HIV AND AIDS SERVICE DELIVERY (IN) EFFICIENCY IN MALAWI: A Transaction Cost Economics Analysis of the Antiretroviral Therapy Programme

A thesis submitted in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY IN HEALTH PROMOTION

of

THE UNIVERSITY OF KWAZULU NATAL

by

BY MURPHY MATHEWS KAJUMI

March, 2025 Revision

DECLARATION

I declare that this thesis titled, **HIV and AIDS Service Delivery (in) Efficiency in Malawi: A Transaction Cost Economics Analysis of the Antiretroviral Programme**, is my own work and has not been previously submitted for a degree or diploma at this or any other university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due acknowledgment is made in the text.

Where applicable, I have cited and referenced all sources, including published or unpublished works, data, quotations, and ideas derived from other researchers. Any contributions made by collaborators or supervisors to the research are explicitly acknowledged within the thesis.

Signed:_________________________
Name: Murphy Mathews Kajumi
Date: April 1, 2025

Certificate of Approval

I certify that this thesis represents the student's own work and effort and has been submitted with my approval.

Name of Supervisor: Prof. Kaymarlin Govender DATE: April 2. 2025

(Signature)

DEDICATION

To Zindaba, Duanne, Sibusiso, Vusizwe, my mother and to the memory of my late father

ACKNOWLEDGEMENTS

First and foremost, I praise God for His grace that has enabled me to complete my PhD studies. Without His divine support, this achievement would not have been possible.

I am deeply grateful to my academic mentors who have guided me through this journey. My heartfelt thanks go to my supervisors, Professors Kaymarlin Govender, Tim Quinlan, and Jim Whitman for their invaluable academic guidance, support, and encouragement. Their expertise and insights have been instrumental in shaping my research and helping me navigate the challenges of my studies. I further wish to acknowledge the suggestions by three reviewers that helped to ensure the thesis was focused. I extend my sincere appreciation to Professor Nana Poku, University of KwaZulu-Natal Vice Chancellor and Principal, and Health Economics and AIDS Research Division (HEARD) Executive Director, for his overall guidance and support throughout my PhD journey. His leadership and vision have provided a strong foundation for my academic pursuits. I would further like to acknowledge the encouragement and support from Dr. Carolien Aantjes, whose words of motivation have kept me inspired and focused. I kept the inspirational post-it note! The entire HEARD staff was supportive and created a friendly and supportive environment during my studies. I am grateful to Dr. Winford Masanjala and Dr. Martin Mwale in the Economics Department at the University of Malawi for support provided during the course my studies.

I further wish to thank the Department of HIV and AIDS in the Ministry of Health and the Malawi National AIDS Commission who kindly allowed me to access databases for the electronic management records. I am grateful to the ART service providers at various health facilities in Malawi for their time and insights. The research was made possible by kind financial support from SIDA. My deepest gratitude further goes to my family for their unwavering support and love. My wife, Zindaba, and my children, Vusizwe, Sibusiso, and Duanne, have been my pillars of strength. Their patience, understanding, and moral support have been my source of inspiration and resilience.

Lastly, I wish to thank my friends for their periodic inquiries about the progress of my PhD. Their interest and encouragement have propelled me to stay focused and determined to complete this journey.

Contents

ACKNOWLEDGEMENTS i

LIST OF TABLES x

LIST OF FIGURES xi

ACRONYMS xiii

ABSTRACT xvi

CHAPTER ONE 1

INTRODUCTION 1

1.1 Introduction 1

1.2 Global Burden of HIV and AIDS and Response 2

1.4 Healthcare and HIV and AIDS Governance Efficiency 6

1.5 HIV and AIDS Governance and Transaction Cost Economics Theory 7

1.6 Research Problem 9

1.7 Research Aim 9

1.7.2 Research Questions 10

1.8 Research Focus and agencies covered 11

1.9 Thesis Organization 13

CHAPTER TWO 15

LITERATURE REVIEW 15

2.1 Introduction 15

2.2 Economic governance and Transaction Costs Economics 15

2.3 Key Concepts 17

2.4 Theoretical Framework 22

2.4.1 The New Institutional Economics 22

2.4.2 The New Institutional Economics and the Neo-Classical Economic Model 23

2.4.3 Branches of the New Institutional Economics 24

2.5 Transaction Cost Economics as Theory of the Firm 27

2.5.1 TCE Behavioural Assumptions 28

2.5.2 Transaction Characteristics 29

2.5.3 Boundary of the Firm under TCE 32

2.5.4 Why do Markets fail, and why does Hierarchy Fail? 32

2.5.5. Discrete Structural Analysis for Transaction Governance Modes Choice 35

2.6 The Relevance of TCE to HIV and AIDS Service Delivery through Public 38

Bureaucracy 38

2.7 Criticism of the NIE and TCE as Theory of the Firm 39

2.8 Empirical Literature on healthcare and HIV and AIDS service delivery efficiency 41

2.9 Research Analytical Framework 44
2.9.1 Analytical Categories and Assumptions 46
2.10 Chapter Summary 51
CHAPTER THREE 53
THE CHANGING LANDSCAPE OF HIV and AIDS FINANCING AND THE EFFICIENCY IMPERATIVE 53
3.1 Introduction 53
3.2 An Enduring Epidemic in the Face of Declining Resources 54
3.2.1 HIV and AIDS Fiscal Burden 56
3.2.2 The HIV and AIDS Fiscal Burden in Eastern and Southern Africa 57
3.3 Changing Global HIV and AIDS and factors that will affect funding Landscape 59
3.3.1 The promise of Treatment as Prevention 60
3.3.2 Changing Donor Funding Priorities and Integration of HIV and AIDS into 61
Health Systems 61
3.3.3 An Era of Shared Responsibility 63
3.3.4 Health Sector and HIV and AIDS Financing in Malawi 65
3.3.5 HIV and AIDS Financing and Efficiency Imperatives for Malawi 66
3.4 Chapter Summary 70
CHAPTER FOUR 71
RESEARCH METHODOLOGY 71
4.1 Introduction 71
4.2 Research Approach 71
4.3 Research Setting 72
4.5 HIV and AIDS Areas of Focus for the Study 73
4.6 Research Design 74
4.7 Research Process 75
4.7.1 Research Operationalisation 76
4.7.2 Pilot Study 77
4.8 Data Collection Methods Description 79
4.8.1 Literature Review 79
4.8.2 Semi-Structured Interviews 81
4.8.3 Database Review and Extraction of Relevant Data 82
4.9 Sampling 82
4.9.1 Inclusion and Exclusion Criteria 82
4.10 Estimating the Financial Inputs in ART Services 84
4.10.1 Estimation of Input Costs 85

4.11 Data Management and Analysis 89
4.12 Responding to the Research Questions 91
4.12.1 HIV and AIDS ART Programme via Data Envelopment Analysis 91
4.12.2 Empirical Model 96
4.12.3 Variables Description 97
4.12.4 Data for Efficiency Analysis 99
4.12.5 Pre and Post Estimation Diagnostics for the second-stage DEA 100
4.13 Reliability and Validity of Results 102
4.14 Ethical Considerations 103
4.15 Research Limitations 104
4.14 Chapter Summary 104
CHAPTER FIVE 106
THE MALAWI HIV AND AIDS ART PROGRAMME (IN) EFFICIENCY AND THE EFFECT OF TRANSACTION COSTS ON PROGRAMME EFFICIENCY 106
5.1 Introduction 106
5.5 Results of DEA Technical Efficiency Estimation 110
5.5.1 Summary of ART DMU Efficiency by District 110
5.5.3 ART Services Technical Efficiency 113
5.5.4 ART DMU (In) Efficiency by Zone and District 114
5.5.5 Health Unit (In)-Efficiency Based on Returns to scale 116
5.6 Second Stage DEA Regression Results 121
5.6.1 Health Units Characteristics 122
5.6.2 The effect of transaction costs on HIV and AIDS Facility Based ART Efficiency 123
5.7 Qualitative Assessment of ART Services Quality 126
5.8 Chapter Summary 129
CHAPTER SIX 131
THE MALAWIHIV AND AIDS ART PROGRAMME CONFORMITY WITH TRANSACTION COST ECONOMICS GOVERNANCE THEORY 131
6.1 Introduction 131
6.2 The Malawi Health Care System 132
6.2.1 Key Actors in the Health Care System 132
6.2.2 Organisation of the Health Care System and the ART Programme 134
6.3 HIV and AIDS ART Services Governance 137
6.3.1 National Level ART Coordination 139
in ART 139
6.3.3 Other Support Services to the ART Programme 140

6.4 HIV and AIDS ART Governance Conformity with Transaction Cost Economics Theory 141

6.4.1 HIV and AIDS ART Transaction Characteristics and Governance instrument alignment 143

6.4.2 Governance instrument and Efficiency in ART service delivery 147

6.4.3 Governance instrument and equity considerations to reach vulnerable populations 151

6.4.4 Observed Gaps and deviation from TCE in HIV and AIDS ART governance 153

6.7 Chapter Summary 155

THE TRANSACTION COST ECONOMICS OF HIV and AIDS ART PROGRAMME IN MALAWI 157

7.1 Introduction 157

7.2 The Transaction Cost Economics Model 158

7.2.1 Governance Choice Under TCE 160

7.3 Distinguishing Production and Transaction Costs in HIV and AIDS ART Governance 166

7.4 Sources of Transaction Costs in HIV and AIDS ART Programme in Malawi 168

7.4.2 Transaction Costs from HIV and AIDS Programme activities 173

7.4.2.1 Types and Sources of Transaction Costs in the ART Programme 175

7.4.3 HIV and AIDS ART transactions and TCE theory prescriptions 190

7.5 Monitoring Transactions Costs and their Incidence in the ART Programme 191

7.6 Measures to Reduce Transaction Costs in ART Service Delivery 192

7.7 Chapter Summary 196

CHAPTER EIGHT 197

DISCUSSION OF RESULTS 197

8.1 Introduction 197

8.2 Results Discussion 197

8.2.1 HIV and AIDS ART Technical (In) Efficiency and how Transaction Cost affect Efficiency 198

8.2.2 HIV and AIDS ART Programme Service Delivery Governance Conformity 204

with TCE Theory 204

8.2.3 Sources of Transaction costs in the Malawi HIV and AIDS ART programme and compliance with TCE prescriptions 207

8.3 Research Limitations and Further Research 209

CHAPTER NINE 210

CONCLUSION 210

9.2 Methodology 211

9.3 Results 211

9.4 Policy Recommendations 213
10. REFERENCES 216
11. APPENDIXES 238

LIST OF TABLES

Table 1: Distinguishing Attributes of Market, Hybrid and Hierarchy Governance Structures............37
Table 2: AIDS to GDP Ratio for Selected Countries Heavily Impacted by AIDS in Eastern and Southern Africa............57
Table 3: OLS assumptions............95
Table 4:Variables description............97
Table 5:Multicollinearity Test Results using VIF............100
Table 6: ART efficiency summary by district............110
Table 7: ART health facility levels of efficiency............114
Table 8: ART Health facilities Efficiency by Health Zone............115
Table 9: ART Health Units Characteristics............122
Table 10: Determinants of facility-based ART service delivery Efficiency Regression............123
Table 11: Distinguishing Attributes of privatisation Regulation and Public Agency in public sector organisation............141
Table 12: Transaction Sources in ART Governance............176
Table 13:Existing frameworks to proxy HIV and AIDS ART transaction costs............191
Table 14: ART Programme Transaction Costs estimation............264

LIST OF FIGURES

Figure 1: Research Analytical Framework 45
Figure 2: Inter-temporal and Spatial ART Evolution Conception 49
Figure 3:Trend in International Donor Country Bilateral HIV and AIDS funding-2002 to 2022 55
Figure 4: Map of Malawi showing ART facilities 73
Figure 5: Research Process Flow 76
Figure 6: Facility Based ART & HTC programme cost estimation process 88
Figure 7: Spatial Distribution of ART facilities according to efficiency scores 112
Figure 8: Summary of health facility efficiency 113
Figure 9: ART Health facility returns to scale 118
Figure 10: ART health facility returns to scale by service delivery level 119
Figure 11: Health facility Returns to scale by Rural and Urban location 120
Figure 12: Facility returns to scale by type of administration 121
Figure 13: Malawi HIV and AIDS Service Delivery Mechanism 138
Figure 14: Transaction Cost Economics Model 159
Figure 15: TCE institutional choice model 161
Figure 16: Transaction Cost Curves with different Governance modes 165
Figure 17: Delineating Production and Transaction Costs in ART 167
Figure 18: Illustration of Transaction costs in ART Supply Chain process 186

APPENDIXES

Appendix 1: Sampling frame for qualitative study 238
Appendix 2: Qualitative Data Analysis Themes 243
Appendix 3a: Estimating Transaction Costs in the Malawi ART Programme 244
Appendix 3b: Hastings-Metropolis MCMC Algorithm 268
Appendix 4: Second Stage DEA output……………………………………………………………...269

ACRONYMS

AGYW	Adolescent Girls and Young Women
AIDS	Acquired Immuno-deficiency Syndrome
ANC	Antenatal Clinic
ART	Antiretroviral Therapy
ARV	Antiretroviral drugs
BoD	Burden of Disease
CHAM	Christian Health Association of Malawi
CMST	Central Medical Stores Trust
COVID-19	Corona Virus Disease 2019
CRS	Constant Returns to Scale
CSO	Civil Society Organisation
DAH	Development Aid for Health
DEA	Data Envelopment Analysis
DEVPOL	Statement of Development Policies
DFID	Department for International Development
DHAMIS	Department of HIV and AIDS programme Management Information System
DHIS 2	District Health Information System
DHO	District Health Office
DHSS	Director of Health and Social Services
DLT	Decentralized Ledger Technology
DMU	Decision Making Unit
DRS	Deceasing Returns to scale
EGPAF	Elizabeth Glazer Paediatric Foundation
EHP	Essential Health Package
EMR	Electronic Medical Records
FY	Financial Year
GAM	Global AIDS Monitoring
GAVI	Global Alliance on Vaccine Initiative
GDP	Gross Domestic Product
GFTAM	Global Fund to Fight AIDS and Malaria
GoM	Government of Malawi
HBP	Health Benefit Package
HAD	HIV Diagnostic Assistant
HEARD	Health Economics and AIDS Research Division
HAS	Health Surveillance Assistant
HCAC	Health Centre Advisory Committees
HCMC	Health Centre Management Committee
HIV	Human Immuno-deficiency Virus
HTC	HIV Testing and Counselling
HTS	HIV Testing Services
ICT	Information Technology & Communication
ILO	International Labour Organisation
IDA	International Development Association

