BJAR

Vol. 1, No. 1

March 2026

Hamilton, Bermuda
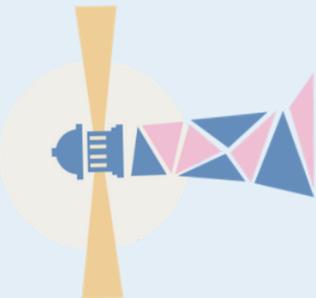
Front matter images from top to bottom:

1. Elizabeth Rollins-Scott

2. Dion Greene

3. Dion Greene

Contents

A Note from the Chief Editors

We founded the *Bermuda Journal of Academic Research* (BJAR) in September 2024. Our vision for the journal was to create an accessible platform for multidisciplinary research to connect Bermuda's community with regional and global conversations in academia. We focused on promoting and celebrating local research while also attracting insightful submissions from academics worldwide. With the cultural flair captured by local creatives, we hope to illustrate how research and Bermudian identity may coalesce.

There are two sections in this journal: peer-reviewed submissions and aspiring researcher submissions. Our peer-reviewed submissions were evaluated by experts in their respective fields, ensuring a high level of academic rigour. Submissions in the aspiring researcher category come from emerging scholars and were reviewed and edited by our editorial team to support their development.

We are proud to share the inaugural edition of the *Bermuda Journal of Academic Research* with the community. Disciplines present in this issue include international economics, development economics, law and public policy, and gender studies. Themes include cannabis legalization, partition violence, rural labour markets, sovereign debt dynamics, and trade conflicts. Thank you for joining us in celebrating knowledge creation, and we look forward to continuing this work and supporting the growth of scholarship in Bermuda for years to come.

Kyaida and Ywione
Chief Editors
BJAR

Acknowledgements

We are thankful for everyone who contributed to this first edition and endeavour, specifically our Editorial Board, Editorial Team, Peer Reviewers and Editors, creatives, sponsors, and most importantly the authors.

CANNABIS PROHIBITION: RACE,
EQUALITY, AND THE CASE FOR
REFORM

Ryan Robinson Perinchief

PEER REVIEW

Cannabis Prohibition: Race, Equality, and the Case for Reform
An Analysis of Democratic Equality Theory and the Disproportionate Impact of Cannabis Law & Policy on Black and Minority Ethnic Groups

Ryan Robinson Perinchief

University of Durham

Abstract

The original intent for the war on cannabis is heavily linked to racial prejudice and discrimination. The consequences of these tainted origins of drug laws can be traced through its impacts on particular racial groups—specifically black and ethnic minorities—including through targeted enforcement and policies, leading to more arrests for drug offences and, subsequently, disproportionate criminalisation and social exclusion. These practices have been used to justify a repeating cycle of discrimination; institutionalised racism and systemic inequality upheld by the law. This dissertation contends that the current regime of cannabis prohibition does not justify its disproportionate and negative effects, which are contrary to democratic equality. Furthermore, in consideration of shifting attitudes in global cannabis policy and emerging trends towards legalisation and commercialisation, an equality-based approach must be adopted to remedy injustices directly inflicted by prohibition and restore full enjoyment of democratic equality.

1. Introduction

The original intent for the war on cannabis is heavily linked to racial prejudice and discrimination. The consequences of these tainted origins of drug laws can be traced through its impacts on particular racial groups—specifically black and ethnic minorities—including through targeted enforcement and policies, leading to more arrests for drug offences and, subsequently, disproportionate criminalisation and social exclusion.[1] These practices have been used to justify a repeating cycle of discrimination; institutionalised racism and systemic inequality upheld by the law. This present situation is contrary to egalitarian principles and theories of justice—often regarded as the 'first virtue of social institutions'[2]—of which the political state and its prevailing systems, laws and regulations, and the administration of law and order are included. Specifically, Elizabeth Anderson's 'Democratic Equality' theory seeks to eradicate oppression and establish a society in which individuals stand in relations of equality to each other.

Accordingly, this dissertation contends that the current regime of cannabis prohibition does not justify its disproportionate and negative effects, which are contrary to democratic equality. Furthermore, in consideration of shifting attitudes in global cannabis policy and emerging trends towards legalisation and commercialisation, an equality-based approach must be adopted to remedy injustices directly inflicted by prohibition and restore full enjoyment of democratic equality.

Chapter I will examine the intrinsic connection between prohibition and racial discrimination by outlining how the historical justifications for cannabis laws in the US and subsequently, the UK and various jurisdictions were not founded on any scientific analyses; rather, where discussions on cannabis did occur, they centred mostly around racial prejudice and rhetoric linking cannabis to criminality. It will also bring to light how these flawed justifications spread globally through international drug policy, culminating in disproportionate enforcement, criminalisation and incarceration, social exclusion and inequality on a widespread scale.

Having provided the underlying background on the history and impact of prohibition, Chapter II will undergo an analysis of Anderson's democratic equality theory to demonstrate its incompatibility with the status quo, thereby demanding a new equality-based approach that will not only prevent further criminalisation, but mitigate the inequality imposed by the principally flawed, outdated and unjust present regime.

Finally, Chapter III will examine potential approaches to cannabis reform using democratic equality theory. Consideration

---

[1] This is discussed extensively in Chapter I.
[2] John Rawls, *A Theory of Justice* (Harvard University Press 1971) 3.

will be given to decriminalisation, legalisation, and affirmative action, including the expungement of criminal records, to demonstrate the applicability of democratic equality as an appropriate response to the devastating impacts of cannabis prohibition.

2. The Tainted Origins of Cannabis Laws and Subsequent Impacts on Racial Minorities

2.1. Chapter Summary

The purpose of this chapter is to descriptively outline the history of cannabis laws and their racial origins, before tracing the disproportionate impact of prohibition on black and ethnic minorities. Analysis will centre on a wide range of jurisdictions which have taken conventional approaches to cannabis prohibition, largely based on homogenised global drug policy—such as the US, UK, Australia and the majority black British Overseas Territory of Bermuda—to underscore the widespread impact of cannabis enforcement against minorities. Finally, brief consideration shall be given to how the institution of law intersects with race to legitimise oppression—relevant for discussion of social exclusion and democratic equality in Chapter II.

2.2. The History of the War on Drugs and Global Cannabis Prohibition

The present system of worldwide drug control is based upon three UN conventions,[3] largely traceable to the historical influence of the US on drug policy which made cannabis a prohibited drug internationally. In the US, there was no national public policy on cannabis prior to its inclusion in the *Uniform Narcotic Drug Act* in 1932 and the passage of the *Marihuana Tax Act 1937*.[4] Hitherto, cannabis was largely unheard of, its use considered a new phenomenon. Whilst it might be assumed that justifications for the first cannabis laws would have centred on health-based concerns, or perhaps a proven link to crime, an analysis of media and legislative accounts has led legal historians to conclude that 'the most prominent [justification] was racial prejudice'.[5] States in the southwest generated support for prohibition by associating cannabis with 'the growing Mexican-American minority', whilst others blamed criminal activity on 'its use among blacks in urban ghettoes'.[6]

It was the Mexicans who first introduced recreational cannabis to the US, immigrating to the southwest after the Mexican Revolution of 1910; it is for this reason that cannabis came to be commonly known as '*marijuana*'—US lawmakers propagated this Mexican terminology to associate cannabis with the unwanted surge of immigrants.[7] A New York Times editorial asserted, 'Marijuana...is a direct by-product of unrestricted Mexican immigration'.[8] Although by the 1920s most Americans still had not heard of cannabis, its popularity began to grow among blacks in urban communities, media reports denouncing

the phenomenon as 'reflective of Negro, jazz and degenerate bohemian subcultures'.

> Thus, not only did few middle-class Americans know about marijuana and its use, but also what little "information" was available provoked an automatic adverse association of the drug with Mexican immigration, crime and the deviant life style in the Black ghettos.[9]

Overall, 'no state undertook any empirical or scientific study of the effects of the drug. Instead they relied on lurid and often unfounded accounts of marijuana's dangers as presented in what little newspaper coverage the drug received'.[10] These prejudices were exploited nationally over the following decade to build a campaign to criminalise cannabis federally, then internationally.

> Apparently, legislators in these states found it easy and uncontroversial to prohibit use of a drug they had never seen or used, and which was associated with ethnic minorities and the lower class.[11]

One prominent figure, Harry Anslinger, first commissioner of the Federal Narcotics Bureau, was responsible for the formation of drug policy from 1930-1962 and spearheaded the *Marihuana Tax Act*,[12] before serving another two years as Representative to the UN Narcotics Commission where he further championed the prohibitionist agenda.[13] In a 1937 'calculated effort to silence opposition' to the bill, Congress held only one hearing at which Anslinger testified that the majority of cannabis users comprised of 'Negroes, Hispanics, Filipinos and entertainers' whose 'Satanic music, jazz and swing, result from marijuana use' and 'cause white women to seek sexual relations with Negroes'.[14] His justifications could not be more explicit:

> The primary reason to outlaw marijuana is its effect on the degenerate races. Marijuana is an addictive drug which produces in its users insanity, criminality, and death.[15]

Despite the objections of the American Medical Association, which declared cannabis a relatively innocuous drug, the Act passed without a recorded vote. As 'the public perception of marijuana's ethnic origins and crime-producing tendencies often went hand in hand', naturally, cannabis, as had the earlier state legislation, 'became entangled with society's views of these minority groups'.[16]

Congress stiffened drug penalties via the *Boggs Act*[17] in 1951; introducing mandatory imprisonment ranging from two to twenty years and placing cannabis in the same category as heroin regarding the seriousness of its effects and penalties for possession and sale. To date, cannabis is designated a Schedule I controlled substance—a category graver than cocaine, OxyContin, and reserved for drugs with the highest potential for abuse, no medicinal or safe method of use—despite not satisfying the three

---

[3] These include the 1961 Single Convention on Narcotic Drugs as amended by the 1972 Protocol, the 1971 Convention on Psychotropic Substances, and the 1988 Convention against Illicit Traffic in Narcotic Drugs and Psychotropic Substances.
[4] 9A ULA 182 (1932) (US); IRC § 39(a) (1937).
[5] Richard Bonnie, Charles Whitebread, 'Forbidden Fruit and the Tree of Knowledge - An Inquiry Into the Legal History of American Marijuana Prohibition' (1970) 56 VaLRev 971, 1016.
[6] ibid.
[7] Prior to this, the plant, 'hemp' had been widely used for industrial and medicinal purposes. Carole Shapiro, 'Law v. Laughter: The War against the Evil Weed and Big Screen Reefer Sanity' (2004) 29 Oklahoma City ULRev 795, 808.
[8] Bonnie and Whitebread (n5) 1036.

[9] ibid 1035.
[10] ibid 1016.
[11] ibid 1021.
[12] n4.
[13] Arthur J. Lurigio, Mikaela Rabinowitz, and Justyna Lenik, 'A Century of Losing Battles: The Costly and Ill-Advised War on Drugs in the United States' (2009) 6 Justice Policy Journal 1, 8.
[14] ibid.
[15] ibid.
[16] ibid.
[17] 21 USC §§ 174 (1951).

tenets for such a categorisation.[18] The only drug not referenced by its scientific name in the schedule, cannabis remains listed under its pejorative, *marihuana*—an ironic allusion to its racist origins.[19]

It is evident that the extent of attention given to cannabis in the US primarily centred around racial prejudices and misinformation, sanctioned by federal policymakers and institutions. But the significance of such is underscored by the fact that the US was instrumental not only in prevailing social attitudes toward cannabis, but legally—heavily influencing international drug policies which remain the standard today.[20]

For example, the UK's first cannabis laws were passed in response to international pressures rather than domestic concerns.[21] When it was first proposed that cannabis be included under the *International Opium Convention*,[22] there was no notable consumption by the general public; however, the little public attention that was received locally also made use of racial prejudices. Notably, the passage of the *British Nationality Act 1948*, which resulted in the migration of half-a-million colonial residents to the UK, was alleged to be the cause of the growth of the domestic cannabis market. This culminated in a 1951 declaration by the head of the Drugs Branch, FW Thornton, that 'the invasion of unemployed coloured men threatened Britain with a serious hashish smoking problem'—further promoting the perception of cannabis as a threat arising from ethnic minorities.[23]

Nevertheless, the influence of the US regarding international drug policy resulted in nearly all countries agreeing to a unified control regime, outlawing cannabis under the *Single Convention on Narcotic Drugs* in 1961 to create a universal system of drug control.[24] The Transnational Institute noted, 'the idea of having a Single Convention was an initiative of the United States, determined to impose a hard line on drugs on the rest of the world'.[25] As the Convention expanded existing control measures to cover the cultivation of plants from which narcotics are derived, in many cases cannabis was lumped into national legislation without discussion—and by 1970 was criminalised in all member countries under the goal of creating 'a drug-free world'.[26] Stressing the main legal innovations from 1925-1964 were in response to 'international pressures, not domestic problems', the UK Drug Policy Commission acknowledged these conventions continue to impose limits on national drug legislations, 'still most powerfully upheld by the USA'.[27]

Thus, to date, the global approach to cannabis remains largely homogenous as a result of these Conventions. Specifically, under the 1961 Convention, cannabis is subject to the general control measures including article 4(1)(c) limiting the production, use and possession of drugs, article 33 outlining that parties must not permit the possession of cannabis 'except under legal authority'.[28] This was introduced into UK domestic law via the *Dangerous Drugs Act 1964*, which criminalised the cultivation of cannabis. In 1971, US President Nixon declared the "War on Drugs", aligning with the 1971 Convention containing a general obligation under article 7 to prohibit THC—the main psychoactive chemical in cannabis; In the same year, the UK passed the *Misuse of Drugs Act*, providing the domestic basis for controlling illicit drugs. This act criminalised, *inter alia*, the possession, usage and cultivation of cannabis, and remains the primary UK legislation concerning cannabis to date.

2.3. The Disproportionate Impact of Cannabis Prohibition

With such evidence demonstrating the racially prejudiced origins of cannabis laws, perhaps it comes as no surprise that enforcement has since been proven to disproportionately affect racial minorities. However, as it is arguable that the mere presence of prejudice as a basis for the formation of a law or policy is not sufficient in and of itself to necessitate its removal, ramifications of the law in practice must be considered to determine whether reform is truly needed.

2.3.1. Drugs, Race, and Crime

In 1971, critics of the Misuse of Drugs Bill predicted it would be impossible to enforce, leading to selective enforcement.[29] Similarly, planners of the "War on Drugs" in the US anticipated their proposed policies would disproportionately target African-Americans and scapegoat minorities.[30] Indeed, such has been the case, with 'longstanding concerns over the disproportionate impact of prohibition enforcement on certain ethnic minority groups'.[31]

It is therefore notable that since enactment, cannabis laws have resulted in the disproportionate investigation, searching, detention, arrest, charging, conviction, incarceration, disenfranchisement and social exclusion of blacks and minorities.[32] Statistics throughout all convention signatories which have collected data in these areas—including the UK, Canada, Australia, New Zealand, and Bermuda—indicate these racially disproportionate impacts have remained consistent worldwide.

In the US, blacks are arrested nearly three times as frequently as whites for cannabis offences, despite user rates standing close to equal.[33] A study of over two million cases identified 'significant racial disparities in the implementation of marijuana enforcement',[34] whilst nearly one in three African-American men in their twenties are under criminal justice supervision due to the "War on Drugs".[35] The issue is not confined to men, as the largest increase in women's incarceration rates consists of black women who, often being 'low level players' in drug offences, are unable to

---

[18] Cannabis (1) has a low potential for harm and abuse; (2) appears to have therapeutic benefits, and (3) the American College of Physicians suggests it may be used safely under appropriate conditions. Eric Blumenson, 'Liberty Lost: The Moral Case for Marijuana Law Reform' (2010) 85 Indiana Law Journal 279, 299.

[19] 21 USC § 812(b)(1) (2006).

[20] Thomas Moran, 'Just a Little Bit of History Repeating: The California Model of Marijuana Legalization and how it Might Affect Racial and Ethnic Minorities' (2011) 17 Wash&Lee J Civil Rts & Soc Just 557, 569.

[21] UK Drug Policy Commission, *An Analysis of UK Drug Policy* (London, UKDPC April 2007) 14.

[22] International Opium Convention (adopted 19 February 1925, entered into force 25 September 1928) 81 LoNTS 319.

[23] James Mills, *Cannabis Nation: Control and Consumption in Britain, 1928-2008* (OUP 2012).

[24] Single Convention on Narcotic Drugs (adopted 30 March 1961, entered into force 13 December 1964) 520 UNTS 151.

[25] Martin Jelsma and Amira Armenta, *The UN Drug Control Conventions* (Transnational Institute, October 2015).

[26] David Bewley-Taylor, 'The Contemporary International Drug Control System: A History of the UNGASS Decade,' in John Collins (ed), Governing the Global Drug Wars (London: LSE IDEAS Special Report, 2012), 49.

[27] UK Drug Policy Commission (n21) 14.

[28] n24 art 33.

[29] J Young, *The Drugtakers: The Social Meaning of Drug Use* (London, Paladin 1971).

[30] Michael Tonry, *Malign Neglect—Race, Crime, and Punishment in America* (OUP 1995) 83-104.

[31] UK Drug Policy Commission (n21) 51.

[32] Kenneth Nunn, 'Race, Crime and the Pool of Surplus Criminality: Or Why the "War on Drugs" Was a "War on Blacks"' (2001) 6 Journal of Gender Race & Justice 381.

[33] R Ramchand, R Pacula and M Iguchi, 'Racial Differences in Marijuana-Users' Risk of Arrest in The United States' (2006) 84 Drug and Alcohol Dependence 264.

[34] Amanda Geller and Jeffrey Fagan, 'Pot as Pretext: Marijuana, Race, and the New Disorder in New York City Street Policing' (2010) 7 Journal of Empirical Legal Studies 591.

[35] D Coker, 'Addressing the Real World of Racial Injustice in the Criminal Justice System' (2003) 93 The Journal of Criminal Law and Criminology 827, 832.

provide useful information to authorities—thus receiving higher sentences.[36]

These figures are even more disproportionate in the UK, where blacks are seven times more likely to be arrested and fourteen times more likely to be imprisoned for drug offences than whites, despite their rate of drug use being no higher.[37] New Zealand's Human Rights Commission acknowledged 'evidence of bias at different points throughout the system from apprehensions to sentencing,[38] revealing Maori/Pacific persons accounted for 42% of possession convictions and were arrested at three times the rate of white/non-Maori users.[39] This disparity remains in jurisdictions where non-whites comprise a numerical majority: In Bermuda, blacks comprise 54% of the population but account for 90% of all possession charges; a government drug reform committee determined 'Blacks are more likely targeted through profiling and structural racism'; 'while racial profiling is evident in charges, the demand for cannabis is non-discriminatory'.[40] This indicates the disproportionate impacts of prohibition are structural—resulting not simply from individual prejudices; but laws and policies themselves, institutions and methods used to enforce them, and categorisations perpetuated throughout society at large.

The use of stop-and-search has been a major contributor to this disproportionality. This is largely because, as 'drug use is a consensual activity, it rarely comes to light through victim or witness reports; police must rely on more proactive methods for finding cannabis offences than other forms of street crime'.[41] A study on 'Race, drugs and law enforcement in England and Wales' determined all 43 police forces target blacks at a higher rate than whites, with drugs searches accounting for 60 per cent of all stop-and-searches.[42] Blacks in the UK are stopped and searched eight times more than whites, although it is less likely for drugs to be found.[43] Additionally, police were found to target suspected cannabis users—with nearly 600 recorded incidences of officers conducting stop-and-searches on the grounds of 'being able to smell cannabis' alone.[44]

This disproportionality is mirrored worldwide. Bermuda's Drug Reform Committee reported 'evidence of racial profiling', citing as many as 17,000 stop and searches in one year—a significant figure, considering the island's mere population of 65,000:

> Of those stopped and searched, 90% were male and 85% were Black. Nearly two thirds were males between the ages of 18 and 36. When considering the black male population in this age group, the figures were such that it was possible to search every one of them four times in the year.[45]

In consideration of these statistics, it could be alleged the true cause of disproportionality lies solely in enforcement practices,

such as racial profiling in stop-and-search. However, as analyses of revised enforcement methods have revealed, efforts to alter policing to comply with the *Race Relations Act 1965* have been counterproductive:[46] 'disproportionality has increased as the use of stop and search has declined, indicating that the remaining use of the powers is more heavily concentrated on black and minority ethnic groups'.[47]

> A comprehensive new analysis of the criminal justice system reveals such reform has failed to address racial bias. Rather, it documents a disquieting increase in ethnic disparity in the policing and prosecution of drug offences.[48]

Overall, arrests for drugs as a result of stop-and-search fell by 52% for whites between 2010-11 and 2016-17, but did not fall at all for blacks.[49] Additionally, the analysis demonstrates that the over-policing of black and minority communities cannot be fully explained by the concentration of law enforcement resources in deprived areas. It also shows how ethnic disparities introduced earlier in the process are perpetuated by sentencing decisions in court.[50]

Beyond enforcement, studies have further revealed that systemically, 'black people are treated more harshly when they are found in possession of drugs': blacks are more likely to be formally dealt with for cannabis possession than whites, and more likely to be charged rather than cautioned.[51] In 2017, 'more black people were prosecuted and convicted of cannabis possession than for the supply of Class A and B drugs combined. For white people the balance was reversed.'[52]

Accordingly, despite making up less than 4% of the UK population, blacks comprise a quarter of those convicted of possession, also sentenced to immediate custody for drug offences at 9.1 times the rate of whites.[53] This has been termed 'an affront to justice' and 'a conveyer belt for the criminalisation of young black people for low-level offending, while treating white people more leniently for the same offences'.[54] Other convention signatories such as the US, Canada, and Australia have reported similarly, with Bermuda acknowledging the issue extends beyond disparities in police interactions as 'Blacks bear the brunt of most possession charges while whites are far less likely to be charged'.[55]

Overall, evidence supports the argument that globally, states' approaches to drug enforcement in general stem partly from wider perceptions 'weighted towards a fictitious narrative that drug use is especially prevalent among black and minority ethnic groups'.[56] The implications of these disproportionate effects of cannabis laws extend beyond the criminal justice system—such as social exclusion resulting from stigma over drugs and crime.[57]

[36] Phyllis Goldfarb, 'Counting the Drug War's Female Casualties', (2002) 6 Journal of Gender, Race & Justice 277.

[37] UK Drug Policy Commission (n21) 51.

[38] Human Rights Commission, *A fair go for all? Addressing Structural Discrimination in Public Services* (HRC July 2012) 34 (NZ).

[39] Sally Abel, 'Cannabis Policy in Australia and New Zealand' (2009) 16 Drug and Alcohol Review 421, 425.

[40] Cannabis Reform Collaborative, *An Analysis of Cannabis Reform in Bermuda* (CRC April 2014).

[41] M Shiner et al, *The Colour of Injustice: 'Race', drugs and law enforcement in England and Wales* (LSE, Stopwatch & Release 2018).

[42] ibid.

[43] ibid.

[44] Tom Embury-Davis, 'Police shouldn't stop and search people just because they can smell cannabis, says report' (Independent, 12 December 2017).

[45] CRC (n44) 33.

[46] UK Drug Policy Commission (n21).

[47] Shiner (n45).

[48] B Bowling and C Phillips, 'Disproportionate and Discriminatory: Reviewing the Evidence on Police Stop and Search' (2007) 70 ModLRev 936.

[49] Catherine Wylie, 'Black people "nine times as likely" as whites to be stopped and searched by police in England and Wales' (Independent, 14 October 2018).

[50] Shiner (n45).

[51] UK Drug Policy Commission (n21) 51.

[52] Shiner (n45).

[53] N Eastwood, M Shiner, and D Bear, *The Numbers in Black and White: Ethnic Disparities In The Policing And Prosecution Of Drug Offences In England And Wales* (Release & Manheim Centre for Criminology at LSE 2013).

[54] Shiner (n45).

[55] CRC (n44) 38.

[56] M Townsend, 'May vowed to reform stop and search – but it's getting worse for black Britons' (Guardian, 13 October 2018).

[57] See discussion in Chapter II.

2.3.2. Criminality as Racial Subordination

But what is the importance of the connection between these tainted origins of cannabis prohibition and their devastating institutionally racist effects? Nunn argues that the impact of the drug war on black communities is simply 'a prominent example of the central role both race and the definition of crime play in the maintenance and legitimization of white supremacy'.[58] Race and crime, as two significant social phenomena, are linked in an endless cycle of oppression: What is defined as crime determines who is oppressed in society and simultaneously legitimates that oppression; Race provides the contours of a discourse of threat that supplies the social phenomenon of crime with power and political significance.[59] In this way, law is used as a mechanism for legitimising inequality and racial subordination[60], whilst crime can mask oppression by allowing it to be represented as a legitimate response to wrongdoing. At the same time, 'pool of surplus criminality' theory holds that when faced with a perceived drug problem, minorities are considered scapegoats for threats perceived by the dominant community; consequently, the historic link between drugs, racial prejudice, and the predictable consequences of prohibition are all evidence of the use of power and violence against marginalised groups, resulting in social exclusion and 'outcast' status in the construction of social order.[61] Such ideas explain longstanding biases in perceptions of criminality—such as white survey respondents overestimating the proportion of crime committed by blacks by up to 30 percent, suggesting harsher punishments when a crime is perceived to be a "black crime", and consistently perceiving white criminality in terms of individual failing whilst categorising black criminality as a result of failings of the entire group.[62]

2.4. Conclusion

Overall, the legal system is breeding criminals with inappropriate cannabis legislation that entangles offenders in court rooms, excluding and marginalising them from the mainstream, taints individuals with a moral stigma, takes away their hope of future success in the community, and bars effective societal engagement.[63] As Chapter II will demonstrate, this situation is contrary to equality principles. The wider societal impact reveals the issue is too far gone for selective concessions and band-aid alterations in enforcement; the prohibition regime must be overhauled—with comprehensive reform incorporating measures which eradicate the outstanding ramifications of the law urgently needed.

3. Democratic Equality and the Case for Cannabis Reform

3.1. Chapter Summary

Chapter I reveals the current regime of cannabis prohibition is, foundationally, the result of sustained efforts associating the drug with criminality and minority groups, with no meaningful dialogue or scientific analysis and a disregard for the disproportionate outcomes of prohibition and enforcement. Having placed the origins of cannabis laws and their lasting impacts in their rightful context, Chapter II goes on to critique the status quo based on the theory of democratic equality as espoused by Elizabeth Anderson—which demands a society in which all persons are free from oppression, stand in positions of equality to each other and share equal access to effective participation at all levels. The chapter will further discuss potential conflicts which may arise considering the illegality of cannabis, before concluding that democratic equality is an appropriate framework by which to consider cannabis reform; the ultimate objectives of which are culminated in Chapter III.

3.2. Democratic Equality: A Society of Equals

Democratic equality as espoused by Elizabeth Anderson is an extension of egalitarianism, concerned with principles of equal rights and opportunities for all individuals in society. However, Anderson criticises the works of mainstream egalitarianists, including Dworkin,[64] Arneson,[65] and Cohen,[66] whom she claims have all lost sight of the meaning of equality by placing too much emphasis on theorising how to compensate people for 'undeserved bad luck' or 'cosmic injustice' rather than fighting inequality.[67] According to Anderson, this doctrine of 'luck egalitarianism' has too narrowly targeted 'victims of brute luck from human affairs' as exemplary beneficiaries of egalitarian concern, to the detriment of the true victims of oppression—arising from hierarchies of race, gender, class and caste throughout society.[68] In this regard, Anderson befittingly reasserts that 'the true meaning of equality' must be to eradicate oppression, creating a society of equals.

> The proper negative aim of egalitarian justice is not to eliminate the impact of brute luck from human affairs, but to end oppression, which by definition is socially imposed. Its proper positive aim is not to ensure that everyone gets what they morally deserve, but to create a community in which people stand in relations of equality to others.[69]

Such a view recognises the role of egalitarianism in challenging oppression—encompassing a social order involving systematic exploitation of power by dominant groups, including through marginalisation, status hierarchy, exploitation and cultural imperialism—and occurring where one holds power over another through social institutions such as laws, customs and norms.[70] These unequal relations generate and justify inequalities in the distribution of freedoms, resources, and welfare.[71]

Democratic equality is therefore the application of egalitarian objectives which work negatively towards the eradication of 'forms of social relationship which dominate, exploit, marginalize, demean, and inflict violence upon others', and positively, recognising that the democratic state is nothing more than citizens acting collectively—to secure a social order in which persons

---

[58] Nunn (n36) 385.

[59] R Delgado, 'Rodrigo's Eighth Chronicle: Black Crime, White Fears on the Social Construction of Threat' (1994) 80 VaLRev 503.

[60] Cheryl Chambers, 'Institutional Racism: Is Law Used as a Tool to Perpetuate Racial Inequality?' (PhD Sociology, North Carolina State University 2008) 1.

[61] Nunn (n36) 385.

[62] National Academies of Sciences, Engineering, and Medicine, *The Criminal Justice System and Social Exclusion: Race, Ethnicity, and Gender: Proceedings of a Workshop—in Brief* (National Academies Press 2018).

[63] Glenn Williams, 'The Criminalization of Recreational Marijuana Use in Canada: A Scientific, Social, Legal and Philosophical Analysis Based on the Work of Douglas Husak' (MA Public Ethics, University of Ottawa 2010) 43.

[64] Ronald Dworkin, 'What Is Equality? II. Equality of Resources' (1981) 10 Philosophy and Public Affairs 283, 285.

[65] Richard Arneson, 'Equality and Equality of Opportunity for Welfare' in Louis Pojman and Robert Westmoreland (eds), *Equality: Selected Readings* (OUP 1997) 231.

[66] G Cohen, 'On the Currency of Egalitarian Justice' (1989) 99 Ethics 906, 922–23, 930–31.

[67] Elizabeth Anderson, 'What Is the Point of Equality?' (1999) 109 Ethics 287, 288.

[68] ibid.

[69] ibid.

[70] Kimberle' Williams Crenshaw, 'Race, Reform and Retrenchment: Transformation and Legitimation in Anti-Discrimination Law' (1988) 101 HarvLRev 1331.

