applying the plough instrument at the individual level requires datasets with ethnicity information, which are not always available. Third, because individual-level femicide data is difficult to obtain, we would need to rely on victimization records or perpetrator data (e.g., inmates), which introduces potential measurement issues.

6.2. Time Series Analysis

An additional extension is to examine the dynamics of the relationship between gender norms and violence over time. The cross-sectional evidence identifies a persistent correlation between historical plough adoption and contemporary femicide rates, but does not capture how this relationship may evolve during different stages of economic development.

The key hypothesis is that the interaction between womens labor force participation and male responses to changing gender roles is not static. In the early stages of development, as women begin to enter the labor market in larger numbers, traditional norms may be challenged. Men, unaccustomed to female economic participation, may perceive this change as a threat and react with violence—a pattern consistent with

the *male backlash* hypothesis. However, as development proceeds and new cohorts grow up in households where mothers are active in the labor market, men become increasingly accustomed to women working outside the home. This intergenerational transmission gradually normalizes female labor participation, reducing the perceived threat and thereby lowering violence against women.

A time-series perspective therefore allows us to trace the dynamic adjustment path: from an initial period of heightened violence during the transition, to a subsequent decline once gender norms adapt. Identifying these non-linear dynamics is crucial to understand whether femicide is primarily a transitional phenomenon linked to norm shifts, or a persistent outcome of entrenched patriarchal systems. Future work combining panel data on femicide, womens labor market participation, and institutional development would allow us to test these dynamic hypotheses more directly.

7. Limitations

While our analysis provides novel insights into the relationship between historical plough adoption and contemporary femicide rates, several limitations must be acknowledged.

Data and Measurement Issues

First, cross-national regressions inevitably rely on a relatively small number of observations, which limits statistical power and makes results sensitive to sample composition. Second, measurement of the dependent variable—femicide—is subject to underreporting and inconsistent classification across countries. Although we harmonize multiple sources, remaining discrepancies in reporting standards may attenuate our estimates.

Controls and Possible Omitted Variable Bias

A second concern relates to incomplete controls. While we include a wide set of historical, geographic, and contemporary covariates, missing data on potentially relevant variables (e.g., law enforcement effectiveness, cultural attitudes toward reporting) may introduce omitted variable bias. Controlling for average homicide rates helps capture the general level of violence, but residual confounding remains possible. Similarly, the inclusion of national fixed effects may not fully absorb unobserved heterogeneity in institutional capacity or social norms.

Limitations Inherited from the Plough Instrument

Finally, our empirical strategy inherits some of the limitations of Alesina2013. In particular, the validity of the exclusion restriction requires that agro-climatic suitability for plough-positive crops affects femicide rates only through historical plough adoption and not through alternative channels. Although we include a comprehensive set of geographic controls, this assumption remains untestable. Moreover, the identification strategy relies on a cultural persistence mechanism: the idea that ancestral norms are transmitted intergenerationally to shape contemporary outcomes. Migration and demographic change may weaken this persistence, complicating the interpretation of results.

In sum, while our instrumental variable approach provides a credible framework for identifying the causal impact of historical plough adoption, these limitations highlight the importance of interpreting our findings with caution and motivate future work using more granular regional and longitudinal data.

8. Discussion and Policy Implications

The findings of this study present a significant challenge to the conventional understanding of the relationship between historical agricultural practices and contemporary gender-based violence. Contrary to the prevailing notion that societies with a history of plough adoption exhibit higher rates of femicide due to entrenched patriarchal norms, our analysis reveals a surprising inverse relationship. This paradox raises critical questions about the mechanisms through which historical gender norms influence present-day outcomes.

The male backlash hypothesis posits that in societies where men perceive a threat to their traditional rolesoften exacerbated by economic modernizationthere may be a violent response aimed at reasserting control over women. However, our results suggest that societies with a legacy of plough agriculture, which are often characterized by conservative gender norms, paradoxically report lower rates of femicide. This finding implies that the cultural persistence of these norms may foster a protective environment for women, counteracting the potential for violence that could arise from economic shifts.

From a policy perspective, these insights underscore the importance of considering historical and cultural contexts when designing interventions aimed at reducing gender-based violence. Policies that aim to challenge deeply rooted gender norms without an understanding of their historical underpinnings may inadvertently provoke backlash and increase violence. Therefore, a nuanced approach that recognizes the complexities of gender dynamics in different cultural settings is essential for effective policy formulation.