IMF	International Monetary Fund
IRS	Increasing Returns to scale
LIMC	Low and Middle Income Countries
m2m	mother2mother
MACRO	Malawi AIDS Counselling Resource Organisation
MAM	Moslem Association of Malawi
MDA	Ministries and Departmental Agency
MLE	Maximum Likelihood Estimation
MoH	Ministry of Health
MPF	Malawi Partnership Forum
MSF	Médecins Sans Frontières
NAC	National AIDS Commission
NASA	National AIDS Spending Assessment
NFPPA	Not-for-profit –private agency
NIMART	Nurse Initiated Management of Antiretroviral Therapy
NGO	Non-Governmental Organisation
NHRL	National HIV and AIDS Reference Laboratory
NIE	New Institutional Economics
NMT	Nurse Midwife Technician
NSF	National Strategic Framework
NSO	National Statistical Office
NSP	National Strategic Plan
OIE	Old Institutional Economics
OOPE	Out of Pocket Expenditure
OPC	Office of the President and Cabinet
OTE	Overall Technical Efficiency
PEPFAR	President's Emergency Preparedness for AIDS Response
PITC	Provider Initiated Testing and Counselling
PIU	Project Implementation Unit
PLWA	People living with AIDS
PLWHIV	People Living with HIV
PMTCT	Prevention of Mother to Child Transmission
PrEP	Pre-Exposure Prophylaxis
PSI	Population Services International
PTE	Pure Technical Efficiency
RM	Resource Mapping
SACS	State AIDS Control Society
SAP	Structural Adjustment Programme
SDGs	Sustainable Development Goals
SE	Scale Efficiency
SLA	Service Level Agreement
SRHR	Sexual and Reproductive Health and Rights
STI	Sexually Transmitted Infection
TB	Tuberculosis

TCE	Transaction Cost Economics
THE	Total Health Expenditure
TWG	Technical Working Groups
UBR	Universal Beneficiary Register
UKAID	United Kingdom for International Development
UNAIDS	United Nations AIDS Commission
UNICEF	United Nations Children's Fund
USAID	United Sates Agency for International Development
VHC	Village Health Committees
VMMC	Voluntary Male Medical Circumcision
VRS	Variable Returns to Scale
WHO	World Health Organisation

ABSTRACT

The gains from antiretroviral therapy have been remarkable over the years, but the global decline and flattening of aid has heightened the concerns for the sustainability of the diagnosis and treatment programmes, and the global HIV and AIDS response in general. Consequently, the heavy dependence on external financing, and limited scope for increased domestic funding in low- and middle-income countries heavily impacted by HIV and AIDS has heightened calls for efficient governance of the response. Concurrently, the ART services embedded in the public healthcare governance show inefficiencies, with high transaction or coordination costs as a key factor contributing to inefficiency. Conventional healthcare economic evaluation approaches have tended to ignore the role that transaction costs minimisation can potentially play towards efficient ART service delivery and contribute towards sustainability. Using a Transaction Cost Economics (TCE) lens, the aim of this research is an analysis of the HIV and AIDS ART governance in Malawi to identify the types of inefficiency, sources of transaction costs, their effect on ART programme efficiency, and the ART programme governance conformity with TCE theory. The specific objectives of the research are: a) to assess the (in) efficiency of the HIV and AIDS ART programme in Malawi and determine the types of inefficiency in the programme; b) to assess the effect of transaction costs on the HIV and AIDS ART programme efficiency in Malawi; c) to assess the degree to which the HIV and AIDS ART programme in Malawi conforms with the TCE economic governance theory, and, d) to determine the sources of transaction costs in the HIV and AIDS ART programme in Malawi, and determine whether these costs are due to non-compliance with the prescription of transaction cost economics theory.

The study used quantitative and qualitative approaches to collect data and analyse the application of TCE in the Malawi HIV and AIDS ART programme governance. Inefficiency estimation uses data envelopment analysis (DEA), a non-parametric linear programming efficiency estimation technique, econometric and statistical methods based on data extracted from the Ministry of Health Electronic Management Records (EMR). Qualitative data were collected from HIV and AIDS policy and programme managers, and service providers, which was analysed using thematic analysis to report on the key features of the HIV and AIDS ART programme governance in Malawi and its conformity with TCE theory.

The results show that the HIV and AIDS ART programme in Malawi operates at inefficient levels with an estimated overall technical efficiency score of 92%, suggesting output can be increased by 8% with existing inputs. It displays increasing and decreasing returns to scale for 23% and 69% of the health facilities providing ART, suggesting that the facilities are operating at too small and too big a scale to be able to take advantage of economies of scale, respectively. Some border district facilities display high technical efficiency levels, largely due to high client volumes, but this may hide the strain on service delivery due to extension of services to foreign country citizens. The results further show that transaction costs are negatively correlated with health facility efficiency, suggesting that their minimisation should contribute to efficient facility-based ART provision in Malawi. Additionally, the design and implementation of the ART programme is found to conform with TCE theory prescriptions, and harbours the governance instruments designed to reduce transaction costs that arise from ART transaction asset specificity, uncertainty and complexity. The main sources of transaction costs are programme coordination and administration and ART financing administration. An excessive ART aid administration safeguards framework by donors and a parallel healthcare supply chain management structure tend to increase transaction costs, and thus, violate the TCE theory efficiency principles.

While ART governance design conforms to TCE theory prescriptions, the ART programme is assessed to be inefficient. Policy interventions that increase the use of excess capacity and align demand with facility allocation should assist to improve HIV programme efficiency in Malawi. An explicit focus on the sources and incidence of transaction costs would further assist to devise measures to minimise them and contribute towards efficient HIV and AIDS ART service delivery.

CHAPTER ONE

INTRODUCTION

1.1 Introduction

The HIV and AIDS epidemic remains one of the most prominent health challenges globally. An estimated 39.0 million people are infected with HIV, with the Eastern and Southern Africa region the most heavily impacted by the epidemic as it accounted for 54 % of people living with HIV globally in 2021 (UNAIDS, 2022a; UNAIDS, 2022b). In Malawi, although the HIV and AIDS disease burden declined by 12.35% between 2010 and 2019, it remains the second highest cause of disability life adjusted years after maternal and neonatal mortality (Institute of Health Evaluation Mterics, 2022; Government of Malawi (GoM), 2023a). Based on previous population-based estimates, the epidemic is characterised as generalised in the heterosexual population (GoM, 2020). Overall, 1.1 million people are estimated to be living with HIV in a country of approximately 21 million people. While the response to HIV and AIDS has seen an increase in financial resources since 2001 globally, recent trends show stagnation and a decline in funding available to fight the epidemic (UNAIDS, 2023). For Malawi, this trend implies particular vulnerabilities as over 97% of the financial resources for HIV and AIDS are from external sources (GoM, 2022). Furthermore, Malawi has a high HIV and AIDS financial resources requirement to Gross Domestic Product (GDP) ratio, estimated at 0.41 in 2016 (Atun, Chang, Ogbuoji, *et al*, 2016). This is due largely to the limited capacity to generate sufficient domestic resources to finance the fight against the epidemic.

This research focuses on the management or governance of HIV and AIDS ART services and starts from the premise that the current governance of these services is inefficient, largely due to a limited focus on the application of transaction costs economics theory that can contribute to efficient ART service delivery efficiency. Achievement of efficiency in the delivery of general health and HIV and AIDS services has become a key theme as countries imapcted by HIV and AIDS begin to increasingly look to themselves to finance the response in the face of dwinlding global HIV and AIDS development assiatnce (GoM, 2023a; GoM, 2023b). As a result, the governance

of general healthcare and HIV and AIDS governance that aim at cost minimisation in the face of declining development aid for HIV and AIDS (Kharsany & Karim, 2016) has become an important area of focus. In particular, HIV and AIDS governance mechanisms that reduce transaction costs have recently received growing attention (Khanakwa & Mbonigaba, 2022, 2024). Transaction costs are the costs incurred when carrying out economic exchanges or performing a transaction or unit of activity (Wallis and North, 1986; Coase 1937; Lawson, 2009). They include the information costs of searching for potential transaction partners, costs for defining and negotiating contracts, and enacting and enforcing contracts. In the context of HIV and AIDS service delivery, these costs may be internal, arising from the interactions between managers and employees within the agency delivering HIV and AIDS services, or external as the agency interacts with end users and funders or HIV and AIDS aid agencies. Under HIV and AIDS service delivery, these costs may extend to the costs for monitoring and compliance, research and financial audits for HIV and AIDS programmes.

This chapter introduces and provides a background to and rationale for the research. It highlights the global and Malawi country HIV and AIDS burden and efforts at responding to the pandemic. It further introduces the Transaction Cost Economics (TCE) theory and how it relates to health and HIV governance. The chapter also presents the research aims.

1.2 Global Burden of HIV and AIDS and Response

The HIV and AIDS epidemic has been the most widespread and prolonged of the global epidemics that the world has had to contend with. Since AIDS was discovered, an estimated 75 million people have been infected, of which 35 million have already died. In 2022, there were 39 million people living with HIV and AIDS globally. Of these, 54 percent were in Eastern and Southern Africa (UNAIDS, 2022a, 2022b). Women and girls bear the brunt of epidemic, as they account for nearly 53% of adults living with HIV and AIDS globally. Although total new infections have declined globally, they remain high,with 1.3 million people newly infected with HIV in 2022 (UNAIDS, 2023). Further, since the disease is largely spread through sexual intercourse among the heterosexual population in Eastern and Southern Africa, demographic trends are likely to be a key factor in the perpetuation of the epidemic. This is because the epidemic in this

region is largely perpetuated through unprotected heterosexual sex in the face limited access to protection measures (Atun, *et al*, 2016). These trends confirm the continuing global health challenge that the epidemic poses despite current efforts to contain it.

The past forty years since the discovery of HIV and AIDS have been characterised by considerable global resource mobilisation and efforts to fight the epidemic. Global and country level responses to HIV and AIDS have seen the number of people accessing ART worldwide increase from just 300,000 in 2003 to 29.8 million in 2022 (UNAIDS, 2023). This is due to scientific advances in the biomedical field and the mobilisation of funding to significantly increase the reach of ART services. These advances heightened optimism about the end of HIV and AIDS as a global public health threat by 2020, and inspired the ambitious Fast Track agenda championed by the UNAIDS (UNAIDS, 2014). The UNAIDS (2023) shows that over the thirteen years since 2010, the number of people accessing antiretroviral therapy nearly quadrupled globally from 7.7 million to 29.8 million by December 2022. Deaths due HIV and AIDS have declined by 69% since they peaked in 2004, and by 51% since 2010 (UNAIDS, 2023). This is all thanks to treatment scale up efforts by international funders and governments in countries heavily impacted by HIV and AIDS (Menzies, *et al*, 2012; UNAIDS, 2023).

From the forgoing text, it is clear that AIDS related deaths show a declining trend with increasing numbers of people that are accessing ART globally. Nevertheless, 9.2 million people are still reported not to have access to the life prolonging ARVs globally as of 2022 (UNAIDS, 2023). This suggests the need for a sustained response, especially in keeping people already on ART on the life prolonging drugs and bringing others that are newly infected on stream. Thus, while the gains made so far in the fight against HIV and AIDS are widely acknowledged, it is also realised that the job is far from over as the epidemic continues to be a global health threat. This is further confirmed by the new 95-95-95 targets[1] that aim to end AIDS as a health threat by 2030. The new targets have implications for the availability of financial and other resources to increase and sustain current efforts, and the achievement of the targets for countries heavily impacted by the epidemic. Consequently, each year, governments and their funding partners have to set aside resources for continued provision of HIV and AIDS services. These AIDS funding

[1] Diagnose 95% of all people living with HIV, 95% of the HIV positive people start and are retained on ART, achieve viral suppression for 95%of all HIV positive people.

obligations represent a fiscal liability, measured in terms of HIV and AIDS resource requirements to a country's gross domestic product. This ratio has been termed the HIV and AIDS- to- GDP- debt ratio (Atun *et al*, 2016). Among the countries in Sub-Saharan Africa that are heavily impacted by the epidemic, Malawi records the highest AIDS–to-GDP-debt ratio at 0.4. Further, Malawi's National HIV and AIDS Strategic Plan covering the period 2015 to 2020 estimated external aid requirement for the response at 8% of its GDP (GoM, 2015), making it the country with the highest external aid requirement to fight the epidemic in the world at the time. This trend does not appear to have changed.

1.3 The HIV and AIDS Burden and Response in Malawi

Malawi is listed among HIV and AIDS hyper-endemic countries alongside Botswana, Lesotho, Mozambique, Namibia, South Africa, Swaziland, Zambia and Zimbabwe (Stover, *et al*, 2016; GoM, 2020). The country has a generalised epidemic in the heterosexual population (Go M, 2020). The first AIDS case in Malawi was documented in 1985. As with other countries heavily affected by the epidemic, HIV and AIDS has evolved to become one of the major diseases burdening society in Malawi. The HIV incidence is high among key populations, with an estimated 7 % of the 33,000 new infections in 2019 recorded among female sex workers, bisexual men and men who have sex with men (Government of Malawi, 2020).[2]. Further, the co-habiting discordant couples previously thought stable appear to be a key risk group, contributing nearly half of all new infections (Go M, 2014). Although the HIV and AIDS outcomes in the country show improvement, the epidemic is one of the leading causes of Disability Adjusted Life Years (DALY) (GoM, 2023a). It accounts for 12.9 percent of the burden of disease (BoD) in Malawi, second only to maternal and neonatal diseases that account for 13.75 percent of the BoD in the country. Other leading causes of DALY in Malawi include respiratory infections and TB (11.75%), Non-Transmittable Diseases and malaria (7.77%) and Enteric infections (7.12%). Not surprisingly, HIV and AIDS accounts for a correspondingly large share of the overall public health spending in the country. The Ministry of Health (GoM, 2022) finds that the share of HIV and AIDS expenditure is 40 percent of the total health expenditure. According to the NASA (GoM,

[2] Includes police officers, primary and secondary school teachers, fishermen, estate workers, female cross-border traders, sex workers, men wo have sex with men and truck drivers

2021), the main funders of the response in Malawi currently are the Global Fund to Fight Tuberculosis, AIDS and Malaria (GFTAM) (54%), Government of Malawi (27%) and the United States Government (18%). The persistence of the epidemic implies the need for current efforts to fight it to continue, and that resources will be required to fund the response. In the face of declining global financial resources available for HIV and AIDS, achievement of efficiency has thus become a key imperative, as more has to be achieved with what is available in financial resources. But there are real ART sustainability concerns due to the changing global funding landscape for HIV and AIDS. More concretely, the Global Fund and PEPFAR, two of the largest funders of HIV and AIDS in Malawi, are likely to reduce funding over the coming years under the expectation that low and middle income countries (LMICs) like Malawi increase the domestic share of HIV funding. While policy discussions are currently under way on how to sustain the response in Malawi, including use of innovative financing to finance the response, they appear to be inconclusive.

In Malawi, the first National HIV and AIDS Policy was developed in 2003. The policy, which provided guiding principles for HIV and AIDS programmes and interventions in the country, was reviewed and a new policy prepared in 2012. In 2005, a Department to deal with HIV and AIDS and Nutrition issues was created in the Office of the President and Cabinet (OPC). The department was to be responsible for AIDS policy guidance and resource mobilisation. A more recent development likely to impact the organisation of HIV and AIDS governance is the folding of both the Department of Nutrition and HIV and AIDS (DNHA) and the NAC under the Ministry of health. Although it is expected that the NAC and the DNHA will retain their roles under the new governance arrangements, this change implies that public policy in the country regards AIDS as a disease that can be controlled largely through biomedical means. This may also have implications on the way the response is coordinated, as well as on funding allocation levels for HIV and AIDS programmes.