[71] Anderson (n71) 312.

stand on equal terms in a democratic community and citizens are obliged to secure for each other 'effective access to the social conditions of their freedom at all times'.[72]

These social conditions comprise 'a package of capabilities sufficient for standing as an equal over the course of an entire life'; guaranteed to individuals in a democratic society without restriction based on social identity or membership. At the primary level, it includes access to basic conditions of human agency, such as autonomy and freedom; More comprehensively, equality within society necessitates the exercise of citizenship, involving functioning not only as political agents, but participating equally within wider society—'the sphere of social life that is open to the general public and is not part of the state bureaucracy and the administration of laws'.[73] This extends to public institutions, services, private enterprise, and the right to receive education and fair value for one's contributions; Any situation whereby a group is subjected to discrimination on the basis of social identity and restricted from access to equal participation in institutions and areas of civil society would violate democratic equality, as they have effectively been relegated to 'second-class citizenship'—even if they maintain political rights in law.[74]

Finally, capable functioning requires not just the ability to exercise political rights such as freedom of speech and voting, but also 'to participate in the various activities of civil society more broadly', such as in economics and other interactions.[75] Crucially, this entails the social conditions of being accepted by others, including 'the ability to appear in public without shame' and not being ascribed 'outcast status'. These capabilities are unalienable, irrevocable and untradeable.[76]

3.3. Why Cannabis Laws are 'an Equality Issue'

It thus follows that democratic equality is an appropriate theory by which to examine cannabis prohibition and its disproportionate results. As evidenced in Chapter I, from inception, global drug policy and cannabis laws were largely implemented based on ascribing outcast status based on social identities such as race and ethnicity. Further, when placed into its correct legal historical context, it is evident that the laws prohibiting cannabis do not simply yield disproportionate racial results by some inadvertent mishap; rather, these laws, from their origin to the present day are built upon an intention to disproportionately target and marginalise ethnic minorities. The tangible, immediate effects of these targeted laws in relation to enforcement and criminalisation have already been demonstrated in Chapter I. But further analysis reveals the wider implications necessary for consideration regarding the various capabilities needed to participate effectively as an equal member of society, which have been violated by the current regime.

3.3.1. Beyond Criminalisation: Social Exclusion and the Capabilities Approach

Beyond the immediate legal ramifications of cannabis offences such as criminalisation and incarceration, prohibition has severely impacted individuals' ability to engage equally and effectively throughout society. Simply put, criminal laws stigmatise and punish people. The resulting stigma works by labelling people as deviant, wicked, or a threat to wider society. Many cannabis users are looked down upon in society; The criminal label placed on otherwise law-abiding citizens can create 'long-lasting effects

on the individual that can severely disrupt their way of life'[77] by limiting individuals from opportunities that societal members otherwise enjoy:

> There can be an instant impact on the individual's relationship with their family, friends and co-workers. There can be a significant financial burden on the individual and relatives if forced to go through the court system with a respectable lawyer. Opportunities and freedoms must often be given up in light of the legal consequences of the offense, such as employment opportunities or international travel.[78]

This not only negatively affects an individual's social and personal life; it can actually lead one to commit further crimes, as 'the "illegalness" of cannabis may increase criminal behaviour in society by virtue of labelling': the stigma associated with criminal sanctioning 'alienates the individual from conventional society, promoting contact with deviant referent groups and enhancing the likelihood of future deviance befitting the label—a self-fulfilling prophecy'.[79]

Crucially, the consequences of what is largely a victimless crime extend to restrictions upon many of the capabilities deemed necessary under democratic equality. Particularly, access to travel, employment, participation in the political process and other areas of engagement have been severely hampered due to cannabis prohibition and social exclusion. Those convicted of drug offenses may be prevented from obtaining financial aid, welfare, or public housing;[80] the high rate of incarceration has contributed to higher poverty rates and lower labour force participation among blacks specifically.[81] An Australian study on individuals who received their first criminal conviction for a minor cannabis offense found that whilst the conviction had little impact on subsequent cannabis use, a significant number reported further complications with the law, and employment, accommodation, relationships and travel opportunities.[82] Bermuda, which relies on travel connections through larger hubs like Canada and the US, notes that convictions for small quantities of drugs comprise the majority of individuals on the 'stop list'—whereby persons deemed 'illegitimate' are given a lifetime ban on international travel—resulting in denied access to airports, ports of entry, US residency, citizenship, and overseas hospital treatment.[83] In some instances, individuals have reported being denied travel based solely on mere suspicion of engaging in drug activities or prior arrest, despite no convictions.[84] A legislative report declared, 'upon being released from prison, several black males have found it extremely difficult to find employment and to earn enough money to support their families, which creates a cycle of poverty, anger and frustration'.[85]

---

[72] ibid 289.
[73] ibid 317.
[74] ibid.
[75] ibid.
[76] ibid 318.

[77] Williams (n67).
[78] ibid.
[79] Robert J MacCoun, Peter Reuter, *Drug war heresies: learning from other vices, times, and places.* (CUP 2001) 91.
[80] Coker (n39) 841.
[81] Jeff Grogger, 'Arrests, Persistent Youth Joblessness, and Black/White Employment Differentials' (1992) 74 Rev Econ & Stat 100, 105.
[82] Simon Lenton, Penny Heale, 'Arrest, Court and Social Impacts of Conviction for a Minor Cannabis Offense under Strict Prohibition' (2000) 27 Contemp. Drug Probs. 805.
[83] Jonathan Bell, 'No stop list, but crimes will prevent entry to US' (Royal Gazette, 19 April 2012).
[84] Raymond Hainey, 'Hope for Bermudians on US stop list' (Royal Gazette, 29 July 2017).
[85] Horton et al, *Executive Summary Joint Select Committee on the causes of violent crime and gun violence in Bermuda* (House of Assembly 2011).

Cannabis prohibition and its unequal enforcement has also had devastating results on the ability of large segments of populations to exercise citizenship and effectively engage in the political process. In the US, many states impose a lifetime ban on voting for persons with felony convictions. Despite comprising only 13% percent of the population, blacks comprise 38% of all Americans who have been stripped of their voting rights, with more than 2.2 million of the 6.1 million ineligible to vote in the 2016 election being African-American; Presently, one in 13 African-Americans cannot vote due to prior convictions—a rate four times higher than non-black Americans.[86] In Kentucky, this rate is so high that 1 in 3 black males are banned from voting for life.[87] The majority of these felony records stem from minor drug offences, many of which were committed prior to adulthood; this means that many individuals barred from voting have never in life been eligible to vote at all.[88] Considering the racially targeted enforcement practices cited in Chapter I, felony disenfranchisement has thus been recognised as 'a potentially effective means to neutralize threats from African-American voters'.[89] Although there is no lifetime franchise ban in the UK, the *Representation of the People Act 1983* also means that imprisoned convicts cannot vote in elections; this disproportionately affects blacks and ethnic minorities due to their higher rates of criminalisation for the same offenses. As if the stakes could not be any higher, 'drug enforcement policies are also the justification given by the US Attorney General's Justice Department for the unequal number of African-Americans charged with federal death penalty eligible crimes'—imposing a fundamental threat to life and liberty itself.[90]

Overall, the effects of continued prohibition extend far beyond mere acknowledgement that it 'has severely inhibited the growth of racial and ethnic minority communities by disproportionately arresting and imprisoning their youth'.[91] Mere identification as a cannabis user or offender can produce conceptions of self consistent with the stigmatising label, have a cumulative effect on criminal identity, and exclude large numbers of individuals from equal access to opportunities in society, including employment, the economy, travel, voting, social benefits, and overall effective engagement.[92] Most notably, these consequences are all exacerbated where the individual falls within the black or ethnic minority identity.

It is therefore evident why drug laws are a matter of egalitarian concern: Cannabis laws have had the effect of disproportionately 'othering' black people especially—casting them as deviant; disproportionately restricting a significant membership of an entire group from access to several opportunities, and outcasting them beyond the bounds of society through consequences ranging from enforcement and incarceration to social exclusion.

3.3.2. An Exception to Democratic Equality?

However, in full examination of democratic equality, it is necessary to consider the extent to which the denial of the outlined 'capabilities' for equal functioning might be justified under the label of criminality. Whilst a crucial tenet of democratic

equality holds that these rights are 'unalienable, irrevocable and untradeable', Anderson cites a sole exception to the guarantees of democratic equality whereby 'the commission of a crime can justify taking away a person's basic liberties and status as an equal in civil society'.[93]

It is thus necessary to discuss the potential claim that democratic equality is not applicable to individuals who commit drug offences, considering its illegality: Does a conviction for a cannabis offence therefore justify ramifications which would otherwise amount to a denial of equality?

Indeed, it could be proposed that individuals who have chosen to commit drug offences should not be romanticised as victims of injustice; that stigma, social exclusion, and denied political participation are simply deserved consequences of a decision to violate the law.

However, this writer posits that Anderson's statement is more nuanced than a mere declaration that democratic equality is contingent upon legal innocence. Surely, no egalitarian would argue that a criminal record amounts to a blanket justification for unequal treatment; whilst a literal interpretation of Anderson's statement might call for some form of restriction upon the full guarantees of democratic equality, it would be implausible to suggest this was intended to extend to all illegal actions.

For example, it would appear unreasonable to demand that an individual convicted for a mere traffic violation be denied the right to vote or function in society free of 'outcast status' as a consequence for committing such a common, largely innocuous strict liability offence. Further, it would not be justified for democratic equality to fall away if the state were to criminalise, eg, criticising the government and individuals were arrested for this offence. Surely, these guarantees of democratic equality are sufficiently important such that they take precedence over many trivial offences not considered threatening to the freedoms of others: Anderson would not have placed such emphasis on the importance of equal society if its guarantees were so fragile that they could be so easily stripped away for the many individuals who dare to indulge in *the killer marijuana*. To suggest otherwise would not only be to justify the subjection of a large section of the population to wide-reaching restrictions on various social, economic and political capabilities considered fundamental in a democratically equal society—it would also signify a willingness to ignore the fact that these restrictions have never applied equally to all individuals engaging in such acts; as these consequences have mostly affected individuals from black and minority ethnic groups. It is therefore evident that this exception would not be justifiable by the offence: the detrimental impact it would have would result in oppression, leading to further social inequality—contradicting the objectives of democratic equality outlined by Anderson in the first place.

Instead, it is likely that the justifications for such conditions are based on the idea that certain behaviours by individuals may pose a threat to the enjoyment of democratic equality for others. Perhaps a more accurate interpretation of Anderson's exception would hold that the instances in which an exclusion from equal status in society may be warranted, are those whereby an individual has committed an offense deemed serious enough to present a threat to the enjoyment of equality by others. In this sense, Anderson's exception for criminal conduct is qualified by the extent to which said actions constitute an infringement upon the rights of other citizens in a society of equals, such that certain exclusions are justified to safeguard the guarantees of equality for

---

[86] C Uggen, L Larson & S Shannon, *6 million lost voters: State-level estimates of felony disenfranchisement* (Washington, DC: The Sentencing Project 2016).

[87] Michael Wines, 'Why So Many Kentuckians Are Barred From Voting on Tuesday, and for Life' (New York Times, 4 Nov 2018).

[88] ibid.

[89] A Behrens, C Uggen & J Manza, 'Ballot manipulation and the "menace of Negro domination": Racial threat and felon disenfranchisement in the United States, 1850–2002.' (2003) 109 American Journal of Sociology 559-605.

[90] Coker (n39) 829.

[91] Moran (n20) 560.

[92] Wodak A, 'The abject failure of drug prohibition' (2014) 47 Australian & New Zealand Journal of Criminology 190.

[93] Anderson (n71) 327.

all. Such conduct would likely include more serious offences such as murder, rape, burglary, fraud, and acts which infringe upon the personal and proprietary rights of others and their ability to participate in society free from oppression. Offences which arguably remain established in criminal law based on rationales of paternalism,[94] such as cannabis laws, would therefore fall outside of this scope; as the means of democratic equality would otherwise be self-defeating regarding its aims.

Such an interpretation appears to fall more in line with the ethos of democratic equality, with Anderson acknowledging that 'even convicted criminals retain their status as equal human beings'—recognising that 'democratic equality guarantees effective access to the social conditions of freedom to all citizens, regardless of how imprudently they conduct their lives'.[95]

> The counterpart to an individual's inalienable right to the social conditions of her freedom is the unconditional obligation of others to respect her dignity or moral equality. [...] This implies that there are some things one may never do to other people, even if one has their permission or consent.[96]

This view appears to be backed in egalitarianism as well as in law. The UN High Commissioner for Human Rights has stated, 'Too often, drug users suffer discrimination, are forced to accept treatment, marginalized and often harmed by approaches which over-emphasize criminalization and punishment while under-emphasizing harm reduction and respect for human rights'.[97]

Thus, although democratic equality does not indemnify individuals against all losses due to imprudent conduct, it does 'guarantee a set of capabilities necessary to functioning as a free and equal citizen and avoiding oppression'.[98] Whilst individuals still maintain personal responsibility, the consequences of their decisions should not lead to radically different outcomes; doing so would be antithetical to the aim of ensuring equal societal relations. If a law is racially targeted or results in the disproportionate revocation of access to capabilities of equal functioning, this would not be defensible in a society committed to ensuring democratic equality. On this basis, it is reasonable to conclude that the mere illegality of cannabis would not be sufficient to justify the resulting oppression inflicted by the present regime.

3.4. Conclusion

In consideration of the effects of prohibition on blacks and ethnic minorities, it is evident why the status quo is contrary to the guarantees of democratic equality: From formation to implementation, cannabis laws have created a caste of people who are locked out of social, economic and political influence and opportunities in society. The cumulative effects of prohibition mean that a significant proportion of members within black and ethnic minority groups cannot participate effectively or equally in relation to others. Cannabis laws therefore violate equality principles and, given this fact, the purported aims of prohibition do not justify the consequential subordination of black and ethnic minority groups. In seeking the construction of a community of equals, democratic equality integrates principles of distribution

with the expressive demands of equal respect, charting an appropriate framework for eradicating the contributors of socially imposed oppression. Accordingly, it directly follows that democratic equality is an appropriate framework by which cannabis law and policy can be examined, whilst simultaneously serving as the basis of reform—to work towards the guaranteeing of effective participation of all, in a society of democratic equals.

4. Examining Equality-Based Responses to the Impact of Cannabis Laws

4.1. Chapter Summary

Having outlined the context of the present cannabis law and policy framework in Chapter I, and its incompatibility with Anderson's Democratic Equality in Chapter II, this Chapter examines potential socio-legal and policy approaches which can be justified under democratic equality. Consideration will be given to decriminalisation, legalisation, and the application of redistributive justice measures such as affirmative action and the expungement of criminal records, in order to demonstrate the applicability of Anderson's theory of democratic equality as an appropriate response to the disproportionate effects of prohibition on black and minority ethnic groups.

4.2. Visions for Reform

According to the Global Commission on Drug Policy, the global war on drugs has been a massive failure. Citing 'repressive measures' which have failed to reduce demand and devastating consequences for individuals and wider society, the Commission declared that fundamental reforms are urgently needed:

> End the criminalization, marginalization and stigmatization of people who use drugs but who do no harm to others. Challenge rather than reinforce common misconceptions about drug markets, drug use and drug dependence. Encourage experimentation by governments with models of legal regulation of drugs to undermine the power of organized crime and safeguard the health and security of their citizens.[99]

The Commission emphasised that 'this recommendation applies especially to cannabis' but encouraged 'other experiments in decriminalization and legal regulation that can accomplish these objectives and provide models for others'.[100] In a similar thread, a European Commission report declared:

> Enforcement of drug prohibitions has caused substantial unintended harms; many were predictable. The challenge for the next ten years will be to find a constructive way of building on these lessons so that the positive benefits of policy interventions are increased, and the negative ones averted.[101]

Having undergone a thorough analysis of these devastating effects in relation to the demands of democratic equality in previous chapters, it is evident that the aim of democratic equality is to establish a society whereby individuals are guaranteed the capabilities necessary to engage in society as equals. Fundamentally, such theory requires the eradication of

---

[94] Joel Feinberg, 'Legal paternalism.' (1971) 1 Canadian Journal of Philosophy 105-124.

[95] Anderson (n71) 327.

[96] ibid 319.

[97] Global Commission on Drug Policy, *The War on Drugs* (Global Commission on Drug Policy 2011) 2.

[98] Anderson (n71) 287.

[99] Global Commission on Drug Policy, *The War on Drugs* (Global Commission on Drug Policy 2011) 2.

[100] ibid.

[101] P Reuter and F Trautmann (eds), *A Report on Global Illicit Drugs Markets 1998-2007* (Brussels, European Commission 2009) 17.

oppression, inclusive of the systems and institutions underlying societal interactions, and therefore must be free to implement reforms in wider areas such as economics and politics. Accordingly, the desideratum of comprehensive cannabis reform must be to construct a regime which shall: acknowledge the historical context of racial prejudice within cannabis policy; objectively consider potential justifications for prohibition as opposed to legalisation or civil regulation of cannabis; and, in light of resulting inequality, instrument the steps required to ameliorate present-day inequities in criminalisation, enforcement, incarceration and social exclusion. Moreover, considering emerging trends indicating more relaxed attitudes toward cannabis and recreational use generally; to prevent further disenfranchisement of racial minorities mitigate and future harm.

4.3. Jurisdictional Responses and Possible Remedies

4.3.1. Decriminalisation

The use of decriminalisation[102]—typically for minor possession only—has increased in recent years, currently effective in jurisdictions such as Australia, Bermuda and some Caribbean countries, Portugal, and thirteen US states. A variety of explanations have been given for this trend, including the high costs of enforcement, economic recession, and shifts in public attitude.[103] However, Yankah contends that the key drivers of the decriminalisation movement are fundamentally attributable to 'a convergence of principles across philosophical starting points' in which both ends of the academic and legal-political spectrum have begun to recognise the dangers of overcriminalization and a criminal law system growing out of control, which has seen ever-increasing numbers of persons imprisoned for ever-increasing lengths of time, and 'threatens to criminalize so much behaviour as to leave the actual exertion of power over the citizen all too often at the whim of the state':

> This power has all too often been focused on racial minorities and the poorest among us. With so much behaviour increasingly falling under the government's criminal law powers, unrestrained and discriminatory enforcement discretion threatens to undermine the rule of law itself.[104]

It is subsequently argued that the criminal law is only justified if it 'prevents a nontrivial harm, suppresses wrongful behaviour, and punishes those who deserve to be punished.'[105] In the case of cannabis, such punishments can be considered unjust 'not only because they exceed any justifiable response to the criminal conduct, but often simply because they are levelled at conduct that ought not be criminalized at all'.[106] Accordingly, both liberals and non-liberals alike have come to an agreement regarding the decriminalisation of cannabis, with the understanding that 'whatever the right answer to the difficult question of addictive drugs, there is no question that the current war on drugs, besides leading to the arrest and conviction of millions, has visited pain in deeply racialized lashes, yoking countless young people with a felony record that often prevents them from seeking an education,

holding a job, or in some cases, obtaining food in times of need'.[107] In this sense, the criminal sanctioning of cannabis must be put to a halt if societies are to even begin to address the many violations of democratic equality inflicted upon minorities.

Such a position seems to have been supported, even if not vigorously, by international bodies. Most notably, a UN briefing document leaked in 2015 affirmed that criminal sanctions for minor drug possession have 'induced negative consequences for safety, security, and human rights', whilst 'the heavy emphasis on criminalization has fuelled high levels of discrimination'.[108] Taking into consideration existing international conventions, the UNODC not only confirmed its internal position that decriminalising drug use and possession for personal consumption is 'consistent with international drug control conventions'; it further suggested that it may very well 'be required to meet obligations under international human rights law':[109]

> Worldwide, millions of people are imprisoned for minor, non-violent drug-related offences, in spite of the international drug control conventions' provisions permitting to apply alternatives to conviction in cases of a "minor nature". [...] A disproportionate share of those incarcerated is from the most marginalized groups such as people who are poor, and racial or ethnic minorities. [...] Incarceration, in turn, fuels poverty and social exclusion, as having a criminal record can negatively affect access to future employment, education, housing, and child custody and also exercising civil rights such as voting.[110]

Ultimately, the practice of criminalisation has provided the rationale for policies allowing police to stop, search, and arrest, and for courts to incarcerate, disproportionate numbers of ethnic minorities, whilst the associated aggressive over-policing and the concentration of incarceration in poor minority neighbourhoods has diminished the liberty and dignity accorded all members of society. The removal of criminal sanctions would likely result in a significant reduction of the number of individuals who are disproportionately restricted from capabilities under democratic equality—particularly those committing minor cannabis offences. Indeed, a study on depenalisation in England and Wales found that the declassification of cannabis from a class B to a class C drug in 2004 effectively made cannabis possession less risky due to the likelihood of a prosecution and maximum penalties being lowered.[111]

However, under the demands of democratic equality, decriminalisation may still be insufficient. Whilst reducing the high number of criminalised minor offenders, decriminalisation would likely place added pressure on minorities due to the fact that most drug networks 'predominately target minority ethnic neighbourhoods'.[112] As such, a heightened focus on enforcement against drug supply networks would disproportionately impact minority communities; Notwithstanding the number of people

---

[102] Decriminalisation means 'leaving production and distribution of the drug entirely illegal, but removing criminal penalties-that is, the threat of arrest and trial, though not necessarily the threat of monetary penalty-for possession for personal use'. Mark Kleiman, *Against Excess: Drug Policy for Results* (Basic 1992) 268.

[103] EN Yankah, 'A Paradox in Overcriminalization' (2011) 14 New Criminal Law Review: An International and Interdisciplinary Journal 1.

[104] ibid 33-34.

[105] Douglas Husak, *Overcriminalization: The Limits of the Criminal Law* (OUP 2007) 65.

[106] Yankah, (n107) 26.

[107] EN Yankah, 'Good Guys and Bad Guys: Punishing Character, Equality and the Irrelevance of Moral Character to Criminal Punishment' (2004) 25 Cardozo LRev 1019, 1031.

[108] United Nations Office on Drugs and Crime, *Briefing paper: Decriminalisation of Drug Use and Possession for Personal Consumption* (United Nations, 2015).

[109] ibid.

[110] ibid.

[111] Braakmann N, Jones S, 'Cannabis Depenalisation, Drug Consumption and Crime – Evidence from the 2004 Cannabis Declassification in the UK' (2014) 115 Social Science & Medicine 29, 31.

[112] Moran (n20) 575.

sentenced to custody for production in the UK increased after depenalisation,[113] the problem of racially disproportionate drug enforcement remains in that as the use of heavy-handed policing tactics decline, their residual use targets black and ethnic minorities even more.[114] Further, as decriminalising minor cannabis offences would increase demand without completely disabling criminal organizations absorbing more profits; 'minorities would be faced with not only the same criminals besetting their communities, but financially strengthened ones'. In this sense, some argue that 'decriminalization represents the worst of all possible policies'.[115]

Finally, as cannabis would still be prohibited under decriminalisation, the element of illegality would prevent the legal basis for acknowledgement of the failed regime of prohibition and its historical impact on minorities; it would not address the problem of disproportionate racial targeting in drug enforcement, nor shield non-minor and non-personal possession cannabis offenders from disproportionate severe punishments, and do nothing to mitigate the effects of criminalisation, stigmatisation and social exclusion to date. Accordingly, whilst 'some important features in the decriminalization of marijuana are critical to the success of any sustained effort to stem the tide of overcriminalization',[116] decriminalisation alone would not meet the guarantees of democratic equality. At best, from a rights-based standpoint it would constitute a bare minimum response—only marginally reducing the explosive growth of criminalisation under what is already universally recognised as a devastatingly draconian regime, without regard to the reprehensibility of prohibition altogether.

4.3.2. Legalisation and Commercialisation

In reality, democratic equality requires more than decriminalisation, because decriminalisation would not create a sufficiently substantive form of equality that democratic equality requires. It is therefore necessary to consider more extensive remedies on the basis that cannabis prohibition has failed entirely.

The current regime has seen the prohibition of a drug that 'at least physically speaking, does not put individuals at high risk, nor mark them out as noticeably different than nonusers'.[117] Yet, from inception, prohibition has concentrated on racial minority neighbourhoods, deepening poverty and social disorganization, thus increasing the vulnerability of residents to more crime; Its enforcement efforts and heavy surveillance increases apprehension of criminal offending; the concentration of individuals on probation increases the range of offenses for which persons can be charged.[118] It therefore appears that the impact of enforcing prohibition has caused more harm than the drug itself.

> That the increasing number of arrests has had ostensibly no diminishing effect on use rates also demonstrates the failures of current marijuana prohibition law.[119]

In many jurisdictions, cannabis is the third most popular drug, behind only alcohol and tobacco, despite having less physical health risks; in the UK, it is the most widely used illegal drug.[120]

And whilst there are many moral, health-related and economic arguments for legalisation, including a rejection of paternalism in favour of individual autonomy, economic stimulus, a more humane medical outlook, and growing evidence contradicting previous health-concerns posed by cannabis,[121] in any case, here we are concerned with whether the status quo is satisfactory regarding the demands of democratic equality.

In this regard, it seems nearly impossible to justify the consequences prohibition has inflicted on black and minority ethnic groups. The Marijuana Policy Project argues that whilst drug abuse is a real problem for minority communities, prohibition hasn't helped; instead, regulating cannabis similar to alcohol would weaken the criminal market and allow resources to be redirected elsewhere whilst keeping minorities out of jail.[122]

Subsequently, a growing number of users and nonusers alike now see little justification for infringing an individual's decision to engage in the use of an apparently harmless euphoriant.[123] Uruguay became the first country to legalise cannabis in 2011, followed by Canada as recently as October 2018. In the UK, doctors have been allowed to prescribe cannabis products for medical use since November 2018, whilst cannabis has been legalised in 10 US states. In light of the recency of such approaches, some have argued that it would be 'irresponsible' for member states to change their cannabis policies from a prohibitive regime in a hurry due to the experimental nature of legalisation policies and their unintended consequences;[124] Others have highlighted the gravity of the risk in that 'legalisation would be virtually irreversible if it goes badly wrong'.[125]

But it is hard to imagine a regime which could go more wrong than the one at present. And as Husak notes, the disaster of cannabis prohibition 'is not merely an abstract problem for the intellectual grist of legal theorists, but one that produces excessive and unjustifiable amounts of punishment to be visited on real people'.[126] Under a lens of democratic equality which has seen millions of black and minority individuals unable to effectively engage in a society of equals, *justice delayed is justice denied*.

Accordingly, legalisation is the only way to ensure the prevention of future disproportionate treatment of black and minority individuals and its devastating effects resulting from the inherently racialised war on cannabis.

4.3.3. Affirmative Action and Beyond

Whilst acknowledging that legalisation would be a positive step toward substantive equality, this too would do nothing to remedy past injustices: previously convicted individuals would still be considered criminals, face restrictions on democratic participation and be denied opportunities for employment, travel, etc. due to the legacy of criminalisation. Indeed, it has now been widely acknowledged that 'prohibition, combined with all too condemnable criminal behaviour among minority communities, has hollowed out neighbourhoods, gutted families, and all but guaranteed failed communities and generations'.[127] After decades of racialised cannabis policies, the broad painting of so many with the stigma of criminality has lead naturally to the creation of a despised criminal class; the 'association in the communal mind

---

[113] Braakmann (n115).
[114] Shiner (n45).
[115] Moran (n20) 576.
[116] Yankah (n107).
[117] Moran (n20) 589.
[118] ibid.
[119] ibid 571.
[120] UK Drug Policy Commission (n21).

[121] Bonnie and Whitebread (n5) 971.
[122] Moran (n20) 576.
[123] Bonnie and Whitebread (n5) 971.
[124] C Fijnaut, 'Legalisation of Cannabis in Some American States: A Challenge for the European Union and its Member States?' (2014) 22 EurJCrimeCrLCrJ 207, 217.
[125] Mark Kleiman, *Against Excess: Drug Policy for Results* (Basic 1992) 268.
[126] Douglas Husak, *Legalize This! The Case for Decriminalizing Drugs* (Verso 2002) 135.
[127] Yankah (n107) 9.

the image of the young black or brown man and the petty drug dealer' contributing to perceptions 'that criminal punishment is in part based on a person's criminal character'.[128] In this sense, there remain deeply embedded racist stereotypes of black criminality and dangerousness in public perceptions of crime.[129] This stigma and resulting social exclusion will likely endure for some time, regardless of any legislative reform. Accordingly, Husak suggests that 'repairing negative attitudes about law and authority among blacks is among the foremost challenges facing criminal justice policy in the twenty-first century'.[130] Whilst 'ending prohibition would be a major step toward alleviating racism in the criminal law'[131] and help prevent future harms, legalisation would offer no relief to the millions affected by prohibition who remain incarcerated at present or are denied access to a variety of opportunities post-conviction.

Therefore, it is finally worth considering whether affirmative action and redistributive justice measures can serve as appropriate tools for ameliorating the devastating legacy of prohibition.

Expungement One potential response has been the expungement of criminal records for cannabis offences. Such has the potential to free millions from prison whilst clearing records to restore access to opportunities such as financial grants, employment, travel, and voting. California has already adopted such a policy, with existing law authorising the District Attorney to recall or dismiss a sentence, seal or redesignate convictions for which a lesser or no offense would be imposed under the *Adult Use of Marijuana Act 2016*.[132] Reports on former convicts who have been pardoned to date indicate that although expungement cannot erase the damage done to their personal life, it 'can herald a whole shift in self-identity and a deluge of opportunities once barred to those whose records had been smeared with felony convictions'.[133]

This concept of expungement is not unfamiliar in the UK. Following the famous pardon posthumously granted in 2013 to war-hero Alan Turing, the *Policing and Crime Act 2017* contains a provision extending this to individuals who were convicted under historical legislation outlawing homosexual acts. Prior to this, there was already a procedure for living persons to apply to the Home Office for expungement—the justification being to remedy injustices resulting from convictions for past offences no longer considered punishable by society.[134]

Some critics have argued that statistical racial disparities alone should not justify rule changes, as 'crime is one area that we must keep absolutely off limits to concepts of affirmative action'.[135] Brolick proposes that if minorities are disproportionately arrested for crimes, they are also perhaps even more disproportionately victims of crimes; thus, affirmative action would elevate 'the so-called civil rights of criminals over the far more fundamental civil rights of their victims'.[136] However, it is Brolick himself who provides an exception for the application of affirmative action in criminal law: in the case of 'a dispositive showing

that a rule was adopted for discriminatory purposes'[137]—and the discriminatory basis of cannabis law and policy has long been demonstrated. Furthermore, an individual's choice to engage in cannabis-related activity is fundamentally a private one; Those who have suffered unjustly at the hands of the state through disproportionate enforcement and incarceration are the real victims—of a legally sanctioned regime which has destroyed communities and barred access to the exercise of actual rights and guarantees under democratic equality. Moral arguments against cannabis users are simply outdated in the wake of a clear commitment to remedying the problem of overcriminalisation and securing justice for the many individuals who have been convicted for acts which no longer justify criminal prosecution today.