8.1. Interpretation

The unexpected inverse relationship between historical plough adoption and contemporary femicide rates invites a re-examination of the male backlash hypothesis. While traditional economic theories suggest that societies with more conservative gender norms would exhibit higher rates of femicide, our findings indicate that these norms may also provide a stabilizing effect that reduces violence against women. This interpretation aligns with the notion that cultural factors, rather than purely economic ones, play a pivotal role in shaping gender-based violence.

Moreover, the use of an instrumental variable strategy, as proposed by Alesina et al. (2013), allows for a more robust identification of the causal mechanisms at play. By leveraging the suitability of plough agriculture as an instrument, we can isolate the impact of historical gender norms on contemporary femicide rates, mitigating concerns regarding endogeneity. This methodological rigor enhances the credibility of our findings and highlights the importance of employing appropriate identification strategies in social science research.

The implications of this interpretation extend beyond academic discourse; they challenge policymakers to rethink their approaches to gender violence. Instead of solely focusing on

economic empowerment as a means to reduce femicide, it may be equally important to engage with and transform cultural narratives surrounding gender roles.

8.2. Broader Relevance

The broader relevance of our findings extends to various domains, including development economics, gender studies, and public policy. Understanding the historical roots of gender norms and their implications for contemporary social issues is crucial for addressing gender-based violence globally. Our study contributes to a growing body of literature that emphasizes the need for interdisciplinary approaches to tackle complex social phenomena.

Furthermore, the paradox of plough adoption and femicide rates raises important questions about the role of cultural persistence in shaping societal outcomes. It suggests that interventions aimed at promoting gender equality must consider the historical and cultural contexts in which they are implemented. For instance, in societies with a strong legacy of plough agriculture, efforts to empower women economically must be accompanied by initiatives that address cultural perceptions of gender roles and violence.

Additionally, our findings have implications for international development organizations that seek to implement gender-focused programs in diverse cultural settings. A one-size-fits-all approach may not be effective; instead, tailored strategies that account for historical legacies and cultural norms are necessary for achieving meaningful progress in reducing gender-based violence.

In conclusion, the interplay between historical agricultural practices and contemporary gender norms is a complex and multifaceted issue that warrants further exploration. Our study opens avenues for future research that can deepen our understanding of the dynamics between culture, gender, and violence, ultimately informing more effective policies and interventions.

9. Conclusion

The findings of our study present a compelling narrative that challenges conventional wisdom regarding the relationship between historical plough adoption and contemporary femicide rates. Contrary to the prevailing assumption that conservative gender norms, rooted in agricultural practices such as plough adoption, would correlate with higher rates of femicide, our analysis reveals a surprising inverse relationship. Societies with a history of plough adoption, which are often characterized by more entrenched patriarchal norms, exhibit significantly lower rates of femicide today. This paradox invites a deeper exploration into the dynamics of cultural persistence and the male backlash hypothesis.

Utilizing an instrumental variable (IV) strategy as outlined by Alesina et al. (2013), we have integrated a robust dataset that spans cross-country and regional analyses. Our methodology not only leverages historical data on plough suitability but also incorporates ethnographic insights that illuminate the cultural underpinnings of gender norms. The robustness of our findings is further supported by extensive controls, placebo tests, and an examination of regional variations that affirm the validity of our conclusions.

The male backlash hypothesis posits that in societies where gender norms are historically conservative, the perceived threat to male dominance may paradoxically lead to protective behaviors that reduce violence against women. This theoretical framework provides a nuanced understanding of how cultural legacies can shape contemporary social outcomes, particularly in the context of gender-based violence. Our results suggest that the mechanisms of cultural persistence may operate in complex ways, where the very norms that are expected to perpetuate violence instead foster environments that mitigate it.

In summary, our research contributes to the growing body of literature that interrogates the intersections of culture, gender, and violence. By elucidating the unexpected relationship between historical plough adoption and femicide rates, we encourage further inquiry into the socio-cultural dynamics that inform gender relations across different contexts. Future research should continue to explore these themes, employing diverse methodologies and theoretical frameworks to deepen our understanding of the factors that influence gender-based violence in contemporary societies.

References

[Alesina et al.(2013)] Alesina, A., Giuliano, P., & Nunn, N. (2013). *On the Origins of Gender Roles: Women and the Plough*. The Quarterly Journal of Economics, 128(2), 469-530.

[Goldin(2006)] Goldin, C. (2006). *The Quiet Revolution That Transformed Women's Employment, Education, and Family*. The American Economic Review, 96(2), 1-21.