The HIV and AIDS prevalence in people aged 15-64 years has steadily declined from 16.4% in 1999 to 9.2 % in 2022, but remains high at around 8% (GoM, 2017a; GoM, 2020). The overall rate of new HIV infections per year has shown a declining trend, but is still high. New HIV infections per year have declined from a maximum 111,000 in 1992 to an estimated 33,000 in 2019 (GoMand UNAIDS, 2019). While efforts at the

establishment and expansion of the antiretroviral therapy (ART) program and its consequent positive effects are acknowledged, nearly 3% of those eligible and in need of ART do not access it in Malawi (GoM, 2022). Further, according to the Spectrum model estimates, 93% of people on ART were estimated to have attained viral suppression, while 94% of PLHIV knew their status (GoM, 2022). It is notable that Malawi has met one of the 95-95-95 targets (with 97% of people that are HIV positive enrolled on ART), which is likely to have contributed to the decline in new infections.

1.4 Healthcare and HIV and AIDS Governance Efficiency

Academic literature on general healthcare and HIV and AIDS governance highlights the importance of cost minimisation in the face of declining development aid for HIV and AIDS (Khanakwa & Mbonigaba, 2022, 2024; Kharsany & Karim, 2016). Consequently, there is a growing body of academic literature in the economic evaluation of health care and HIV and AIDS that focuses on efficiency analysis in health service delivery as a means to free up resources for HIV and AIDS services. These studies estimate the efficiency of facility-based HIV and AIDS service delivery (Bautista-Arredondo, 2015; Abimbola, *et al*., 2015; Lépine, Vassall, & Chandrashek, 2015). An investigation of technical efficiency for a large HIV and AIDS programme (the Avahan) in India using data envelopment analysis (DEA) showed high inefficiency levels, with scope for reduction of current inputs by 43% without affecting the level of outputs (Lépine, Vassall, & Chandrashek, 2015). This points to the possibility of freeing up resources through efficient governance. Further, in Rwanda, a technical efficiency assessment of the impact of Community Based Health Insurance (CBHI) on efficient HIV and AIDS services provision showed that CHBI improved the technical efficiency of service delivery, including total factor productivity of health facilities (Bautista-Arredondo *et al*,2015). Furthermore, some studies use the unit costs of healthcare service delivery to assess service delivery efficiency (Accorsi, et al., 2010) ; Adesina, Couture, & McGinn, 2015). These stuidies highlight the contribution of measures such as task shifting and service integration in contributing towards efficient healthcare service delivery. In addition, other studies focus on the decentralisation of health care services and how it affects efficient allocation and technical efficiency of healthcare (Cuenca, 2020; Twea, Manthalu, & Mohan, 2020). Nevertheless, most studies that focus on technical efficiency and cost effectiveness tend to use single outputs and outcomes, and ignore the

application of TCE and the incidence of transaction costs and their effects on service delivery efficiency. This means that there is a gap in the estimation and monitoring, particularly of transaction cost levels within HIV and AIDS service delivery. This gap further implies that it is difficult for policy makers and managers to control transaction costs towards efficient service delivery in the face of ever dwindling resources.

A further strand of studies specifically highlights the role of Transaction Cost Economics theory in HIV and AIDS governance of healthcare and HIV and AIDS, and the implications for efficient service delivery (Munthali, 2010; Guiness, 2011; Bautista-Arredondo *et al*, 2014; Bautista-Arredondo, *et al.*, 2018; Steinecker & Czirfuszová, 2021; Khanakwa & Mbonigaba, 2022; 2024). These studies provide evidence of comparative HIV and AIDS governance modes according to TCE theory and their implications for efficient delivery. Nevertheless, one of the shortfalls of this literature is that it has not quantified levels of transactions costs in HIV AND AIDS service delivery and their implications on programme efficiency.

Insights from this research will contribute towards closing the knowledge gap on the empirical evidence regarding the application of TCE in ART services delivery, with a focus on the role of transaction costs minimisation as a means towards efficient ART services. These insights may further have wider applicability to the broader healthcare governance framework in Malawi. Further, understanding the sources of transaction costs and their effect on the (in) efficiency of HIV and AIDS programmes will shed light on the types of organisational designs that may guarantee efficiency in service delivery. This knowledge will in turn, indicate policy and institutional actions necessary for secure an efficient framework for HIV and AIDS ART service delivery.

1.5 HIV and AIDS Governance and Transaction Cost Economics Theory

The research draws on insights from the New Institutional Economics (NIE), particularly the transaction costs economics (TCE) theory to assess the HIV and AIDS ART programme management and efficiency implications. The NIE is chosen over the neoclassical economic theory as the theoretical framework for the analysis of the ART programme governance, and the implications of transaction costs on efficiency of the HIV and AIDS programmes. This choice is because the NIE specifically focuses on the role of institutions that yield positive transaction costs in economic organisation that are

assumed away under the classical economic theory, thereby, rendering the latter unrealistic with regard to HIV and AIDS ART governance analysis. More concretely, the role of institutions does not feature as an analytical tool under the traditional or neo-classical economic theory (Williamson, 2000; Munthali, 2010; Jagwe *et al*, 2010; Fernandez, 2016). This renders neo-classical economic theory insufficient for the analysis of transaction costs and efficiency of organisations, including in HIV and AIDS governance.

According to Williamson (1979), the NIE is concerned with the *sources, incidence* and *implications* of transaction costs on efficiency outcomes of firms or decision-making units (DMUs). In private sector economic governance, Transaction Cost Economics theory suggests that the existence of positive transaction costs associated with the market gives rise to alternative modes of the organisation of transactions, such as through firms or other types of bureaucracy to facilitate exchange, production or service delivery in the quest for efficiency (Coase, 1937; Williamson, 1985; Wallis and North, 1986). Under TCE, organisation for the purpose of achievement of social goals is seen as being shaped by, among other things, the need to reduce transaction costs and achieve efficiency. Consequently, governance mode choice, whether under private or public organisation is informed by looking at modes that are capable of reducing transaction costs.

Governance structures under private ordering are distinguished between the market (resource allocation through the price mechanism), vertical integration (through firms or public bureaucracies) and hybrid modes combining elements of market and vertical integration (Williamson, 1979). Each type of governance structure (market or firm) is associated with transaction costs, and hence, a major prediction of Transaction Costs Economics theory is that the governance structure that reduces transaction costs and achieves efficiency will be preferred and adopted over others (Marinescu, 2012).

In Malawi, as in many other low-income countries, the HIV and AIDS programmes are delivered either through public sector or the public in combination with private and not-for-profit bureaucracies. But a relevant question that the research addresses is whether public governance of the ART programme is the most efficient among the available solutions. Thus, analysis of the efficiency of the HIV and AIDS ART programme is likely to show areas of efficiency and inefficiency, and might thus be useful in informing resource allocation within the national HIV and AIDS ART response.

1.6 Research Problem

Inefficiency in HIV and AIDS ART programmes is a key concern at a time when global HIV and AIDS financial assistance is declining or flattening, in the face of an enduring epidemic and growing ART financial needs. Countries with a heavy HIV burden have extremely limited fiscal space to significantly increase funding levels to respond to the epidemic. Although the data point to improved ART coverage and outcomes, the sustenance of these outcomes is critically dependent on the availability of adequate financial resources to finance the AIDS response. The continuing HIV and AIDS burden in a context of declining financial resources and high AIDS to Gross Domestic Product (GDP) debt ratios implies the need to minimise costs and achieve efficiency in HIV and AIDS service delivery. Among the key factors widely acknowledged in the literature that affect healthcare efficiency and HIV and AIDS Programme is the level of transaction costs. Currently, in Malawi, estimates of transaction costs in the healthcare service delivery where ART services are provided suggests that they are in excess of USD1.0 for every USD5.0 allocated to primary health care and HIV and AIDS. This likely impacts efficient service delivery.

Yet, despite the recognition of the role of high transaction costs and their potential contribution to the inefficiency of HIV and AIDS programme, conventional economic evaluation has ignored the specific role that transaction costs play in reducing costs and hence, contribute to efficiency. More concretely, there is limited research on how the existing ART governance modes generate transaction costs, their incidence and effect on the technical efficiency of HIV and AIDS service delivery. This has been due to the limited application of Transaction Cost Economics theory towards efficient HIV and AIDS programme management. This research contributes to closing this gap by applying TCE theory to the Malawi ART programme and its implications on (in) efficient facility-based HIV and AIDS ART service delivery.

1.7 Research Aim

The aim of the research is to analyse the Malawi HIV and AIDS ART programme using the TCE governance theory that considers the sources, incidence and impact of

transaction costs on the ART programme technical (in) efficiency. The specific objectives of the research are as follows:

a) To assess the (in) efficiency of the HIV and AIDS ART programme in Malawi and determine the types of inefficiency in the programme;
b) To assess the effect of transaction costs on the HIV and AIDS ART programme efficiency in Malawi;
c) To assess the degree to which the HIV and AIDS ART programme in Malawi conforms with the TCE economic governance theory;
d) To determine the sources of transaction costs in the HIV and AIDS ART programme in Malawi, and determine whether these costs are due to non-compliance with the prescription of transaction cost economics theory;

1.7.1 Research Hypothesis

The research hypothesis is: The (in) efficiency of the HIV and AIDS ART programme in Malawi is a consequence of inadequate application of transaction cost economics (TCE) theory, which in turn leads to limited monitoring and control of internal transaction costs to inform (in) efficiency considerations in ART policy, programme planning and implementation.

1.7.2 Research Questions

The thesis is anchored on the inefficiency problem of the HIV and AIDS ART programme in Malawi in the face of dwindling funding to support the HIV and AIDS programmes. Consequently, the study seeks to understand whether there is a lack of or limited internal administrative control and monitoring or governance as proposed by transaction cost economics theory. Further, it is concerned with whether these inadequacies in, or the lack of control lead to inefficiency in the ART programme. The central research question for the study is, therefore:

> ***Are inadequacies in the design and internal control and monitoring of costs (economic governance) in the HIV and AIDS ART programme as offered by Transaction Cost Economics resulting in the inefficiency of the programme in Malawi?***

To answer this central question, focused on the ART programme, the following sub-research questions were formulated:

a) Are HIV and AIDS programmes in Malawi inefficient, and what type of inefficiency do they display?
b) Are transaction costs affecting the overall efficiency of the HIV and AIDS ART programme in Malawi?
c) How does the Malawi HIV and AIDS ART programme governance comply with the governance theory (transaction costs economics theory)?
d) What are the main sources of transaction costs in the current governance of the HIV programme in Malawi? Are these due to non-compliance with the prescription of transaction cost economics theory?

1.8 Research Focus and agencies covered

The research focused on analysis of (in) efficiency related to the HIV and AIDS ART service delivery. The ART thematic area in the HIV and AIDS response covers treatment, care and support, while the HIV Testing Services (HTS) services are the gateway to treatment, care and support. These areas have been selected based on the share of expenditure in the Malawian National HIV and AIDS response. Except for fiscal year 2017/ 18 when it accounted for 51.8% of the total HIV and AIDS spending, the ART programme has accounted for an average of 60% of national HIV and AIDS expenditure since FY2015/ 16 (GoM, 2021). The decline in the ART expenditure in fiscal year 2017/18 is attributed to a reduction in funds available from the Global Fund to Fight AIDS and Malaria for the period (GoM, 2021). The research further focuses on key actors in either the financing or implementation of the ART services to gauge insights into the extent to which (in) efficiency considerations by means of transaction costs monitoring and control are included in the ART financing, policy and programming. The agencies chosen are the Department of HIV and AIDS in the Ministry of Health, the Central Medical Stores Trust, the National AIDS Commission, President's Emergency Preparedness for AIDS Response (PEPFAR), Population Services International (PSI), Elizabeth Glazer Paediatric Foundation (EGPAF) and mother2mother (m2m).

For ART services, the Department of HIV and AIDS in the Ministry of Health was the focus of the research since it is responsible for the overall governance of the HIV and

AIDS ART programme. Treatment and care are further likely to be the main drivers of costs in the near to long term in view of the new UNAIDS 95-95-95 treatment targets towards ending AIDS as a public health threat by 2030. Procurement, distribution of drugs and actual delivery of services thus represent large expenditure items. Hence, although the Central Medical Stores Trust (CMST) is not directly responsible for the acquisition of all of the HIV and AIDS response drugs requirements, it is included because it is the official Government agency responsible for the healthcare supply chain governance in the national health system. Moreover, until 2019, it provided the storage of all the ART drugs and commodities in the country[3]. Population Services International (PSI), now Family Health Services (FHS), was the focus of the research as it is heavily involved in the acquisition of test kits and delivery of HIV Testing and other prevention interventions. The PEPFAR was included because it is one of the largest donors in the Malawi HIV and AIDS service delivery, while EGPAF was added due to the support it receives from PEPFAR for the ART programme, including data management for the electronic management records for the ART programme, as well as with respect to Expert Clients who support the provision of RT services, including follow up of defaulting clients. M2M was included to gain insights into the support provided to selected ART sites, including the role of Expert Clients.

Research Contribution

This research builds on previous empirical and academic evidence on HIV and AIDS governance using Transaction Cost Economics where cost minimisation towards efficient service delivery is a key aspiration. While previous research has explored various aspects of ART programme implementation, linking transaction costs with programme technical efficiency has remained a gap. This study contributes to the literature on HIV/AIDS ART programme governance by examining the relationship between transaction costs and the technical efficiency of ART service delivery. Using Malawi as a case study, this study bridges this gap by analysing how transaction costs influence the technical efficiency of ART programmes. Further, by assessing the extent to which ART programme governance conforms to TCE theory, this study extends the application of TCE in the HIV and AIDS governance, which may have further applications to the broader health sector, providing insights into optimising ART service delivery in Malawi and similar resource-constrained settings.

[3] Key Informant Interview, July 2020

1.9 Thesis Organisation

After this introduction, the thesis is organised into 8 chapters comprising of a general literature review, an overview focused on the global HIV and AIDS the changing funding landscape, a methodology chapter and three analytical chapters (Chapters 5 to 7) that focus on addressing the thesis research questions. It further includes a results discussion and a conclusion chapter.

Chapter 2 presents a narrative literature review which provides the conceptual and theoretical background, and sets the analytical basis for the research. It covers the New Institutional Economics theory, focused on Transaction Cost Economics as the theoretical framework for the research. It further presents empirical literature on healthcare and HIV and AIDS efficiency analyses in Malawi and other countries. The chapter further describes the analytical framework and the scope of the research.

Chapter 3 aims at describing the global HIV and AIDS financing changing landscape and how it imposes an efficiency imperative on HIV and AIDS governance on low and medium income countries heavily impacted by HIV and AIDS to secure a sustainable response.