Redistributive Justice Beyond the blackletter law, Anderson also declares that to be capable of functioning as an equal citizen also includes effective participation in the economy.[138] Whereas Dworkin suggests an equal society would ensure its members have an equal share of resources,[139] democratic equality justifies a more positive duty to achieve outcomes in which minorities can participate in society as true equals. Anderson argues that substantive equality may require different levels of resources for individuals from different circumstances, to 'avoid being oppressed by others and to function as an equal in civil society':

> Because of differences in their internal capacities and social situations, people are not equally able to convert resources into capabilities for functioning. They are therefore entitled to different amounts of resources so they can enjoy freedom as equals.[140]

In this sense, democratic equality goes beyond, eg, Rawls's maximin criterion of distributive equity,[141] by declaring that equality in the space of capabilities may justify an equal division of resources to accommodate the disadvantaged.[142]

Subsequently, it must be noted that in recent years, cannabis has proven to be a profitable cash crop in states that have legalised it—representing a multibillion-dollar state-sanctioned industry to a newly minted class of entrepreneurs.[143] Yet, Black and Latino victims of the drug war are noticeably absent from current markets, with less than 12 of the 3,000 US storefront dispensaries owned by blacks.[144]

> After a long history of pervasive discrimination in housing, employment, and education, Black and Latino Americans are far less likely to be able to raise the money necessary to start marijuana businesses. The interlocking systems of inequality leading to the racial wealth gap have made many of the black and Latino victims of marijuana prohibition unlikely to capitalize on the newly permissive environment.[145]

Amid predictions that the market for legal cannabis will reach $18 billion by 2020, it is particularly concerning that the profits of

---

[128] ibid.

[129] Tonry (n34).

[130] Husak (n130).

[131] Tonry (n34).

[132] AB-64 Control, Regulate and Tax Adult Use of Marijuana Act 2016.

[133] Madison Margolin, 'How Legalization is Helping Cannabis Felons Move on With Their Lives' (Civilized, 15 Nov 2018).

[134] Ministry of Justice, 'Thousands officially pardoned under "Turing's Law"' (UK Gov, 31 January 2017).

[135] Clint Bolick, 'Civil Rights and the Criminal Justice System' (1997) 20 Harvard Journal of Law & Public Policy 391.

[136] ibid.

[137] ibid.

[138] Anderson (n71) 317.

[139] Dworkin (n68) 283.

[140] Anderson (n71) 320.

[141] Joshua Cohen, 'Democratic Equality' (1989) 99 Ethics 727.

[142] ibid 729.

[143] Harvard Law Review, 'Drug Policy — Marijuana Justice Act of 2017 — Senator Cory Booker Introduces Act to Repair the Harms Exacted by Marijuana Prohibition.' (2018) 131 HarvLRev 926.

[144] Amanda Chicago Lewis, 'How Black People Are Being Shut Out of America's Weed Boom: Whitewashing the Green Rush', (BuzzFeed News , 16 March 2016).

[145] Harvard Law Review (n147).

legal sale have disproportionately gone to middle and upper-class white Americans; additionally, as 'the highest concentration of drug dealers is found in lower income, urban environments prone to minority dwelling', the commercialisation of cannabis 'might represent another tool of economic oppression' if minorities are excluded from the legal industry due to the current flow of income being diverted elsewhere.[146] Such discrepancies would threaten to further entrench inequalities exacerbated by prohibition.

Accordingly, in addition to remedying the impacts of criminalisation, democratic equality must also demand that 'future legislation committed to dealing with the harm exacted by prohibition' consider providing 'a regulatory framework that addresses the racially disparate distribution of the wealth generated by the market for legal marijuana'.[147]

Perhaps the most comprehensive attempt to address the legacy of prohibition in recent times was evidenced via the proposed *Marijuana Justice Act 2017*.[148] Citing the 'collateral consequences' of arrest and incarceration and other inequalities perpetuated by prohibition, the bill attempted to expunge all convictions for cannabis use or possession offences to 'help curb the myriad collateral consequences facing those convicted of marijuana-related felonies'; Additionally, it aimed to establish a Reinvestment Fund to create grant programs 'to reinvest in communities most affected by the war on drugs' by providing job training, re-entry services, health education, community centres, and youth programs.[149] The act followed earlier attempts at redistributive justice, such as California's *Proposition 64*,[150] which established an equity permit program giving preference in the cannabis permit process to residents convicted under prohibition, to secure better economic participation for victims of prohibition.

Such reforms are provided for by democratic equality under the recognition that one's capabilities of engaging as an equal in society 'are a function not just of one's fixed personal traits and divisible resources, but of one's mutable traits, social relations and norms, and the structure of opportunities, public goods, and public spaces'.[151] Ultimately, the longstanding results of prohibition on blacks and minorities extend much further than the mere allocation of resources. But without such provisions for affirmative action, cannabis legalisation threatens to further entrench inequalities exacerbated by the dreadful history of prohibition. Overall, the call for urgent reform is growing—'the time is ripe to overhaul cannabis laws to remedy the negative effects which have accrued, and chart forward a new path for drug law and policy'.[152]

5. Conclusion

This dissertation has illustrated how global cannabis prohibition has largely been shaped by discrimination; the effects of which have been evidenced from their origins right up to the present day. The initial justifications for cannabis prohibition laws were founded upon racial prejudices and politicised rhetoric, legitimising a cycle of oppression. These policies, globalised via international conventions, fuelled the expansion of disproportionate criminalisation, incarceration, and social exclusion, compounded by public perceptions associating blacks and minorities with violence and criminality. The cumulative effects have resulted in denial of the guarantees of democratic equality, which demand a society that works toward the abolition of oppression in favour of equal relations between individuals; including basic liberties, freedom from outcast status and social exclusion, and effective economic participation.

An appropriate response to this decades-long denial of equality would cease cannabis prohibition and the continuation of a system which disproportionately punishes minorities for an activity conducted no more by blacks than whites. Additionally, measures must be implemented to acknowledge the historical injustices of prohibition and ameliorate its present-day effects, which have predictably and deliberately stigmatised millions of black and minority ethnic individuals. Finally, in light of an emerging multi-billion-dollar legal cannabis industry which has seen an overwhelming majority of white entrepreneurs profit from the precise activities for which blacks and minorities have been disproportionately targeted, redistributive measures must be implemented to eradicate inequality and prevent future harm.

Whilst this dissertation has focused normatively on utopian ideals of a society under democratic equality, it is worth noting that the practical limitations often highlighted in relation to the proposed reforms are not insurmountable. Nevertheless, it is evident that the current regime of cannabis law and policy runs contrary to egalitarian principles; requiring a comprehensive overhaul if it is to stand up to the scrutiny of democratic equality theory.

All in all, the time is always ripe to chart forward a new path for a 21st century global legislative and policy framework of cannabis control—an order founded upon visions of a world which guarantees effective access to mutually shared benefits of modern civil society, with respect for true relations of equality for all.

■ References

[1] Alexander M, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness,* (The New Press 2010)

[2] Anderson ES, 'What Is the Point of Equality?' (1999) 109 Ethics 287

[3] Bonnie RJ, Whitebread CH, 'Forbidden Fruit and the Tree of Knowledge - An Inquiry Into the Legal History of American Marijuana Prohibition' (1970) 56 VaLRev 971

[4] Bowling B and Phillips C, 'Disproportionate and Discriminatory: Reviewing the Evidence on Police Stop and Search' (2007) 70 ModLRev 936

[5] Cannabis Reform Collaborative, *An Analysis of Cannabis Reform in Bermuda* (CRC April 2014)

[6] Global Commission on Drug Policy, *The War on Drugs* (Global Commission on Drug Policy 2011)

[7] Husak D, *Legalize This! The Case for Decriminalizing Drugs* (Verso 2002)

[8] Husak D, *Overcriminalization: The Limits of the Criminal Law* (OUP 2007)

[9] Mills J, *Cannabis Nation: Control and Consumption in Britain, 1928-2008* (OUP 2012)

[10] Nunn KB, 'Race, Crime and the Pool of Surplus Criminality: Or Why the "War on Drugs" Was a "War on Blacks"' (2001) 6 Journal of Gender Race & Justice 381

[11] Rawls J, *A Theory of Justice* (Harvard University Press 1971)

[12] Shiner M, Carre Z, Rebekah D, and Eastwood N, *The Colour*

---

[146] Moran (n20).
[147] Harvard Law Review (n147) 933.
[148] S. 1689 (115th).
[149] Harvard Law Review (n147).
[150] AB-64 CA (Control, Regulate and Tax Adult Use of Marijuana Act) (2016) (US).
[151] Anderson (n71) 319.
[152] Reuter and Trautmann (n105).

of Injustice: 'Race', drugs and law enforcement in England and Wales (LSE, Stopwatch & Release 2018)

[13] Tonry M, *Malign Neglect—Race, Crime, and Punishment in America* (OUP 1995)

[14] UK Drug Policy Commission, *An Analysis of UK Drug Policy* (London, UKDPC April 2007)

[15] Yankah EN, 'A Paradox in Overcriminalization' (2011) 14 New Criminal Law Review: An International and Interdisciplinary Journal 1

ASPIRING RESEARCHERS

Picture Credits : Kian Mooney

Breaking the Silence: Gender, Socio-Economic Status, and Violence in Partition Narratives

Liyah Hannah Anjum

King's College London

Abstract

This paper investigates the intersection of gender hierarchy and socioeconomic status in shaping the male violence that women experienced during the Partition in Punjab. Utilising a discourse analysis methodology, the findings reveal women from all socio-economic statuses faced violence because of their subordinate position in the gender hierarchy. They received little to no support from their families after surviving male violence, and men of middle and lower socio-economic status perpetrated much of this violence; influenced by economic disparity, demoralisation and military power. This paper highlights the inescapable nature of the gender hierarchy in communal and domestic spaces, and underscores the limitations of women's agency. This paper offers new insights into the scale of patriarchal control within the gender hierarchy and how socioeconomic intersections are less significant at explaining male violence women experienced during the Partition as opposed to the gender hierarchy.

1. Introduction

"It was man-made, and women suffered".[1] War and Conflict, orchestrated by men and subjecting women to torture, rape and violence. India and Pakistan often make headlines for the plethora of male violence women endure daily, the past events of Partition are no different as women were violated in savage attacks by men. Why did this male violence occur? I believe critical feminism can provide an explanation for this phenomenon. There is a plethora of critical feminist thought on Partition violence against women in the form of sexual violence and honour killings, however I want to analyse the relationship between gender and socio-economic status with female experiences of male violence during the Partition. Authors such as Chaudhary, Reena et al. and Srivastava extensively studies honour based male violence during the Partition and included a discourse analysis methodology to reveal women's "othered" and marginalised identity.[2] They emphasised spatially constructed identities in regards to nation building and the use of women's bodies as sites of violence in religious conflict. However, they overlooked the role socio-economic status played in shaping these identities of these women and the violence perpetuated by men. Chaudhary et al. briefly mentioned the importance of class, gender, and ethnicity but didn't explore further, while Srivastava only touched on socio-economic class in relation to spatial construction. Socio-economic class is based upon income, education, occupation,

wealth and social prestige, such factors influence one's identity, thus my research will explore this intersection in combination with the experience of communal violence and honour killings. Furthermore, these authors highlighted women's subjugation, dependency on men and vulnerability to violence, yet ignored the gender hierarchy. In particular regards to power, my research will explore the violence women experienced during the Partition in relation to power dynamics, thus the gender hierarchy.

I argue that the interplay of the gender hierarchy and socio-economic status reveals that a women's position in the gender hierarchy overpowers the identity of socio-economic status when constructing a subordinate female identity. I also argue that women of all socio-economic statuses experienced violence because of their subordinate gendered status. Women received no support after surviving male violence from their communities, regardless of socio-economic status due to their subordinate gendered status and men of middle/lower socio-economic statuses perpetuated the most violence due to demoralisation by conflict, economic imbalance and military authority. I will use Lene Hanson's post-structuralist discourse analysis model to analyse my data.[3] My first empirical section looks at the interaction between the gender hierarchy and identity. My first endeavour discusses how the low status of women in the gender hierarchy resulted in a subordinate "othered" identity that used women as sites for nationalistic violence. My second endeavour discusses how survivors of male violence had no agency except in the circumstance where they could change their religion and marry their abusers. My second empirical section "the interaction between socio-economic status and violence" discusses how

---

[1] Amrita Pritam, *India: A people partitioned*, interview by Andrew Whitehead, BBC World Service Radio Series, 1992.

[2] Ankita Chaudhary et al., "Partition Wounded Women Physically and Mentally: A Feminist Perspective," *Social Science Journal* 13, no. 2 (January 2023): 2540–50; Sumit Srivastava, "Revisiting Partition, 1947: Gender, Community and Violence," April 2014.

[3] Lene Hansen, *Security as Practice: Discourse Analysis and the Bosnian War* (Routledge, 2013).

men with a middle and lower socio-economic status were easily mobilised to commit violence in the communal sphere and honour killings in the domestic. I then proceed to explain how women from all socio-economic statuses experienced violence and received no support after surviving violence. I evidence these findings with examples of upper-class women and middle/lower class who survived abduction and rape, and how none of these women were given any familial or state support after surviving such ordeals.

My discussion is a cross sectional analysis between all my research questions. I begin with a synthesis of my findings from the previous sections, before using them to answer my research questions. I argue that only the gendered status of women during the Partition in Punjab impacted the likelihood of them experiencing violence because their gender identity "othered" them as symbols of honour and the nation. I disprove that socio-economic status impacted the likelihood as upper, middle and lower status women were vulnerable to communal violence because of their gender identity. I then again argue that only the gendered status of the community impacted the community response to female victims of male violence because all communities across Muslims, Sikhs and Hindus were patriarchal that subordinated women. This is evidenced by the abandonment of survivors, silence of survivors and marriage to abusers. This is seen in women from high, middle and low socio-economic statuses. I used critical feminist theorists: Peterson, Ward and Phillip to explain my conclusions about gender.[4] My third question argues both gender and socio-economic disparities within communities created power dynamics that were exploited by the perpetuation of violence against women, as honour killings show men abusing power at the top of the gender hierarchy. Furthermore, males with middle/low socio-economic statuses were more easily mobilised to commit violence. I used Cockburn's "Gender Relations as Causal in Militarization and War" to explain this.[5]

I will then explain how the double oppression women faced in the domestic and communal sphere suggesting the gender hierarchy was inescapable, contributes new material to the existing literature. Additionally, I will mention how many of my findings agree and solidify existing literature surrounding honour, national violence and "othering". Furthermore, I will also state how my research has answered the element of socio-economic intersection, by revealing gender identity and gender hierarchy is better at explaining the occurrence of Partition gendered violence. Finally, I will summarise my findings in the conclusion and explain the limitations of my research and its contribution to the Partition literature.

2. Literature Review

After reviewing the literature on gendered violence during the Partition, particularly its causes and consequences, I discovered that numerous authors have contributed to understanding the occurrence of communal violence and how women suffered in

relation to their spatial and national identities during Partition.

Violence against women during Partition was discussed most notably through Menon in "Rehearsing the Partition: Gendered Violence in 'Aur Kitne Tukde'", Virdee in "Remembering Partition: Women, Oral Histories and the Partition of 1947" and Butalia in "From the Other Side of Silence".[6] These foundational texts highlighted bias as stories such as "Buta Singh's" did not focus on the women's point of view despite them being abducted.[7] Butalia brings light to the lack of female agency in storytelling. Furthermore, Virdee emphasised women's social identity during Partition as limited to their relationships with men, noting that "their presence in male storytelling is largely limited to references to their relationships as daughters, wives and mothers."[8]

These literatures have been the building blocks for other authors to analyse and interpret empirical evidence and apply critical feminist theory. Mann takes inspiration from Butalia's and Menon's writings when analysing "the gendered rape and silence of the national body".[9] Mann wrote about the nationalistic symbolism women's bodies had in Partition as rape served as an instrument of national humiliation.[10] These findings were corroborated by Mukherjee in "Reading Women's Journey Through the Debris of Indian Partition in the 'Charnel Ground of History'", who explored how women's bodies became "carriers of national honour" during Partition.[11] There is a lack of intersectionality in the authors' arguments as class and religion are not discussed in relation to the violence experienced as this article focuses mostly on the legacy of Partition in India/Pakistan.

This element of religion is however explored in Deepa Narasimhan-Madhavan's "Gender, Sexuality and Violence: Permissible Violence against Women during the Partition of India and Pakistan".[12] This author analysed how religious nationalism enabled sexual violence as women's identities were reduced to instruments of political and religious messaging.[13]

The concept of honour is explored in many articles surrounding violence against women during Partition. Robert Hayden explains the concept in "Rape and Rape Avoidance in Ethno-national Conflicts: Sexual Violence in Liminalized States".[14] Women's bodies are sexed beyond objectification as the honour of the man, family and nation lie in the preservation of her honour.[15]

I have found this concept of honour and violence against women in Punjab during the Partition an integral concept in explaining the occurrence of such violence within the familial sphere and during communal violence. However, I want to further explore this concept of honour with respect to socio-economic status as it intersects with the concept of power and identity construction. Munawar et al. in "Female Sexuality as Carrier

[4] V. Spike Peterson, "Gendered Identities, Ideologies and Practices in the Context of War and Militarism," in *Gender, War, and Militarism: Feminist Perspectives* (California: ABC-CLIO, 2010), 17–29; Shannon Philip, "Making Men and Masculinities Visible: A Macro Level Enquiry into Conceptualizations of Gender and Violence in Indian Policies," *NORMA* 10, no. 3-4 (November 30, 2015): 326–38; Jeanne Ward, "It's Not about the Gender Binary, It's about the Gender Hierarchy: A Reply to 'Letting Go of the Gender Binary,'" *International Review of the Red Cross* 98, no. 901 (April 2016): 275–98.

[5] Cynthia Cockburn, "Gender Relations as Causal in Militarization and War," *International Feminist Journal of Politics* 12, no. 2 (June 2010).

[6] Jisha Menon, "Rehearsing the Partition: Gendered Violence in 'Aur Kitne Tukde,'" *Feminist Review*, no. 84 (2006): 29–47; Pippa Virdee, "Remembering Partition: Women, Oral Histories and the Partition of 1947," *Oral History* 41, no. 2 (2013): 49–62; Urvashi Butalia, "From 'the Other Side of Silence,'" *Manoa* 19, no. 1 (2007): 41–53.

[7] Butalia, "From 'the Other Side of Silence,'" 47.

[8] Virdee, "Remembering Partition," 51.

[9] Harveen Mann, "South Asian Partition Literature and the Gendered Rape and Silence of the National Body," *South Asian Review* 22, no. 1 (December 2001): 3–22, at 4.

[10] Ibid., 5.

[11] Anuparna Mukherjee, "Reading Women's Journey through the Debris of Indian Partition in the 'Charnel Ground of History,'" *Rocky Mountain Review* 66 (2012): 93–94.

[12] Deepa Narasimhan-Madhavan, "Gender, Sexuality and Violence: Permissible Violence against Women during the Partition of India and Pakistan," *Hawwa* 4, no. 2 (November 1, 2006): 396–416.

[13] Ibid., 414–415.

[14] Robert M. Hayden, "Rape and Rape Avoidance in Ethno-National Conflicts: Sexual Violence in Liminalized States," *American Anthropologist* 102, no. 1 (2000): 31.

[15] Ibid., 31–32.

of Masculinity" further support this concept by arguing that women's sexuality was used to construct and reinforce masculine identity during Partition.[16] Additionally, Dey in "The Female Body as the Site of Male Violence during the Partition of India" explores how women's bodies became territories of contestation between communities.[17]

The only Partition violence literature I found that mentions violence women experienced during Partition and how it may intersect with socio-economic status is Srivastava's "Revisiting Partition, Gender, Community and Violence." Srivastava mentions class briefly, noting that "the relationship between violence and class cannot be overlooked"[18] and that "classes interact with violence in complex ways."[19] However, this was not explored further in the analysis.[20] Chaudhary et al. also briefly acknowledged that "class, gender, and ethnicity are important intersections" but did not provide a detailed exploration.[21]

Upon reading various literature mentioned, I found a common intersection of gender and socio-economic status to be a reoccurring gap. I found that in the literature reviewed there was the element of spatial identity based on nationalism and honour, but no sufficient exploration of the social and economic dynamics between women's experience of violence and those that perpetuated it.

3. Methodology

My research question and sub-research questions:

"What does the interplay of gender hierarchy and socio-economic status reveal about the male violence women experienced during the Partition in Punjab?"

- Did the status of women during the Partition in Punjab impact the likelihood of them experiencing violence?
- Did the status of the community impact the community response to female victims of male violence?
- Did disparities within communities create power dynamics that were exploited, leading to the perpetuation of violence against women?

I have decided to use discourse analysis to analyse and interpret data. Many authors in this area of research have used discourse analysis. Discourse analysis looks at how language can construct meaning within socio-cultural contexts.[22]

The spectacle of violence by Dasgupta uses discourse analysis successfully to analyse Partition fiction that focus on the sexual violence committed against Hindu and Sikh women from Muslim men.[23] She analysed the text and found repetitive signs that constructed women as sexualised victims, Hindu and Sikh women as pure chaste victims and Muslim men as barbaric predators.[24]

The data that will be collected will be oral testimonies, written testimonies and memoirs, as they are unrefined and convey intricate detail in comparison to accounts and fictional short stories. Using content analysis I intend to discover repetitive patterns and binaries in order to systematically reach a conclusion. I have chosen to analyse written testimonies and oral histories from Partition because they provide primary accounts of the violence that occurred. Partition is often regarded as "forgotten" as it happened approximately 80 years ago, thus data is limited.[25] Data sources include Kavita Puri's *Partition Voices* BBC Radio 4 series,[26] Devika Chawla's *Home, Uprooted*,[27] Andrew Whitehead's BBC World Service Radio Series *India: A People Partitioned*,[28] and Ian Talbot and Darshan Singh Tatla's *Amritsar*.[29]

As Talja outlines, qualitative discourse analysis examines how language within texts constructs and negotiates meaning within socio-cultural contexts.[30] My analysis will not include quantitative data regarding religion, income or literacy rates as this information was not recorded in 1947. Therefore, to determine socio-economic class of witnesses I will classify them based on their living arrangements, occupation, and education levels as described in the testimonies.

Furthermore, thematic analysis could be argued instead, however I do not intend to simply identify and report patterns within my analysis of qualitative data.[31] In order to discern the interaction between the gender hierarchy, socio-economic status and violence, I need to go beyond the identification of themes and codes within the cultural context of Punjab and analyse how language used in the data set constructs these oppositional male and female identities.

Srivastava also used discourse analysis to uncover why violence against women occurred during the Partition. The author found a women's religious and gendered identity to be the main reasoning behind the violence they suffered.[32] Although Srivastava's research was limited because it focused on religion and specifically stated she did not want to intersect socio-political causes into her analysis.[33] This weakened her argument, as communities can be influenced by socio-economic situations. Furthermore, honour and violence against women during Partition was influenced by socio-economic imbalances and the mobilisation of middle and lower class men in conjunction with religious influence, as I will discuss.

The literature previously mentioned applies discourse analysis, which found women were reduced to bodies they had no ownership over during the Partition, and that those bodies were mobilised to represent and serve the nation. My analysis intends to explain why this occurred; was it only the construction of gendered identity that facilitated violence or did socio-economic status play a role.

This explains why poststructuralist discourse analysis is the best option as it can explain how identities are constructed based on social contexts of the case study. For my research, the identities constructed are gendered ones: male and female. Hansen's model focuses on how security discourse constructs identity through

[16] Riffat Munawar, Asma Yunus, and Shahzad Mushtaq, "Female Sexuality as Carrier of Masculinity: A Feminist Critique of History of Subcontinent Partition (1947)," *European Academic Research* 1, no. 8 (November 2013): 2167–75.

[17] Arunima Dey, "The Female Body as the Site of Male Violence during the Partition of India in Bapsi Sidhwa's 'Ice-Candy-Man,'" *Complutense Journal of English Studies* 26 (November 13, 2018): 27.

[18] Srivastava, "Revisiting Partition," 123.

[19] Ibid., 124.

[20] Ibid., 126–127.

[21] Chaudhary et al., "Partition Wounded Women," 2548.

[22] Hansen, *Security as Practice*, 41.

[23] Shumona Dasgupta, "The Spectacle of Violence in Partition Fiction: Women, Voyeurs and Witnesses," *Journal of Postcolonial Writing* 47, no. 1 (February 2011): 30–41.

[24] Ibid., 32.

[25] Virdee, "Remembering Partition," 51.

[26] Kavita Puri, "BBC Radio 4 – Partition Voices, Series 1, Aftermath," BBC Radio Four, August 7, 2017.

[27] Devika Chawla, *Home, Uprooted* (Fordham University Press, 2014).

[28] *India: A People Partitioned*, interview by Andrew Whitehead, BBC World Service Radio Series, 1992–2008.

[29] Ian Talbot and Darshan Singh Tatla, *Amritsar* (Seagull Books, 2007).

[30] Sanna Talja, "Analysing Qualitative Interview Data," *Library & Information Science Research* 21, no. 4 (November 1999): 471.

[31] Victoria Clarke and Virginia Braun, "Thematic Analysis," *The Journal of Positive Psychology* 12, no. 3 (December 9, 2016): 297.

[32] Srivastava, "Revisiting Partition," 124.

[33] Ibid., 124.

a system of binary oppositions.[34] Hansen's model looks for the "linking" and "differentiation" between self and other in texts. Gender and socio-economic status are part of these self/other representations. However, as I previously mentioned, this is only one element of the identity construction as religion and socio-economic status also have an impact. Discourse analysis does have drawbacks; it can struggle to account for structural factors such as the role of institutions, policies, or economic systems that may influence social constructions. Therefore it is important to keep in mind that discourse analysis alone cannot fully capture the structural forces such as colonial legacies, that also shaped violence during the Partition.[35]

This research will be methodologically post-positivist because my theory is critical feminist theory. As my research question looks at the gender hierarchy, a critical feminist lens is the most appropriate as it analyses the power imbalance between males and females in society. As explained by Ward, the gender hierarchy places masculinity as superior to femininity in a social system where gendered power is unequal.[36] This theory explains why men engaged in violence against women during the Partition, as they were at the top of the hierarchy, and women were at the bottom. Peterson complements this, and also extends it by explaining how socio-economic status was utilised as a vector for "othering" women.[37] She argues that women are objectified as "markers of group boundaries" and "bearers of group honour."[38] This theory explains how women's identity was categorised as subordinate in the gender hierarchy that made them vulnerable to violence. I also believe this theory explains why men engaged in violence to protect their honour as well as their women. If the women in a man's family were attacked by the "other" religion, their identity as patriarchal protectors was threatened and this showed they could not protect their women, thus lowering their position in the gender hierarchy. This explains why men felt the need to engage in violent vengeance of the "others" women, as it would rectify this and save their masculine identity. This is evident in the concept of honour killings, which will be further explored in my analysis.

Furthermore, I believe this theory regarding the gender hierarchy can intersect with Cynthia Cockburn's theory that economic, ethnic/national and gender power all intersect to construct the identity of those who are violent during war/conflict.[39] I believe this is exemplified in the violence men committed during the Partition. As men of all socio-economic statuses, religions and ethnicities committed violence to protect their honour, I argue these three intersections can explain the violence men committed; specifically, demoralisation by conflict, being feminised by lowered socio-economic status, and the threat to their position atop the gender hierarchy angered them and pushed them to regain it by committing violence. I will use these theories to explain my findings.

Philip proposes that in patriarchal Indian society masculinity is defined in opposition to femininity, and that men's social standing depends on their capacity to control women's behaviour and sexuality.[40] Cockburn also notes how men were "demoralized and angry" during conflict, which supports how the feminisation

of their identity due to their women being abused pushed them into the notions of vengeance to recover their emasculated gender identity.[41]

4. Empirical Sections

The data I have collected and analysed will be presented as my main findings in the proceeding two sections titled "Gender Hierarchy and Identity" and "Socio-Economic Status and Violence". These sections will look at the interaction between the two variables.

4.1. Gender Hierarchy and Identity

Having analysed my data, I have found that women during the Partition in Punjab occupied a subordinate identity to men and a low status in the gender hierarchy, which was further lowered by being a survivor. My data revealed, men recalled the violence against women in their stories. For example, in *Amritsar* by Talbot and Tatla, many men recount events of violence against women such as "I saw a young girl being raped," "Women were stripped and paraded naked."[42] This language indicates that women were seen as an extension of the men representing purity and honour. This suggests protection of women was vital, as a failure of such would result in a failure of self protection.

I also found this subordinate identity was symbolised; as the violence these women experienced or witnessed illustrated how women were not centred, and that their identities were spatially constructed. Poonan Josh was 17 years old when she migrated from Pakistan to Amritsar, India. She recalls how a group of Muslim men attacked her convoy and killed her family. The male violence she witnessed was targeted and deliberate, and she described the attackers as "huge, wild-looking men."[43]

My second revelation regarding gender hierarchy and identity was that women who survived male violence sometimes chose to marry and stay with their kidnappers or rapists. This seemed to be a choice because they viewed their kidnappers as saviours, perhaps in comparison to the violence they had witnessed in their own community. In *Amritsar*, Kaur recounts how some women "refused to go back" and "wanted to stay with their new husbands."[44] This presented an element of agency — albeit limited and constrained — that women exercised within the rigid gender hierarchy.[45]

Another element to this revelation includes a unique perspective that these women preferred life with their kidnapper as opposed to their own family. Kaur tells of one girl: "I get milk twice a day. My mother-in-law does not order me to get out of bed and sweep."[46] "Does not order me" suggests that in her previous home she was ordered around, perhaps forced to do household chores. This goes beyond simply doing chores as it exemplifies this notion that women's identity within the familial sphere was that of a subordinate — a servant. The gender hierarchy was thus inescapable, both in the communal and domestic sphere.[47] This also shows how women could exercise the little agency they had by marrying their kidnapper. Most likely by doing this, girls could avoid permanent marginalisation from the family and community by taking on a new religious identity in a new community.

[34] Hansen, *Security as Practice*, 43.
[35] Senem Aydin-Düzgit and Bahar Rumelili, "Discourse Analysis: Strengths and Shortcomings," *All Azimuth: A Journal of Foreign Policy and Peace*, 2018, 4–5.
[36] Ward, "It's Not about the Gender Binary," 275.
[37] Peterson, "Gendered Identities, Ideologies and Practices," 27.
[38] Ibid.
[39] Cockburn, "Gender Relations as Causal in Militarization and War," 150.
[40] Shannon Philip, "Making Men and Masculinities Visible," 327.