[Duflo(2012)] Duflo, E. (2012). *Women Empowerment and Economic Development*. Journal of Economic Literature, 50(4), 1051-1079.

[Akerlof(2007)] Akerlof, G. A. (2007). *The Missing Motivation in Macroeconomics*. The American Economic Review, 97(1), 1-24.

[Heckman(2006)] Heckman, J. J. (2006). *Skill Formation and the Economics of Investing in Disadvantaged Children*. Science, 312(5782), 1900-1902.

[Jayachandran(2015)] Jayachandran, S. (2015). *The Roots of Gender Inequality in Developing Countries*. Annual Review of Economics, 7, 1-30.

[Fajnzylber et al.(2002)] Fajnzylber, P., Lederman, D., & Loayza, N. (2002). *What causes violent crime?*. European Economic Review, 46(7), 1323–1357.

[Cullen(2023)] Cullen, C. (2023). *Method matters: The underreporting of intimate partner violence*. The World Bank Economic Review, 37(1), 49–73.

[UN Women(2022)] UN Women. (2022). *Gender-related killings of women and girls (femicide/feminicide)*. UN Women.

[Campbell et al.(2017)] Campbell, J. C., Webster, D., Koziol-McLain, J., Block, C., Campbell, D., Curry, M. A., Gary, F., Glass, N., McFarlane, J., Sachs, C., et al. (2017). *Risk*

factors for femicide in abusive relationships: Results from a multisite case control study. In *Domestic violence* (pp. 135–143). Routledge.

[Loinaz et al.(2018)] Loinaz, I., Marzabal, I., & Andrés-Pueyo, A. (2018). *Risk factors of female intimate partner and non-intimate partner homicides*. European Journal of Psychology Applied to Legal Context, 10(2), 49–55.

[Sorrentino et al.(2022)] Sorrentino, A., Cinquegrana, V., & Guida, C. (2022). *Risk factors for intimate partner femicide–suicide in Italy: an ecological approach*. International Journal of Environmental Research and Public Health, 19(16), 10431.

[Dawson & Carrigan(2021)] Dawson, M., & Carrigan, M. (2021). *Identifying femicide locally and globally: Understanding the utility and accessibility of sex/gender-related motives and indicators*. Current Sociology, 69(5), 682–704.

[Soares(2004)] Soares, R. R. (2004). *Development, crime and punishment: accounting for the international differences in crime rates*. Journal of Development Economics, 73(1), 155–184.

[EIGE(2023)] European Institute for Gender Equality. (2023). *Gender Norms*.

[Lomazzi(2023)] Lomazzi, V. (2023). *The Cultural Roots of Violence against Women: Individual and Institutional Gender Norms in 12 Countries*. Social Sciences, 12(3), 117.

[Flood & Pease(2009)] Flood, M., & Pease, B. (2009). *Factors Influencing Attitudes to Violence Against Women*. Trauma, Violence, & Abuse, 10(2), 125–142.

[Knickmeyer et al.(2010)] Knickmeyer, N., Levitt, H., & Horne, S. G. (2010). *Putting on Sunday Best: The Silencing of Battered Women Within Christian Faith Communities*. Feminism & Psychology, 20(1), 94–113.

[Tarar & Pulla(2014)] Tarar, M. G., & Pulla, V. (2014). *Patriarchy, Gender Violence and Poverty amongst Pakistani Women: A Social Work Inquiry*. International Journal of Social Work and Human Services Practice, 2(2), 56–63.

[Saunders et al.(2022)] Saunders, D. G., Jiwatram-Negrón, T., Cardenas, I., & Nanasi, N. (2022). *Patriarchys Link to Intimate Partner Violence: Applications to Survivors Asylum Claims*. Violence Against Women, 29(11), 1998–2021.

[UN Women(2024)] UN Women. (2024). *Five Essential Facts to Know About Femicide*.

[Segrest & Domke-Damonte(2003)] Segrest, S. L., & Domke-Damonte, D. J. (2003). *Exploring the role of machismo in gender discrimination: a comparison of Mexico and the US*. Equal Opportunities International, 22(1), 13–31.

[Cislaghi & Heise(2019)] Cislaghi, B., & Heise, L. (2019). *Gender norms and social norms: differences, similarities and why they matter in prevention science*. Sociology of Health & Illness, 42(2), 407–422.

[Pfau-Effinger(1998)] Pfau-Effinger, B. (1998). *Gender cultures and the gender arrangement a theoretical framework for crossnational gender research*. Innovation: The European Journal of Social Science Research, 11(2), 147–166.