Chapter 4 describes the methodology for the study, justifying the multi-method approach designed to answer each research question and objective more comprehensively. In particular, the research uses both qualitative and quantitative data collection and analysis methods. Data were collected from different sources, including the Ministry of Health HIV and AIDS electronic medical records (EMR), consultations with health policy and programme managers and service providers involved in HIV and AIDS governance at different levels.

Chapter 5 addresses the research question on whether the HIV and AIDS ART programme in Malawi is inefficient or not, and what types of (in)efficiency it displays. The chapter further addresses how transaction costs affect the efficiency of the HIV and AIDS ART programme. It uses a two-stage data envelopment analysis (DEA) that combines non-parametric efficiency estimation and regression analysis. The first stage is a non-parametric linear programming based efficiency analysis that is used to estimate the overall technical (in) efficiency of the programme, as well as the individual health

facility technical (in) efficiency. The DEA estimates technical, allocative and scale efficiency, which are used to classify the types of inefficiency displayed by the ART programmes. The effect of transaction costs on ART facility service delivery (in) efficiency is achieved via econometric analysis using the (in) efficiency analysis results from first stage DEA as the dependent variable, regressed on transaction costs and other control variables that include the period a facility has been offering ART services, type of ownership (private, public, private not-for-profit or for profit), location and level of service delivery (primary, secondary and tertiary).

Chapter 6 reports on the conformity of the HIV and AIDS ART programme governance instruments in Malawi. It answers the research question on the HIV and AIDS ART programme governance conformity with the Transaction Cost Economics governance theory. It seeks to highlight whether the inefficiencies in the ART programme are the result of non-conformity with TCE theory. It is a qualitative analysis based on literature review and interviews with actors in ART policy and service delivery. It assesses whether or not the design and implementation of the ART programme governance instruments, including public agency and partnership arrangements comply with TCE in terms of minimising transaction costs due to asset specificity, uncertainty and complexity in ART transactions.

Chapter 7 addresses the research question on the sources and incidence of transaction costs in the HIV and AIDS ART programme in Malawi. This is done via a qualitative analysis of the Transaction Cost Economics of the programme to detail the rationales for governance instrument choice and programme features that unveil the transaction costs sources, incidence and their implications on (in) efficiency in the HIV and AIDS ART service delivery.

Chapter 8 highlights and discusses the main findings of the research. It summarises how each research question has been addressed, as well as the implications for policy regarding HIV and AIDS ART governance in Malawi. Chapter 9 concludes the thesis and presents policy recommendations emanating from the research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter provides the conceptual and contextual background to set the scene for the concepts used in the thesis and their possible linkages, which will help to understand the results and empirical analysis presented in later chapters. Unlike chapter one, which sets the scene for the entire thesis, the chapter presents the contextual and conceptual background to understand the analysis in subsequent chapters. This literature review aims, firstly, to contribute towards addressing a gap in the literature by focusing on the nature and sources of transaction costs in HIV and AIDS service delivery and how they affect efficiency. Although transaction costs are acknowledged in HIV and AIDS management, there is limited application of transaction costs economics in HIV and AIDS and health management (Khanakwa & Mbonigaba, 2022). A transaction cost economics analysis of the HIV and ADS governance would thus contribute to an understanding of where attention needs to be focused to achieve efficiency in the HIV and AIDS programme. Secondly, the literature review extends existing literature on the characteristics of transaction costs in HIV and AIDS governance that may have wider relevance to health governance in resource constrained settings such as Malawi. In addition, the literature review will add to the literature on the efficiency of HIV and AIDS programmes.

The chapter starts with a general introduction, which is followed by an overview of New Institutional Economics (NIE). This extends to a presentation of a synopsis of the main strands of the NIE, before ultimately presenting the Transaction Cost Economics (TCE) theory, one of the NIE branches, as the theoretical framework for the research. The review further highlights the relevance of the TCE in the analysis of HIV and AIDS governance. Empirical evidence and the gaps on efficiency of healthcare, HIV and AIDS service delivery and the application of a TCE lens to HIV and AIDS governance is further presented. A summary of the chapter concludes.

2.2 Economic governance and Transaction Costs Economics

The goal of economic governance or exchange is achievement of efficiency in the way transactions are coordinated. This goal is reflected in governance choice or preferences in the modes of economic transactions, with organisational modes that are calculated to be efficient by economic agents preferred over those deemed less so (Williamson, 2010). Currently, there is an efficiency imperative in HIV and AIDS economic governance owing to the flattening and dwindling development aid for HIV and AIDS. In particular, organisational forms that facilitate delivery of HIV and AIDS services need to be subjected to the efficiency economic criterion given the substantial resources that HIV and AIDS attracts within the national health budget allocations, as well as the growing need to achieve more with less. This is further against the background that commitments already made, such as in the area of ART provision need to continue in the face of declining global HIV and AIDS assistance.

The current research uses Transaction Cost Economics theory, which is located within the broader frame of New Institutional Economics (NIE) to analyse the economic governance of the HIV and AIDS ART programme in Malawi. Transaction Cost Economics (TCE) is one of the approaches that can potentially be used to address efficiency issues, but appears to have been applied only sparingly to health governance in general, and HIV and AIDS governance in particular (Khanakwa & Mbonigaba, 2022). This reflects a knowledge gap in health economics research. Transaction costs are relevant to research in Malawi because they are a potential source of efficiency gains, especially considering that over 80% of funding to the response is externally resourced through development assistance for HIV and AIDS. This external funding comes with additional layers of monitoring in the form of safeguards, which tend to increase transaction costs. In terms of economic governance, the NIE in general, and the TCE in particular departs from the neoclassical economic model of the firm through the introduction of positive transaction costs and a relaxation of the neoclassical economics rigid assumption of costless transactions in exchange, perfect rationality and complete information among economic actors. Whereas the neo-classical economic theory of the firm does not take into account institutions, the New Institutional Economics (NIE) actively integrates the role of institutions in economic governance analysis. With particular reference to transaction costs, Transaction Cost Economics (TCE), a branch of the New Institutional Economics (NIE), posits that all feasible organisational forms (market, hybrid, private bureaucracy, public bureaucracy) are flawed (Williamson,

1995). Consequently, each organisational mode should be presumed efficient until the transaction costs associated with its attributes during both the *ex-ante* and *ex-post* periods have been comparatively assessed, in line with the nature of the transactions that are performed. In Malawi, the *de facto* governance choice for HIV and AIDS services appears to be the public bureaucracy and regulation (out-sourcing or hybrid), with public health facilities as well as private-for-profit and private-not-for-profit as key sites of ART service delivery. This means that at the outset the current ART governance framework should be presumed efficient until an efficiency analysis is completed, considering the mandate and financial, human, capital resources available to the public bureaucracy that delivers ART.

2.3 Key Concepts

There are a number of key concepts that underlie economic governance that the research uses. They include institutions, governance, transaction, transaction costs and efficiency. These concepts are briefly reviewed below.

Transaction

At the centre of the NIE and TCE analysis is the concept of a *transaction*. An early conception of *a transaction* is offered by Commons (1932) who defined a transaction as "a unit of activity—a transaction, with its participants, J.R. Commons, 1932, p 4". Williamson (2010) has suggested that the Transaction Cost Economics theory is premised on this definition of a transaction. This definition highlights the importance of an exchange activity and the role of economic agents in the activity. It also implies the need for a mechanism to coordinate the activity and its participants, and hence, some form of rules to mediate the exchange to resolve conflict and achieve mutual gain (Williamson, 2010). Wander (2013) has referred to a transaction as the transfer of property rights from one party to another. Transactions can be simple, such as spot markets, or complex requiring formal contracts that take time to prepare, negotiate and conclude, and may involve long term relationships between economic agents that also require monitoring and enforcement (Vasiliauskiene and Snieska, 2009). This study adopts the definition of a transaction as a unit of activity with its participants in line with the Commons (1932), as it can be applied to different types of economic activity, including the arrangements for coordinating health and HIV and AIDS service delivery

activities, and their cost and efficiency implications. The requirement for means to coordinate the transaction is kept in view (Williamson, 2010). Thus, it has been suggested that for transaction economising purposes, there are three critical dimensions of a transaction- *complexity, asset specificity* and *disturbances* to which the transaction is subjected (Williamson, 2010). These characteristics are further discussed below.

Institutions

Institutions are defined as a set of formal and informal rules of conduct that guide coordination of relationships between individuals or groups to achieve social goals (Kherallah and Kirsten, 2002; Gräbner & Ghorbani, 2019). Formal rules include laws and regulations such as contracts, which govern the political, economic and social life of a society, which extend to the legal system, political system and markets. Informal rules include norms, traditions, customs, shared values and religion which govern social behaviour (Kherallah and Kirsten, 2002). The implicit purpose of institutions is thus to guide social interactions to achieve common goals (Kherallah and Kirsten, 2002; Hodgson, 2006)). These institutions influence the design of governance mechanisms by imposing restrictions on economic agents, leading to the creation of transaction costs in economic exchange. They shape the content of organisations to deal with societal issues (Jan *et al*, 2007; Dickson and Buse, 2008; Boyce, *et al*, 2011; Guinness, 2011; Parker, 2012). They may further influence technology choice through specification of standards. Property rights, or the rules governing alienation and acquisition of property shape how transactions ought to be carried out between individuals or organisations in a society (Birner and Wittmer, 2004; Easter and McCann, 2010). These rules may impose costs that do not contribute directly to the physical production of goods or provision of services via negotiation and ex *post* compliance monitoring. In the context of HIV and AIDS services, these 'rules of the game' may be considered at the agency, country and international levels.

Williamson defines institutions in terms of means to achieve order and mitigate conflict in economic exchange to achieve mutual gain (Williamson, 1985). This implies a pre-agreed framework that sets the parameters for engagement among transactors or participants to achieve mutual gain in economic exchange. This definition is consistent with North (1990) who defines institutions as "the 'rules of the game', the humanly devised constraints that structure political, economic and social interactions. A common

thread running through the definitions is the existence of rules to guide exchange or social relationships. When considered in the context of coordination of transactions, the institutions play the key role of setting the rules of the game to mediate conflict and facilitate societal gain. In the context of health or HIV and AIDS service delivery, these rules govern how the services are delivered as well as the coordination of resources, including labour and finances to achieve efficiency

This study adopts the notion of institutions as humanly devised constraints or rules to govern relationships for the achievement of socially desired goals. This is considering that in the context of HIV and AIDS, there are rules at the macro level governing global and national health governance to deliver HIV and AIDS services. Similarly, at the more micro level, different economic agents engage with their partners using formal agreements or contracts that define the rules of engagement. In addition, service delivery between service providers and ART clients adheres to prescribed standards. These rules extend to those for the acquisition of inputs to deliver the services, including human resources. This facilitates analysis of how the HIV and AIDS services are coordinated and what the rules imply for transaction costs in ART governance.

Governance and Governance Mechanism

An additional concept under the research used to inform this analysis is *governance.* Governance as used in the institutional economics literature refers to the means for coordinating resource allocation to guarantee their most efficient use in the production and exchange of goods and services (Williamson, 2000; Wallis and North, 1986; Edwards & Guilfoos, 2020). A governance mechanism is defined as a mode chosen to coordinate transactions (Coase 1937; Williamson, 1981; Tadelis and Williamson, 2012). The market is one mode of governance and uses the price mechanism (Williamson, 2010). Transactions may also be organised through non-market means to coordinate transactions or allocate resources towards achievement of efficiency. The price mechanism is the primordial mechanism of efficient resource allocation where transactions are organised through the market system. The non-market means of organizing transactions or resource allocation represent *hierarchy* as an alternate mode of economic governance to markets (Williamson, 2011; Zenger *et al*, 2011). In contrast to the market, the operation of private firms in an economy uses orders, rather than the

price mechanism, by managers to deploy resources within the firm to uses that are calculated to yield maximum value for the owners of the firm. Similarly, non-private sector entities such as governments, religious agencies and not-for-profit organisations engage in economic activities involving deployment of resources through fiat. In the context of organisation for HIV and AIDS services delivery, governance refers to particular modes of organisation such as public, private and not-for-profit bureaucracies designed to guarantee efficient or competent delivery of the services. This definition facilitates analysis of transaction coordination through the public bureaucracy and its implications for transaction costs and efficiency and/ or competency implications in HIV and AIDS service delivery.

Transaction Costs

In general, *Transaction costs* are the costs incurred in concluding an economic exchange or performing a transaction or unit of activity (Coase 1937; Williamson, 1981; Wallis and North, 1986; Vandeninden and Paul, 2012). While there are costs of production and distribution, completing the exchange or trade involves costs. These costs include the costs of finding partners, looking for information, negotiating deals, contract preparation, and monitoring and enforcement of contracts (Coase, 1937; Williamson, 1981; Wallis and North, 1986; Zenger *et al*, 2011; Tadelis & Williamson, 2012). This definition implies that beyond production costs in an exchange or transaction, there are other costs related to the effort exerted to find partners and complete the transaction or units of activity. In analysing transaction costs in the context of aid delivery, Lawson (2009) has suggested that: "aid transaction costs might be defined as the costs necessary for an aid transaction to take place but which add nothing to the actual value of that transaction", (Lawson, 2009, p.10)". This definition points to how transaction costs can be observed empirically. Transaction costs are important because they vary irrespective of the competitive price of a good or service (Robins, 1987), and because they add to the costs of production and distribution of services or products, and, hence contribute to economic efficiency outcomes. The definition further reinforces the notion that while activities such as monitoring service delivery standards and targets are useful to services such as ART, they do not add value to the service itself. This implies that transaction costs have an important role in efficient HIV and AIDS service delivery. There is no unified theory on what transactions costs are and how they are measured under the New

Institutional Economics (Wang, 2003). Nevertheless, this study adopts the stance that irrespective of whether they are measured at the margin or viewed in their totality for a given governance mode, they are ultimately cumulative and add to the costs of production and exchange or delivery of a product (Williamson and Tadelis, 2012). Moreover, the ultimate unit cost of delivery of a particular product or service such as HIV and AIDS ART reflects the total costs of the transaction, that is, the total of production and transaction costs. This has implications on the efficiency of the delivery of the service or product.

Economic Efficiency

Economic efficiency refers to the optimal allocation of resources such that goods and services are produced, or exchange is achieved at the lowest possible cost while maximising output and satisfaction, ensuring that there is no waste of resources (Pindyck & Rubinfeld, 2018). This definition captures the first order economising principle regarding efficient use of resources. In economic governance literature, *efficiency* is commonly considered as consisting of technical and allocative efficiency (Coelli, 1996). Technical efficiency is the ability of a firm to derive maximum output from a set of inputs (Farrel, 1957; Thomas & Callan, 2007; de Ferranti *et al*, 2008). Technical efficiency is achieved through economising on resources used in production. It is achieved when output (quantity and quality) is attained from a minimum deployment of inputs such as labour, capital and technology (Thomas & Callan, 2007; de Ferranti *et al*, 2008). On the other hand, allocative efficiency is defined as a firm's ability to deploy its production inputs in optimal proportions given their prices (Farrel, 1957; Thomas and Callan, 2007). This conception is particularly relevant for HIV and AIDS service delivery considering the dwindling global financial resources. There is a further debate as to whether the analysis of efficiency is appropriate for public sector bureaucracy. This is because the focus of transaction costs economics and efficiency analysis has largely been on the private sector bureaucracy under the theory of the firm (Pérard, 2009; Friese, Heimeshoff, & Klein, 2018). However, Williamson (1999, 2010) suggests that analysis of efficiency in public sector bureaucracy is justified on the basis of both *cost* and *competence* as some transactions are best suited for public sector bureaucracies, while others are not. The implication is that once the issue of competence of a bureaucracy or governance mode to deliver a specific mandate is established, it is plausible to assess its efficiency based on the cost criterion and *vice versa*.