[41] Cockburn, "Gender Relations as Causal in Militarization and War," 152.
[42] Talbot and Tatla, *Amritsar*, 150.
[43] Kavita Puri, "BBC Radio 4 – Partition Voices, Series 1, Aftermath," BBC Radio Four, August 7, 2017.
[44] Talbot and Tatla, *Amritsar*, 20.
[45] Ibid., 27.
[46] Ibid., 26.
[47] Ibid., 30.

4.2. Socio-Economic Status and Violence

The first theme that frequented my data was men with a middle to lower socio-economic status committed the most violence as they were more easily mobilised. Discourse analysis revealed that themes of honour, binary opposition between Hindus/Sikhs and Muslims, and demoralisation were ever-present in the testimonies of those who witnessed and experienced violence.

Swaran Singh Riat, a 15 year old Sikh was involved in an attack in 1947, he recalls how he was told "one man from each house was required to attack a Muslim village. To avenge the recent slaughter of Sikhs."[48] He recalls feeling compelled to participate despite being young, demonstrating how middle and lower status males were mobilised into violence through community pressure and notions of collective honour.

Anant Kaur, an upper class Hindu deputy officer witnessed violence against women during the Partition. Her job was to reunite kidnapped girls, and she recalls girls from low class families and "well to do families" who were kidnapped.[49] She recalls how the kidnappings transcended class boundaries, emphasising that all women were vulnerable regardless of socio-economic status.[50]

Furthermore, upon analysing Angela Araha's interview from *India: A People Partitioned*, she evidences this further. Angela was born into a Christian upper-class family in Pakistan. She recalls groups of men shooting nuns and killing her father at a church. Despite being from an upper-class family, her experiences of violence were devastating.[51]

Furthermore, a final example of this mobilisation of lower/middle status men to commit violence is illustrated in Kartar Kaur's and Bisham Sahni's Partition story. Kartar Kaur migrated from Pakistan to Firozpur Punjab and can be classified as middle class based on her descriptions of her household and community. She experienced violence firsthand and her testimony reveals how men from similar backgrounds were the primary perpetrators.[52] Men in her community questioned the masculinity of those who did not participate, asking "why not kill all your young women and girls."[53] "Why not" shows their questioning of the men's decisions, thus their ability to lead and most importantly their ability to protect their honour. As theorised, the identity of these men was threatened by conflict and the feminisation of their inability to protect women, driving them to violence.

Another theme that emerged from my findings was that all women who experienced violence regardless of socio-economic status, were not granted sufficient familial or state support. Re-occurring patterns of stigma, dishonour, shame and marginalisation were present in all testimonies.[54]

In *Home, Uprooted* by Devika Chawla, she tells the story of a well-educated upper-class woman.[55] It can be seen she was upper-class as she was educated, lived in a lavish house and her sister was a doctor.[56] Despite coming from a higher socio-economic status, she survived brutal violence as she was struck on the head

multiple times with an axe by Muslim men. After her rescue, she found that despite her class status, the community response was the same as for women of lower status: shame, silence, and lack of support.[57] The family tried to suppress her story as it brought dishonour.[58]

Anant Kaur's stories can be exemplified again. As previously mentioned Anant Kaur was an upper-class officer that reunited girls with their families for the abandonment act. Although she did not experience any violence herself being upper class she suffered emotionally from witnessing the abandoned girls. She tells of a girl who was "rejected by her own family" after being recovered.[59] Regardless of socio-economic status, she was rejected by her own family, marginalised and physically sent away. This illustrates the subordination these women endured within the communal and familial sphere, as being a survivor resulted in facing consequences.

This lack of support after surviving such ordeals is further evidenced in Kaur's retellings of "the girls were adamant not to go back."[60] She also mentions how as an officer "we took her out, but she cried, she would not go back to Pakistan" and how other officers stated "bring her." These examples illustrate the lack of female agency after surviving violence as they were forced against their will to return. "Bring her" encapsulates the attitude society had towards these women, as the brash tone indicates women were not respected. The abandonment act made it clear that women were subordinated and their wishes and wellbeing came second to honour, the nation and the patriarchal demands of the state.[61]

4.3. Cross-Sectional Analysis

This section contains a cross analysis between my findings, as I answer my research questions and address my thesis. I will then use critical feminist theories to explain these findings before situating them within the literature.

I have found that the status of women during the Partition in Punjab did impact the likelihood of them experiencing violence. I found only the gendered status of women impacted the likelihood, as socio-economic status did not. My findings show that women were the primary targets of communal violence because of their identity as "carriers" of familial and national honour, as discussed in my analysis. Chaudhary et al. found that "women's bodies became battlegrounds for communal violence" during Partition.[62] Men's patriarchal identity was constructed in opposition to these feminised subjects; men wanted to prevent dishonour and being "othered," "feminised" and emasculated.[63] The subordinate position of women in the gender hierarchy unfortunately made it extremely likely they would experience violence during the Partition as their identity was spatially constructed to represent the honour of the group.

These findings and conclusions reinforce revelations made by Dasgupta, Chavda, Mann and others in Partition literature as they discuss these nationalistic and honour focused representations extensively. Dasgupta discusses how women's bodies represented the honour of the community and became "spectacles of violence."[64] Mann argues that the "national body" was gendered,

[48] Puri, "BBC Radio 4 – Partition Voices, Series 1, Division," BBC Radio Four, July 31, 2017.

[49] Talbot and Tatla, *Amritsar*, 126.

[50] Ibid., 154.

[51] Angela Araha, *India: A people partitioned*, interview by Charles Haiviland, BBC World Service Radio Series, February 8, 2003, 4–7.

[52] Talbot and Tatla, *Amritsar*, 124.

[53] Ibid., 127.

[54] Bhisham Sahni, *India: A people partitioned*, interview by Andrew Whitehead, BBC World Service Radio Series, September 9, 1996.

[55] Chawla, *Home, Uprooted*, 45.

[56] Ibid., 51.

[57] Ibid., 35.

[58] Ibid., 48.

[59] Ibid., 127.

[60] Ibid., 27.

[61] Talbot and Tatla, *Amritsar*, 36.

[62] Chaudhary et al., "Partition Wounded Women," 2543.

[63] Peterson, "Gendered Identities, Ideologies and Practices," 27.

[64] Dasgupta, "The Spectacle of Violence in Partition Fiction," 35.

and that the rape of women during Partition served as a symbolic attack on the nation itself.[65] Chavda explores how women's sexual identity was forcibly stripped as a mechanism of communal violence.[66]

In regards to my second research question, my findings revealed that the gendered status of communities did impact the community response to female victims of male violence. I found that the gendered status did impact the community response to survivors because all communities were patriarchal. I found that socio-economic status did not impact the community response because my observations reveal that communities of all socio-economic statuses failed to adequately support survivors of male violence. My findings unveil that survivors were often abandoned, silenced or faced further violence upon return. This subordination was cross-cutting across all religious and socio-economic lines.

My findings also show women often feared returning to their families after surviving male violence as they were either sent away or killed. This shows that women were further subordinated in the gender hierarchy because they could no longer reproduce the purity and honour of the family, having been "violated." Philip proposes that in patriarchal Indian society, marriage is the only institution that gives women legitimate protection.[67] This in combination with the discovery that they faced violence in their families and communities contributes to the understanding that the gender hierarchy was inescapable in both the domestic and communal spheres.

The occurrence of women and girls who didn't want to leave their kidnappers or "husbands" after Partition can be explained by their mobilisation to depend on male protection. My conclusions regarding the abandonment act reinforce findings by Chaudhary, as she extensively covers the patriarchal state's control and oppression of abducted women.[68] My observations revealing that socio-economic status doesn't impact the community response to survivors fills this gap in my literature as no author specifically explores how the community response to survivors intersects with socio-economic status. It reveals that gender identity is a better determinant than socio-economic status of community attitudes towards survivors of male violence. This also contributes new material to the literature surrounding the agency women possess when it comes down to gender identity. Furthermore, my findings regarding the small amount of agency women exercised by marrying their abductor, also contributes new information to the wider debate as it exposes how truly subordinated these young girls were, and one of the few circumstances they could exercise agency.

Furthermore, my observations found that gendered disparities and socio-economic disparities within communities did create power dynamics that were exploited, leading to the perpetuation of violence against women. My observations detail how men with a middle and lower socio-economic status were more easily mobilised to commit violence than men of upper class status. This is because they were demoralised by conflict, economically disadvantaged and were given authority by the military.

I believe that my conclusions surrounding middle/lower status men perpetuating the most violence can be explained by Cynthia Cockburn's "Gender Relations as Causal in Militarization and War". She proposes that relative power (economic power, gender power and ethnic/national power) intersects to create the identity of those who become violent during conflict.[69] Cockburn notes how men were "demoralized and angry," supporting how the feminisation of their identity due to their women being abused pushed them into the notions of vengeance to recover their emasculated gender identity.[70]

I also believe this theory can explain why army men were found to be raping and abusing women. I believe they were exploiting their power because their authoritative position was constructed by a superior position in the gender hierarchy, an increase in economic/class power from conflict and a national/ethnic power. This discovery fills the socio-economic gap in my literature as it shows how socio-economic and gender disparities created a power gap that was exploited by male violence.

5. Conclusion

To conclude, this paper has found that the interplay of the gender hierarchy and socio-economic status reveals that women of all socio-economic statuses experienced violence because of their subordinate gendered status. Women received no support after surviving male violence from their communities, regardless of socio-economic status due to their subordinate gendered status and men of middle/lower socio-economic statuses perpetuated the most violence due to demoralisation by conflict, economic imbalance and military authority. This paper addressed three sub-research questions:

- Did the status of women during the Partition in Punjab impact the likelihood of them experiencing violence?
- Did the status of the community impact the community response to female victims of male violence?
- Did disparities within communities create power dynamics that were exploited, leading to the perpetuation of violence against women?

I found the status of women during the Partition in Punjab influenced their vulnerability to violence irrespective of socio-economic status. Kidnapping, rape, mutilation, and honour killings were common, impacting women of all socio-economic backgrounds. The gender hierarchy played a more significant role than socio-economic status in determining women's vulnerability.

The gendered status of communities significantly influenced responses to female victims of male violence, irrespective of religious affiliation. All communities failed to effectively support survivors, regardless of socioeconomic-status. Survivors suffered oppression and subordination, often finding agency in marriage to their abductors, yet were rejected and stigmatised regardless of status. Fear of returning to their families further illustrates the inescapable gendered dynamics of post-violence experiences.

Gender and socio-economic disparities within communities fuelled power imbalances, perpetuating violence against women during Partition. Middle/lower-status individuals, including army officers, villagers, and tribal members, were the main culprits. Cockburn's theory of gender relations and conflict explains how alterations in power dynamics shape aggressive masculine identities, as seen during Partition. Lower-status males, demoralised by conflict and economic imbalance, were easily mobilised to commit violence. Army members exploited their authoritative position, highlighting how intersections of gender, economic, and national power perpetuated violence.

[65] Mann, "South Asian Partition Literature," 5.
[66] Yogini Chavda, "Unveiling Identity and Sexuality: Analyzing Sexual Violence, Trauma, and Alienation in Amrita Pritam's *Pinjar* as a Lens on the Partition of India," *Vidhyayana – An International Multidisciplinary Peer-Reviewed E-Journal*.
[67] Philip, "Making Men and Masculinities Visible," 329.
[68] Chaudhary et al., "Partition Wounded Women," 2545.

[69] Cockburn, "Gender Relations as Causal in Militarization and War," 152.
[70] Ibid., 150.

I believe my research was limited by the small sample of my data. This was due to a data breach at the British Library thus an inability to obtain more information. Additionally, narratives on Partition violence are rare to find due to the taboo nature of the subject. Further research would benefit from a broader collection of testimonies and a wider geographical scope beyond Punjab. This would also ensure less bias and misinformation as it would not rely on interviews recalling the past.

I believe my study sheds light onto the extent of which patriarchal control was responsible for violence against women during the Partition. I also believe it provides an insight into the role gender identity plays in South Asian society and how heavily it influences community behaviour and responses to violence. This paper's analysis of socio-economic intersections with gender hierarchy contributes new material to the Partition literature and underscores the need for further research into the structural forces that enabled, perpetuated, and normalised gendered violence across all levels of society.

■ References

[1] Ahmed, Akbar S. "The Politics of Ethnicity in Pakistan Society." *Asian Affairs* 21, no. 1 (February 1990): 20–35.

[2] Araha, Angela. *India: A people partitioned*. Interview by Charles Haiviland. BBC World Service Radio Series, February 8, 2003.

[3] Aydin-Düzgit, Senem and Bahar Rumelili. "Discourse Analysis: Strengths and Shortcomings." *All Azimuth: A Journal of Foreign Policy and Peace*, 2018.

[4] Brass, Paul R. "The Partition of India and Retributive Genocide in the Punjab, 1946–47." *Journal of Genocide Research* 5, no. 1 (March 2003): 71–101.

[5] Butalia, Urvashi. "From 'the Other Side of Silence.'" *Manoa* 19, no. 1 (2007): 41–53.

[6] Chaudhary, Ankita, Reena, Gaurav Sharma, Baijanti Mala, and Sapna Goel. "Partition Wounded Women Physically and Mentally: A Feminist Perspective." *Social Science Journal* 13, no. 2 (January 2023): 2540–50.

[7] Chavda, Yogini. "Unveiling Identity and Sexuality: Analyzing Sexual Violence, Trauma, and Alienation in Amrita Pritam's *Pinjar* as a Lens on the Partition of India." *Vidhyayana – An International Multidisciplinary Peer-Reviewed E-Journal.*

[8] Chawla, Devika. *Home, Uprooted.* Fordham University Press, 2014.

[9] Christensen, Ann-Dorte and Palle Rasmussen. *Masculinity, War and Violence.* Routledge, 2018.

[10] Clarke, Victoria and Virginia Braun. "Thematic Analysis." *The Journal of Positive Psychology* 12, no. 3 (December 9, 2016): 297–98.

[11] Cockburn, Cynthia. "Gender Relations as Causal in Militarization and War." *International Feminist Journal of Politics* 12, no. 2 (June 2010).

[12] Daiya, Kavita. "'Honourable Resolutions': Gendered Violence, Ethnicity, and the Nation." *Alternatives: Global, Local, Political* 27, no. 2 (April 2002): 219–47.

[13] Dasgupta, Shumona. "The Spectacle of Violence in Partition Fiction: Women, Voyeurs and Witnesses." *Journal of Postcolonial Writing* 47, no. 1 (February 2011): 30–41.

[14] Dey, Arunima. "The Female Body as the Site of Male Violence during the Partition of India in Bapsi Sidhwa's 'Ice-

Candy-Man.'" *Complutense Journal of English Studies* 26 (November 13, 2018): 27–45.

[15] Hansen, Lene. *Security as Practice: Discourse Analysis and the Bosnian War*. Routledge, 2013.

[16] Hayden, Robert M. "Rape and Rape Avoidance in Ethno-National Conflicts: Sexual Violence in Liminalized States." *American Anthropologist* 102, no. 1 (2000): 27–41.

[17] Hill, K., W. Selzer, J. Leaning, S. J. Malik, and S. S. Russell. "The Demographic Impact of Partition in the Punjab in 1947." *Population Studies* 62, no. 2 (2008): 155–70.

[18] Mann, Harveen. "South Asian Partition Literature and the Gendered Rape and Silence of the National Body." *South Asian Review* 22, no. 1 (December 2001): 3–22.

[19] Menon, Jisha. "Rehearsing the Partition: Gendered Violence in 'Aur Kitne Tukde.'" *Feminist Review*, no. 84 (2006): 29–47.

[20] Mookerjea-Leonard, Debali. "Quarantined: Women and the Partition." *Comparative Studies of South Asia, Africa and the Middle East* 24, no. 1 (2004): 33–46.

[21] Mukherjee, Anuparna. "Reading Women's Journey through the Debris of Indian Partition in the 'Charnel Ground of History.'" *Rocky Mountain Review* 66 (2012): 93–105.

[22] Munawar, Riffat, Asma Yunus, and Shahzad Mushtaq. "Female Sexuality as Carrier of Masculinity: A Feminist Critique of History of Subcontinent Partition (1947)." *European Academic Research* 1, no. 8 (November 2013): 2167–75.

[23] Narasimhan-Madhavan, Deepa. "Gender, Sexuality and Violence: Permissible Violence against Women during the Partition of India and Pakistan." *Hawwa* 4, no. 2 (November 1, 2006): 396–416.

[24] Navarro-Tejero, Antonia. "Sacks of Mutilated Breasts: Violence and Body Politics in South Asian Partition Literature," 2019.

[25] O'Loughlin, Kelly, Donna Seger, and Faculty Advisor. "Sexual Violence as a Crime against Humanity as Seen in the Indian Partition of 1947," 2018.

[26] Peterson, V. Spike. "Gendered Identities, Ideologies and Practices in the Context of War and Militarism." In *Gender, War, and Militarism: Feminist Perspectives*, 17–29. California: ABC-CLIO, 2010.

[27] Philip, Shannon. "Making Men and Masculinities Visible: A Macro Level Enquiry into Conceptualizations of Gender and Violence in Indian Policies." *NORMA* 10, no. 3-4 (November 30, 2015): 326–38.

[28] Pritam, Amrita. *India: A people partitioned*. Interview by Andrew Whitehead. BBC World Service Radio Series, 1992.

[29] Puri, Kavita. "BBC Radio 4 – Partition Voices, Series 1, Division." BBC Radio Four, July 31, 2017.

[30] Puri, Kavita. "BBC Radio 4 – Partition Voices, Series 1, Legacy." BBC Radio Four, August 14, 2017.

[31] Puri, Kavita. "BBC Radio 4 – Partition Voices, Series 1, Aftermath." BBC Radio Four, August 7, 2017.

[32] Roy, Kaushik. "Partition of British India: Causes and Consequences Revisited." *India Review* 13, no. 1 (January 2, 2014): 78–86.

[33] Sahni, Bhisham. *India: A people partitioned*. Interview by Andrew Whitehead. BBC World Service Radio Series, September 9, 1996.

[34] Saxena, Chandni. "Dimensions and Dynamics of Violence during Partition of India." *Proceedings of the Indian History Congress* 74 (2013): 909–20.

[35] Shafer, Kevin and Spencer L. James. "Gender and Socioeconomic Status Differences in First and Second Marriage Formation." *Journal of Marriage and Family* 75, no. 3 (May 20, 2013): 544–64.

[36] Srivastava, Sumit. "Revisiting Partition, 1947: Gender, Community and Violence." April 2014.

[37] Swenden, Wilfried. "The Territorial and Non-Territorial Management of Ethnic Diversity in South Asia." *Regional & Federal Studies* 22, no. 5 (December 2012): 613–31.

[38] Talbot, Ian. "The 1947 Partition of India and Migration: A Comparative Study of Punjab and Bengal." *Removing Peoples. Forced Removal in the Modern World*, 2009, 321–47.

[39] Talbot, Ian and Darshan Singh Tatla. *Amritsar*. Seagull Books, 2007.

[40] Talja, Sanna. "Analyzing Qualitative Interview Data." *Library & Information Science Research* 21, no. 4 (November 1999): 459–77.

[41] Virdee, Pippa. "Remembering Partition: Women, Oral Histories and the Partition of 1947." *Oral History* 41, no. 2 (2013): 49–62.

[42] Ward, Jeanne. "It's Not about the Gender Binary, It's about the Gender Hierarchy: A Reply to 'Letting Go of the Gender Binary.'" *International Review of the Red Cross* 98, no. 901 (April 2016): 275–98.

[43] Wilcox, Lauren B. *Bodies of Violence: Theorizing Embodied Subjects in International Relations*. Oxford University Press, 2015.

The Bermuda Journal
of Academic Research

Family Comes First: How Agricultural Households Operating Under Market Failures Absorb a Labour Supply Shock

Ana Van Der Ree

University of Oxford

Abstract

I use the agricultural household model's separation property as a framework to analyse how agricultural households absorb a large shock in a context where markets function poorly. Under complete markets, farm production decisions should depend solely on prices and technology, and are separable from household characteristics like the household's labour endowment. I show that Covid-19 caused a large increase to households' labour endowments in rural Ethiopian farm households. I use this finding to examine whether the labour supply shock was so large that imperfect labour markets were unable to fully absorb it, causing households to increase total labour utilisation on the farm in response. The results show that agricultural households were operating under market failures before and after the shock. I find no evidence that the labour input misallocation consequences of market failures were statistically larger after Covid-19. I test the propositions by estimating first-differences specifications for two distinct panels of farm households in rural Ethiopia, a pre-Covid panel (2011–2015) and a Covid panel (2018–2021) using LSMS-ISA data.

1. Introduction

Poor households constitute a large proportion of the population in low-income countries. Across Sub-Saharan Africa, many of the rural poor are reliant on farming for their livelihoods and consumption needs (Suri & Udry 2022). In these settings, labour is the household's primary endowment. A sudden shock to labour supply, therefore, prompts households to re-optimize their labour allocation decisions across tightly linked production, consumption and labour markets. Understanding how agricultural households absorb a shock to labour supply provides an important insight to the functioning of markets in rural settings.

Covid-19 introduced a sudden disruption to mobility and migration patterns, economic activity and health attitudes. Consequently, many rural households retained members at home that may have otherwise migrated away, or received returning members back to the home. This represents a large shift outwards of the labour supply curve in rural labour markets. If rural labour markets function imperfectly and wages do not adjust fully, this results in an above-clearing market wage. Wage 'stickiness' has potentially important consequences for employment and output, with a wage set above market-clearing level leading to higher unemployment, in theory.[1] Kaur (2019) finds that nominal

wages do not adjust downwards in response to (negative demand) shocks in a setting for casual agricultural labour. Downwards wage rigidity causes employment reductions borne by poorer individuals—landless and small landholders—and when rationed out of the market, landholders increase labour supply to their own farms (Kaur 2019). This suggests there is vast potential for the labour supply shock from Covid-19 leading to "disguised unemployment" in agricultural settings if wages are rigid. In this paper I use the agricultural household model as a framework for a set of labour market institutions through which to analyse how rural Ethiopian households absorb a large positive shock to labour supply.

The agricultural sector is an interesting context to study labour (mis)allocation because of evidence that low agricultural productivity and high share of agricultural employment can explain—at least in part—a large fraction of cross-country dispersion in aggregate productivity (Gollin & Udry 2021; Restuccia et al. 2008). Therefore, understanding the labour allocation decisions of the agricultural household is also key to understanding what drives productivity growth. Since the demand for agricultural labour is highly seasonal and many African countries display low population density, agricultural labour markets may be highly constrained, making it difficult for farms to source labour at the right times (Suri & Udry 2022). On the other hand, farmers may be "poor but efficient" and learn not to repeatedly misallocate their (few) resources at their own peril (Schultz 1964). This would suggest that labour markets

---

[1] While this is a prediction of classic labour demand-supply interactions, evidence for it is mixed in the literature. There is little direct evidence that wage rigidity actually affects employment in the labour market in any setting, with Card (1990) as a notable exception.

function smoothly for the most part. The completeness of labour markets will determine how households absorb a labour supply shock, thus motivating my investigation based on a test for market completeness.

Labour markets in rural settings are often interconnected with markets for consumption, output and credit, as small family-run farms produce agricultural goods both for consumption and sale. The agricultural household model captures this interdependence (Singh et al. 1986). Assuming markets are complete and the household is a price taker in all markets, farm production decisions are treated as if the utility-maximising household operates a profit-maximising firm. This implies input choices only depend on prices and farm technology. The model yields a testable prediction of separation: consumption decisions are made in the second stage after any household income from farm profits has been pinned-down in the first stage. As the model's prediction relies on the assumption that markets are complete, rejection of the separability hypothesis is evidence for markets failures. This has prompted a literature which tests for complete markets using the separation hypothesis (Benjamin 1992; Dillon & Barrett 2017; Dillon et al. 2019; LaFave & Thomas 2016; Udry 1996).

Tests of the agricultural household model's assumption are well rehearsed in the literature. I aim to use the model's prediction as a framework to analyse how agricultural households absorb a large shock in a context where markets function poorly.

I employ a test for separation based on labour demand on the farm, as follows. If the family farm behaves as a profit-maximising firm in complete and competitive markets, then optimal production decisions—such as input utilisation—do not depend on any household characteristics—such as labour endowment. I regress the demand for farm labour on household labour endowment to test if a relationship exists. If labour endowments affect the farm's labour input decisions, resources are allocated sub-optimally; households are operating under market failures.

My first proposition is that Covid-19 caused a large positive shock to household labour endowment because mobility restrictions and business closures prompted members to stay or return home. My second proposition is that the outwards shift of the labour supply curve was so large that the shock was not fully absorbed, so households increased total labour utilisation on the farm in response. This is a failure to separate household characteristics from production choices. My third proposition is that before the shock, markets were already functioning poorly and the large supply shift led to a deterioration in the consequences of market failures. My fourth proposition is that more 'sophisticated' households were better able to absorb the shock by placing household members in the formal wage market.

I test the propositions on two panels of households—before and after Covid-19—from rural Ethiopia using 10 years of agricultural household surveys spanning from 2011–2022. Ethiopia's Living Standards Measurement Surveys and Integrated Surveys on Agriculture (LSMS-ISA) consists of an initial panel of households that was tracked from 2011–2015 and a different refreshed panel tracked from 2018–2021. Longitudinal data is key for the test to ensure results are not contaminated by unobserved heterogeneity between households. Better management practices or soil quality, for instance, are related to farm labour utilisation and can be misattributed to changes in household labour endowment if these unobserved characteristics are not controlled for. Using variation within the household across time improves the identification of

the test because such heterogeneity is held constant. A limitation of Ethiopia's LSMS-ISA data is that the same households were not observed continuously before and after the Covid shock, which means observed differences may partly reflect cohort differences. But it also has a key advantage: it enables the baseline surveys (2011 and 2018) to represent two groups that face similar labour-market conditions and life-cycle stages. I explore this further in the data description.

Firstly, I show the panels of households are similar at baseline across a range of household and farm characteristics, confirming they are comparable for investigation. This includes that average household labour endowment is almost exactly the same at baseline. Secondly, I show that households exposed to Covid-19 experienced a larger increase in labour endowment on average between the baseline and follow-up, supporting my first proposition. This is primarily driven by a greater retention of ageing children in the household. Thirdly, I find evidence to reject complete markets both before and after Covid-19. Households increase labour utilisation on the farm in response to increases in their labour endowment, consistent with my second proposition that households were unable to fully absorb the shock. The elasticity almost doubles for households exposed to Covid-19 compared to the earlier panel, suggesting market failures led to greater labour misallocation consequences for farm households who experienced the shock. When testing this formally in a pooled sample including all households and a binary indicator if the household was part of the Covid panel, the Covid-differential in the elasticity fails to reach statistical significance. This is qualitatively consistent but quantitatively inconsistent with the third proposition. I explore whether more sophisticated households with experience in the formal wage market absorb the shock better by using their market experience to adjust on that margin. I find no significant evidence for heterogeneity in household responses.

The rest of the paper proceeds as follows. In Section 2, I discuss the implication of complete markets (Subsection 2.1), the labour effects of Covid-19 in Ethiopia (Subsection 2.2) and the literature on the separation test (Subsection 2.3). The Agricultural Household Model is presented in Section 3. Section 4 outlines the identification strategy and Section 5 summarises the data. The key results are discussed in Section 6 followed by a heterogeneity analysis in Section 7. I conclude in Section 8.

2. Background

2.1. Complete markets

A market is complete when arbitrary units of the good can be bought at a specific price and there are no information asymmetries or transaction costs. Evidence for incomplete markets suggests agricultural households are operating under market failures and are constrained in achieving profit-maximisation. This may be contrasted with other explanations for their behaviour—such as operating in non-competitive markets, entirely missing markets, or behavioural constraints such as social norms—which would require different policy (Dillon & Barrett 2017). Markets can be incomplete for several reasons: asymmetric information leading to family members being used on the farm more than hired workers; imperfect substitutability of family and hired labour; heterogeneity in the price paid to farmers and the market price of the good; rationing from limits on the number of hours family members can work outside the household; weak farm contract enforcement.

It is not possible with a general test for separation to identify the sources of market failures. Specifically, non-separation is evidence for market failures in at least two factor or output markets. LaFave and Thomas (2016) point out this is not always a disadvantage due to the nature of rural settings having inter-linked markets. In fact, no single binding constraint explains low African agricultural productivity and it is likely that a combination of market failures operate in multiple markets simultaneously (Suri & Udry 2022). Therefore, while I implement a test based on labour utilisation on the farm, it is not strictly evidence for labour markets failures. If credit markets function poorly, credit constraints can limit a household's ability to hire labour even if labour markets are complete.[2] However, because of the nature of the Covid shock causing a large shift in labour endowments, I direct my attention to labour markets as this is primarily where the shock will be absorbed.

2.2. The Covid-19 Pandemic in Ethiopia

The Covid-19 pandemic hit Ethiopia hard, ranking it 6th in Africa for most total infections.[3] Even if the health shock was modest (arguably, it was not), the economic shock was large. The government imposed a six-month long State of Emergency in 2020 which banned public meetings, restricted travel and forced school closures (Ayele & Fessha 2021). In a comprehensive policy brief, D. Harris et al. (2021) find that employment and incomes declined across the board during the pandemic and, while they bounced back by 2021, much of this represented a move to lower-quality and temporary jobs. They observe that casual labourers were less likely to find jobs and recover their incomes compared to small business owners even long after the onset of the pandemic. 'Young Lives' also documents a shift to working in the agricultural sector and towards self-employment (Ford et al. 2021). Part of this may be explained by lay-offs in factories and export-oriented production. Garment worker lay-offs in the Hawassa Industrial Park—one of several major public industrial infrastructure projects—resulted in many workers having to migrate out of the city, which could be reflective of a broader return of industrial workers to rural areas (D. Harris et al. 2021). Although schools re-opened, by the end of 2020 around two-fifths of 19-year-olds still in education had not engaged in any form of learning since school closures, with those from the poorest households and rural areas affected most (D. Harris et al. 2021).

The descriptive findings together support the proposition that the pandemic caused a positive labour supply shock to agricultural households. The moderate shift towards agricultural "low-quality" work may also have disproportionately affected agricultural households by further constraining rural labour markets. Nolte et al. (2022) find that Ethiopian respondents state the reduced availability of hired labour as an important factor in preventing the pursuit of 'normal' agricultural activities after Covid-19, suggesting agricultural labour markets exhibited market failures.

2.3. Previous Literature

The agricultural household model was first formally developed in the foundational work by Singh et al. (1986) while Bardhan and Udry (1999) brought it to the dynamic setting. The core analytical concept of integrating consumption and production decisions was influenced by Sen (1962)'s and Chayanov (1991)'s observations

that a family farm's production volume is likely related to its family size. In a seminal paper, Benjamin (1992) found no evidence for incomplete markets in rural Indonesia by estimating a cross-sectional relationship between farm labour demand and household composition. A source of concern in the cross-section is the potential bias arising from unobserved variation across households in the sample. LaFave and Thomas (2016) return to the same setting two decades later with extremely granular panel data to address the potential bias. They exploit within-household variation to rule-out fixed unobserved heterogeneity between households driving the results and come to the opposite conclusion of Benjamin (1992).