[Yu(2005)] Yu, W.-H. (2005). *Changes in Women's Postmarital Employment in Japan and Taiwan*. Demography, 42(4), 693–717.

[Savat(2023)] Savat, S. (2023). *Research highlights gender bias persistence over centuries*. The Source, March 13. Retrieved from https://source.washu.edu/2023/03/research-highlights-gender-bias-persistence-over-centuries/.

[Bredtmann et al.(2020)] Bredtmann, J., Otten, S., & Rulff, C. (2020). The intergenerational transmission of gender role attitudes: Evidence from immigrant mothers-in-law. *Journal of Economic Behavior & Organization*, 179, 101–115. https://doi.org/10.1016/j.jebo.2020.08.021.

[Giuliano(2020)] Giuliano, P. (2020). OUP accepted manuscript. *Oxford Review of Economic Policy*, 36(4), 944–961. https://doi.org/10.1093/oxrep/graa044.

[Lowes(2023)] Lowes, S. (2023). Kinship structure and the family: Evidence from the matrilineal belt. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4236471.

[Robinson and Gottlieb(2019)] Robinson, A. L., & Gottlieb, J. (2019). How to close the gender gap in political participation: Lessons from matrilineal societies in Africa. *British Journal of Political Science*, 51(1), 1–25. https://doi.org/10.1017/s0007123418000650.

[Gneezy et al.(2009)] Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5), 1637–1664. https://doi.org/10.3982/ecta6690.

[Mulawa et al.(2017)] Mulawa, M. I., Kajula, L. J., & Maman, S. (2017). Associations between peer network gender norms and the perpetration of intimate partner violence among urban Tanzanian men: A multilevel analysis. *Prevention Science*, 19(4), 427–436. https://doi.org/10.1007/s11121-017-0835-8.

[Masters et al.(2020)] Masters, S. L., Ford, J. L., & Wittenborn, A. K. (2020). Perceptions of gender norm violations among middle school students: An experimental study of the effects of violation type on exclusion expectations. *The Journal of Early Adolescence*, 41(4), 527–549. https://doi.org/10.1177/0272431620931193.

[Chen et al.(2025)] Chen, L., Liu, Y., & Zhang, H. (2025). Composition of peer mothers and gender norms: Class randomization and short-run effects. *Journal of Economic Behavior & Organization*, 234. https://doi.org/10.1016/j.jebo.2025.107008.

[Wolter et al.(2015)] Wolter, I., Glüer, M., & Hannover, B. (2015). Reading is for girls!? The negative impact of preschool teachers traditional gender role attitudes on boys reading related motivation and skills. *Frontiers in*

Psychology, 6, Article 1267. https://doi.org/10.3389/fpsyg.2015.01267.

[Seluman et al.(2024)] Seluman, I., Musa, H., & Adeyemi, T. (2024). Stereotypical portrayal of gender in mainstream media and its effects on societal norms: A theoretical perspective. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(1). Retrieved from https://www.allmultidisciplinaryjournal.com/uploads/archives/20240711205605_A-24-142.1.pdf.

[Galdi et al.(2013)] Galdi, S., Maass, A., & Cadinu, M. (2013). Objectifying media: Their effect on gender role norms and sexual harassment of women. *Psychology of Women Quarterly*, 38(3), 398–413. https://doi.org/10.1177/0361684313515185.

[Boserup(1970)] Boserup, E. (1970). *Womans role in economic development*. London: George Allen and Unwin Ltd.

[Alesina et al.(2013)] Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2), 469–530. https://doi.org/10.1093/qje/qjt005.

[Giuliano(2014)] Giuliano, P. (2014). The role of women in society: From preindustrial to modern times. *CESifo Economic Studies*, 61(1), 33–52. https://doi.org/10.1093/cesifo/ifu019.

[Pryor(1985)] Pryor, F. L. (1985). The invention of the plow. *Comparative Studies in Society and History*, 27(4), 727–743. https://www.jstor.org/stable/178600.

[Uberti and Douarin(2022)] Uberti, L. J., & Douarin, E. (2022). The feminisation U, cultural norms, and the plough. *Journal of Population Economics*, 36, 5–35. https://doi.org/10.1007/s00148-022-00890-5.

[Hazarika(2018)] Hazarika, G. (2018). The plough, gender roles, and corruption. *Economics of Governance*, 19(2), 141–163. https://doi.org/10.1007/s10101-018-0202-7.

[Alesina et al.(2011)] Alesina, A., Giuliano, P., & Nunn, N. (2011). Fertility and the plough. *American Economic Review*, 101(3), 499–503. https://doi.org/10.1257/aer.101.3.499.