2.4 Theoretical Framework

2.4.1 The New Institutional Economics

As noted before, the study adopts TCE as the theoretical framework for the analysis of ART governance in Malawi. The TCE is branch of the New Institutional Economics. The NIE seeks to expand neoclassical economics[4] to include the role of institutions or social norms and rules that underpin economic activity (Obińska-Wajda, 2016; Todorova, 2016). Inspired by Ronald Coase's seminal paper on the 'The Nature of the firm' (Coase, 1937), there has been a growing body of literature on the importance of institutions in economic organisation (Ritchter, 2005;Williamson, 2010; Vitola & Senfeld, 2015; Ménard, 2018). This has resulted in the integration of transaction costs economics into the analysis of the theory of the firm – its existence, boundaries and internal organisation (Foss and Klein, 2008; (Deitrich, Krafft, & McHardy, 2015).

The NIE argues that societal rules shape economic organisation and should be included in the analysis of organisational performance as they produce positive transaction costs (Coase, 1937; Trachtman, 1997). Rather than replace the neoclassical model, as was the aim of the old institutional economics, the NIE seeks to enhance economic governance analysis by considering institutions as an endogenous variable to the analysis of the firm (Coase, 1937; Deitrich, Krafft, & McHardy, 2015). The NIE assumes friction in economic exchange, and, hence, positive transaction costs and reflects more closely the real-world firm than the neoclassical model firm. Thus, while the neo-classical approach primarily focuses on the power of the price mechanism to achieve efficient resource allocation with costless transactions, the New Institutional Economics framework extends the analysis to include the role of institutions, and assumes positive transaction costs in production or service delivery (Coase, 1937& 1960; Arrow, 1969; North, 1971; Williamson, 1985). This further reflects a shift to viewing transactions and resource allocation through a *contract,* rather than the orthodox *choice* lens under the neo-classical model (Tadelis & Williamson, 2012).

In the NIE framework, markets are only one mode in the organisation of transactions from a set of possibilities that include the firm or hybrids of markets and firms. A

[4] Mainstream economics commonly found in textbooks that assumes complete rationality, perfect competition, frictionless transactions and complete information

consequence of the positive transaction costs assumption in the market as a mode of transactions organisation is that managers or entrepreneurs will seek alternative modes, such as hierarchy (organisation through a bureaucracy, public or private), to coordinate transactions if the market is calculated to be inefficient. This also explains the co-existence of firms and markets (Coase, 1937), reflecting the fact that some transactions are efficiently coordinated through the market while others through hierarchy.

2.4.2 The New Institutional Economics and the Neo-Classical Economic Model

An appreciation of the differences in the assumptions between the neoclassical economic model and the NIE model helps to clarify the role that new institutional economics, and in particular, transaction costs economics plays in informing policy on efficient economic governance. The NIE starts from the neoclassical economic model as a frame of reference, with the market as the primordial economic institution for efficient exchange. The neoclassical model constitutes the mainstream economics that predominates teaching economic text books (Zimbauer, 2001; Agboola, 2015). The neoclassical model does not include institutions in its analysis of economic governance, as it assumes frictionless transactions (Rindfleisch, 2019). The NIE on the other hand is a heterodox discipline, extending the neoclassical analysis to include the role of institutions in economic governance, and by so doing addresses a key shortfall in mainstream economics (Williamson & Ghani, 2012).

Among the key assumptions of the neoclassical model is the perfect rationality or hyper-rationality of economic agents with perfect knowledge, such that it is possible to allocate resources in ways that achieve efficient production and exchange. The NIE relaxes these rather rigid, and somewhat unrealistic assumptions. The NIE introduces notions of bounded rationality (limited rationality versus hyper-rationality) and information asymmetry (differing and limited levels of information available to agents) to reflect more realistically the limitations of economic man (Agboola, 2015). The aim is to reflect a more realistic empirical representation of the empirical firm and governance modes (Zimbauer, 2001).

The neoclassical model further assumes perfect knowledge of the market and actors, such that appropriate strategies can be created to gain competitive advantage over competitors without cost (Agboola, 2015). It further assumes perfect competition in the

market, with many actors such that there are no transaction costs. Under the NIE and transaction costs economics in particular, these assumptions are relaxed to reflect the real-world firm. Under the NIE, emphasis is placed firstly on incomplete information and, hence, uncertainty. Secondly, incomplete contracts due to limited rationality and capability to anticipate and include all future disturbances and provide for them in contracts, is an important consideration. Thirdly, the need to enforce contracts during implementation due to the opportunism of agents (self-seeking interest with guile) and intertemporal transaction disturbances is a key assumption under the NIE (Buitelaar, 2004). These yield 'frictional' costs that are assumed away under the neoclassical model of the firm (Coase, 1937; Williamson 1985, North and Wallis, 1986). The NIE has direct relevance to the governance of health and ART services. The organization of service delivery through public or private bureaucracy is subject to the same factors that generate transaction costs in the market and private firms. For instance, the coordination of the HIV and AIDS ART programme service delivery requires preparation and enforcement of incomplete contracts with both staff (internal to the public bureau) and contracted private and Civil Society Organizations that provide health or HIV and AIDS ART services. Further, the governance of HIV and AIDS financing involves transaction costs to deliver aid, including fiduciary safeguards which need to be accounted for in efficiency analysis.

2.4.3 Branches of the New Institutional Economics

The NIE has several branches, although a common theme is the relevance of institutions and positive transaction costs that affect efficiency in economic organisation. Further, while some of the literature has focused on macro-institutions such as analysis at the national (geopolity) or international levels (North, 1990; Acemoglu, Jonson & Robinson, 2001; Todorova, 2016), others have focused analysis at the more micro-institutional level such as sectors and individual firms (Coase, 1937 & 1960; Williamson, 1975). The literature that analyses the macro-institutions such as property rights and transaction costs in a country has implications for development policy, while those that focus on micro-level institutions such as firms analyse micro-institutional level transaction costs, and have implications for efficient organisation within firms or hierarchy at the margin (Claude, 2018). The analysis of transaction costs and efficiency adopted for the study fits

within this micro-analytic framework as it focuses on ART service delivery through health facilities.

The NIE as an overarching theory for economic organisation became prominent in the 1970s, with three[5] main strands characterising it. These strands are agency theory (Holstrom & Milgrom, 1991; 1994), property rights theory (Hart & Moore,1990) and transaction costs economics theory (Coase, 1937; Williamson, 1981; North & Wallis, 1986). Although they all acknowledge the role of transaction costs in economic exchange and the need to include them in the analysis of economic organisation, they have tended to emphasize different aspects in relation to the boundary and size of the firm. The agency theory frames the problem of economic organisation in terms of principal-agent relationships, where owners of factors of production (principals) and those hired to work on behalf of the owners of factors of production (agents) relate with differing objectives (Weinstein, 2012).

The agency theory is based on two behavioural assumptions. First, that individuals seek to maximize their utility and second that they will take advantage of incomplete contracts to increase their individual pay-off from a transaction. Hence, the objectives of the agents and principals are rarely likely to coincide, with implications on measures to minimize harm to the firm through investments in oversight activities. This yields transaction costs (Zogning, 2017), which have implications on efficiency as they raise the unit costs of a product or service. The policy advice emanating from agency theory is the creation of incentive structures in contracts that guarantee that agents benefit in the enterprise. It is argued that this reduces the shirking behaviour and transaction costs, and hence, contributes to efficiency in a firm. For instance, making employees shareholders in a company is expected to increase employee performance and create positive incentives for greater effort and protection of the business by employees (Siberberg & Suen, 2001). In the context of public sector bureaucracy service delivery, such as ART service delivery, the higher authority representatives or managers assume the responsibility of the principal as they enforce values and performance standards while staff at service delivery points are the agents. Nonetheless, probity hazards that are

[5] Other strands include evolutionary economics, constitutional choice, collective action theory, public choice theory, economic contract theory, new institutional approach to economic history and modern Austrian economics (see Ritchter, 2005)

common to the public sector require investment in oversight and administrative law that also yield transaction costs.

Property rights theory stems from Coase (1937 & 1960), although later, under the NIE it has been closely associated with Hart and Moore (Hart & Moore, 1990).The basic argument under property rights theory is that where there exists a strong property rights framework and its protection, there will be better development outcomes (Acemoglu, Johnson & Robinson, 2008; Besley & Ghatak, 2009; Locke, 2013). Hence, a key prediction of the property rights theory is that the existence of robust property rights will lead to efficient coordination of transactions to create value (Coase, 1960; Zenger *et al*, 2011; Vitola & Senfelde, 2015). Indeed, Acemoglu, Jonson and Robinson (2008), argue that in contrast to developed countries where the property rights framework is strong, the persistence of poverty in most developing countries can be explained by weak economic institutions, which are in turn the result of weak property rights frameworks. Since property rights theory focuses on the macro economy, it has limited direct relevance to the more micro-level health and ART service delivery.

A third main strand of the NIE is Transaction Cost Economics which draws from a number of disciplines, including organisation theory and contract law (Williamson, 1993; Williamson, 2007). It argues that firms or organisations exist because there are costs associated with the use of the price mechanism (Coase, 1937), and that those costs are higher than when hierarchy (firm or non-profit agency bureaucracy) is used to coordinate transactions. The creation of alternative modes of organisation further occurs when the market fails (Marinescu, 2012). The market as a mode of efficient transactions organisation fails when the assumptions of perfect knowledge of the economic agents, complete information, and costless transactions fail. Hence, firms or public agencies are set up as alternative governance modes to the market to provide particular services that would attract less than optimal transaction costs if offered through the market (Coase, 1937; Williamson, 1988; Williamson 2000).

Transaction Cost Economics further argues that in the organisation for achievement of social goals, efficiency is achieved when transaction costs are reduced or are at the optimal levels (Williamson, 2007). This implies that in every mode of organisation designed to coordinate transactions (through the firm or other types of bureaucracy, or

through the market), there is an optimal level beyond which additional transactions brought under the organisation will lead to inefficiency (Coase, 1937;1960; Williamson, 2000; Zenger *et al*, 2011; Tadelis & Williamson, 2012;). When considered in the context of HIV and AIDS governance, the public bureaucracy through which these services are organised has to be evaluated by looking at how transaction costs affect efficiency in service delivery.

This study uses the TCE to analyse the efficiency of HIV and AIDS programme in Malawi, focusing on facility-based ART service delivery. The TCE is preferred out of the three strands highlighted above because it directly problematises transaction costs as they relate to efficient economic governance based on the mode of governance chosen. Hence, the TCE will further be explained as a theory of the firm to derive results that inform transaction costs and efficiency in ART HTS service delivery through a public bureaucracy. The rest of the NIE strands highlighted will thus be suppressed henceforward.

2.5 Transaction Cost Economics as Theory of the Firm

The unit of analysis under Transaction Cost Economics is a transaction (Shervani, Frazier, & Challagalla, 2007; Williamson, 2010). The completion of each transaction carries costs, and TCE argues that the overriding objective of a firm is achievement of efficiency in the coordination of transactions (Williamson, 2014). In the context of the theory of the firm-its boundaries and size, Transaction Cost Economics is considered a theory of organisational efficiency (Ketokivi, 2020; Nagle, *et al*., 2020). It seeks to find the most efficient mode of coordinating transactions so as to minimise waste (Ketokivi & Mahoney, 2017; Mugwagwa, *et al*., 2020). In general, while the neoclassical model conceives the firm as a production function, the TCE views the firm as a governance mode, that has to be evaluated against other forms of governance in pursuit of the economic efficiency objectives (Williamson, 2014). Thus, instead of looking at the firm as a technological device, TCE focuses on how transactions are coordinated to achieve efficiency. Because the firm is one mode among many for the organisation of transactions, the manager-entrepreneurs tend to engage in a comparative institutional analysis by continuously searching for the most efficient mode of transaction coordination (Tadelis & Williamson, 2012).

Transaction Cost Economics argues that the main organisational challenge in governance choice is one of adaptation for complex transactions that are recurring and facing uncertainty. Hence, managers engage in a process of *discriminating alignment* involving decisions about which activities should be organised within the firm, which ones to co-produce and which ones to let the market organise where the market-hierarchy-interface exists (Ketokivi & Mahoney, 2017). The discriminating alignment hypothesis as described under TCE is stated as follows:

> *"different kinds of transactions are more efficiently governed under different modes of governance" (Tadelis & Williamson, 2012, p.8)*

The implication of this hypothesis is that governance choice to coordinate transactions is likely to be informed by a comparative analysis of organisational modes, considering levels of transaction costs. In the context of HIV and AIDS governance, the implication is that the current ART organisation through the public bureaucracy reflects the most efficient arrangement based on this calculation. Hence, it should be open to the efficiency criterion assessment.

2.5.1 TCE Behavioural Assumptions

Transaction Cost Economics is projected as a theory to explain adaptation arising from uncertainties or disturbances that affect long-term incomplete contracts at the margin. According to Williamson (1981), the transaction costs approach focuses analysis at three levels, namely, agency or organisation-wide level, at department or unit level within the organisation, and a specific activity in the production of a product or a service.

As an economic organisation theory, TCE is based on two behavioural assumptions. First, TCE assumes that individuals have *bounded rationality*. This means that contracts for complex transactions are bound to be incomplete as it is not humanly possible to include all contingencies *ex ante*. As a consequence, preparation of contract documents (through legal counsel) and enforcement is likely to be costly (Williamson, 1981; Shervani, *et al.*, 2007). In HIV and AIDS management, this may include the costs of gathering information about beneficiaries or identification of potential partners and

assessment of their capacities to deliver HIV and AIDS services. This result is important because it implies the need for *ex post* adjustments and design, as well as implementation of safeguards, which further imply transaction costs with a bearing on the efficiency in service delivery.

Second, TCE assumes that economic agents (individuals, firms) are *opportunistic* (Williamson, 1995; Castano and Mills, 2012). The TCE approach recognises that while most people would act in good faith, some are bound to display self- seeking interest behaviour that may be detrimental to the transaction or the contract (self-seeking interest with guile). There is no guarantee, and no way of knowing beforehand that either principals or agents will behave in ways that will only work to benefit themselves, and at the detriment of the firm. An investment in safeguards to protect the investor is thus prompted, that may include monitoring and enforcement costs. In the HIV and AIDs management, intra- organisation management of staff to ensure they are performing their contracted duties to expected standards and quality would be an example. Further, in an HIV and AIDS sub-granting funding arrangement by a donor to another agency such as an NGO or a public sector agency, or district level government agency, these costs may relate to investments in monitoring the use of resources and efforts to achieve agreed targets by the funder. The safeguards may include conducting financial and performance audits. In addition, and in the context of ART service delivery through health facilities, headquarters and District Hospital monitoring and compliance visits and proper ART input use constitute sources of transaction costs.