Udry (1996) brings the test to the African context to bridge the knowledge gap on characterisations of African rural markets. He uses plot-level panel data from Burkina Faso and Kenya to test for separation and rejects market completeness in both settings. The lack of rich data on plot quality limits Udry (1996) as he cannot rule-out unobserved variation in plot characteristics as confounders. This concern is echoed by more recent findings in Gollin and Udry (2021) which point out that unobserved heterogeneity and measurement error can lead to misallocation being overstated by three-fold. This signals to a broader point that the margin of productive inefficiency in Africa is possibly inflated in the literature, which motivates this investigation.

Dillon and Barrett (2017) use cross-sections of LSMS-ISA data to find evidence for market failures, estimating elasticities of household labour endowment to farm labour demand within the range of 0.32–0.53 for Sub-Saharan African countries including Ethiopia. A few years later, Dillon et al. (2019) return to the exploration by exploiting LSMS-ISA panels and, similarly, find no evidence for separation. Dillon et al. (2019) use the first two waves of Ethiopia's LSMS-ISA surveys. To my knowledge, the separation property has not been previously used to investigate how farm households operating under market failures adapt their farm labour allocations in response to a large labour supply shock.

3. Theoretical Framework

3.1. A dynamic model of the agricultural household

The agricultural household model captures the interdependence of consumption and production decisions by illustrating a profit-maximising family-run farm (firm) that operates within a household maximisation problem (Singh et al. 1986). This replicates the characteristic common in many developing countries in which households produce crops both for consumption and market sale. The model shows that when the household is a price taker for all goods which it consumes and produces and markets are complete, optimal farm inputs are determined independently of leisure and consumption choices. Then, given the income derived from production, consumption decisions are made. Therefore, decision-making is sequential.

In this simplified version of the model the household gets utility from consumption of market and agricultural goods (denoted for simplicity as one good, c_t) and leisure, l_t. The infinitely lived household maximises the expected present discounted value of current and future utility:

$$\max \ \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t U(c_t, l_t)\right] \tag{1}$$

The household can produce the good/crop on its farm using inputs land, A_t, and labour, L_t, according to its production

---

[2]Without a functioning credit market, the household's shadow value of cash (used to pay hired labour) is higher than the interest rate. The households will hire less labour than optimal.

[3]When ranking in terms of most infections per population, Ethiopia falls to the 37th in Africa.

technology:

$$F(L_t, A_t) \tag{2}$$

Labour used on the farm can be sourced from family labour, L_t^F, or hired labour from the marketplace, L_t^H. Land inputs can come from the farm's own land, A_t^F, or rented-out land, A_t^M:

$$L_t = L_t^F + L_t^H \tag{3}$$

$$A_t = A_t^F + A_t^M \tag{4}$$

Although in practice other inputs such as fertilizer will be used and the farm will produce various crops, these features are abstracted away to simplify the analysis and do not change the model's predictions. The critical part of the model is that households face complete markets: they take as given the price of the good, p_t, land rental price, a_t, and wage for market work, w_t. Importantly, this implies the market wage is the shadow wage for farm work, which is therefore the same shadow wage for hired and family labour. This underpins the assumption that hired and family labour are perfect substitutes. The model also abstracts away any market frictions, such as matching frictions in hiring, time lags between seeking and obtaining a job, labour supply shortages, or limited availability of jobs. This has implications for the assumptions regarding how Covid-19 affected labour markets, which I unpack in light of my results in Section 6.

The household's time endowment, E_t^L, is divided between working on the farm, L_t^F, working in the market, L_t^M, and leisure, l_t. Similarly, the household's land endowment, E_t^A, can be used for their own farming activity, A_t^F, or rented out to other households, A_t^M:

$$E_t^L = L_t^F + L_t^M + l_t \tag{5}$$

$$E_t^A = A_t^F + A_t^M \tag{6}$$

Because this is a dynamic model, farmers can smooth consumption across time by borrowing or saving in complete credit markets with interest rate r_t. This comes together in the inter-temporal budget constraint, where Ω_t is the household's wealth at the beginning of period t:

$$\Omega_{t+1} = (1+r_t)\left[\Omega_t + w_t(L_t^F + L_t^M) + \{p_t F(L_t, A_t) - w_t L_t - a_t A_t\} - p_t c_t\right] \tag{7}$$

Wealth in the next period, Ω_{t+1}, is equal to the interest earned on wealth in the current period net of wage earnings from labour work $w_t(L_t^F + L_t^M)$, farm profits (total revenue less total costs: $p_t F(L_t, A_t) - w_t L_t - a_t A_t$), and consumption expenditure ($p_t c_t$). Substituting (5) we obtain an expression that includes l_t:

$$\Omega_{t+1} = (1+r_t)\left[\Omega_t + w_t(E_t^L - l_t) + (p_t F(L_t, A_t) - w_t L_t - a_t A_t) - p_t c_t\right] \tag{8}$$

With complete and competitive markets, households choose the quantities of c_t and l_t which maximise (1) subject to the budget constraint (8), resource constraints (3)–(6), and non-negativity constraints:

$$l_t, \, c_t, \, L_t^M, \, L_t^F, \, L_t^H \geq 0 \tag{9}$$

In the first stage, the household chooses inputs L_t^* and A_t^*

to maximise expected farm profits. In the second stage, the household maximises utility, conditional on expected income from the farm. This sequentiality is what makes the model recursive. The solution to the problem is, therefore, characterised by:

$$\Pi^* = \max \, \mathbb{E}\left[p_t F(L_t^*, A_t^*) - w_t L_t^* - a_t A_t^*\right] \tag{10}$$

$$A_t^* = A_t(a_t, w_t, p_t) \tag{11}$$

$$L_t^* = L_t(a_t, w_t, p_t) \tag{12}$$

where total labour demand on the farm, L_t^*, and total land demand on the farm, A_t^*, depend only on the prices of inputs, crop price, the farm's productive technology and shadow prices. Crucially, any household preferences or characteristics that affect consumption decisions do not affect production choices. This leads to a neat result: while it initially appeared that the agricultural household solved a joint problem where consumption and production decisions were intertwined, with complete markets, production decisions are separable from consumption choices. The opposite is not true: consumption still depends on the amount of profits realized from production as this enters the budget constraint. If all markets but one are complete—for instance, if land cannot be traded—then separation still holds because relative prices can adjust to accommodate the non-tradable good (Feder et al. 1985).

The separation result has a testable prediction that, under complete and competitive markets, farm labour demand should not be affected by household labour endowment. Intuitively, if a household's labour endowment falls—for instance because a young male leaves to find work in the city—demand for labour on the farm should remain unaffected because family labour is substitutable with hired labour at the same shadow wage. Similarly, excess supply of family labour can be sold in the market at the same wage so should have no impact on production.

3.2. How Covid-19 affects the separation prediction

Proposition I: My first proposition is that Covid-19 caused a large positive shock to household labour supply. There are two possible mechanisms:

- *Retention:* Household members who otherwise may have chosen to leave to urban areas for employment or education delay departure due to uncertainty. For instance, young men in the household who would have migrated for better work opportunities stay in the household longer.
- *Return:* Members living outside the household return home during the pandemic. For instance, adults moving back with family for food security reasons, and youth returning from urban areas due to school or business closures.

This represents an increase in the household's labour endowment. If markets clear, additional labour will be fully absorbed by the labour market and will have no effect on input decisions. For instance, extra family labour will be sold to the local market at the shadow wage due to the perfect substitutability of labour, or hired farm labour will contract and be replaced by the new family member. If markets do not clear, the positive shock to labour endowment may not be fully absorbed. Farms may respond by expanding total labour utilisation on the farm.

Proposition II: My second proposition is that markets are not fully able to absorb the labour supply shock. The outwards shift

of the labour supply curve in the agricultural labour market may be so large that wages do not fully adjust if wages are 'sticky'. Therefore, households respond to the increase in their labour endowment by employing a higher quantity of total labour on the farm. For instance, farms may add more family labour to their input mix without reducing hired labour inputs even though this need not be optimal for L_t^*. This is a failure to separate household characteristics from production choices. In other words, family may come first.

Proposition III: There are two possibilities regarding the state of market completeness: the first scenario is that before the Covid shock, markets were clearing. The increase in labour supply prevented what was previously characterised as market completeness. More probably, however, before the shock, markets were already functioning somewhat poorly. My third proposition is, therefore, that the large supply shift led to greater consequences from market failures. This is equivalent to the intensive margin of non-separation.

In addition to the Propositions, I provide examples for three other mechanisms via which Covid-19 can have affected households and markets:

Changing the household's demographic composition. Returning household members bring with them new spouses or children into the household; increased illness or death in the household; elderly individuals encouraged to move in with their family or children. Such changes in household demographics can increase care-taking burdens, particularly for women, and crowd out time availability for farm work. It should have no effect on farm labour demand if the model's assumptions hold because only technology, prices and inputs determine L_t^*. Under separation, the care burden is accounted for in the utility function $U(\cdot)$ via a fall in leisure and the household would compensate by adjusting its consumption-leisure trade-off. The adjustment to the care shock happens in the second stage after profit-maximising decisions have been pinned down in the production process. If instead, care burdens shift labour demand on the farm, then demographic factors are "leaking" into production decisions and separation is violated.

Increasing labour market frictions. Infection risk or increased scepticism about Covid-19 increases transaction costs of hiring labour onto the farm; travel restrictions reduce supply of hired labour; lower availability of non-agricultural jobs; information frictions from difficulty of knowing who is available or eligible to be hired onto the farm. The model technically assumes there are no frictions in the labour market. In practice, frictions are likely to exist in rural settings. Covid-19 also most likely intensified frictions. This may limit the suitability of the model's assumptions for a shock context like Covid-19, which I explore further in light of the results in Section 6. Frictions may contribute to the "family comes first" effect by making it more difficult to hire labour or more desirable to use family labour on the farm.

Increasing credit constraints. Tighter household budgets or credit, for instance due to lower remittances from urban areas, make hiring costly even if labour is available; increased perception of risk by local banks or moneylenders lowers loan availability or increases interest rates. The various market failures may strengthen one another's impact on farm labour utilisation.

I do not explicitly test for the effects of changing demographic composition, market frictions or credit constraints on farm labour demand in my identification. However, they serve as an important backdrop to the results and can shed light on additional mechanisms operating in the household.[4]

4. Identification

4.1. Separation test

I test for separation in the agricultural household model through the household's demand for labour in (12). Following Dillon et al. (2019), I estimate the following empirical specifications for labour utilized by the farm household h at time t:

$$L_{ht} = \beta_0 + \beta_1 E_{ht} + \gamma A_{ht} + \gamma_2\, demog_{ht} + \eta_t + \kappa_{ea} + \epsilon_{ht} \quad (13)$$

$$\Delta L_{ht} = \beta_0 + \beta_1 \Delta E_{ht} + \gamma \Delta A_{ht} + \gamma_2 \Delta demog_{ht} + \epsilon_{ht} \quad (14)$$

Equation (13) is a pooled cross-section model where the variation is across households in the sample and equation (14) is a household fixed effects model in first differences where the variation is within-households over time. The latter is the main specification of interest. The variables L_{ht}, E_{ht} and A_{ht} enter as logs, following the literature (Dillon et al. 2019; LaFave & Thomas 2016).

L_{ht} is the total number of person days of labour used on the farm h in period t and E_{ht} is the household's labour endowment in period t. Labour demand may be related to other farm characteristics, such as land size, which is captured by the cultivated acreage of the household in each period, A_{ht}. Farm labour demand may be affected by other household demographics, such as the share of working-age or elderly adults. $demog_{ht}$ accounts for this as a vector of controls including the share of prime-age males, prime-age females, elderly males and elderly females in each household h in each period t, following Dillon et al. (2019).[5]

η_t represents time fixed effects at the survey wave/year level. This absorbs any wave-specific shocks—such as a bad harvest year, national policy or general macro trends in labour use—which influence labour demand across all households. In equation (13) η_t enters as levels and in (14) it drops out because all time-fixed effects are absorbed by first differencing. κ_{ea} are village fixed effects that absorb any time-invariant village-specific characteristics, such as elevation or distance to urban cities, in order to compare households within the same village (enumeration area, EA) by holding this heterogeneity constant. Village fixed-effects disappear in first-differencing.[6] ϵ_{ht} is an error term. We might expect the error terms of households in the same village to be correlated if there are village-level shocks that affect labour demand. To tackle the intra-cluster correlation, I allow errors in each EA-wave combination to be correlated with one another but independent of other EA-wave cells by clustering standard errors at the EA-wave level.

Under the null that separation holds and markets clear, household labour endowment E_{ht} will have no impact on farm

---

[4]I do not explore heterogeneous responses to the labour shock across gender lines due to time constraints. There is evidence that gender-differentiated plots matter in Western Africa, while such differentiated allocations are not as marked in Eastern Africa. An interesting addition would be to explore whether return of males compared to females affects farm labour demand. Dillon et al. (2019) find no such evidence in the same context of agricultural Ethiopia.

[5]Shares are calculated as the number of individuals in that category divided by the total number of individuals in the household. Prime-age is defined as 15–60 years old, elderly is defined as 61 years and older.

[6]Time-varying village spillover effects may be a concern. There are possible general equilibrium effects if many households in a village experience return migration, such as wage depression or lower hiring from other villages. While I do not explicitly address this, it remains as a valid consideration to keep in mind for the results.

labour demand and β_1 will be zero. However, identification in the pooled cross-section is contaminated by unobserved household-specific characteristics such as managerial skills or preference for working on one's own farm. For instance, a more productive farmer will demand more labour on her farm.[7] These characteristics tend to be time-invariant or slow-moving across time. As these characteristics are unobserved, in (13) they will be captured by ϵ_{ht} and bias the estimate if they are correlated with labour endowment. One solution is to add these characteristics into the specification as in Benjamin (1992). However, as they are unobservable and difficult to measure, this has often been a struggle in the literature that uses cross-sectional data. Specification (14) addresses this by first-differencing out all time-invariant heterogeneity in each household, allowing identification to come solely from within-household changes in labour endowment across time. This is the key advantage of using panel data.

Including the vector $demog_{ht}$ helps isolate the effect of labour endowment itself by holding constant the underlying demographic composition of the household. The number of children or elderly individuals in the household may impact farm labour demand in practice if separation fails. Therefore, by omitting this vector from the specification, it risks attributing the effects that actually come from household structure (i.e. who is available within the household) to labour endowment (i.e. how many workers are available within the household). Separating these effects leads to a cleaner test on ΔE_{ht} with a clear mechanism.

4.2. Identifying the effect of the Covid-19 shock

To examine how households absorbed the Covid-19 shock to labour endowment I estimate the specifications in two different time periods:

(i) **Panel I:** a pre-Covid period representing counterfactual market conditions,

(ii) **Panel II:** a Covid period representing market conditions after the shock.

Comparing the estimated elasticity across the panels suggests whether the shock amplified non-separable behaviour. This approach provides a comparative test across the two regimes revealing how households absorbed the shock—however, it does not constitute a formal differences-in-differences design as the same households are not observed continuously across all four waves. A larger elasticity in the Covid period suggests greater separation failure.

However, the comparison is treated cautiously because Covid-19 may have altered market conditions—such as by introducing market frictions—in ways that are important for the cleanness of the separation test. I discuss this in Section 6. The two panels, of course, differ in more ways than only Covid exposure—other shocks or cohort effects could also drive differences in the estimated slopes. I examine differences in the baseline data for each of the panels in the next section and find that the data suggests the two panels are comparable across a range of farm and household characteristics.

5. Data

5.1. Overview of the data

I test the model's predictions using four waves of comprehensive household data from Ethiopia's LSMS-ISA, surveys conducted by Ethiopia's Central Statistics Agency and the World Bank. They collect extremely rich survey data on over 3000 nationally representative rural households on a wide range of topics necessary for the separation test including farm production, consumption, labour and individual demographics. The granularity of the agricultural variables and ability to track individuals across waves are key advantages of this dataset.

The surveys were conducted in 2011, 2013, 2015, 2018, 2021.[8] The panel of households was refreshed in the 2018 wave to ensure the sample was representative for Ethiopia's eleven regions and for rural and urban areas. Therefore, the 2018–2021 households represent a new panel, not a follow-up of the 2011–2015 sample.

Ideally a differences-in-differences design would follow the same households through all waves, so that pre- versus post-Covid comparisons come entirely from within-household changes. Using data from the two different panels might seem a disadvantage because changes observed between panels could in principle arise from cohort or survey differences. But it also has a key advantage: it enables the baseline surveys (2011 and 2018) to represent two groups that face similar labour-market conditions and life-cycle stages. If the same individuals had been tracked from 2011 to 2021, the 18–24 year-olds in the first wave would be 28–34 years old in the fifth wave, which encapsulates an entirely different life-cycle labour decision.[9] In contrast, by comparing differences between 2011–2015 to differences between 2018–2021, the two panels each include (i) the baseline survey for their sample, and (ii) span a relatively similar interval of three or four years.[10] Comparison is corroborated by the feature that both panels are nationally representative and households are selected from the same regional states.

Since the separation test relies heavily on changes in household size, it is important that selective attrition does not contaminate inferences. For the pre-Covid panel, 91.2 percent of the rural households from 2011 were re-surveyed in 2015, representing relatively low attrition. The Covid panel suffers from a larger response rate problem: the Tigray conflict in northern Ethiopia between 2020–2022 resulted in the Tigray region being excluded from the 2021 survey. This meant the 398 rural households (or 54 rural enumeration areas) from the Tigray region were dropped in the Covid-panel. To minimise the potential influence of this discrepancy on the results, I run all key regressions of the earlier panel both with and without Tigray households and find that the results remain unaffected (see Addendum A1).

5.2. Variables and survey details

The questionnaires begin with a household roster including basic demographic data (age, sex) on the individuals living in the household. I use this list of household members to create a variable for the labour endowment of household h in year t,

---

[7]A more skilled manager extracts more output from each worker-day (better task allocation, fewer mistakes) so that each additional worker has a higher marginal product. Under profit-maximization, the manager demands labour up to the point where the marginal product of labour equals its (shadow) wage cost. Thus, a more productive farmer will demand a higher quantity of farm labour, ceteris paribus.

[8]The surveys are conducted over the course of 9–12 months starting in September, so technically the waves are 2011–2012, 2013–2014, 2015–2016, 2018–2019, 2021–2022. For simplicity, I will refer to the waves as the year in which the surveying began.

[9]For instance, households from the 2011 sample that have children of the relevant age for this investigation would have moved beyond that phase by 2021 and been replaced by characteristics that were not relevant for the test in 2021 but were in 2011.

[10]The 2021 wave was likely originally planned for 2020 following the consistent 2-year increment in the LSMS surveys but delayed due to the pandemic. Therefore, the interval between waves 4–5 is three years.

E_{ht}, by counting the number of working-age individuals. I had to choose an age cut-off at which someone enters or exits the working-age. Following Dillon et al. (2019), I include all adults (including senior citizens) in the count for E_{ht} and allow children to gradually age into E_{ht} using an equivalence scale: 10-year olds and younger do not count, 11-year-olds count as 0.2 adults in the workforce, 12-year-olds as 0.4 adults, and so on where a 15-year-old counts as one full adult. The results are robust to different specifications of labour endowment including without the phase-in of children. I also use the household roster to construct the vector of demographic controls, $demog_{ht}$, which is the share of prime-age (15–60 years-old) and elderly (61+) males and females. To construct the cultivated acreage, A_{ht}, I use the area reported by plot holders in the farm or, in cases where area was not recorded, I use the GPS-measured area from the survey.

The agricultural survey modules are granularly measured at the holder-parcel-field-crop level. There is evidence that within-household allocation of labour inputs across plots, including across gender lines, matters for misallocation and productivity (Duflo & Udry 2004; Goldstein & Udry 2008; Kilic et al. 2015; Udry 1996). Nevertheless, while acknowledging this literature, I aggregate all variables to the household-level because this is the relevant unit for my analysis; variation at the plot level is not necessarily informative for my predictions about farm behaviour in response to the Covid-19 shock. It is also difficult to link disaggregated plots across survey years as these units are sometimes noisily measured, so aggregating also smooths these measurement errors (Gollin & Udry 2021).

Because the analysis explores the household's demand for labour on the farm, my sample consists of only land-holding households that cultivated in at least one survey year—i.e. excluding households that display no demand for labour on their farm ever.

Accurate measurement of labour used on the farm is crucial for this research. The surveys distinguish between two farming seasons—cultivation (also known as post-planting) and harvest—and collect comprehensive data on family and hired labour in each season on the farm.[11] Surveys conducted in informal settings have struggled with accurate measures of labour supply: Gollin and Udry (2021) emphasize the prevalence of measurement error in similar settings exploring misallocation in African agriculture.

While LSMS-ISA surveys measure the quantity of hired workers, the survey structure may overlook nuances in how labour is exchanged in Ethiopian informal settings. In an analysis of his Ethiopian fieldwork, Aspen (1993) discusses how financially-unremunerated exchange labour for agricultural tasks like weeding and harvesting is common in rural Ethiopia, known as *däbo* or *wänfäl*. This involves a mutual exchange where the host serves food and beer in exchange for labour, and participation is expected to be reciprocal. Importantly, exchange labour is not solely a production decision: participation is also motivated by the desire to remain invested in the social network and feel a sense of loyalty. The LSMS-ISA do not clarify whether mutual exchange labour is included in the respondent's count of hired labour. I find that between 13–16% of individuals in the samples said they provided free exchange labour to, on average, 4–5 other households. This indicates exchange labour is prevalent. More detailed data tracking households that received exchange labour,

or the number of person-days exchanged, is lacking.

5.3. Summary statistics

Table 1 presents summary statistics of the data for each year. Only households tracked between survey waves are included so that the baseline characteristics represent the samples used in the empirical tests. Tigray households are excluded in both waves of Panel II.

The most important statistic of Table 1 lies in the first row: in both panels the average household has an endowment of roughly 3 working age adults at baseline (with significant variation across households). This is one measure that validates the comparability between panels for this investigation. We also see in 2021, average household labour endowment is considerably larger than at baseline, which is suggestive of Proposition I.

The pattern is robust to different measures of household labour endowment that exclude children. Average shares of demographic groups in the household are balanced across panels. Farm characteristics are reported in the table. In baseline years the average farm cultivates roughly 6 acres of land, of which around two-thirds are owned. What remains consistent across the panels is that the vast majority (80–90 percent) of labour is provided by family members. There is an active labour market with 35% of households hiring labour on their farms.

5.4. Shock to labour endowment

There are two additional features necessary in the data to motivate the paper: exogenous variation of household labour endowment is required for identification and a large (positive) shock to household labour supply from Covid-19 is the premise for Proposition I.

Table 2 reports average changes in household labour endowment between the baseline and follow-up years. The statistics are highly consistent with Proposition I that Covid-19 caused a positive shock to labour supply. Average ΔE in 2018–2021 is much larger and positive (0.64) compared to in 2011–2015 (0.09).

All the flows of working-age individuals in Panel II compared to Panel I suggest that Covid-19 caused an expansion in the average household's labour endowment: average move-outs of the household are notably attenuated in Panel II; average move-ins rise in Panel II—possibly consistent with return migration; and the labour endowment boost from retaining ageing children is markedly stronger in Panel II. There is identification of the key regressor as Table 2 shows that 75% of households in Panel I and 70% in Panel II experienced any net ΔE from one survey to the next. I argue the variation is primarily exogenous, especially in Panel II because of the nature of Covid-19 being an extreme shock.

6. Empirical Results

6.1. Elasticity of farm labour demand to household size

This section presents the estimated results for the model of demand for farm labour. Table 3 presents the estimates of β_1 which, under the null of separation, should be zero.

6.1.1. Cross-section model and first differences model

The first column in Table 3 reports estimates of β_1 pooling all households and treating the sample as a cross-section as in specification (13), separately for each panel. In both panels, separability is strongly rejected in the cross-section. The elasticity of farm labour demand to household labour endowment is around 0.4.

---

[11] The household labour questionnaire asks the total number of weeks each household member worked, the average days per week and the average hours per day. The hired labour questionnaire asks the total number of men/women/children hired and the total number of days they were hired. These units are aggregated to create person-day variables for hired and household labour.

Table 1. Summary statistics

Variable	Panel I: Pre Shock		Panel II: Shock	
	2011	**2015**	**2018**	**2021**
Household characteristics				
Labour endowment, with kids	3.02 (1.41)	3.03 (1.45)	2.91 (1.30)	3.24 (1.53)
Labour endowment, no kids	2.72 (1.27)	2.72 (1.32)	2.65 (1.17)	2.95 (1.40)
Age of head (years)	44.47 (15.21)	47.55 (15.41)	44.33 (14.98)	46.54 (14.72)
Education of head (years)	1.57 (2.82)	1.84 (3.21)	2.28 (3.60)	2.53 (3.86)
Months away (hh average)	0.16 (0.49)	0.11 (0.48)	0.34 (0.96)	0.31 (0.81)
Share of prime age males	0.23 (0.17)	0.24 (0.19)	0.26 (0.19)	0.26 (0.18)
Share of prime age females	0.26 (0.16)	0.27 (0.19)	0.28 (0.17)	0.28 (0.18)
Share of elderly males	0.03 (0.10)	0.04 (0.13)	0.03 (0.11)	0.03 (0.10)
Share of elderly females	0.03 (0.12)	0.05 (0.18)	0.04 (0.14)	0.04 (0.14)
Share of children	0.43 (0.23)	0.48 (0.33)	0.40 (0.24)	0.45 (0.34)
Farm characteristics				
Land cultivated (acres)	6.10 (34.02)	5.11 (36.63)	6.03 (38.79)	9.19 (63.84)
Land owned (acres)	4.22 (17.89)	3.61 (16.18)	4.00 (18.58)	3.15 (14.03)
Farm demand characteristics (person-days)				
Total Labour Demand	246.92 (513.01)	230.29 (752.19)	159.04 (194.94)	130.84 (158.68)
Family supplied labour	223.76 (427.08)	179.70 (232.77)	134.21 (160.41)	110.70 (138.49)
Hired labour	23.16 (268.28)	50.59 (702.58)	24.83 (76.67)	20.13 (63.96)
Labour demanded for…				
Cultivation	175.89 (469.02)	141.92 (661.19)	100.90 (143.13)	66.81 (92.23)
of which hired	17.19 (264.07)	37.80 (630.14)	14.94 (54.54)	8.57 (33.70)
Harvest	71.03 (107.86)	88.37 (253.72)	58.14 (80.22)	64.03 (91.65)
of which hired	5.97 (34.03)	12.79 (223.10)	9.89 (42.23)	11.57 (47.94)
Observations	2777	3130	1688	1612

Notes: Standard deviations are in parentheses. "Labour endowment with kids" uses adult equivalence scale for children aged 11–15. "Prime" demographic groups are those aged 15–60. Total labour demand is the sum of hired and family labour in both cultivation and harvest seasons.

Table 2. Inter-annual changes in number of members and labour endowment

Changes between waves:	Panel I: Pre Shock 2011 & 2015	Panel II: Shock 2018 & 2021
Δ Number of members	−0.299	0.590
Δ Labour endowment (E)	0.091	0.642
ΔE from ageing children	−0.114	0.471
ΔE from move-ins	0.185	0.281
ΔE from move-outs	−0.628	−0.110
Any net Δ in E (=1)	0.753	0.704
Increase in E (=1)	0.455	0.650
Decrease in E (=1)	0.298	0.054

Notes: All entries are household-level averages that phase ageing children into the labour endowment except for the first row which counts each member as one person. Children aged 11–14 count as (Age −10) × 0.2 in the labour endowment. Δ Labour endowment (E) is the sum of the three categories above.

The key results are reported in Column (2) of Table 3 for the main specification of interest (14). In the counterfactual pre-shock sample, separation is rejected. This suggests market failures were already operating before the effect of the Covid shock on households. In the Covid-19 panel, separation is also rejected. The Covid sample estimate is larger and more significant than the counterfactual estimate, roughly doubling from 0.33 to 0.66. Households that experienced the positive shock to their labour supply increased labour inputs on their farm more than households that did not experience the shock; increased farm labour utilisation helped partly absorb the larger household labour supply.

Comparing the cross-sectional and fixed effects estimates within each panel, in Panel I the elasticity falls by a quarter when accounting for time-invariant household factors. This indicates unobserved characteristics are important for explaining non-separation in the 2011–2015 panel, but cannot explain it alone as I still find evidence for non-separation when including household fixed effects. However, the cross-sectional bias in Panel II is in the opposite direction: the elasticity increases by almost half. I suggest this is because the large sweeping shock to labour endowment weakens the correlation between the unobserved household qualities and household size; there is now a stronger driver of household size.

6.1.2. Farm demand for family labour and hired labour

Columns (3) and (4) in Table 3 disaggregate the dependent variable into demand for family labour and hired labour, respectively. The insignificant estimated elasticities in column (4) are not evidence in favour of complete markets; column (2) is the key variable of interest for the separation test. Rather, statistically insignificant elasticities of hired labour demand suggest the farm is not substituting efficiently in response to increases in its family size. Furthermore, farms expand family labour on the farm in response to larger household labour endowments, and do so significantly more in 2018–2021 (looking at Column (3), elasticities are 0.38 vs 0.70). This is consistent with households operating under labour market failures and absorbing the shock by expanding family labour utilisation on the farm (Proposition II).

Columns (3) and (4) indicate that households do not shed hired labour when family labour rises; if anything, they hold on to (or even slightly increase) hired labour alongside their own labour. This could be due to informal insurance agreements with other households in the village to help each other out with financial hardship from Covid-19. There is evidence in the literature that social norms play a potentially big role in governing rural labour markets (Breza et al. 2019). This motivates the intuition that "family comes first": when households get more internal labour, they use it—rather than shedding external workers—and expand total farm production to absorb that labour.

6.1.3. Confounding time-varying shocks

First-differencing eliminates time-invariant household heterogeneity, but time-varying shocks could still confound the estimates. A shock that may confound the estimates is government support to households. Ethiopia's Productive Safety Net Program (PSNP) is a large government program that provides food and cash transfers and remunerated public work to help relieve food insecurity of the rural poor. I find the results in Table 3 are robust to the exclusion of households that participated in the PSNP. Another concern is that Covid-19 impacted ΔFarm Labour directly, not just through ΔE, for instance by lowering crop prices. In practice, I attempt to mitigate this concern by controlling for cultivated acreage and household demographic composition.