[Alesina et al.(2020)] Alesina, A., Brioschi, B., & La Ferrara, E. (2020). Violence against women: A cross-cultural analysis for Africa. *Economica*, 88(349). https://doi.org/10.1111/ecca.12343.

[Anderson & Umberson(2001)] Anderson, K. L., & Umberson, D. (2001). *Gendering Violence: Masculinity and Power in Mens Accounts of Domestic Violence*. Gender & Society, 15(3), 358380. https://doi.org/10.1177/089124301015003003

[Bicchieri(2006)] Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.

[Giuliano(2017)] Giuliano, P. (2017). *Gender: A Historical Perspective*. Oxford Academic. https://doi.org/10.1093/acrefore/9780190625979.013.138

[Taccini & Mannarini(2023)] Taccini, F., & Mannarini, S. (2023). *An Attempt to Conceptualize the Phenomenon of Stigma toward Intimate Partner Violence Survivors: A Systematic Review*. Behavioral Sciences, 13(3), 194. https://doi.org/10.3390/bs13030194

10. Appendix

10.1. Data Sources and Methodology

This study employs an instrumental variable (IV) strategy to explore the relationship between historical plough adoption and contemporary femicide rates. The primary data sources include historical agricultural records, contemporary crime statistics, and ethnographic studies that provide insights into gender norms across various societies. The instrumental variable used is the suitability of land for plough agriculture, which serves as a proxy for historical plough adoption rates.

10.2. Robustness Checks

To ensure the validity of our findings, we conduct several robustness checks. These include alternative specifications of the regression model, the inclusion of additional control variables such as GDP per capita, education levels, and urbanization rates. We also perform placebo tests by examining the relationship between plough adoption and other forms of violence that are not gender-specific, to confirm that our results are not driven by confounding factors.

10.3. Ethnographic Insights

Ethnographic data plays a crucial role in understanding the cultural persistence of gender norms. We analyze case studies from various regions, highlighting how historical agricultural practices have shaped societal attitudes towards gender roles. This qualitative evidence complements our quantitative analysis, providing a richer context for interpreting our results.

10.4. Regional Variations

Our analysis also considers regional variations in femicide rates and gender norms. By disaggregating the data by region, we can identify patterns that may differ across cultural contexts. This allows us to assess the generalizability of our findings and to explore how local factors may interact with historical plough adoption to influence contemporary outcomes.

10.5. Limitations

While our study provides significant insights, it is important to acknowledge its limitations. The reliance on historical data may introduce measurement error, and the cross-sectional nature of our analysis limits causal inference. Future research could benefit from longitudinal studies that track changes in gender norms and femicide rates over time, as well as experimental designs that test the impact of interventions aimed at altering gender norms.

10.6. Machine Learning Specification

We follow Baiardi and Naghi (2024) who adopt the DML methodology introduced by Chernozhukov (2017, 2018). The

average treatment effect θ_0 in is obtained from a residual on residuals regression:

$$(Y - \mathbb{E}[Y|X]) = (D - \mathbb{E}[D|X])\theta_0 + W, \quad (1)$$

where $\mathbb{E}[(D - \mathbb{E}[D|X])W] = 0$. The conditional means $\mathbb{E}[Y|X]$ and $\mathbb{E}[D|X]$ are obtained by the different supervised machine learning model.

Cross-fitting approach is adopted to reduce potential bias stemming from regularization and overfitting, as in Chernozhukov et al. (2018). Chernozhukov et al. (2018) extends the DLM to an IV setting where the residuals $D - \mathbb{E}[D|X]$ are instrumented by the residuals of a regression of instrumental variables on the covariates.

Table 1 shows results for both baseline and IV estimates. Results are overall comparable to the original empirical strategy.

	Lasso	Trees	Boosting	Forest	Nnet	Ensemble	best	OLS
Median ATE	-1.274	-0.793	-0.883	-0.903	-1.136	-1.021	-0.904	-0.931
se	0.336	0.287	0.320	0.309	0.331	0.330	0.311	0.378

Table 1. Note: Analysis of the main robustness checks using DML. Column 8 reports the baseline and OLS estimates and IV. Standard errors adjusted for variability across splits usingthe median method are reported for the DML estimates. Robust standard errors are reported in column 8. The number of covariates does not include the treatment variable. All specifications control for current income squared, economic complexity, political hierarchies, a dummy for tropical climate, the use of large animals and agricultural suitability.