2.5.2 Transaction Characteristics

The Transaction Cost Economics framework is based on three characteristics of transactions as observed above. The three[6] characteristics of transactions that lead to transaction costs are *complexity*, *uncertainty* and *asset specificity (*Riordan & Williamson, 1985; Tadelis and Williamson, 2012; Coggan, *et al*, 2013)). Rather than focus on the measurement of transaction costs, the TCE framework explains the transaction costs economising behaviour of agents, which has a bearing on transaction governance mode choice and efficiency (Shervani, Frazier, & Challagalla, 2007). According to Williamson (1979), transactions further have to be recurring to be the

[6] A fourth characteristic is *frequency, but is usually suppressed from analysis because it is tautological*

subject of efficiency analysis, since one-off transactions may not provide a basis for comparison of organisation through different governance modes.

Uncertainty occurs when it is not possible to predict outcomes in the environment that changes quickly, largely due to bounded rationality and opportunism by actors (firms or individuals) in the contracting arrangements. This leads to *ex ante* incomplete contracts that take time to prepare. These contracts further require to be monitored and enforced *ex post* to ensure adherence. Further, *uncertainty* of transactions happens when it is difficult to measure performance of the transactions, or when the parameters to measure performance are not present (Shervani, Frazier, & Challagalla, 2007). In health and HIV and AIDS ART service delivery, uncertainty occurs due to the opportunism among actors (patients, service providers, contractors) involved in service delivery. Further, it may be extremely difficult to measure the performance of individual healthcare or HIV and AIDS ART service providers.

Complexity obtains when goods or services cannot be easily defined. In such circumstances, coordination of transactions may involve information search, negotiation and *ex post* monitoring and coordination within the context of the governance mode chosen-firm or market. This is contrasted with a spot market, where the good or service is clearly defined and the price is known (such as the purchase of a loaf of bread from the super market). This leads agents or transactors to invest in the monitoring of performance, leading to the creation of transaction costs. The multiplicity of actors and diverse geographical coverage of health and ART services mean that these transactions are complex, requiring negotiation, information search and strong coordination. Further, and with regard to HIV and AIDS ART financing, a probity hazard prompts the setting up of safeguards, including independent Financial Management Agents by the financiers to address these risks. This yields transaction costs in HIV and AIDS ART governance.

The third characteristic, *asset specificity* occurs when specialised investments are required to deliver services or goods. These may be location specific, such as plants to manufacture products or components, or specialised skills (Williamson, 1981; Williamson, 2011; Zenger et al, 2011; Sautter, 2021). Once these investments are made and contracts concluded, it is impossible to switch them to other uses without incurring loss in their value. Further, abandoning the contract relations for other more efficient

ways of getting the same goods or services efficiently may be difficult due to prior commitments. This leads to the lock-in or hold up problem in contracting, further raising the prospect of transaction costs. In the context of HIV and AIDS services, an illustration of asset specificity may be HIV and AIDS ART services that are location-specific (selected health facilities that meet certain requirements). The location of these services in specific locations may imply that it would be extremely difficult to discontinue the services from certain locations even when they have clearly become inefficient. Abandoning ART service delivery would thus be socially undesirable as it may mean that ART clients that accessed services from the facilities deemed inefficient travel long distances with societal cost implications. A further example could be key population specific programmes that have to be targeted to this sub-population, their locales and preferences, which might not reflect efficient allocation but is justified on the basis of equity.

Given the above characteristics of transactions and the adjustment costs that attend them, TCE predicts that a firm will decide which transactions to be coordinated through the market or through the firm (hierarchy) with the efficiency objective in mind. Transaction Cost Economics argues that transactions that are continuously repeated over time and require complex coordination are likely to be organised through firms, while simple transactions are likely to be organised through the market (Tadelis and Williamson, 2012; Coggan, *et a*l, 2013). Transactions that are highly uncertain, or have uncertainty of at least an intermediate nature are likely to be organised through firms or hierarchy than through the market (Tadelis and Williamson, 2012). Finally, transactions requiring investment in assets that are highly specific are likely to be coordinated through firms or hierarchy, while those that do not require specific investments are likely to be organised through the market. According to Williamson (1979) these decisions are achieved through a comparative institutional assessment regarding levels of transaction costs at the margin. Transaction costs economizing decisions at the margin will thus entail continuous adjustments until a governance mode deemed efficient in organizing the transaction is reached (Williamson, 1981; Loppo & Zenger, 1998; Cordella, 2006; Tadelis & Williamson, 2012). Nonetheless, it is noteworthy that within the TCE framework, consideration of total costs, that is, both firm production and transaction costs and, hence, efficiency assessments can be determined only after the fact. A key implication of this is that the size of the firm cannot be determined *a priori*, but will

depend on this continuous assessment and alignment of transactions. This framework is open to criticism because it would appear that everything can be justified *ex post,* and, therefore, weakens the TCE's predictive capabilities.

2.5.3 Boundary of the Firm under TCE

The theory of the firm with transaction costs is a study of the boundary of the firm (Langlois, 1992). The essential question posed by Coase was why did firms exist side by side with market governance (Coase, 1937)? As noted above, under the institutional economics perspective, and transaction costs economics in particular, choosing between governance modes in economic organisation is a comparative institutional analysis endeavour of the available governance modes (Williamson, 1985; Agafonow, 2017)). Indeed, Coase's (1937) question about why firms existed alongside markets is at the heart of this comparative institutional analysis. The discriminating alignment hypothesis highlighted above implies that agents continuously weigh the advantages and disadvantages of the available governance modes against efficiency objectives with regard to transaction coordination.

2.5.4 Why do Markets fail, and why does Hierarchy Fail?

As observed before, some transactions are best organised either through the market or firms because both market and hierarchy are prone to failure. Faced with this reality, economic agents will seek to switch between the two governance mechanisms in the organisation of transactions to achieve efficiency. A central question in the economic governance debate is, thus, why do markets fail to prompt the establishment of hierarchy? And why does hierarchy fail so as to call on the powers of the market? These questions are discussed below.

a) Advantages of the market over hierarchy

When compared with hierarchy (firm or other types of bureaus), the advantages of the market include the existence of high-powered incentives due to competition (Williamson, 1995). Because of this competition and the need to achieve viability, actors quickly adjust to disturbances. Secondly, the markets are more efficient at aggregation of information for business decision making than hierarchy due to the multiplicity of agents

that are simultaneously calculating costs and benefits, and expressing preferences (Chadad, 2009). This interaction and calculation are embedded in the prices of goods and services. Further, the market has superior matching capacities when compared with hierarchies due to their capacity to facilitate the interaction of multiple buyers and sellers for heterogeneous goods and services (Williamson, 1995).

Nevertheless, markets fail where transactions are specialised and require excessive processes for searching, negotiation, contracting and contract enforcement. These yield costs that render the market inefficient. In addition, markets fail when assets become specific or novel, due to rents that emerge that can be appropriated by both sellers and buyers, leading to *ex post* negotiation and haggling (Zenger *et al*, 2011; Chassagnon, 2014). Further markets become inefficient at the coordination of transactions when there is need for complex coordination and inter-dependence. Because of human beings' incapability to devise complete contracts, alongside opportunistic proclivities, coordination requires close monitoring, rather than the passive coordination appropriate for simple and specific transactions. With respect to HIV and AIDS service delivery, an example would be the services of a nurse or physician in the provision of HIV and AIDS services. Whereas, in theory, it is possible to acquire their services to provide ART services from the market each time there is need for such a service, it would be cost efficient to recruit them under a contract as employees and coordinate their input under a bureaucracy or a firm.

b) Advantages of Hierarchy over Markets

In hierarchies, organisation of transactions is by orders, whereas in markets organisation is through the unregulated price mechanism (Williamson, 1995). A key feature of hierarchies is the role played by entrepreneur- managers that seek to allocate resources or assets in a manner that will maximize value or profit. Hierarchies appear better at adaptation than markets since they are able to direct use of resources in exactly the manner that is calculated to reduce transaction costs, and hence, contribute to greater value creation (Tadelis and Williamson, 2012). Furthermore, coordination of transactions that entail highly specialised investments is better done through hierarchies, given their capacity to coordinate and adapt to changes post contracting (Zenger *et al.*, 2011). Second, compared with markets that have a multiplicity of sellers, hierarchies have a high capacity to shape social identity and informal organisation (Zenger *et al.*,

2011) that may be useful for firm viability. Hence, in instances where assets have to be organised in a manner that enhances the promotion of social identity and informal organisation, hierarchies should be preferred over markets. Third, hierarchies have an enhanced capacity to shape knowledge sharing in a way that is focused on how best to improve products within a firm (Wahab, Rose, & Osman, 2012). This implies that specific processes may be created that guarantee knowledge sharing within firms is leveraged towards improved production or service delivery than would be the case under market coordination of transactions.

Nevertheless, hierarchy has its own limitations. First, Zenger *et al.,* (2011) point to the influence and political costs to organizing transactions through firms or hierarchies. These costs arise from intervention by central authorities, including the costs of monitoring and intervention. Where intervention from central authorities is high, inefficiencies in production or service delivery are likely to be high, especially where intervention by central authorities is driven by the opportunistic behaviour by employees or agents. Second, there is the problem of social attachment that comes with costs. This occurs when, due to personal relations that develop between firms engaged in inter-agency exchange, it is not possible to switch to more efficient exchange relationships because doing so would damage expectations for future exchange relations (Gulati & Nickerson, 2008). Because of this, there are costs associated with inefficient relationships that do not make optimal contributions to the creation of value for the firm (Zenger *et al.*, 2011). In contrast, the market has a great capacity to flexibly get rid of inefficient assets or activities through its high-powered incentives that arise from competition. Third, within hierarchies, there are higher costs and inefficiencies arising from the social comparison process. Because agents (employees in this case) engage in comparison of prices and rewards, and where they perceive that their rewards are less favourable compared to others within the sector or industry, they tend to discard or withdraw all or part of their efforts or skills towards value creation (Obloj & Zenger, 2015; Contreras & Zanarone, 2018). This may cause disruption in the production process and hence, lead to costs to the organisation. Social comparison thus implies an increase in costs to the organisation. Consequently, the economising capacities of hierarchies in the face of social comparison and social costs is constrained, and leads to hierarchical failure, and hence, provides a justification for use of the market as mode of governance (Nickerson & Zenger, 2008).

From the foregoing discussion, and given the strengths and limitations of different organisational modes, it is evident that the firm will use different modes for transactions coordination. While markets and hierarchies represent the extreme opposites of transaction coordination modes, hybrids (combining firms and market) can also be deployed to manage transactions in the quest for achievement of efficiency and value creation. Within hierarchy, this may extend to choices between the types of bureaucracy-public, private-for- profit and non-profit. Indeed, with regard to health care services, it has been argued that both the market and the state perform poorly when compared with the Non-Government or Civil Society Organisation (CSO) bureaucracy (Leonard, 2002; Lieberherr, 2009; Spithoven, 2012). As noted before, entrepreneur-managers will consider the cost of transactions at the margin and decide whether a particular transaction should be internalized into the hierarchy or it should be coordinated through the market based on cost comparisons for the two governance modes (Nickerson & Yen, 2018). More complex transactions with specific assets that require greater coordination over time are likely to be organised through hierarchy, while the less complex, easily definable transactions with low asset specificity are likely to be coordinated through the market. Thus, the boundary or size of the firm is defined by the point at which it is no longer efficient to internalize transactions. The rise of procurement units in bureaucracies stem mainly from this result to facilitate acquisition of services or products that a firm deems would be efficiently produced outside the firm. In the context of HIV and AIDS governance in Malawi, it is observed that similar transactions in procurement, distribution and warehousing are either coordinated by the public sector agencies or contracted out to private agencies.

2.5.5. Discrete Structural Analysis for Transaction Governance Modes Choice

Discrete structural analysis in TCE projects the notion that there exists scope for differing governance modes on a continuum of governance modes, with the market and hierarchy as polar opposites of each other under the private sector organisation. A similar notion is applicable to the public sector organisation where privatisation of service delivery and provision through a public agency are extreme opposites of each other (Williamson, 1991). In between, there are other distinguishable governance modes such as hybrids and regulation in the private and public sector, respectively.

a) Private sector organisation

As noted before, agents will perform comparative institutional analysis to choose and deploy the most efficient means of transaction coordination. In the private sector, comparative analysis of economic governance is between the market and hierarchy, although in practice, this strict demarcation where markets and hierarchy are extreme opposites does not exist as there are other modes of transactions coordination, including hybrids of market and hierarchy (Williamson 1991; Chassagnon, 2014; Reimers, Guo, & Li, 2018; Nagle, Seamans & Tadelis, 2020). Using this framework, the co-existence of the market and hierarchy suggests that a market-hierarchy institutional comparative analysis is achieved through a comparison of alternative governance mechanisms. Williamson (1995) has suggested that managers engage in a discriminating alignment endeavour, where, at the margin, determinations are made as to whether some transactions are coordinated through hierarchy or through the market in a bid to reduce transaction costs and create greater value, all towards achievement of efficiency. In an analysis of comparative economic organisation, Williamson (1999) affirms the existence of discrete structural alternatives to the coordination of transactions in the private sector, with extensions to public sector economic governance[7].

[7] Williamson(1999) distinguishes between private and public sector ordering of the governance of private and public sector transactions

The following table (Table 1) summarizes the alternative economic governance instruments and their attributes as applied to the private sector organisation.

Table 1: Distinguishing Attributes of Market, Hybrid and Hierarchy Governance Structures

Attributes	Governance Structure		
	Market	*Hybrid*	*Hierarchy*
I*nstruments*			
Incentive intensity	++	+	0
Administrative control	0	+	++
Performance Attributes			
Adaptive Autonomy	++	+	0
Adaptive Integrity	0	+	++
Contract Law			
Employment Relation			
Executive Autonomy	++	+	0
Staff security	0	+	++
Legalistic dispute settlement	++	+	0

Key: ++=strong; +=semi strong; 0= weak; Source: Williamson, O. E. (1999)

In the above framework, economic agents achieve first order economising by reducing waste through adoption of governance instruments that are efficient. This implies that the design of governance instruments in the private sector would be assumed to first consider instruments that reduce waste towards efficient delivery of mandates. Furthermore, allocative efficiency (the allocation of resources to their most efficient use) is derived from the first order economising (Reimers, Guo, & Li, 2018). This implies that after design of governance instruments that are deemed to be waste reducing *ex ante,* there is need for *ex post* adjustments to the governance instruments in response to

disturbances that may affect efficient delivery in the long term. This is in response to the need to adapt to uncertainty both internally (within the firm) and externally (outside the firm) in the coordination of transactions. Internal uncertainty is caused by the opportunism of individuals employed by the firm whose interests may not be fully aligned with those of the owners of the firm. This TCE discriminating alignment framework to distinguish the attributes of governance instruments can also be applied to public sector organisation, with privatisation, regulation (hybrid) and pure public agency as distinct governance modes for service delivery-*see chapter 6*.