6.2. Did the estimated elasticity increase post-Covid?

Taking stock, I have found that Covid-19 increased labour supply in the household which is consistent with Proposition I. Next, I found evidence for Proposition II that households were unable to fully absorb the shock, which resulted in household characteristics leaking into input allocation decisions. This raises the important question of whether the labour supply shock led to greater consequences from market failures, statistically speaking. Therefore, Proposition III is focused on the intensive margin of market failures.

In Table 4 I pooled both panels of households into one sample and interacted all right-hand-side regressors with a Post-Covid dummy variable equal to 1 if the household was surveyed in the 2018–2021 panel. This allows a formal test of whether the separation is statistically weaker after the Covid shock or if the difference in estimates is due to sampling fluctuation.

Although the estimate of the interaction term (0.327) is economically large—implying the elasticity roughly doubled in the Covid panel—it is measured imprecisely (standard error of 0.25) and fails to reach statistical significance. Therefore, while Table 3's estimates suggested a substantially bigger farm input response to ΔE during Covid, Table 4 shows the difference is not statistically distinguishable from sampling variation. I conclude that the evidence is qualitatively consistent but quantitatively inconsistent with Proposition III.

Between the follow-up survey of Panel I to the follow-up of Panel II (2015 to 2021), Ethiopia's rural markets likely became increasingly integrated and evolved, following the general trend of agricultural markets deepening over the last quarter century (Suri & Udry 2022). Against this backdrop of deepening markets, the persistent finding of non-separation—especially a potential intensification during the Covid-19 period—becomes more notable. It suggests that Covid-19 likely re-introduced or exacerbated imperfections that had otherwise been diminishing over time.

7. Heterogeneity in household sophistication

How well the system of agricultural labour markets absorb the shock to household labour endowment may depend on how 'sophisticated' the household is. Sophisticated households in the rural setting are households with higher wealth or consumption

Table 4. Estimation of Covid-19 impact on labour endowment elasticity

	Δ Farm labour
ΔE	0.329**
	(0.131)
ΔE× Covid-panel	0.327
	(0.252)
Covid-panel	0.131
	(0.141)
Controls	Yes
Observations	4,350

Notes: Standard errors in parentheses, clustered at the EA-wave level. Covid takes value 1 for waves 4–5 and 0 otherwise. Controls include Δ ln(acres) and demographic shares. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

levels, higher human capital, participation in the wage market, formal salaried jobs or more land. The Covid-19 shock may have affected them differently because these households could better smooth the consequences of a labour supply shock. I put forward Proposition IV that sophisticated households absorb the Covid shock better than non-sophisticated households by placing more family members into the formal wage market.

Table 5 presents results for the heterogeneous responses by households in the wage market along two measures of sophistication: previous experience in the wage market and higher human capital. The statistically significant estimates of the first interaction term (2nd row of each panel) show that more sophisticated households have more household members working in the wage market. Previous market employment is more important than having an educated household head for getting members into formal work.

The "experience" benefit that sophisticated households exhibit disappears with Covid. The triple interaction terms are negative for both measures of sophistication and larger in magnitude than the "experience" benefits. This suggests the Covid shock was structural and affected households across the spectrum, even sophisticated households with pre-established networks in the market. Taken together, the estimates of Table 5 are inconsistent with Proposition IV. Sophisticated households were not better able to absorb the shock, suggesting all households faced market

Table 5. Heterogeneous sophistication of the household

	Δ Number of members employed in wage market
A. Employment in wage market at baseline	
ΔE	−0.005
	(0.013)
ΔE× Had market employment	0.351***
	(0.132)
ΔE× Covid-panel	0.161
	(0.125)
ΔE× Had market empl. × Covid-panel	−0.445*
	(0.246)
B. Educated household head at baseline	
ΔE	−0.005
	(0.021)
ΔE× Educated	0.116**
	(0.052)
ΔE× Covid-panel	0.243*
	(0.132)
ΔE× Educated × Covid-panel	−0.226
	(0.175)

Notes: Pooled sample of both panels with a binary indicator "Covid-panel" equal to 1 if the household was part of the 2018–21 panel. Standard errors in parentheses, clustered at the enumeration area for each wave. All regressions include controls for changes in log of cultivated acreage and changes in demographic shares. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. Effect of household labour endowment on farm labour demand

Dep. variable:	A. Pooled	B. Household FE in first differences		
	Farm labour (1)	Δ Farm labour (2)	Δ Family labour (3)	Δ Hired labour (4)
Panel I: Pre Shock				
Labour endow. (E)	0.412***			
	(0.057)			
ΔE		0.329**	0.381***	−0.188
		(0.131)	(0.133)	(0.181)
N. Households	5,561	2,777	2,777	2,777
Panel II: Shock				
Labour endow. (E)	0.463***			
	(0.083)			
ΔE		0.657***	0.700***	0.254
		(0.215)	(0.210)	(0.368)
N. Households	3,262	1,574	1,574	1,574

Notes: Standard errors in parentheses, clustered at the enumeration area for each wave. All regressions include controls for changes in log of cultivated acreage, changes in demographic shares, and wave fixed effects. Labour endowment E and all dependent variables are measured in logs. All dependent variables are measured in person-days. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

failures in the same degree and the Covid-19 shock was structural. I find this result is robust to using other proxies for household sophistication such as household consumption expenditure per capita and PSNP participation (a proxy for lower 'sophistication'), as shown in Addendum A2.

8. Conclusion

The separation test is a well-established procedure for testing the completeness of rural markets (Dillon et al. 2019). I use it as a framework to analyse how the system of rural markets and agricultural households absorb a shock. I find that rural Ethiopian farm households (i) felt the impact of Covid-19 through an increase in their household size; (ii) were operating in a set of markets exhibiting market failures; (iii) this meant the additional family labour supply could not be fully absorbed by labour markets; (iv) households responded by increasing labour utilisation on the farm, representing a failure to separate production input allocations from consumption characteristics; (v) this effect was structural and affected households at all levels of sophistication.

Separation is a 'portmanteau test' for completeness in at least two factor or output markets (LaFave & Thomas 2016). While it is not technically possible to identify which markets exhibit failures, I take the results as suggestive of imperfections in labour markets because of the nature of Covid-19 as a large unprecedented disruption to labour supply. Furthermore, even if labour markets are complete but separation fails because of other market failures (such as credit), this still represents misallocation on the farm which is important in itself. It highlights that the set of rural labour market institutions are inadequate to help buffer the misallocation and efficiency consequences from a large shock to agricultural households. As these households tend to rely on agriculture for livelihoods and consumption needs, this outcome is relevant for their day-to-day wellbeing.

While the Covid-19 shock represented an increase in the labour supply in agricultural rural markets—potentially causing a temporary labour surplus—it is important to distinguish which constraints are afflicting the labour market. Further research may explore whether it is demand constraints including lack of available jobs, hiring frictions or discrimination; supply constraints of workers not having the right skill-set or employability; or intermediation constraints including matching frictions, wage rigidity, information asymmetries, skills signalling difficulties or lack of trust. Effective policy prescription requires an understanding which constraints are operating in the market. It is also relevant to policy-makers to understand whether the misallocation consequences from the Covid-19 shock persist today.

Suri and Udry (2022) express that the literature is limited on the role that poorly functioning labour markets play in hampering agricultural productivity in Africa. I hope that my paper contributes, in some small way, to address these shortcomings and fits into the broader agenda that urges more research on labour market failures in rural settings and how they reverberate to agricultural productivity in developing economies.

■ Addendum

A1: Excluding Tigray in Panel I

Table 6. Effect of household labour endowment on farm labour demand
(excluding Tigray in Panel I)

Dep. variable:	A. Pooled	B. Household FE in first differences		
	Farm labour (1)	Δ Farm labour (2)	Δ Family labour (3)	Δ Hired labour (4)
Panel I: Pre Shock				
Labour endow. (E)	0.412***			
	(0.059)			
ΔE		0.306**	0.366***	−0.164
		(0.140)	(0.143)	(0.197)
N. Households	4,938	2,466	2,466	2,466
Panel II: Shock				
Labour endow. (E)	0.463***			
	(0.083)			
ΔE		0.657***	0.700***	0.254
		(0.215)	(0.210)	(0.368)
N. Households	3,262	1,574	1,574	1,574

Notes: Standard errors in parentheses, clustered at the enumeration area for each wave. All regressions include controls for changes in log of cultivated acreage, changes in demographic shares, and wave fixed effects. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A2: Additional heterogeneity measures

Table 7. Heterogeneous sophistication of the household (additional measures)

	Δ Farm labour
A. Expenditure per capita at baseline	
ΔE	0.269*
	(0.161)
$\Delta E \times$ High expenditure pc	0.052
	(0.250)
$\Delta E \times$ Covid-panel	0.161
	(0.125)
$\Delta E \times$ High expenditure pc $\times$ Covid-panel	−0.189
	(0.805) *Notes:* Pooled sample
B. PSNP participation at baseline	
ΔE	0.351**
	(0.142)
$\Delta E \times$ Participated in PSNP	−0.199
	(0.290)
$\Delta E \times$ Covid-panel	0.294*
	(0.262)
$\Delta E \times$ Participated in PSNP $\times$ Covid-panel	0.301
	(0.556)

of both panels with a binary indicator "Covid-panel" equal to 1 if the household was part of the 2018–21 panel. Standard errors in parentheses, clustered at the enumeration area for each wave. All regressions include controls for changes in log of cultivated acreage and changes in demographic shares. Significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

■ References

[1] Aspen, H. (1993). Competition and Co-operation. North Ethiopian Peasant Households and their Resource base.

[2] Ayele, Z. & Fessha, Y. (2021). COVID-19 in Ethiopia: A Year in Review.

[3] Bardhan, P. K. & Udry, C. (1999). *Development Microeconomics*. Oxford University Press.

[4] Benjamin, D. (1992). Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models. *Econometrica* 60(2), 287.

[5] Breza, E., Kaur, S. & Krishnaswamy, N. (2019a). Scabs: The social suppression of labour supply.

[6] Breza, E., Kaur, S. & Krishnaswamy, N. (2019b). Social Norms as a Determinant of Aggregate Labor Supply. NBER Working Paper No. w25880.

[7] Breza, E., Kaur, S. & Shamdasani, Y. (2021). Labor Rationing. *American Economic Review*.

[8] Card, D. (1990). Unexpected Inflation, Real Wages, and Employment Determination in Union Contracts. *American Economic Review* 80(4), 669–688.

[9] Chayanov, A. (1991). *The Theory of Peasant Co-operatives*. I.B. Tauris & Co Ltd.

[10] Dillon, B. & Barrett, C. B. (2017). Agricultural factor markets in Sub-Saharan Africa: An updated view with formal tests for market failure. *Food Policy* 67, 64–77.

[11] Dillon, B., Brummund, P. & Mwabu, G. (2019). Asymmetric non-separation and rural labor markets. *Journal of Development Economics* 139, 78–96.

[12] Duflo, E. & Udry, C. (2004). Intrahousehold Resource Allocation in Cote d'Ivoire: Social Norms, Separate Accounts and Consumption Choices. NBER Working Paper No. w10498.

[13] Feder, G., Just, R. E. & Zilberman, D. (1985). Adoption of Agricultural Innovations in Developing Countries: A Survey. *Economic Development and Cultural Change* 33(2), 255–298.

[14] Ford, K., Porter, C. & Pankhurst, A. (2021). A Lost Year of Learning for Girls in Ethiopia: Evidence from the Young Lives At Work Covid-19 Phone Survey.

[15] Goldstein, M. & Udry, C. (2008). The Profits of Power: Land Rights and Agricultural Investment in Ghana. *Journal of Political Economy* 116(6), 981–1022.

[16] Gollin, D. & Udry, C. (2021). Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture.

[17] Harris, D., Baird, S., Ford, K., Hirvonen, K., Jones, N., Kassa, M., Meyer, C., Pankhurst, A., Wieser, C. & Woldehanna, T. (2021). The Impact of COVID-19 in Ethiopia: Policy Brief.

[18] Harris, J. & Todaro, M. (1970). Migration, Unemployment and Development: A Two-Sector Analysis.

[19] Jones, M., Kondylis, F., Loeser, J. & Magruder, J. (2022). Factor Market Failures and the Adoption of Irrigation in Rwanda. *American Economic Review* 112(7), 2316–2352.

[20] Kaur, S. (2019). Nominal Wage Rigidity in Village Labor Markets. *American Economic Review* 109(10), 3585–3616.

[21] Kaur, S. (2024). Labor Markets in Developing Countries. Nobel Symposium on Development Economics.

[22] Kilic, T., Palacios-López, A. & Goldstein, M. (2015). Caught in a Productivity Trap: A Distributional Perspective on Gender Differences in Malawian Agriculture. *World Development* 70, 416–463.

[23] LaFave, D. & Thomas, D. (2016). Farms, Families, and Markets: New Evidence on Completeness of Markets in Agricultural Settings. *Econometrica* 84(5).

[24] Nolte, K., Sipangule, K. & Wendt, N. (2022). Agricultural households in times of crisis. The COVID-19 pandemic, livelihoods and land-use decisions. *Journal of Land Use Science* 17(1), 134–160.

[25] Restuccia, D., Yang, D. T. & Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics* 55(2), 234–250.

[26] Schultz, T. W. (1964). Changing Relevance of Agricultural Economics. *Journal of Farm Economics* 46(5), 1004.

[27] Sen, A. (1962). An Aspect of Indian Agriculture. *Economic Weekly*.

[28] Singh, I., Squire, L. & Strauss, J. (1986). A Survey of Agricultural Household Models: Recent Findings and Policy Implications.

[29] Suri, T. & Udry, C. (2022). Agricultural Technology in Africa. *Journal of Economic Perspectives* 36(1), 33–56.

[30] Udry, C. (1996). Gender, Agricultural Production, and the Theory of the Household.

[31] World Development Indicators. The World Bank.

The Path Not Taken: Could State Contingent Debt Instruments Have Prevented Zambia's 2020 Debt Default?

Njavwa Sanga

University of Oxford

Abstract

This dissertation investigates whether State-Contingent Debt Instruments (SCDIs) could have prevented Zambia's 2020 sovereign default and reduced the adjustment costs from delayed debt restructuring. Using a calibrated dynamic general equilibrium (DGE) model of a small open economy calibrated to Zambia, the study compares the welfare outcomes of standard debt arrangements against SCDI alternatives under adverse climate shocks proxied by Total Factor Productivity (TFP) shocks. Results highlight the potential of SCDIs to preserve fiscal space, mitigate debt distress, and improve economic resilience in the face of external climate shocks.

Corresponding author: Njavwa Sanga *E-mail address:* nsanga1727@gmail.com

1. Introduction

Climate shocks are threatening Zambia's debt sustainability. Zambia is not unique in this regard. Small open economies, particularly climate-vulnerable, developing economies, are disproportionately affected by the ongoing climate crisis.[1] In a particularly recent era of rapidly declining foreign aid and reduced global liquidity, there is mounting pressure on the domestic budgets of these economies to close fiscal gaps and do so as climate risks grow. Climate shocks impose significant challenges for public finance to stabilise the economy and protect firms and households' consumption.

In traditional economic theory, debt is a tool used by agents to smooth consumption across time. For a government, this intertemporal 'trading' of consumption is achieved by borrowing from and lending to the rest of the world using debt and the current account of the balance of payments. In this way, the open economy can 'de-link' current consumption from current production which is going to be particularly important when production suffers an adverse shock. Crucially, consumption need not decline with it.

The issue is that conventional debt contracts are unattractive in the sense that they pass all debt service risk onto the debtor. This leads us to a broader public discussion: are there ways to provide borrowing countries with greater financing flexibility when subject to an adverse shock? An increasingly popular instrument being considered in the global financial architecture is state-contingent debt instruments (SCDIs), which are discussed more in Section 2. These instruments tie debt service payment obligations to the prevailing state of the world. In effect, the debt service burden falls during a 'bad' state, allowing for countercyclical policy space when a country is most likely to struggle to meet its obligatory spending, and rises again during a 'good' state.

This paper considers the use of SCDIs, using a counterfactual analysis, to investigate whether Zambia might have avoided sovereign debt default in 2020 had it had access to them during the 2019 drought. The model used throughout the analysis (an adaptation of Buffie et al., 2012) is designed to provide an integrated assessment of the welfare gains available under various government borrowing practices, while considering the downsides of conventional borrowing practices. It is to support debt sustainability analysis in the context of systematically incorporated TFP shocks. The justification for this line of enquiry is that as climate risks intensify in severity and frequency, their impact on output and the economy's productive capacity could translate into lower government revenue and/or the need for higher spending. It is also likely that climate change will affect the structure of economic systems in ways that are difficult to incorporate into standard debt sustainability analysis (DSA) scenarios.[2] Hence, for a climate-vulnerable, small open economy such as Zambia, it remains pertinent to explore how the design and availability of SCDIs might promote much-needed stabilisation and debt sustainability.

As of March 2025, 9 of the poorest countries in the world were considered to be in external 'debt distress' currently, 26 are at 'high risk' according to the IMF's (2015) DSA. This is defined by the International Monetary Fund (2025) as 'facing severe difficulties in servicing their debt obligations'. Many of these countries' external debt has been deemed unsustainable: they can no longer meet all their current and future financial obligations without recourse to exceptional financing, adjustment measures, and/or

---

[1] Center for Global Development, "The Socioeconomic Impact of Climate Change in Developing Countries in the Next Decades: A Review," Working Paper 681, 2024.

[2] European Central Bank, "The Climate Change Challenge and Fiscal Instruments and Policies in the EU," Occasional Paper 315, 2022.

default. The World Bank (2024a) reported that in 2023, developing countries spent an unprecedented $1.4 trillion on external debt servicing due to steep interest costs. In 2025, interest payments for 56 developing countries exceeded 10% of government revenue on, and for 17 countries, more than 20% of revenue.[3] Prior to default, high debt service payments constituted a large share of Zambia's statutory expenditure as it aimed to reduce its debt overhang. This eventually exceeded spending on critical public services.[4]

External sovereign debt, however, can be and remains a powerful instrument for economic policy. It allows governments to smooth intertemporal consumption by decoupling current expenditures from current revenues, thereby enabling counter-cyclical fiscal policy and long-term investment planning. By altering the feasible trajectory of public spending, debt issuance can enhance aggregate welfare when risks are appropriately shared between debtor and creditor. However, this relies critically on the expectation that debt will be serviced fully and punctually. The risk of debt distress arises when these expectations falter.

For many developing economies, debt distress arises more as a consequence of illiquidity than insolvency.[5] This is exacerbated by the so-called 'original sin' problem: much of their debt is denominated in foreign currencies, limiting the ability to respond via monetary expansion. Unlike advanced economies, they cannot simply print domestic currency to finance short-term gaps. The result is that debt crises in these contexts are frequently liquidity-driven, with implications for how creditors respond and how governments formulate appropriate policy responses.

The aim of this paper is thus to explore the economic outcomes of alternative debt structures on Zambia's debt sustainability compared to conventional borrowing practices by employing a counterfactual analysis. Following a brief review of where this paper fits in the literature, Section 3 provides context on the Zambian economy and its path to default. Section 4 lays out the modelling approach before Section 5 provides context for the policy simulations and Section 6 details specific scenarios. Discussion of the results and policy relevance are found in Section 7. The final section concludes. Details of the model and the baseline calibration, as well as the various policy simulations, are described in a set of appendices.

2. Related Literature

This paper builds on three primary strands of literature. First, it draws from the class of structural macroeconomic models used to analyse public investment, fiscal constraints, and debt sustainability in low-income countries (LICs). Buffie et al. (2012) present a dynamic general equilibrium (DGE) model tailored for fiscal policy analysis in small open economies with limited access to capital markets.[6] This framework was further refined by Adam et al. (2016) to address the specific institutional and economic constraints facing LICs, incorporating public capital accumulation, concessional and non-concessional borrowing, and fiscal adjustment mechanisms.[7] These models underpin the calibration and structure of the model used in this dissertation.

Second, the analysis contributes to the growing body of literature on SCDIs, which aim to improve external sovereign debt sustainability by linking debt service to the prevailing state of the world. Hatchondo et al. (2016) argue that SCDIs can reduce default probabilities by aligning payment obligations with the borrower's capacity to pay.[8] The IMF (2017) has provided practical proposals for sovereign SCDI design, highlighting instruments indexed to GDP, commodity prices, or disaster-related events.[9] The core premise is that such instruments provide automatic fiscal relief during economic downturns or exogenous shocks, facilitating countercyclical fiscal policy and reducing costs borne of delayed adjustment.

The case of Barbados provides early evidence of the feasibility and potential of SCDIs in practice. Following its 2018–2019 debt restructuring, Barbados included natural disaster clauses in most of its newly issued debt, both domestic and external. The IMF's analysis of these clauses highlighted that amortisation could be deferred for up to two years after a climate event, and triggers were tied to the Caribbean Catastrophe Risk Insurance Facility (CCRIF) payouts. Such arrangements created automatic liquidity support during periods of fiscal stress and demonstrate how SCDIs can be operationalised within the global financial architecture.[10]

SCDIs offer several benefits. They might preserve debt stock accounting by reprofiling rather than reducing principal, maintain investor confidence through pre-defined, rule-based mechanisms, and offer a form of climate justice by giving risk-sharing to countries disproportionately affected by exogenous climate shocks that they contributed minimally to. As argued in this paper, SCDIs can also reduce the likelihood and cost of sovereign default by mitigating the "debt overhang" effect—where expected future debt burdens discourage investment and consumption. This echoes early work on sovereign default incentives, including Cohen and Sachs (1985), who explore the dynamics of debt sustainability and risk premia for sovereign borrowing.[11]

Despite their promise, SCDIs face implementation challenges. First-mover moral hazard concerns by the creditor arise under the belief that debtor countries have reduced incentives to invest in resilience-building. Similarly, it is contended that the complexity of contract design, pricing, and the lack of established legal frameworks remain barriers to broader adoption. Nevertheless, proposals such as the G20 Common Framework for Debt Treatments highlight growing international interest in more flexible and adaptive debt instruments, particularly for climate-vulnerable countries like Zambia.

Finally, this research aligns with a third strand of literature that examines macro-fiscal risks from climate change. Recent studies such as UNEP and NIESR (2022) model the effects of climate-induced volatility—such as droughts, floods, and commodity price shocks—through sector-specific TFP shocks.[12]

This paper lies at the intersection of these three strands of research. By interacting climate vulnerability with sovereign debt dynamics, the following analysis contributes to the literature by conducting a structured assessment of adverse TFP shocks (as

[3] UNCTAD, "A World of Debt: A Growing Burden to Global Prosperity," 2024.

[4] Finance for Development Lab and ZIPAR, "The Road to Zambia's 2020 Sovereign Debt Default," 2023.

[5] O. Bjerkholt, "New Approaches to Debt Relief and Debt Sustainability in LDCs," CDP Background Paper No. 5, United Nations Department of Economic and Social Affairs, 2004.

[6] E. F. Buffie, A. Berg, C. Pattillo, R. Portillo, and L.-F. Zanna, "Public Investment, Growth, and Debt Sustainability: Putting Together the Pieces," *IMF Working Paper No. 12/144*, 2012.

[7] C. Adam, D. Bevan, and T. Ohlenburg, "Public Investment, Fiscal Rules, and Growth in Low-Income Countries," IGC Working Paper S-43302-UGA-1, 2016.

[8] J. C. Hatchondo, L. Martinez, and C. Sosa-Padilla, "Sovereign Defaults and Debt Sustainability: The Role of State-Contingent Debt," *IMF Economic Review* 64, no. 4 (2016): 641–664.

[9] International Monetary Fund, "State-Contingent Debt Instruments for Sovereigns," IMF Policy Paper PP/17/10, 2017.

[10] M. Anthony, G. Impavido, and B. van Selm, "Barbados' 2018–19 Sovereign Debt Restructuring—A Sea Change?" *IMF Working Paper WP/20/34*, 2020.

[11] D. Cohen and J. D. Sachs, "Growth and External Debt under Risk of Debt Repudiation," NBER Working Paper w1703, 1985.

[12] UNEP and NIESR, "Economic Impacts of Climate Change: Exploring Short-Term Climate-Related Shocks for Financial Actors with Macroeconomic Models," 2022.

proxies for climate shocks), under different debt arrangements.

3. Stylised Facts: Zambia's Debt Dynamics

Zambia's sovereign debt trajectory has been shaped by episodes of debt relief, rapid market re-entry, external shocks, and structural vulnerabilities. In 2005, Zambia received significant debt relief under the Highly Indebted Poor Countries Initiative (HIPC) and Multilateral Debt Relief Initiative (MDRI), launched by the IMF & World Bank, that reduced their debt stock by around $4.5 billion, from $6.9 billion to $2.3 billion.[13] These initiatives, as stated, aimed to 'bring poor countries' debt burden to a sustainable level', thereby eliminating "debt overhang" and reducing their external debt stocks. This engendered a significant improvement in Zambia's credit rating, stimulating five years of macroeconomic stability, fiscal discipline, and high GDP growth. In 2011, this strong macroeconomic performance prompted the World Bank to upgrade its classification of Zambia from a low-income country (LIC) to a lower-middle-income country (LMIC).[14]

Between 2010 and 2015, Zambia's new administration capitalised on favourable external conditions. It issued its first Eurobond in global capital markets in 2012—a 10-year note worth $750 million at a 5.65% coupon rate, with the intent to invest in public infrastructure.[15] This was quickly followed by further issuances: $1 billion in 2014 at 8.5%, and $1.25 billion in 2015 at 8.97%, bringing Zambia's Eurobond exposure to $3 billion within three years.[16] Investor appetite reflected optimism that Zambia's 8% growth trend since 2004 would continue. Zambia's 2014 return to the Eurobond market coincided with the slowing of the commodity 'supercycle',[17] and they faced a credit rating downgrade (B+ to B) and thus a higher coupon rate of 8.5% on their latest loan. Fiscal policy loosened, and increasing reliance on non-concessional debt from private creditors weakened the country's economic resilience.

Since 2015, Zambia's economic vulnerabilities have become more pronounced. The 2015 and 2019 El Niño-linked droughts severely affected the country's hydro-dependent electricity sector.[18] Daily power outages hindered industrial activity and triggered the government to increase energy imports and subsidies. As a result, GDP growth slowed to 3.4% in 2015, and inflationary pressures surged. A large portion of Zambia's external sovereign debt was US dollar-denominated. Hence, the depreciation of the Kwacha by 40% contributed 10 percentage points to the GDP to the debt stock.

By 2016, Zambia's economic outlook had deteriorated. Public debt doubled from 60.6% to 120% of GDP between 2016 and 2019, driven by a combination of sustained non-concessional borrowing, weak copper export revenues, and rising interest costs, which had risen to 6% of GDP by 2019. This reduced fiscal space hampered social spending: by 2019, Zambia was allocating 1/3 of GDP to

obligatory debt service payments compared to 8.8% to essential public services like health and education.[19]

A joint IMF–World Bank Debt Sustainability Analysis (DSA) in 2019 formally assessed Zambia's external debt as unsustainable and Zambia's credit rating was downgraded to CCC, signalling high default risk. The debt burden had quadrupled since 2014, and Zambia began to fall into arrears. In 2020, it defaulted on a $42.5 million Eurobond coupon payment, becoming the first African country to default on commercial debt during the COVID-19 era.

Zambia's 2020 sovereign default triggered a prolonged and uncertain debt restructuring process under the G20 Common Framework (CF). President Hichilema, elected in 2021 on a mandate of fiscal prudence and growth, in 2022 began the process of debt restructuring with its creditors under the CF, supported by the IMF. Despite the urgency of relief, negotiations persisted for over three years, creating economic uncertainty and fiscal paralysis. According to the World Bank (2023), Zambia's external sovereign debt is increasingly fragmented: 49% is held by private creditors, 27% by China, 10% by the World Bank, 9% by other multilateral institutions, and 5% by other bilateral lenders. This diversification of creditor types, particularly of private and non-Paris Club actors, complicated restructuring efforts. These delays have imposed high adjustment costs: capital inflows have slowed, borrowing costs have risen, and pressure to implement fiscal consolidation has intensified. These challenges have coincided with climate volatility and weaker commodity prices, further straining Zambia's macroeconomic stability.

Zambia is currently in a post-default phase where it is continuing to negotiate and restructure its debt with creditors.

4. The Modelling Approach

The model employed in this paper is an adaptation of the frameworks originally developed by Buffie et al. (2012) and subsequently utilised by Adam, Bevan, and Ohlenburg (2016). Buffie et al. (2012) create a macroeconomic model to evaluate the impact of public investment surges in developing economies. Adam, Bevan, and Ohlenburg (2016) extend this framework to explicitly address public investment efficiency and macroeconomic management in low-income countries. The adaptation used herein maintains the essential structure:[20] it is a standard two-sector model comprising exportable and non-tradeable goods in a real, small open economy setting. Production within each sector depends on inputs of public capital, private capital, and labour, undertaken by competitive, price-taking firms. The model exhibits constant returns to private factors individually and increasing returns when considering all three inputs collectively. Additionally, it assumes steady, uniform trend growth in productivity across both sectors.

To model Zambia, it is extended to include a copper sector that is majority-owned by foreign investors, and employs domestic labour but remits profits net of royalty. It is assumed that all copper production is exported, and therefore, there is no domestic consumption. Instead, it is a major source of export revenues and hence foreign currency, and the government receives royalty payments. Public debt may be domestic, external concessional, or external non-concessional. Interest payable on the latter may include a risk premium that rises with the government's indebtedness. Some households have access to financial markets,

[13] IMF, "Zambia: Enhanced Initiative for Heavily Indebted Poor Countries—Completion Point Document," IMF Country Report 05/137, 2005.

[14] United Nations, "UN Common Country Analysis Update 2024: Zambia," 2024.

[15] M. Kessler, M. Mbewe, T. Humann, M. Kalikeka, I. Masilokwa, and S. Mwamba, "The Road to Zambia's 2020 Default," ZIPAR-FDL, 2023.

[16] C. Adam and A. Musonda, "Copper Mining in Zambia: From Collapse to Recovery," in *Plundered Nations? Successes and Failures in Natural Resource Extraction*, ed. P. Collier and A. J. Venables (Palgrave Macmillan, 2011), 89–114.

[17] A period, often more than 5 years, during which commodity prices trend above their long-term averages.

[18] F. Alfani, A. Arslan, N. McCarthy, R. Cavatassi, and N. Sitko, "Climate-Change Vulnerability in Rural Zambia: The Impact of an El Niño-Induced Shock on Income and Productivity," FAO Agricultural Development Economics Working Paper 19-02, 2019.

[19] Debt Justice, "Zambia Country Profile—Debt Data Portal," 2020.

[20] See Buffie et al. (2012) for a full set of model equations.

while others are credit-constrained and consume their current income. The former maximise an additive intertemporal utility function; for all households, instantaneous utility is a function of both consumption and leisure.

Instead of explicitly modelling an energy sector, climate-induced TFP shocks are differentiated across sectors to approximate the effect of climate-induced productivity declines. The copper sector faces severe TFP reductions due to its reliance on electricity (e.g., hydropower disruptions). Higher energy costs and supply chain disruptions affect the exportables sector indirectly. Finally, the tradable sector is least affected but still experiences spillovers. The copper sector is unique in the sense that it can be a source—or amplifier—of supply shocks. Mining production in Zambia is highly dependent on the reliable availability of electricity. The mining sector alone accounts for approximately 50% of national electricity consumption, and over 80% of Zambia's electricity generation comes from hydropower.[21] This makes the sector particularly vulnerable to climate-related disruptions such as droughts. This vulnerability was clearly demonstrated in 2015, when a severe El Niño-induced drought led to a 7% reduction in national power generation, contributing to a 17% fall in copper and cobalt output. While efforts have since been made to diversify the energy mix (including investments in solar), droughts continue to represent a risk to mining operations.

Concerning the output effects, climate change can affect both supply and demand. For the former, extreme weather events, the intensity and frequency of which are positively correlated with global warming, can reduce the endowments of some input factors (such as land) and destroy physical capital. Hours worked and labour productivity may also decrease in outdoor industries (such as construction) as a result of rising temperatures.[22]

The model enforces budget balance by design; therefore, sovereign default is defined operationally as a situation where the government faces either prohibitively high borrowing costs, a loss of market access, or the need to implement economically infeasible fiscal adjustments. This dynamic approach draws on the theoretical, two-period framework developed by Sachs and Cohen (1986), in which debt default risk arises under uncertainty in second-period income. The key condition governing repayment is the so-called no-default condition, shown in equation 1:

$$y_2 - (1 + r)d_1 \geq y_2(1 - \phi) \tag{1}$$

This stipulates that the borrower will repay if and only if consumption with repayment is at least as high as consumption under default, where ϕ represents the cost of defaulting. Default occurs when second-period income falls below a critical threshold: $y_2 \leq y_2^*$, where $y_2^* = \frac{(1+r_s)d_1}{\phi}$. Hence, the country defaults in 'bad times' and repays in 'good times'. The probability of default is given by:

$$\pi = \Pr(y_2 \leq y_2^*) = F\left(\frac{(1 + r_s)d_1}{\phi}\right) \tag{2}$$

where $F(\cdot)$ denotes the cumulative distribution function of y_2 and is rising in the borrowing cost, r_s.

Rising interest rates, surging debt ratios, or unsustainable tax hikes capture these dynamics, which align with real-world episodes of sovereign distress as defined by the IMF. Within this

framework, the role of SCDIs is assessed by their ability to mitigate these indicators of default risk and improve welfare outcomes relative to the baseline scenario of standard borrowing.

4.1. Scope of the model

Saving and investment are fundamentally inter-temporal concepts about shifting consumption and output through time. So, the correct way to analyse the current account is using dynamic (multi-period) models. The model is a real small open economy model, with three sectors: exportables, non-tradeables, and copper. Production in the first two sectors are a function of public and private capital and of labour, and is carried out by domestic firms that are competitive and price-taking. The majority (80%) of copper production is conducted by foreign private ownership of capital and land share so the sector uses minimal domestic factors. Public debt is disaggregated into domestic, external concessional, and external non-concessional. Interest costs payable on non-concessional debt may include a risk premium that rises with government indebtedness or following periods of debt restructuring to reflect a penalty on future borrowing. Households are characterised as 'Saving households' that have access to financial markets and are able to smooth consumption intertemporally, and 'Rationed households' that are credit constrained and consume income hand-to-mouth.

Taxes are levied on capital and labour incomes and on consumption, with rates adjusting dynamically to finance domestic debt and government spending. As a result, they serve as a proxy for fiscal burden on the population—if they rise under a given debt scenario, it signals fiscal strain or adjustment as the model attempts to plug the fiscal gap. Thus, comparing tax trajectories across different debt scenarios (e.g. standard borrowing vs. SCDI) gives a direct measure of the fiscal cost of each financing structure.

4.2. Limitations

The model is a calibrated policy simulation framework designed to explore the properties of alternative scenarios and so cannot be used for forecasting. It is not a monetary model, so it does not examine inflation or nominal exchange rate dynamics. It is a value-added model, meaning that it cannot examine intermediate input linkages, and it is an aggregated macroeconomic model, so it cannot be used to examine the sectoral composition of public investment. Valid concerns of issues such as tax leakage and congestion effects are assumed away in favour of the assumption of full efficiency of public spending. Though less tractable in reality, this is well-justified in order to isolate the effects of various debt arrangements in the presence of shocks. Apart from aggregate public investment, other public expenditure is not modelled other than as a transfer to private consumers.

Despite the model's ability to examine the effects of changes in the world price of copper on public finances, this paper chooses not to pursue this enquiry. Zambia is a small open economy and hence a price-taker in global commodity markets. This research is focused on welfare outcomes under different debt and financing structures during a negative TFP shock and isolates these changes. It is therefore less concerned with the impact of terms of trade shocks. Further study could explore welfare outcomes and economic gains in the instance of expected rises in international copper prices.

[21] International Energy Agency, "Zambia: Energy Mix," 2024.

[22] M. Dell, B. F. Jones, and B. A. Olken, "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature* 52, no. 3 (2014): 740–798.

5. Calibration and the Context for Policy Simulations

To utilise this model for debt policy analysis and simulation necessitates calibrating the initial baseline solution and each policy simulation. The aim of the calibration is twofold: to ensure that the initial numerical solution of the model replicates the broad macroeconomic contours of contemporary Zambia, and to ensure that the values of key behavioural parameters and elasticities governing the model behaviour generate plausible responses to the simulated policy experiments. For the policy simulations, the paper seeks to characterise a suite of SCDI options that participating governments might plausibly consider in light of an exogenous climate shock to the lending government's economy. These are summarised in Table 1 below. The baseline calibration parameters are reported in Appendix A.

5.1. Baseline model calibration

The baseline scenario captures Zambia's macro-fiscal conditions during much of the 2000s, a period characterised by relatively prudent borrowing practices, limited exposure to commercial debt markets, and only mild exogenous shocks. It assumes a modest TFP shock and access to a concessional sovereign bond with a 20-year maturity, reflecting the long-term concessional financing Zambia primarily relied on before issuing Eurobonds.

The baseline calibration, described in Table 1, is calibrated to reflect how the Zambian economy in equilibrium at Period 0 handles a mild negative TFP shock beginning in Period 1 that serves as a proxy for an exogenous climate shock, namely drought in the case of Zambia. At the onset of the shock, domestic debt is 27% of GDP, while external concessional borrowing is 11% and non-concessional borrowing accounts for 33%. Revenue is raised principally from import taxes (which reflects the relative ease with which these are collected in small open economies as compared to labour income tax) as well as other taxes on expenditure such as excise and VAT/sales tax. Government expenditures are allocated between necessary public investment spending, transfers to households, and recurrent expenditures including debt service costs. The equilibrium state is reflected in tax smoothing and a modest (single-digit) current account surplus. The calibration sets the initial GDP to 100 and normalises all initial prices and wage vectors to 1. It also defines a set of parameters defining production. Moreover, whereas the private sector is assumed to optimise its consumption and production behaviour by construction in the model, the government is not an optimising agent. Instead, it is assumed to operate by making simple exogenous policy choices over borrowing, expenditure and taxation in the presence of adverse TFP shocks. This construction allows for positive analysis of alternative debt financing structures and fiscal management following climate-induced productivity shocks.

5.2. Debt structure simulations

The debt structure simulations outlined in Table 1 and analysed throughout this paper, share a common thread in which policymakers aim to plug the fiscal gap created by the productivity shocks, and output by issuing external debt. This is with exception to Scenario 3, which explores the fiscal response required when Zambia loses complete access to global external credit markets. Following the increase of external debt, authorities increase expenditure of public investment and transfers. Borrowing is undertaken in the present to fund public investment that is expected to yield greater returns than the initial cost of borrowing. Debt sustainability of this borrowing requires the country's external debt (inclusive of interest obligations) be no greater than the discounted present value of future trade surpluses. As a share of GDP, over continuous time, this is given by the fundamental equation of motion for the debt ratio in equation 3:

$$\dot{d}_t = (r - g)d_t - b_t \tag{3}$$

External debt over time, $\dot{d}_t$, is an increasing function of the interest costs of debt r, and the debt stock, d_t, and decreasing in growth and the trade balance.

5.2.1. SCDIs

For the purposes of this model, SCDIs have been designed/configured such that in the instance of an exogenous climate shock, a 'bad' state, the interest rate on non-concessional debt is automatically reprofiled and adjusted downwards towards the concessional interest rate. This aids debt sustainability by reducing r in equation 3. This rate is kept lowered for the duration of the shock but increases to a premium rate of 10% once the climate (TFP) shock has dissipated and the 'good' state returns.

Table 1. Summary of TFP Shock and Response Scenarios

No.	Name	Details
1	Pre-Eurobond	Baseline case before Zambia issued Eurobonds. Low external debt, mild TFP shock.
2	Standard borrowing	10-year non-concessional borrowing under a severe TFP shock. No SCDI or external support.
3	No external debt	Same as Scenario 2 but with no external borrowing allowed. All adjustment done through taxes and domestic debt.
4	SCDI full + premium	SCDI reduces interest rates during the shock, then rises to premium rate. Matches Scenario 2 in all other respects.
5	SCDI full, no premium	Same as Scenario 4 but no interest rate increase after the shock.
6	SCDI full + grant	Same as Scenario 4 but includes extra support from donors.
7	SCDI partial reprofiling	Same as Scenario 4 but interest rate falls by less. Tests whether partial relief is enough to avoid default.
8	High-risk borrowing	Short, large borrowing over 5 years. Tests outcomes when Zambia borrows aggressively and front-loads debt.
9	Consecutive shocks (10-year)	TFP shock occurs twice, using a 10-year bond structure. Tests resilience to repeated external shocks.
10	Consecutive shocks (20-year)	Same as Scenario 9, but with 20-year bond maturities. Tests whether longer debt helps smooth adjustment.

This reflects the novelty premia demanded on new instruments by creditors. The degree to which interest rates are cut will depend on the pre-arranged terms of the debt agreement. It is this pre-shock agreement that allows for automatic reprofiling and thus reduces the time and steep cost of adjustment for the economy.

5.2.2. Grants

The interest cost savings generated during each period of automatic debt re-profiling are treated as equivalent to implicit grants, credited directly to the fiscal balance. This effectively expands fiscal space, providing the government with crucial 'breathing room' to preserve essential public investment and transfers during economic strain. The model also incorporates explicit grants from multilateral institutions such as the IMF and World Bank, which reflect crisis support to Zambia during the 2024 drought.[23] In light of recent reductions in foreign aid budgets—including substantial cuts to programmes like USAID[24]—this interpretation of SCDIs as implicit, rule-based grant equivalents provides a policy-relevant lens through which to assess debt management strategies in developing countries like Zambia.

5.3. Severity of shocks

With the exception of the baseline, all other scenarios are treated with a 'severe' shock. Productivity in the copper, exportable and non-tradeable sectors fall at their lowest by 15%, 10% and 5%, respectively. In comparison, the baseline TFP shocks 10%, 7% and 3% to the three sectors are defined as 'mild'. The classification of the shocks considered reflect the ongoing intensification of climate events as global warming is projected to breach the 1.5-degree limit, beyond which both the frequency and severity of climate shocks will increase. The model attempts to capture this in the design of the scenarios and hence future-proof key results and insights. Drought is endemic to Zambia, partly due to below-average precipitation, particularly during the seasonal rains. The country has a long history of drought years: 1987/88, 1991/92, 1994/95, 1997/98, 2001/03, 2004/05, 2011/12, 2015/16, 2018/2019 and 2024, which was declared a state of disaster by President Hichilema (2024).[25] This sequence implies that the country experiences drought every 4 to 5 years, and the frequency is projected to increase in the future due to climate change.[26]

6. Debt Structure Scenarios, Climate Shocks, and Sensitivity Analysis

6.1. Introduction

The following exercise aims to investigate which debt structures, in the presence of climate-induced productivity shocks, best preserve fiscal space, minimise welfare losses, and ultimately promote long-term debt sustainability by helping Zambia avoid default. The intention is to explore how the design, specifically the magnitude of reprofiling, of SCDIs might protect this economy when triggered by adverse climate shocks. To assess how they might mitigate the uncertainties such shocks pose to medium and long-term external debt sustainability.

Default can be defined as, at its simplest, a broken promise, or a breach of contract. For sovereign debt, this might be a missed payment or a series of them. This is illustrated operationally in the model where sharp tax hikes, which serve as proxies for fiscal adjustment, are required to plug the fiscal gap. A feature of DGE models is that they enforce budget balance across the economy so sudden tax rate increases demonstrate the model struggling to finance the imbalance of default. In this framework, the use of SCDIs is assessed by their ability to mitigate these indicators of default risk.

6.2. The 10 simulations

Baseline Scenario 1: Pre-Eurobond Zambia

The baseline simulation is designed to reasonably reflect Zambia's economy in the years before it began much more aggressive borrowing practices by issuing 10-year Eurobonds. The TFP shock occurs immediately, reducing output disproportionately across the three sectors, as shown in Table 2. The exportable sector, which only recovers in period 7, does so particularly slowly due to the direct effects of the shock and diminished business confidence and investment necessary to recover output.

Table 2. Macroeconomic Indicators Over Time – Baseline

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	98.06	102.93	102.52
Consumption (level rel. to initial)	100	100.40	101.24	97.16
Tax Rate	5	9.92	10.24	14.03
Debt Stock (% of GDP)	71	107.61	121.79	93.74

Note: $t = 0$ denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

In order to support households and protect consumption, authorities introduce transfer payments equivalent to 5% GDP for the first 5 years of the shock when impacts are largest. Payments are then tapered down to 2.5% GDP for the next 5 years as the economy recovers and eventually phased out to their initial pre-shock level. This is helped financed by a 20-year sovereign bond that Zambia has already issued, with a straight-line amortisation rate of 10% at a constant interest rate of 8%. This can be interpreted as classic bilateral borrowing, reflective of Zambia's emerging access to global financial markets, non-urgent external borrowing, and continued reliance on concessional finance.

The government also undertakes what is assumed to be critical public investment across the economy (e.g. health, education), which is financed by a combination of domestic resources and concessional and non-concessional finance. The allocation of fiscal resources between public investment and direct transfers is influenced by several factors, including the government's pre-existing policy priorities, the political cycle[27] (particularly proximity to an election year), and the perceived nature of the shock[28]—whether it is temporary or permanent, or rooted in demand- versus supply-side pressures.

As the bond continues to be paid down and no additional borrowing is undertaken, required total tax revenue must increase to plug the fiscal gap since grants remain constant. The interest cost of non-concessional debt does not breach 5% of GDP, so crucially, this occurs smoothly with no sharp tax hikes. The economy returns to its pre-shock state in period 25 without defaulting.

---

[23] See World Bank (2024b).

[24] In January of 2025, USAID cut 85% of their programmes; see Si (2025).

[25] H. Hichilema, "Statement on the Drought Situation in Zambia," Republic of Zambia, Office of the President, 2024.

[26] European Commission, "Drought Resilience Profile: Zambia," 2021.

[27] See for example Alesina et al. (1997).

[28] See for example Blanchard and Leigh (2013).

Scenario 2: Severe drought with standard borrowing

This scenario illustrates the impact of an unpredictable, exogenous climate event to Zambia's path to default in 2020. Therefore, it serves as the effective point of comparison for subsequent analysis. The severity of the TFP shock increases from 'mild' to 'severe', causing copper production, which relies heavily on climate-vulnerable hydroelectric power, to plummet. Table 3 shows the initial fall in output and consumption. Production of exportables falls, pulling down GDP by around 10% of its initial value. Non-tradeable output rises in period 3 due to a brief real exchange rate appreciation (the price of non-tradeables increases relative to the price of exportables) and continues to rise until period 7 as labour moves out of the other two sectors.

Table 3. Macroeconomic Indicators Over Time – Scenario 2

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	94.36	103.36	102.63
Consumption (level rel. to initial)	100	98.89	88.38	99.72
Tax Rate	5	9.31	19.84	16.04
Debt Stock (% of GDP)	71	115.5	127.94	60.87

Note: t = 0 denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

Since the shock is correctly assumed to be temporary,[29] this prompts authorities to issue a more aggressive 10-year sovereign bond in an effort to stabilise the economy in the short run and fund increased transfers and critical public investment. Given the severity of the shock, it is reasonable to assume that the government would prioritise and allocate relatively more than in the previous scenario to transfer payments (6% then tapered to 3%). The amortisation rates of 10% for the first 9 years plus a 50% payment reflect Zambia's 'bullet'-style Eurobond borrowing terms explained in the previous section. The total debt stock rises between period 1 and 11 as depicted in Table 3.

To afford the principal payment in period 10, the required total tax revenue would need to more than double from 10 to 25% of non-copper GDP—a far too steep adjustment for an economy. Default can be said to occur at period 11 as interest costs of non-concessional debt rise quickly before the TFP shock has fully passed through causing Zambia to miss the payment, indicating the difficulty of making such a payment. Much of this adjustment would have come from increasing import taxes (tariffs) by more than the increase in taxes on non-tradeables and exportables. This is broadly reflective of the relative ease with which developing economies are able to raise revenues from tariffs as opposed to other formal taxes.

Scenario 3: No external debt

All of the assumptions of the severity of the TFP shock in the first scenario remain, except now the financing structure differs: Zambia, in this instance, does not issue any external debt. This scenario may be interpreted either as a 'no action' policy by the government or as Zambia losing some form of access to the global capital markets.

In response, the government must continue to make necessary public investments into key sectors such as healthcare and education, whilst protecting the most vulnerable households in the economy with transfer payments. Except, in the absence of external debt, significant domestic revenue mobilisation is

---

[29] If permanent, consumption adjusts downwards by the full amount of the shock, the current account is rebalanced to 0, and no smoothing occurs.

Table 4. Macroeconomic Indicators Over Time – Scenario 3

Indicator	t = 0	t = 5	t = 10	t = 20
Output (level rel. to initial)	100	94.26	102.30	102.97
Consumption (level rel. to initial)	100	92.72	98.95	102.44
Tax Rate	5	15.62	16.34	6.80
Debt Stock (% of GDP)	71	109.52	100.51	71.89

Note: t = 0 denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

needed immediately in order to plug this fiscal gap. Table 4 reports that tax rates would need to increase threefold by period 5 when consumption and output are unrecovered, to raise the required revenue. Adjustment of this scale and severity places a tremendous burden on firms and households alike. So whilst Zambia incurs no external debt default on its books, welfare across the economy falls as discussed in greater detail in Section 7. This gives insights into the potential benefits of issuing debt, particularly during an exogenous shock.

Scenario 4: SCDI with full reprofiling

We now consider the hypothetical: a world in which Zambia can issue state-contingent debt instruments as earlier defined. The profile of the TFP shock and all other assumptions of Scenario 2 remain unchanged.

As before, external debt finances increased public investment and transfers, and artificially bolsters the economy until period 10, when the principal payment must be paid in full. Under the standard borrowing scenario, drastic and rather unfeasible tax hikes were needed to cover this payment, hence it could be concluded that default occurred in the following period 11.

However, in this scenario, the SCDI is triggered at the onset of the shock and instantaneously reduces the non-concessional interest rate from 9.5% to 2.5%. This rate is upheld until period 8 when production recovers to at least its initial values. These interest cost savings on the external debt are treated as an implicit grant in the fiscal budget, which gives the government sufficient fiscal space to make key investments and transfers, and meet its obligatory debt servicing payments. Once this grace period lapses, Zambia faces a premium on its future borrowing for 8 periods during the 'good' state of the world, signifying lender compensation.

Table 5. Macroeconomic Indicators Over Time – Scenario 4

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	94.28	103.73	102.58
Consumption (level rel. to initial)	100	101.99	89.01	99.71
Tax Rate	5	7.10	17.68	15.68
Debt Stock (% of GDP)	71	108.05	121.24	58.96

Note: t = 0 denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

Scenario 5: SCDI with full reprofiling, premium-free

A natural extension is to simulate an identical SCDI structure but without the imposition of a premium. In this case, the interest rate simply returns to its baseline level of 9.5% in period 8. Interestingly, the macroeconomic dynamics are very similar to those in Scenario 4 (Table 5), suggesting that the timing of debt service relief, rather than the exact cost of future borrowing, is the dominant driver of improved outcomes. This strengthens the

theoretical case that SCDIs operate primarily through liquidity channels.

Scenario 6: SCDI with full reprofiling and grant support

This scenario extends the full SCDI arrangement (Scenario 4) by introducing additional donor-financed crisis support. It simulates a stylised real-world response in which the climate-induced shock triggers not only an automatic market-based reprofiling via the SCDI mechanism, but also draws in multilateral (such as from the IMF or World Bank) or bilateral donor assistance.

The TFP shock remains unchanged, with productivity falling across all sectors. However, in response, Zambia receives an explicit external grant equivalent to 3% of GDP, injected directly into the current account. As a result, consumption is partially de-linked from output: it rises above baseline levels in period 5 despite the contraction in GDP. This reflects the combined effect of automatic payment relief (via the SCDI) and additional liquidity from the grant.

Tax rates remain moderate throughout—peaking below 15%—allowing the government to finance transfers and investment without resorting to sharp fiscal consolidation or defaulting. Notably, the debt trajectory is markedly improved: the debt-to-GDP ratio falls from 71% initially to just 52.7% by period 20. This outcome reflects both the automatic interest relief embedded in the SCDI and the external liquidity provided by the grant. Among all scenarios considered, this design yields the clearest evidence of improved debt sustainability in the aftermath of a climate-induced productivity shock.

Table 6. Macroeconomic Indicators Over Time – Scenario 6

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	98.6	102.1	106.8
Consumption (level rel. to initial)	100	105.56	91.20	100.50
Tax Rate	5	4.19	14.3	12.88
Debt Stock (% of GDP)	71	98.31	111.63	52.73

Note: t = 0 denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

Scenario 7: SCDI with partial reprofiling

To determine whether full reprofiling is a necessary condition for debt sustainability, this scenario simulates a partial cut in non-concessional rates from 9.5% to 4% (instead of 2.5%) during the shock. The movements in the economy track the same path; however, the government's response necessitates that total tax revenue required exceeds 20% since the implicit grant is diminished. This pushes authorities closer to the trade-off between incurring the cost of default versus enacting steep adjustment to meet repayment costs.

Table 7. Macroeconomic Indicators Over Time – Scenario 7

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	94.33	103.57	102.61
Consumption (level rel. to initial)	100	100.64	88.49	99.64
Tax Rate	5	8.03	18.83	16.10
Debt Stock (% of GDP)	71	111.21	124.71	60.24

Note: t = 0 denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

Scenario 8: High-risk borrowing

This scenario simulates a high-risk financing strategy wherein Zambia issues a short-term, 5-year commercial bond in the wake of a negative TFP shock. This aggressive borrowing approach is indicative of weak fiscal discipline or myopic policymaking. While government transfers initially shield consumption from the shock, the benefits are short-lived: by period 5, the debt stock has surged to 125% of GDP—up from 71% in Table 8—and the country must begin repaying principal under non-concessional terms.

Table 8. Macroeconomic Indicators Over Time – Scenario 8

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	94.33	101.65	103.58
Consumption (level rel. to initial)	100	102.89	86.95	103.70
Tax Rate	5	1.05	34.15	8.53
Debt Stock (% of GDP)	71	125.01	118.98	33.80

Note: t = 0 denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

The truncated borrowing period means that interest costs of debt spike to 9% of GDP well before the economy recovers from the shock. Critically, although consumption rises briefly in period 5, it plummets in period 11 due to an abrupt fiscal tightening. Required tax revenue jumps to 34.15%—a level that would be politically and economically unfeasible in a real-world context. Without access to external grants, Zambia defaults in period 6 when the amortisation schedule outpaces fiscal capacity. Despite a recovery in consumption and output by period 20, the cost of front-loaded debt service produces a sharp welfare penalty as discussed in Section 7, confirming the unsustainability of this high-risk strategy.

Scenario 9: Consecutive shocks with 10-year bond

This exercise simulates a world of increasingly frequent climate or TFP shocks in the model. The economy is hit immediately with the first severe shock putting downward pressure on output and consumption. Critically, before the economy can fully recover to its initial values, it is subject to another TFP shock of similar magnitude in period 10.

Authorities again issue a 10-year sovereign bond at the onset of the first shock. The overlap between repayment obligations and the second shock results in acute fiscal pressure. By period 11, consumption drops to 84.4—the lowest across all scenarios—and the debt stock reaches 137.1% of GDP. The government would be forced to raise taxes dramatically, from 5 to 20.7% over the first half of the horizon, with tax rates remaining high through period 20. Although output and consumption recover toward the end of the simulation, the short maturity amplifies adjustment costs in the interim. These results illustrate the dangers of "debt bunching," where principal repayments coincide with adverse economic conditions, triggering inefficient fiscal tightening and deep welfare losses.

Scenario 10: Consecutive shocks with 20-year bond

Finally, to test how sensitive outcomes are in a system of consecutive shocks to a given financing structure, a longer, 20-year bond maturity is considered. The extended repayment window delays the fiscal burden, allowing more space to absorb the shocks. Debt still rises sharply—to 131.7% of GDP by period 11—but tax rates increase more gradually, peaking at 21.6% only by period 20. Crucially, the consumption drop is smaller and

Table 9. Macroeconomic Indicators Over Time – Scenario 9

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	94.64	97.84	105.19
Consumption (level rel. to initial)	100	98.99	84.38	99.94
Tax Rate	5	9.31	20.70	24.87
Debt Stock (% of GDP)	71	114.86	137.14	73.15

Note: $t = 0$ denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

more contained, reaching a low of 91.1 before recovering to 97.5. These outcomes show that while long-term bonds cannot prevent debt accumulation under repeated shocks, they meaningfully reduce the pace and severity of fiscal adjustment. The simulation underscores the value of longer maturities in improving short-run resilience—not by averting default outright, but by softening its economic and welfare costs.

Table 10. Macroeconomic Indicators Over Time – Scenario 10

Indicator	t = 0	t = 5	t = 11	t = 20
Output (level rel. to initial)	100	94.63	97.29	104.62
Consumption (level rel. to initial)	100	98.20	91.07	97.51
Tax Rate	5	9.93	10.68	21.59
Debt Stock (% of GDP)	71	113.92	131.71	107.98

Note: $t = 0$ denotes the initial period before the shock. Tax rates reflect the average across exportables, non-tradeables, and imports.

7. Discussion

7.1. Summary of key findings

The scenario analysis above indicates that when severe shocks occur under an SCDI debt arrangement with full reprofiling plus an explicit grant, Zambia indeed averts default in the model. It is therefore argued that SCDIs could have preserved fiscal space and eased adjustment pressures in 2020. The only other instance that this is true is in the baseline scenario, when the TFP shock is mild and therefore borrowing is less aggressive. The key mechanism at work is that a generous SCDI offers immediate relief on debt interest payments when capacity in the economy is at its lowest as a result of an adverse climate shock. This implicit grant allows for smaller or slower tax rate increases relative to standard borrowing practices, where grants are assumed to be unchanging.

To assess the welfare consequences of alternative debt arrangements, this section evaluates outcomes across three lenses: intertemporal utility, consumption-equivalent variation (CEV), and two social welfare functions (SWF_1 and SWF_2). The analysis provides a comprehensive account of both average welfare improvements and distributional effects, offering a policy-relevant interpretation of how state-contingent debt instruments (SCDIs) compare to conventional borrowing structures. It is assumed that around a third of the population are saving households, s, who are able to allocate consumption across time through borrowing and saving and hence, smooth consumption over their lifetimes. Two-thirds of households are modelled as rationed, h, who consume hand-to-mouth. This distinction drives a wedge in their respective utility outcomes, U_s and U_h.

Table 11 reports welfare outcomes under each scenario. Scenario 2, standard borrowing, is treated as the effective baseline against which the outcomes of the counterfactuals are compared.

In the event of a shock, government transfers are non-discriminantly channelled to both sets of households. The consumption equivalent variation (CEV),

$$\text{CEV} = \frac{PV \; c_{\text{policy}}}{PV \; c_{\text{baseline}}} - 1 \qquad (4)$$

is the present value (PV) of consumption under the specified policy, relative to the present value of consumption under the baseline, where PV is discounted at the subjective discount rate, β. CEV is expressed as percentage gains which offers an intuitive measure of policy desirability. It asks: "What is the equivalent amount of lifetime consumption that households would be willing to forgo to access a particular financing structure compared to the standard borrowing case?" Essentially, it is the willingness to pay in terms of consumption.

The SCDI arrangement with full reprofiling plus an explicit grant offers the highest welfare gains, with an aggregate CEV of 3.26%. This means that households would be willing to forgo more than 3% of the present value of their lifetime consumption to access this policy. Even in the absence of grants, SCDI Full, SCDI Full + Premium and SCDI Partial + Premium yield significant welfare improvements, suggesting that automatic re-profiling alone offers substantial insurance value. In contrast, high-risk borrowing and repeated shock scenarios result in negative CEVs.

Utility is defined as:

$$U_i = \sum_{t=0}^{\infty} \beta^t \frac{(c_t^i)^{1-1/\tau_i}}{1 - 1/\tau_i}, \quad i = s, h \qquad (5)$$

where β is the subjective discount factor, c_t^i is consumption by household type i at time t, and τ_i governs the intertemporal elasticity of substitution.

Disaggregating by household type reveals the extent to which policy impacts are unevenly distributed. Column 4 of Table 11 shows that saving households exhibit modest variation in utility across scenarios compared to rationed households who experience pronounced changes and depend more heavily on transfers and public services. This is explained by the ability to borrow/save offering protection during an adverse shock since it de-links the fall in income from consumption. For rationed households, the SCDI with full reprofiling plus a grant improves their utility by 3.7% over standard borrowing. Conversely, the high-risk scenario is most damaging for credit-constrained households. This difference in utility indicates that rigid debt structures exacerbate vulnerability among credit-constrained groups, and that SCDIs offer both macroeconomic resilience and a pro-poor distributional tilt.