2.6 The Relevance of TCE to HIV and AIDS Service Delivery through Public Bureaucracy

As observed before, significant resources are dedicated to the fight against HIV and AIDS in Malawi and other countries heavily impacted by the HIV and AIDS pandemic. The organisational form chosen for service delivery is largely the public bureaucracy, with extensions to the private-for-profit and private not-for- profit bureaucracies under partnership arrangements. To the extent that HIV and AIDS service delivery in Malawi is largely through the public bureaucracy, there is need to highlight the relevance of TCE as an analytical framework for efficiency analysis for the public agencies. This is considering that Transaction Cost Economics theory has mainly been used to explain economic governance pertaining to the private sector bureaucracy as a theory of the firm rather than public sector ordering. In the private sector, this comparative analysis has largely been between private sector bureaucracy and the market, with limited extension to the public sector. A valid question, therefore, is: how relevant is TCE to the analysis of public sector organisation for service delivery?

Firstly, it has been argued that since non-private bureaucracies also manage activities that are economic in nature or that require efficient use of resources they can be subjected to efficiency analysis (Arrow, 1969). Moreover, to the extent that the public bureaucracy accounts for a huge amount of financial resources in its transactions, including funds, skilled personnel and infrastructure, it needs to be assessed for economic efficiency (Poniatowicz, 2017). This implies that the public bureaucracy through which HIV and AIDS services are delivered can be subjected to transaction cost analysis and how these costs affect efficiency. Secondly, in assessing the place of the

public bureaucracy, that is organisation of services through public agencies, in economic organisation, Williamson (Williamson 1999) suggests that the public bureaucracy should be considered among the alternative governance modes, given that there may be alternative organisational modes to deliver similar services. Thirdly, some transactions are only suitable for the public bureaucracy, while others are not due to their differing attributes. In this connection, all governance modes are flawed and should only be used for the purposes that they have a comparative advantage in to achieve efficient coordination of transactions (Williamson, 1999). Hence, for particular services such as HIV and AIDS, it is plausible to subject the current dominant organisational mode to efficiency analysis via TCE. Fourthly, the *remediableness* principle should be invoked when TCE is applied to the analysis of the public sector, and to an extent, the not-for-profit bureaucracy. According to Williamson (1999), the remediableness principle states that an existing mode of organisation is presumed efficient unless a superior feasible alternative can be described and implemented with expected net gains.

The growing body of empirical studies that focus on public bureaucracy efficiency using transaction cost economics further demonstrate the application of TCE theory to public governance of service delivery. These studies range from assessment of the public bureaucracy efficiency in the management of energy (Friese, Heimeshoff, & Klein, 2018), transaction costs and efficiency in forest management through public buraucracy (Hünecke, *et al.*, 2019) and access to healthcare (Abimbola, *et al.*, 2015). There are further examples of efficiency analysis that specifically focus on the HIV and AIDS sub-sector service provision (Bautista-Arredondo *et al*, 2014; Bautista-Arredondo, *et al.*, 2018; Khanakwa & Mbonigaba, 2022;2024). Specific to Malawi, Adesina, Couture, & McGinn (2015) estimate and compare the unit costs of integrated and non-integrated family planning and HIV facility services, and find that integration is associated with lower unit cost for service delivery. Nonetheless, their study did not particularly isolate nor focus on the role of transaction costs in integrated service delivery.

2.7 Criticism of the NIE and TCE as Theory of the Firm

Although the NIE in general and, TCE in particular have resulted in a major shift in the economic governance literature, a number of its limitations are observed.

First, it has been argued that the NIE is inadequate as it does not adequately explain the emergence of institutions although institutions are a key factor in explaining the economics of governance under it (Ankarloo & Palermo, 2004). However, this criticism, while relevant to the discourse on the NIE theory and methodology, it may only be valid to the extent that the NIE may not fully explain the emergence of institutions as social constructs. Nevertheless, the development of the NIE has a focus on the emergence of institutions of economic governance as arising from the need to mediate relations among the actors involved in the economic exchange (Kherallah and Kirsten, 2002).

Second, the NIE has been criticised for not breaking completely with the neoclassical tradition, and retaining the foundational elements of neoclassical economic theory (Ankarloo & Palermo, 2004). It only relaxes the neoclassical assumptions of instrumental rationality and complete information, and thus, only seeks to improve the neoclassical theoretical and analytical framework by treating institutions as an endogenous factor in economic organisation. Further, marginalism,[8] which figures prominently in the neoclassical economic model is retained as an analytical principle under NIE, including under transaction costs economics. Further, there is a limited possibility of operationalisation of transaction costs as an applicable concept in the real world (Meramveliotakis, 2018). This criticism stems mainly from the NIE's focus on models of optimisation based on individual rationality - individuals, their goals and rational maximising behaviour in economic exchange as the basis for explaining the emergence of institutions (methodological individualism). Under this framework, where institutions are perceived only as mechanisms to mediate relationships, the NIE is viewed to be inadequate. It is argued that because the individual operates within institutions, the NIE has an inadequate basis for a theory on the emergence of institutions (Meramveliotakis, 2018). However, the criticism that the NIE has not fully broken ranks with the neoclassical model may not be fully valid. This is because, at the outset, the NIE's objective is not to replace the neoclassical model, but rather to enhance its analytical power by including institutions as an endogenous variable that explains economic organisation for achievement of efficiency objectives (North, 1993; Williamson, 1999; Coase, 1937). This is in contrast to the Old Institutional Economics (OIE) framework that stood diametrically opposed to mainstream economics of the

[8] The notion that people make economic decisions over specific units or increments of units, rather than based on categorical approaches, i.e. yes or no, all or nothing

neoclassical model, and sought to replace it with an alternative theory of economic rules and objectives based on institutionalism (Galbács, 2017).

2.8 Empirical Literature on healthcare and HIV and AIDS service delivery efficiency

With regard to healthcare and HIV and AIDS service, there is a growing body of academic literature that provides evidence on the status of efficiency in service delivery. Recent academic literature on healthcare in general and HIV and AIDS governance efficiency highlights the importance of cost minimisation in the face of declining development aid for HIV and AIDS (Kharsany, 2016; Khanakwa & Mbonigaba, 2022; 2024). To that end, a number of studies estimate the efficiency of facility-based HIV and AIDS service delivery (Bautista-Arredondo, 2015; Abimbola, *et al.*, 2015; Lépine, Vassall, & Chandrashek, 2015). The evidence from this strand of research suggests that there is scope for the release of additional resources due to efficiency gains in the general healthcare and HIV and AIDS governance. For instance, Lépine, Vassall, & Chandrashek (2015) investigated technical efficiency of a large HIV and AIDS programme in India using DEA, the Avahan. They found that the programme could reduce existing inputs by 43% without affecting the overall output. The study was useful in highlighting the effect of the programme organisational factors such as size of the NGO participating in the programme, previous experience in HIV and AIDS prevention programmes, the number of years an NGO had been involved in the Avahan programme and sound management of the NGO. In addition, research in Sub-Saharan Africa supports the importance of efficiency as a means towards releasing resources from what is available towards HIV and AIDS service delivery. For instance, Bautista-Arredondo et al (2015) assessed the impact of Community Based Health Insurance (CBHI) on efficient HIV and AIDS services provision in Rwanda, and applied the classical economic analysis of efficiency using Data Envelopment Analysis (DEA). The study demonstrated the technical efficiency of HIV and AIDS service delivery, including total factor productivity over time between 2006 and 2007. It however, focused only on the impact of CBHI as a financing mechanism, and did not apply TCE in the management of HIV and AIDS.

Specific to Malawi, Adesina, Couture, & McGinn (2015) estimate and compare the unit costs of integrated and non-integrated family planning and HIV facility-based services,

and find that integration is associated with lower unit cost delivery. This implies that the concentration of different types of care workflows into a single medical care centre reduces transaction costs and ultimately contributes to efficient service delivery. Thus, the role of deliberate administrative measures to organise healthcare workflows towards transaction cost minimisation is further implied (Steinecker & Czirfuszová, 2021). Furthermore, available research evidence highlights the importance of focusing on unit costs in HIV services as key to finding ways of reducing costs and in the process, release funds for service delivery. In this connection, emphasis is placed on the importance of unit costs of service delivery performance and staff productivity towards efficient and sustainable HIV and AIDS service delivery (Accorsi, *et al*., 2010). Nevertheless, these studies do not explicitly isolate transaction costs and their contribution to the observed unit costs reduction, implying a gap in understanding the sources, incidence and their effect on efficiency in HIV and AIDS service delivery management.

Another strand of academic literature focuses on the decentralisation of health care services and how decentralisation affects efficient resource allocation and technical efficiency of health care and HIV and AIDS. Although the decentralisation of healthcare services potentially improves equity in the distribution of services and may contribute to a lessening of the financial burden on users, the evidence shows that decentralised health service delivery is inefficient(Cuenca, 2020). Factors such as fiscal capacity mismatches in the decentralised units, fragmentation of the health delivery mechanisms and parallel health delivery systems are reported to impact negatively on health service delivery efficiency.

In Malawi, Twea, Manthalu, & Mohan (2020), analysed the process for creating and applying an objective resource allocation formula to facilitate efficient distribution of health resources to decentralized health governance units in District Assemblies in the context of decentralisation. This resource allocation formula had the potential to achieve efficient and equitable allocation of resources, and support the implementation of an essential health benefit package (HBP), and contribute towards the attainment of universal health coverage. The decentralised operational budgets would cover ring-fenced ceilings of drugs, personnel emoluments and salaries, and other recurrent transactions. The notion of an objective resource Allocation formula reported by the study reflects the role of administrative allocation of resources to achieve efficiency in

healthcare at the units that deliver health services. However, the study did not use TCE framework and how it would contribute towards efficient allocation of resources in the country. Further, the formula does not appear to have been deployed, as resource allocation for primary and secondary health services continues to be based on specific conditions or disease burden (GoM, World Bank & Global Finacing Facility, 2021) rather than efficiency considerations. For instance, Kaplan & Merson (2002) argue that allocation of HIV and AIDS resources needs to consider cost-effectiveness analyses that target prevention of HIV and AIDS rather than global proportional allocation to achieve both efficiency and equity objectives. Nonetheless, although studies that focus on decentralised public health service delivery describe the governance of health services, they do not explicitly shed light on the role of TCE as an approach to effective economic governance of healthcare delivery. Thus, there is a gap in knowledge regarding the governance of specific transactions, and how that governance affects efficient service delivery.

Transaction Cost Economics and HIV and AIDS Service Delivery

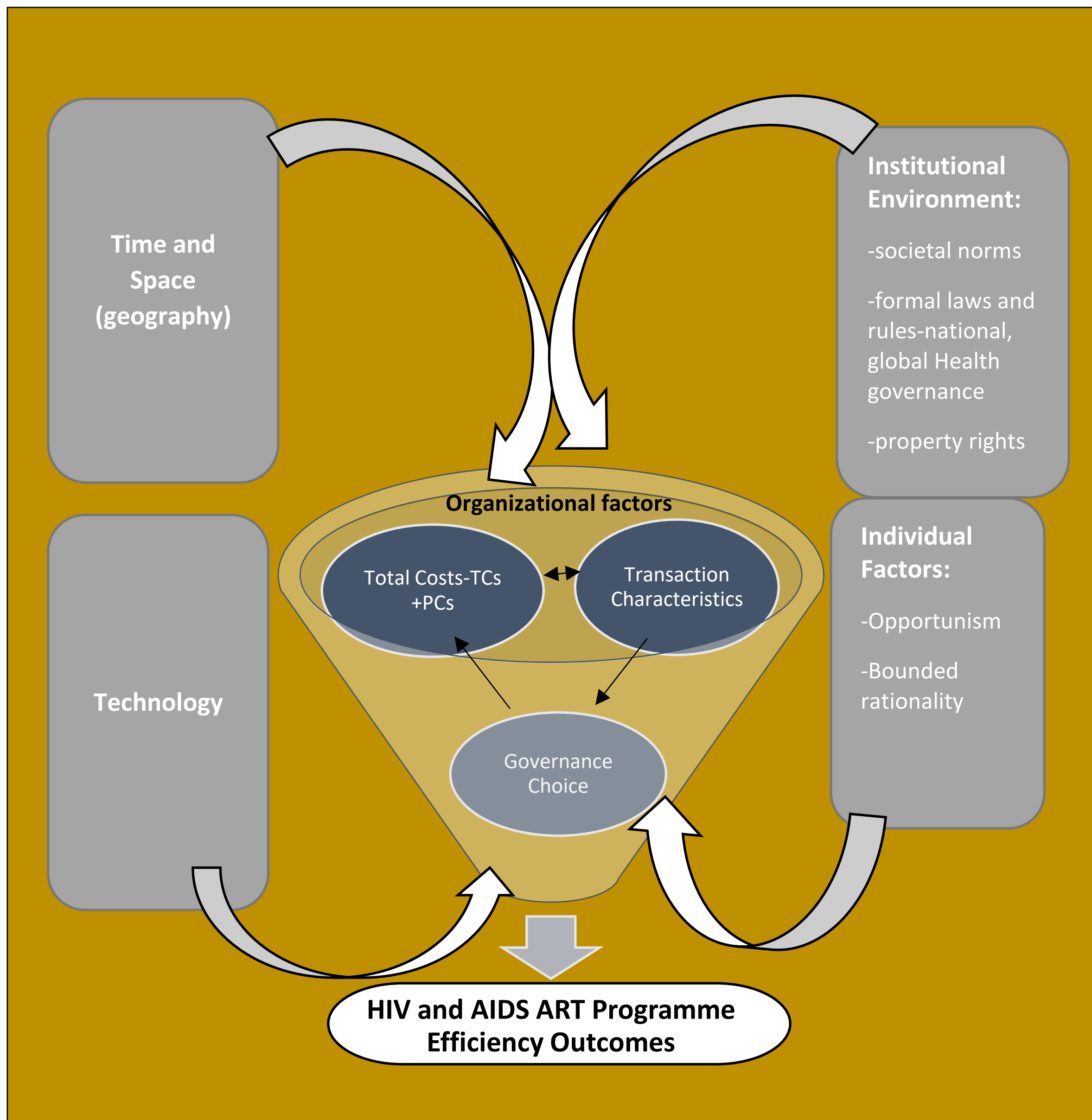
Literature on the application of transaction costs economics within health economics to inform efficient health and HIV and AIDS governance has been somewhat limited. This is despite the acknowledgement that one of the factors that leads to health and HIV and AIDS service delivery inefficiency is transaction costs (Zeng, et al, 2012; Khanakwa & Mbonigaba, 2022; GoM, 2023). Nevertheless, there are a few exceptions with some studies providing empirical evednce of the importance of the application of transaction cost economics governance theory in healthcare and HIV and AIDS delivery (Bautista-Arredondo *et al*, 2014; Bautista-Arredondo, et al., 2018; Khanakwa & Mbonigaba, 2022; Khanakwa & Mbonigaba, 2024). An early application of the TCE theory in HIV and AIDS service delivery shows the potential role that transaction costs economics can play in informing efficient service delivery. In India, transaction cost economics is applied to compare governance instruments in the delivery of a large HIV and AIDS prevention services by India's State AIDS Control Society (SACS), direct contracting of individual NGOs in State X, and contracting a management NGO that coordinated NGOs that delivered the services in State Y Guiness (2011). The evidence points towards differing transaction cost reducing capabilities of different implementation arrangements or governance modes. Nevertheless, while this literature

highlights the importance of TCE as a management approach towards securing efficient HIV and AIDS and general healthcare governance modes, it has remained largely qualitative with limited insights into the incidence and effect of transaction costs on efficient HIV and AIDS service delivery. Thus, there is a gap in information about the magnitude of transaction costs in the current HIV and AIDS governance modes, and how the levels of transaction costs affect efficient service delivery. The current research contributes to addressing that gap by estimating transaction levels and assessing their effect on the efficiency of facility-based HIV and AIDS ART delivery.