The social welfare functions (SWF) are the weighted aggregate of utility across both sets of households in the economy. The first welfare function, SWF_1, is defined as a utilitarian linear combination:

$$\text{SWF}_1 = \omega_s \cdot U_s + \omega_h \cdot U_h \qquad (6)$$

where ω_i is the population weight of household type i. In this model, $\omega_s = \frac{1}{3}$ and $\omega_h = \frac{2}{3}$, reflecting a larger share of liquidity-constrained (rationed) households. This function aggregates household utilities using fixed population weights, thereby giving a population-weighted average welfare level. This makes it sensitive to total welfare but less sensitive to inequality than SWF_2 which introduces inequality aversion as follows:

Table 11. Welfare metrics under different debt arrangements and shocks. Utilities and SWFs are expressed as percentages of baseline saving utility = 0.

Debt Arrangement/Shock	CEV(%)	SWF_1 / SWF_2 (%)	U_s / U_h (%)
Standard borrowing	0.0	0 / 0	0 / 0
Pre-Eurobond	0.93	0.21 / 0.44	0.16 / 1.71
No Debt	−0.03	0.022 / 0.47	0.00 / 0.57
SCDI Full + premium	1.43	0.30 / 0.38	0.26 / 1.62
SCDI Full	1.57	0.34 / 0.5	0.29 / 1.83
SCDI Partial + premium	0.83	0.18 / 0.2	−0.16 / 0.96
SCDI Full + Grant	3.26	0.68 / 1.26	0.59 / 3.71
High-Risk	−0.25	−0.08 / −0.16	−0.06 / −0.55
Consecutive Shocks ($T = 10$)	−0.55	−0.19 / 0.3	−0.15 / −1.3
Consecutive Shocks ($T = 20$)	−0.09	−0.02 / 0.04	−0.04 / −0.4

$$SWF_2 = \left(\omega_s U_s^{1-\phi} + \omega_h U_h^{1-\phi} \right)^{\frac{1}{1-\phi}} \tag{7}$$

Here, ϕ is the Atkinson inequality aversion parameter. A higher ϕ implies that the social planner places greater weight on improving the utility of worse-off groups. This study sets $\phi = 2$, a commonly used benchmark in the development economics literature (e.g. Deaton, 1997) to reflect moderate to strong aversion to inequality.[30] $\phi = 2$ has become standard in applied welfare analysis involving poor and vulnerable populations. Table 11 reports that while both functions identify SCDI Full plus Grant as the socially optimal outcome, the parameterisation of ϕ in SWF_2 magnifies the welfare losses under regressive policies and the gains under equitable ones. For example, SWF_2 assigns a 1.26% improvement to SCDI Full plus Grant, compared to only 0.68% under SWF_1.

Inequality is penalised by assigning greater weight to utility gains accruing to poorer households (here, the rationed group), and discounts equivalent utility gains accruing to already better-off households (the savers). Crucially, SWF_2 reveals improvements in equity that SWF_1 understates and that the greatest social welfare gains under SCDIs come from increasing the welfare of those with the least flexibility under rigid, high-risk debt structures.

The choice between SWF_1 and SWF_2 reflects a trade-off between efficiency and equity, depending on the policymaker's objectives—for example, ϕ might be chosen if it is politically gainful to protect rationed households that comprise most of the economy and hence the voting base. An insightful extension of this welfare analysis would be to introduce overlapping generations (OLG), enabling welfare evaluation across both those who experience the shock and those who bear its long-run fiscal cost.

7.2. SCDIs: Political Feasibility and Implementation

The protracted and uncertain restructuring process under the G20 Common Framework (CF) has exposed critical weaknesses in the global sovereign debt resolution architecture—particularly its lack of timeliness, transparency, and predictability.[31] In contrast, state-contingent debt instruments (SCDIs) offer a rule-based alternative: pre-agreed clauses for temporary payment relief means that they can reduce the costs and uncertainty associated with conventional restructurings.

SCDIs also provide a structural solution to the classic external transfer problem. In cases where governments face external debt obligations but limited access to foreign currency, they may respond by taxing the export sector heavily to appropriate export earnings. This creates a distortionary disincentive for firms to produce for export, worsening balance-of-payments pressures and deepening economic contraction. SCDIs alleviate this by aligning repayment with capacity to pay and preserving incentives for export-oriented production.

From a policy reform standpoint, sovereign debt restructurings present a strategic window to introduce SCDIs at scale. Unlike new issuances, which face market resistance due to the "first-mover" disadvantage, restructurings allow retrofitting of these instruments across the entire renegotiated debt stock. With creditor consent, this can help resolve legacy contract fragmentation and enhance the consistency of debt treatment.[32]

Moreover, SCDIs represent a form of embedded climate justice. They acknowledge the disproportionate exposure of low-income countries to climate shocks—events they did not cause—and offer relief in those countries when fiscal space is most constrained. From the creditor's perspective, SCDIs reduce the likelihood of outright default by allowing payments to adjust in response to negative shocks, thus preserving long-term repayment capacity.

Political feasibility, however, remains contingent on key design features: coordination across diverse creditor groups, credible and transparent trigger mechanisms, and deeper integration into sovereign debt markets. Yet these challenges are not insurmountable. Rising climate risk and the IMF's Global Sovereign Debt Roundtable have renewed calls for innovation in sovereign debt instruments. SCDIs offering contingent interest rate reductions or maturity extensions are increasingly seen as viable tools to prevent adverse shocks from escalating into debt crises—benefiting borrowers, creditors, and the international financial system alike.

Multilateral institutions, particularly the IMF and World Bank, have a critical role to play in normalising SCDIs within the sovereign debt architecture. Their endorsement, combined with technical support and coordinated creditor engagement, can help mainstream the adoption of SCDIs as part of a broader effort to climate-proof public finances and enhance debt sustainability in small open economies such as Zambia.

8. Conclusion

This dissertation set out to explore whether Zambia's 2020 sovereign default could have been avoided had it adopted

---

[30]A. Deaton, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, World Bank, 1997.

[31]D. Grigorian and A. Bhayana, "Zambia: A Case Study of Sovereign Debt Restructuring under the G20 Common Framework," Center for Global Development, 2024.

[32]C. Cohen, S. A. Abbas, M. Anthony, et al., "The Role of State-Contingent Debt Instruments in Sovereign Debt Restructurings," IMF Staff Discussion Note, 2022.

state-contingent debt instruments (SCDIs), and whether such instruments can promote long-run debt sustainability in the presence of adverse climate shocks. Using a dynamic general equilibrium model calibrated to Zambia's macroeconomic and sectoral realities, the analysis simulated the macro-fiscal consequences of various debt arrangements in response to severe, climate-induced productivity shocks. The results provide compelling evidence that well-designed SCDIs—especially when combined with donor support—can offer a powerful, rules-based framework for maintaining debt sustainability and preserving welfare.

Three central findings emerge. First, in contrast to standard borrowing arrangements, SCDIs that automatically reprofile interest rates during economic downturns reduce fiscal strain, prevent abrupt tax hikes, and maintain public investment flows. This timing of relief is critical: liquidity support during "bad states" mitigates the compounding effects of output loss, rising debt service costs, and procyclical austerity. Second, the addition of grants—either explicit or treated as equivalent to interest savings—amplifies the stabilising effects of SCDIs, enabling governments to smooth transfers and investment without jeopardising long-run fiscal sustainability. Third, the welfare gains from such arrangements are not only economically significant but also socially progressive. The strongest improvements in lifetime utility accrue to the most vulnerable households, underlining the distributional advantages of embedding contingency and solidarity into sovereign debt contracts.

The study also highlights that not all alternative debt strategies yield equal benefits. High-risk, short-term commercial borrowing was shown to amplify default risk and reduce welfare—particularly for liquidity-constrained households. Similarly, strategies that forgo external borrowing entirely impose severe adjustment costs and undermine macroeconomic stability, even if they avoid external arrears. Conversely, maturity extension alone cannot prevent crisis, but it can cushion its most acute impacts when shocks are repeated or prolonged.

From a policy standpoint, the findings strengthen the case for the incorporation of SCDIs into the debt architecture of climate-vulnerable developing countries. Their countercyclical features offer a sound path to enhance resilience in an era of intensifying climate volatility and declining grant finance. Future challenges will lie more in their implementation rather than their theoretical desirability. As the global community rethinks sovereign debt governance in the wake of Covid-19 and mounting climate stress, Zambia's experience stands as a timely reminder that how countries borrow can matter just as much as how much.

Ultimately, this dissertation argues that SCDIs are not simply a financial innovation, but a necessary evolution in global debt architecture—one that better aligns the instruments of development finance with the lived realities of an increasingly shock-prone world. For countries like Zambia, whose fiscal futures are closely tied to climate dynamics, adopting such instruments may be the path not taken—but perhaps the path that must be built.

■ References

[1] Adam, C., Bevan, D. and Ohlenburg, T. (2016). Public investment, fiscal rules, and growth in low-income countries. IGC Working Paper S-43302-UGA-1.

[2] Adam, C. and Musonda, A. (2011). Copper mining in Zambia: From collapse to recovery. In P. Collier and A. J. Venables (eds.), *Plundered Nations? Successes and Failures in Natural Resource Extraction*, Palgrave Macmillan, pp. 89–114.

[3] Alesina, A., Roubini, N. and Cohen, G. D. (1997). *Political Cycles and the Macroeconomy*. MIT Press.

[4] Alfani, F., Arslan, A., McCarthy, N., Cavatassi, R. and Sitko, N. (2019). Climate-change vulnerability in rural Zambia: The impact of an El Niño-induced shock on income and productivity. FAO Agricultural Development Economics Working Paper 19-02.

[5] Anthony, M., Impavido, G. and van Selm, B. (2020). Barbados' 2018–19 sovereign debt restructuring—A sea change? *IMF Working Paper WP/20/34*.

[6] Bjerkholt, O. (2004). New approaches to debt relief and debt sustainability in LDCs. CDP Background Paper No. 5, United Nations.

[7] Blanchard, O. and Leigh, D. (2013). Growth forecast errors and fiscal multipliers. *American Economic Review* 103(3), 117–120.

[8] Buffie, E. F., Berg, A., Pattillo, C., Portillo, R. and Zanna, L.-F. (2012). Public investment, growth, and debt sustainability: Putting together the pieces. *IMF Working Paper No. 12/144*.

[9] Center for Global Development (2024). The socioeconomic impact of climate change in developing countries in the next decades: A review. Working Paper 681.

[10] Cohen, C., Abbas, S. A., Anthony, M., Best, T., Breuer, P., Miao, H., Myrvoda, A. and Togo, E. (2022). The role of state-contingent debt instruments in sovereign debt restructurings. IMF Staff Discussion Note.

[11] Cohen, D. and Sachs, J. D. (1985). Growth and external debt under risk of debt repudiation. NBER Working Paper w1703.

[12] Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. World Bank.

[13] Debt Justice (2020). Zambia country profile—Debt data portal.

[14] Dell, M., Jones, B. F. and Olken, B. A. (2014). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature* 52(3), 740–798.

[15] European Central Bank (2022). The climate change challenge and fiscal instruments and policies in the EU. Occasional Paper 315.

[16] European Commission (2021). Drought resilience profile: Zambia.

[17] Finance for Development Lab and ZIPAR (2023). The road to Zambia's 2020 sovereign debt default.

[18] Grigorian, D. and Bhayana, A. (2024). Zambia: A case study of sovereign debt restructuring under the G20 Common Framework. Center for Global Development.

[19] Hatchondo, J. C., Martinez, L. and Sosa-Padilla, C. (2016). Sovereign defaults and debt sustainability: The role of state-contingent debt. *IMF Economic Review* 64(4), 641–664.

[20] Hichilema, H. (2024). Statement on the drought situation in Zambia. Republic of Zambia, Office of the President.

[21] IMF (2005). Zambia: Enhanced initiative for heavily indebted poor countries—Completion point document. IMF Country Report 05/137.

[22] IMF (2017). State-contingent debt instruments for sovereigns. IMF Policy Paper PP/17/10.

[23] IMF (2025). World economic outlook database, April 2025: Zambia.

[24] IMF and World Bank (2019). Zambia: Joint World Bank-IMF

debt sustainability analysis.

[25] International Monetary Fund (2025). List of LIC DSAs for PRGT-eligible countries as of March 31, 2025.

[26] International Energy Agency (2024). Zambia: Energy mix.

[27] Kessler, M., Mbewe, M., Humann, T., Kalikeka, M., Masilokwa, I. and Mwamba, S. (2023). The road to Zambia's 2020 default. ZIPAR-FDL.

[28] Si, A. K. (2025). Which countries could be most affected by US aid cuts? *Devex.*

[29] UNCTAD (2024). A world of debt: A growing burden to global prosperity.

[30] UNEP and NIESR (2022). Economic impacts of climate change: Exploring short-term climate-related shocks for financial actors with macroeconomic models.

[31] United Nations (2024). UN common country analysis update 2024: Zambia.

[32] World Bank (2023). Zambia's external debt composition, 2023.

[33] World Bank (2024a). International debt report 2024.

[34] World Bank (2024b). World Bank steps up support to increase Zambia's resilience to natural disasters. Press Release, December 18, 2024.

Appendix A: Baseline Calibration Parameters

Parameter	Value	Definition
g	3.0	Trend per capita growth rate (% p.a)
a_{ratio}	1.75	Ratio rationed/non-rationed households
i_{zy}	0.10	Public investment rate (% GDP)
h_{xo}, h_{mo}, h_{no}	1.5, 10, 3.5	Consumption tax rate (% consumption)
θ_{lo}	5%	Labour tax rate (% wage bill)
θ_{no}, θ_{xo}	10, 9	Profit tax rates (% sector-specific profits)
θ_{ro}	7.5	Royalty tax rate (% copper profits)
$\lambda_{hx}, \lambda_{hm}, \lambda_{hn}$	0.2, 0.4, 0.4	Long-run tax financing shares (x, m, n sectors)
η_g	1.0	External risk premium parameter
ω	2.00	Adjustment cost scaling parameter
τ_s, τ_h	0.80, 0.50	Inter-temporal elasticity of substitution
ϵ	0.7	Elasticity of substitution in consumption
ι_s, ι_h	0.4, 0.4	Frisch elasticity (saving, rationed households)
η	0.01	Interest elasticity of foreign bonds
$\alpha_x, \alpha_n, \alpha_{cu}$	0.45, 0.3, 0.76	Capital shares (exportables, non-tradables, copper)
$\delta_x, \delta_n, \delta_z$	0.05	Depreciation rates (x, m, n sectors)
ρ_x, ρ_m	0.25, 0.35	Consumption shares (exportables, imports)
prem	0.01	Premium for future borrowing
R_{zo}	0.35	Initial return to public capital
$r_o, r_{do}, r_{dco}, r^*$	0.09, 0.0075, 0.09, 0.015	Initial interest rates
f_{div_o}	0.24	Return to capital/land in copper
$\text{share}_b, \text{share}_d, \text{share}_{dc}$	0.27, 0.11, 0.33	Initial debt (dom., concess., non-concess.)
$\text{share}_{\text{remit}}, \text{share}_{\text{grants}}$	0.05, 0.015	Init. private remittances, grants (% GDP)
y_o, q_{cu_o}, w_o	100, 14, 1.0	Initial GDP, copper share in GDP, nominal wage

Note: Data used in calibration sourced from Bank of Zambia reports, Ministry of Finance Zambia, IMF Article IV reports, and World Bank reports.

The Bermuda Journal
of Academic Research

What Factors Matter in Assessing the Economic Costs of Trade Conflicts?

Jude Snelling

University of Exeter

Abstract

This paper critically analyses the economic costs of trade conflicts through the lens of the 2018–19 US-China trade war. Drawing on a range of academic literature, it examines the factors that matter in assessing the welfare consequences of protectionist measures, including complete tariff pass-through to consumers, the absence of terms-of-trade gains, supply chain redistribution, and the ruinous effects of retaliatory tariffs. The findings underscore that protectionist policies, while politically motivated, impose significant costs on national social welfare, international trade networks, and global cooperation.

1. Introduction

Trade conflicts have become the block in which the world economy has been built upon in recent years, with widespread consequences for global social and economic welfare. Of these disputes, the 2018–19 US-China trade war outlines as having the broadest implications due to its unprecedented effects on global trade dynamics. The imposition of protectionist measures and corresponding retaliation between these leading powers led to price increases, supply chain disruptions, as well as alterations in consumer welfare; analysing these trends aids one to answer the research question, 'What factors matter in assessing the economic costs of trade conflicts?' Understanding the dynamics of trade conflicts and their economic costs is crucial in assisting policy makers to achieve socially efficient outcomes and mitigate the adverse effects of protectionist policies. It is also a groundbreaking event in terms of understanding the link between real policy decisions and economic trade theory.

The case of the United States and China is especially interesting because its outcomes contradict traditional theory regarding the terms of trade effects when a globally dominant economy imposes a tariff. The tariffs, enacted by the US from July 2018 and into 2019, affected $362 billion worth of Chinese imports at rates between 10% to 25%, leading to a rise in the US duty on imports from 1.6% to 5.4%; additionally the absence of a decline in global export prices contradicts the theory's expectations (Amiti et al. 2020). Traditionally, domestic prices are expected to rise and world prices to fall with the introduction of a tariff by a large global economic player. Nonetheless in 2018, world prices did not fall and US consumer prices experienced a passthrough of the tariffs of 100% (Amiti et al. 2019). This paper critically analyses studies regarding the US-China trade war, determining their contributions to comprehending the factors to be considered when deriving the economic costs of trade conflicts. The ultimate

purpose of this detailed examination is to provide insights on the broader implications of trade wars and how such effects are determined.

The context in which this trade war erupted was one of a growing sense of economic nationalism, resulting in increased tensions between the US and China. The US measures on Chinese imports were met with retaliatory tariffs on $112 billion of US exports, targeting key American sectors such as agriculture and industrial goods (Benguria and Saffie 2019). The trade war impacted supply chains all over the world, as the innumerous firms reliant on imports from China witnessed their costs increase, and US consumers experienced extortionate price increases on consumer goods. Trade diversion occurred as businesses sought out new origins of supply with countries like Vietnam replacing the foregone US-China trade; the influence on global trade dynamics emphasises the widespread political and economic implications of the conflict (Amiti et al. 2020).

This paper seeks to explore the insights provided by a range of academic literature on the economic costs of trade conflicts to consumers, with particular attention on supply chain redistribution and the retaliatory measures that are so ruinous.

2. What Factors Matter in Assessing the Economic Costs of Trade Conflicts?

Firstly, to determine the effects of trade conflicts on consumers, as well as the consequences of them, one can analyse the U.S.-China trade war. Research by Amiti, Weinstein and Redding displays that the tariffs imposed by the US in 2018 and 2019, between 10% and 25% on various goods, were thrust entirely onto U.S. consumers, as businesses increased markups to compensate for higher costs. This is striking, especially considering that in principal, the increase in domestic prices as a result of a tariff could be combatted by foreign exporters reducing the price they

charge for the targeted goods. There is little evidence of this improvement post implementation for the US in their terms of trade, suggesting that the tariff is solely incident on the US consumer. Supporting this is the fact that large increases were witnessed for the goods that were targeted by tariffs, with the prices of those goods increasing by 10% to 30% (Amiti et al. 2019). Similarly, Amiti et al. (2020) found through a differences-in-differences analysis that 100% of import taxes of certain goods are transferred onto importers and consumers in the US, and they have not had the desired effect of altering foreign export prices while at the same time greatly negatively impacting US import volumes.

Additionally Fajgelbaum et al. (2020) found prices of targeted imports did not fall in the US, once again demonstrating the complete pass-through of these tariffs to consumers. There is an exception in the steel industry in fact, where foreign exports incurred most of the cost of the tariff, resulting in the value of steel imports falling relatively less than in other industries (Amiti et al. 2020). The explanation for this could be due to the fact the production of steel is capital intensive and the US have a higher relative factor intensity in this regard, supporting the Heckscher-Ohlin theory of international trade. Traditional theory pertaining to terms of trade effects states that when a large economy such as the US introduces tariffs, with reference to Figure 1 as discussed by Amiti et al. (2019), foreign firms would lower prices to P^* in order to meet the new US consumption demand with the imposition of a tariff being included. The absence of price reductions by Chinese exporters in the case of the US imply that it may have, in the short run at least, a perfectly elastic export supply curve, meaning that the price of those goods remain unchanged despite the reduced import quantities (refer to Figure 2).

In Figure 1 the government revenue collected by the tariff is represented by regions A and C, the transfer from consumers and foreign producers to the US government. The terms of trade, or difference in a country's gain is measured by the area $C - B$. The consequence of the short run export supply curve being horizontal is that the amount of government revenue gained from the tariff's implementation reduces to region A, and the overall terms of trade is negative as it only incorporates the deadweight welfare loss region B, as shown by Figure 2. Amiti et al. (2019) quantified this deadweight welfare loss to the US as $6.9 billion, holding only with the assumption that the government uses the revenue generated from the tariff to produce social welfare benefits of an equivalent amount. Not doing so increases the potential losses to the taxpayers by as much as the gain of the entirety of the tariff payments, meaning that it could possibly rise another $12.3 billion to a total of $19.2 billion (regions $A + C$ of Figure 2) according to Amiti et al. (2019). Fajgelbaum, Goldberg, Kennedy and Khandelwal found that the varieties of imports encompassed by US tariffs fell on average 31.7%, which inherently reduces social welfare, as well as the fact that there is no differential change in the pre-tariff import prices when comparing targeted countries like China with untargeted countries exporting the same product. This is a further indication that one cannot reject the existence of horizontal foreign export supply curves in the short run for the US.

Cavallo, Gopinath, Neiman and Tang also provide a detailed analysis on the events of the US-China trade war, concluding yet again that the incidence of the tariffs has largely affected the US with tariffs almost being passed through entirely to the prices US importers paid; a 20% tariff is linked to an increase in import prices of 18.5%. The consequence of protectionist measures is other countries respond in retaliation, which is what Canada, China, the EU and Mexico did in response to the American policies enacted in 2018 (Cavallo et al. 2021). What is intriguing is that Cavallo et al. (2021) found that the post-tariff prices dictated by foreign exporters to the US remained stable whereas US exporters notably reduced their prices as a result of these retaliatory tariffs. Not only does this confirm the previous notion discussed—the existence of elastic foreign export supply to the US that relegates the incidence of the tax onto the US consumer—but it also suggests that the retaliatory measures by angered trading partners are achieving their intended outcomes in reducing US welfare.

Furthermore, the retaliation from China is shown by Waugh to increase consumption and employment losses in the US, reducing the country's total exports and negatively affecting the labour market. The most vulnerable of regions to Chinese tariffs experienced a relative 0.75 percentage point decline in employment; reduction in US consumption is related to these employment implications (Waugh 2019). This paper helps to answer the research question by understanding what the relevant factors are that determine the economic costs of trade wars. The first being that there is a reduction in welfare for all consumers in the US due to the tariff on Chinese goods increasing prices and reducing variety; Waugh (2019) is consistent with Fajgelbaum et al. (2020) in providing evidence regarding this. Secondly, it is discussed that those most impacted by Chinese retaliatory tariffs—the industries in which the relatively higher comparative advantage was now lost to China—were directly affected and experienced higher welfare and consumption reductions (Waugh 2019). This notion is confirmed by Benguria and Saffie through evidence provided detailing the heterogeneous nature across industries of the negative impact of foreign tariffs on US exports. The Chinese retaliatory tariffs affected $112 billion worth of US exports, or 87% of the total US exports to China, and a ten percentage point increase in Chinese tariffs resulted in a 38% decline in imports of industrial supplies and a 16% decline in agricultural goods (Benguria and Saffie 2019). This implies that the US consumers employed in the industrial sector will experience relatively more hardship in terms of reduced welfare when compared to those in the agricultural sector. Importantly Benguria and Saffie (2019) also identified that the targeted imports affected by US tariffs led to reduced exports in the industries using those imports in their production, exposing the factors that calculate the costs of rearranging supply chains as US producers seek to obtain alternative inputs.

To illustrate the impact of trade conflicts on supply chains, one can turn to the analysis provided by Luo, Kang, Hu, Su and Dai which interestingly looks at the dual effects of the US-China trade war and the COVID-19 pandemic on US imports. The reaction of US importers in response to global lockdowns and the trade war was to relocate sections of their chain of production away from China to avoid the tariffs to parts of Southeast Asia (Luo et al. 2023). The Southeast Asian exporters are a relatively attractive substitute due to their low costs of production, but as pointed out by Benguria and Saffie (2019), their production is centred around Chinese suppliers. The widespread pandemic seems to have reversed this trend as it is shown that the US imported an increased number of goods directly from China, and indirectly via Southeast Asian suppliers importing from China (Luo et al. 2023). The potential implications of this reversal is that the US will have increased welfare losses; the costs incurred by US consumers as a result of the complete pass-through of the tariff onto consumer prices, as previously mentioned.

This research represents the immediate losses to consumers as a result of protectionist measures, and the long-term consequences of reduced consumer welfare has the possibility of stagnating economic growth as consumption drives GDP. This angst amongst the general public could lead to more indignation regarding trade liberalisation, further exacerbating the already lasting negative implications from protectionist policies that reduce international cooperation. Supply chains are disrupted from protectionist measures and consumer welfare is lost to great proportions as a result of retaliation to such measures. Trade conflicts inflict direct costs on consumers via higher prices, especially in the case of the United States, and the tax revenue generated is not sufficient enough to combat these negative implications. The redistribution of supply chains is costly and takes time, and the lost consumer welfare as a result of this redistribution also contributes to the costs of trade conflicts; the government should be exploring policy measures that take into account wider societal impacts to reduce these negative effects.

3. Conclusion

The case study of the US-China trade war in 2018–19 provides an insight into the economic and social consequences of trade conflicts, underscoring their profound impacts on consumers and supply chains. These conflicts emphasise the challenges of protectionist measures, which, while politically aimed at domestic economic protection, often impose immense costs on national social welfare, international trade networks, and global cooperation.

The short-run impact of the trade war on US consumers was detrimental with tariffs being passed entirely onto consumers as well as increasing the prices of imports by 10% to 30% (Amiti et al. 2019; Amiti et al. 2020). Cavallo et al. (2021) further highlighted the correlation between tariff rates and consumer costs by concluding that a 18.5% increase in import prices was caused by a 20% increase in tariffs. The stability of prices of foreign exporters, contrary to traditional trade theory, accentuated these effects and led to significant welfare losses. Amiti et al. (2019) quantified the potential losses to consumer welfare in certain industries up to $19.2 billion, reflecting the inefficiencies of these protectionist policies. Regions reliant on imports as inputs to provide exports experience the most hardship, as demonstrated by Waugh (2019).

Regarding supply chains, the trade war completely disrupted existing agreements, forcing US firms to source alternative producers in Southeast Asia (Luo et al. 2023). However, as emphasised by Benguria and Saffie (2019), these alterations came at a high cost. These rising production costs reduced the competitiveness of US exports, thus the tariffs counterintuitively harmed industries it set out to protect. Luo et al. (2023) further explained how fragmentations of supply chains when affected by tariffs in combination with the Covid-19 pandemic revealed vulnerabilities that introduced inefficiencies.

The US-China trade war offers valuable intel to policy makers. While seeming to address short-term objectives, protectionist measures have the potential to impose significant long-term costs on social welfare, trade networks and global stability. The areas that need to be targeted by policy makers include increased cooperation across countries, the strengthening of supply chains, and support for vulnerable industries and communities affected most by trade conflicts. Through development of this international collaboration, countries can alleviate the negative effects of trade conflicts and preserve the benefits of globalisation.

■ Appendix
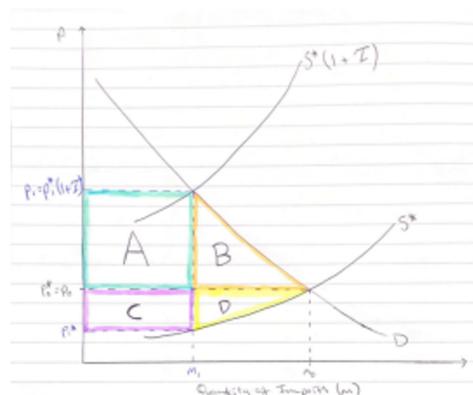
Figure 1. Impact of a Tariff on Prices. Government revenue is represented by regions A and C; the terms of trade gain is measured by $C - B$. Under standard assumptions, the tariff drives a wedge between domestic price $P_t = P^*(1 + \mathcal{T})$ and world price P^*.
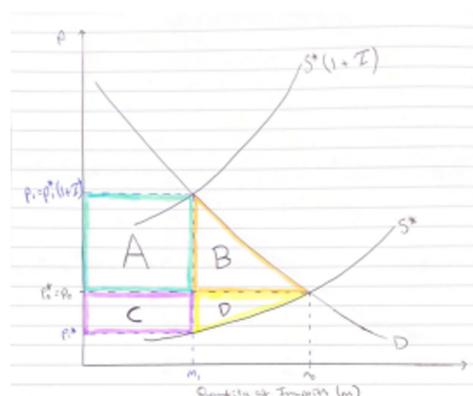
Figure 2. Impact of a Tariff on Prices with Perfectly Elastic Export Supply. With a horizontal foreign export supply curve S^*, domestic price rises to $P_t = P^*(1 + \mathcal{T})$ but world price P^* remains unchanged. Government revenue reduces to region A only, and the terms of trade effect is strictly negative, consisting solely of deadweight loss region B.

■ References

[1] Amiti, Mary, Stephen J. Redding, and David E. Weinstein (2020). "Who's Paying for the US Tariffs? A Longer-Term Perspective." *AEA Papers and Proceedings* 110, 541–546.

[2] Amiti, Mary, David Weinstein, and Stephen J. Redding (2019). "The Impact of the 2018 Trade War on U.S. Prices and Welfare." *NBER Working Paper Series*, No. 25672.

[3] Benguria, Felipe and Felipe Saffie (2019). "Dissecting the Impact of the 2018–2019 Trade War on US Exports." Available at SSRN 3505413.

[4] Cavallo, Alberto, Gita Gopinath, Brent Neiman, and Jenny Tang (2021). "Tariff Pass-Through at the Border and at the Store: Evidence from US Trade Policy." *American Economic Review: Insights* 3(1), 19–34.

[5] Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal (2020). "The Return to

Protectionism." *The Quarterly Journal of Economics* 135(1), 1–55.

[6] Luo, Wei, Siyuan Kang, Sheng Hu, Lixian Su, and Rui Dai (2023). "Dual Effects of the US-China Trade War and COVID-19 on United States Imports: Transfer of China's Industrial Chain?" arXiv preprint arXiv:2309.02271.

[7] Waugh, Michael E. (2019). "The Consumption Response to Trade Shocks: Evidence from the US-China Trade War." Technical report, National Bureau of Economic Research.