2.9 Research Analytical Framework

Based on the review of literature, the research analytical framework is constructed with an institutional economics backdrop. The analytical framework guides the research through a description of factors to be explored, and their relationships and ensuing outcomes (Grant & Osanloo, 2014; Coral & Bokelmann, 2017). For the current research, the analytical framework focuses on a number of factors that combine to influence HIV and AIDS ART (in) efficiency outcomes. It further relates the various factors that influence the HIV and AIDS programme efficiency. These include organisational factors, which may also be viewed as internal factors as they can be directly influenced by managers to affect efficiency outcomes. Other factors are the institutional environment that consist of formal and informal rules that affect service delivery, time and geography or space that may have a bearing on efficiency outcomes as the HIV and AIDS programme or enterprise evolves, and, individual factors (opportunism or shirking) that may further affect transaction costs and, hence, efficiency.

The figure below (Figure 1) depicts the analytical framework that the study adopts.



Note: PC=production costs; TC=transaction costs

Figure 1: Research Analytical Framework

Source: Adaptation from Williamson, 1995 and Corel and Bokelmann, 2017

As can be observed from the figure, the funnel depicts factors that are within the ambit of managers to affect efficiency outcomes in the HIV and AIDS governance. The resulting decisions that lead to governance choice and total costs in turn affect the (in)

efficiency of the services, when all other things are held constant. However, there are external factors such as technology, individual factors, time and geography that interact with and influence the outcome of the internal factors to affect HIV and AIDS ART service delivery (in) efficiency.

2.9.1 Analytical Categories and Assumptions

The key internal organisational factors are transaction characteristics, decisions about governance mode to coordinate transactions and total decision-making unit or firm costs.

Transaction Characteristics

The managers or policy makers assess the characteristics of transactions, on the basis of which they make decisions about governance choice. In the context of HIV and AIDS ART efficiency assessment, the study assumes that policy makers assessed the nature of transactions leading to the current ART governance choice, that is, through the public bureaucracy. In particular, all HIV and AIDS transactions related to ART fulfil the TCE criteria to qualify for assessment according to the Williamsonian TCE framework (Williamson, 1995). Firstly, these services persist over time in the continued fight against HIV and AIDS, and with lifetime treatment requirements for people on ART. Whole national ART programmes have thus been integrated into the health systems to deliver these services. Secondly, they are complex, involving bundles of contracts that require complex coordination structures. This complexity is reflected in the interrelationships among the various financiers of these services and national governments, between principal recipients of grants and sub-grantees, between the central level and district hospitals, between the Government and Civil Society Organisations that provide ART services, between principals (Government) and individual staff, and ultimately between ART clients and service providers. Although the research does not go beyond the assessment of relationships with clients of ART, a contractual arrangement may also be assumed between clients and service providers based on the state-citizen contract relationships. Thirdly, they are asset specific considering that they require specialised staff skills (trained service providers), systems and designated health facilities in specific geographical locations that offer these services. The location of the services in selected health facilities and in specific geographic locations may also constitute elements of asset specificity under the ART programme. The research describes the transaction characteristics in the HIV and AIDS governance in Malawi, and highlights their implications for efficiency.

Governance Choice

An assessment of transaction costs characteristics informs governance choice. Governance mode will be either the market, hierarchy or, a combination of market and hierarchy in a hybrid scheme in private sector organisation. For private- for- profit entities, the mode for coordinating transactions is based on the characteristics of transactions and the costs that attend them, with the overriding objective of profit maximisation. On the other hand, public sector and not-for-profit bureaucracies governance choice is likely to be informed by efficiency, competence and equity considerations. Either way, each governance choice implies a level of transaction costs that has a bearing on efficiency. Additionally, for services such as ART , agency competence, or appropriateness with equity objectives is a key consideration in the decision-making framework. In the context of HIV and AIDS transactions, governance of transactions is largely through hierarchy, but of a public sector kind. It has to be noted nonetheless, that some of the health units offering ART are private sector and civil society bureaucracies “contracted” by the public sector to provide the services free of charge at the service delivery point. Thus, a hybrid mode of ART service delivery is discernible. Consequently, the choice of HIV and AIDS ART governance mode is among privatisation, regulation (hybrid) and pure public agency provision.

Total Costs

The total costs consist of the costs required to provide the service (production and distribution costs), including drugs, reagents and trained staff. Total costs further include the costs of coordinating the HIV and AIDS ART transactions, which extend to the cost of buildings, promotion of the services, human resources with supporting functions such as Health Surveillance Assistants, warehousing and follow up of clients that default. When compared to the outputs of these services, the levels of total costs will determine the level of efficiency of the ART programme. This efficiency can be determined at the decision-making unit level such as a health facility or the total HIV and AIDS Programme in the country. The study sought to estimate the costs associated with HIV and AIDS service delivery as inputs into the calculation of efficiency.

Efficiency

The management of transactions in HIV and AIDS service delivery will ultimately aim to achieve the efficiency economic criterion. It is measured variously, but under the current study, *technical efficiency*, the relation of inputs to outputs produced, is used to evaluate the performance of health facilities that provide ART services in Malawi. It is further used to assess the effect of transaction costs on the efficiency of ART provision.

External Factors

There are a number of external factors that interact with the firm and affect costs and efficiency outcomes. These include the institutional environment, technology, individual factors and time and space or geography

Institutional Environment

The institutional environment consists of various social norms and rules, written and unwritten that shape transactions contribute to transaction costs. Certain rules may prescribe technology choice and how transactions are coordinated. This extends to property rights that influence the costs of inputs such as drugs and HTS reagents. For instance, donor rules and the general global governance framework may influence country level HIV and AIDS programme design and implementation, and hence yield particular governance modes that have differing transaction costs levels, which in turn affect efficiency. To this end, the existence of donor agencies or what are referred to as development partners in the HIV and AIDS programme is key to the research focus. Although under TCE the external environment is acknowledged, it does not directly enter into the analysis of efficiency at decision making unit level. These have been termed shift variables that affect transactions and efficiency (Williamson, 1995). Changes in the external rules may either reduce or increase levels of both transaction and production costs in HIV and AIDS programmes, with implications on the efficiency of the programme.

Individual Factors

Another external factor, also assumed to be given, and considered a shift variable is the individual factors. The individual factors relate to opportunism and self-serving interest by individuals that requires mechanisms or rules to constrain behaviour so that transactions are efficient, also called safeguards. This entails coordination costs with a

bearing on total costs, and, hence, efficiency. *Technology* has implications on the deployment of resources in the production process. This has further implications on the costs of production, which affect service delivery efficiency. In the context of HIV and AIDS, the science related to drugs and approaches to delivery of the services depict the state of technology, and hence the efficiency of the services. Although the research assumes this to be given, it was alert to the fact that the current state of technology regarding HIV and AIDS services has a bearing on production costs that have to be accounted for in the efficiency evaluation. This is analysed through the organisation of the ART programme and the institutional economics of HIV and AIDS.

Time and Geographical Space

Time and geography have been included as analytical categories in line with the research's assumption that the HIV and AIDS services are offered in a complex environment which requires a systems analysis (Coral & Bokelmann, 2017). The figure below (Figure 2) depicts a schematic of how changes to the HIV and AIDS governance may have evolved, and are likely to evolve over time and spatially.

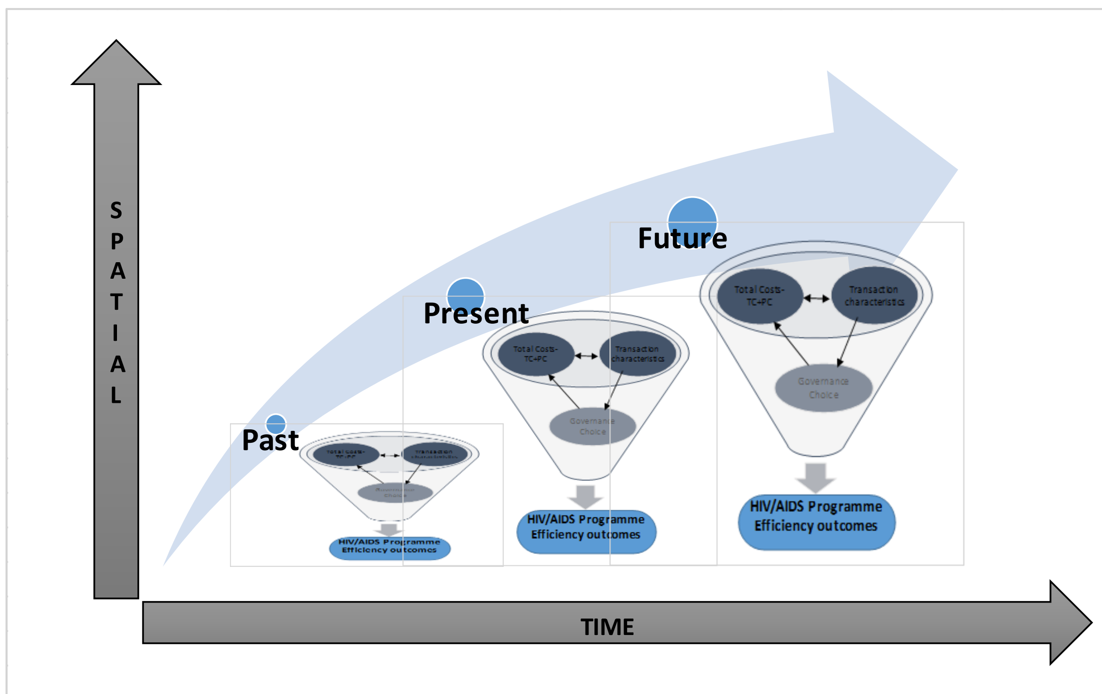


Figure 2: Inter-temporal and Spatial ART Evolution Conception

Source: Author based on Literature review

Inclusion of a temporal variable is important because the organisation of the programme is likely to have changed with time as greater knowledge was acquired regarding aspects of the HIV and AIDS response that are effective. Further, the response has evolved from a few service provision points in the early 2000s to the current levels involving 754 sites in fiscal year 2022. This has a bearing on the efficiency of the programme as the reach of the programme means outputs such as numbers of clients reached has increased. But this further implies an increase in the costs of both providing the services and monitoring and coordinating transactions. The caveat is that it is also possible to have fewer sites of ART provision in future if more efficient alternatives to facility-based ART service delivery can be identified. This may further arise from changes in the institutional environment (global health governance rules, technology, national laws, norms).

Research Contribution

Using Malawi as a case study, this research contributes to the existing literature on HIV/AIDS antiretroviral therapy (ART) programmes in low-income countries by examining the underexplored relationship between transaction costs and the technical efficiency of such programmes. While prior research has evaluated ART programme performance, financing, and health outcomes, few studies have systematically analysed how transaction costs arising from contracting, monitoring, and service coordination affect programme efficiency. This study addresses this gap by empirically examining the transaction costs incurred in ART service delivery and their effect on the technical efficiency of the HIV and AIDS ART programmes. In addition, this study extends knowledge on the application of TCE theory by assessing the conformity of HIV and AIDS ART programme governance structures in Malawi with the core principles of TCE theory. Understanding this alignment is critical to the identification of potential sources of inefficiency and informing evidence-based policy decisions regarding HIV and AIDS programme design and implementation. The research further suggests a methodology for tracking and quantifying transaction costs within HIV/AIDS programmes. This methodology is likely to contribute to the understanding of the specific costs associated with HIV and AIDS governance that can be classified as transaction costs. By applying this methodology, the study provides practical insights for HIV and AIDS programme managers and policy makers seeking to optimise resource allocation and improve the overall efficiency of ART programme in Malawi and similar resource-constrained settings.

2.10 Chapter Summary

This chapter has reviewed the available literature related to economic organisation using the New Institutional Economics theory to provide the analytical basis for the thesis. It has presented the main concepts under the NIE used in the study, and outlined the main strands of the NIE. The chapter has further presented the Transaction Cost Economics theory under the NIE as the theoretical framework for the study. In particular, it has highlighted the behavioural assumptions of the TCE and the characteristics of transactions that determine whether a transaction is coordinated through the market or hierarchy. A discriminating alignment process has been highlighted as the focus of assessment of each transaction and its associated costs at the margin, on the basis of which governance choice is made. It has noted that where the costs for coordination of transactions in the market is higher than through hierarchy, transactions are likely to be organised through hierarchy, and vice versa. Since the real world of economic governance is not sharply divided between markets and hierarchy, the chapter has noted that the TCE further recognises the existence of hybrid governance modes that occupy the place between markets and hierarchy. Shortfalls of the TCE as a theory of the firm identified in the literature are recognised, but the study further notes that they do not render TCE inappropriate in explaining governance choice for efficiency goals. The chapter has further highlighted the relevance of TCE to the organisation of transactions in the public sector towards efficiency goals. It has noted that an increasing body of literature applying the TCE to analyse efficiency in the public bureaucracy continues to emerge. The chapter has further presented the analytical framework for the study, including key internal and external factors that need to be explored when evaluating the efficiency of HIV and AIDS programmes. In particular, internal factors such as transaction characteristics, governance choice, costs and efficiency have been noted. External factors include the institutional environment, technology, individual factors and time and geography.

The chapter has further presented empirical literature related to facility-based healthcare and HIV and AIDS service delivery. While empirical evidence highlights the importance of transaction costs and efficiency in healthcare and HIV and AIDS service delivery, the application of TCE as a management theory appears limited. In particular, a focus on transaction costs incidence and its effect on service delivery efficiency is limited. Thus,

the research contributes to closing these knowledge gaps by applying the TCE theory to Malawi's HIV and AIDS ART governance.

The next chapter presents the approach and methodology of the research.

CHAPTER THREE
THE CHANGING LANDSCAPE OF HIV and AIDS FINANCING AND THE EFFICIENCY IMPERATIVE

3.1 Introduction

Understanding the changing landscape of HIV and AIDS is critical to the sustenance of HIV and AIDS programmes, especially in the context of the declining global funding for HIV Programmes. Over the years, global investment in the HIV and AIDS response have significantly contributed to the reduction of the global HIV and AIDS burden, increasing ART and improving health outcomes, including the prolonging of life. Nevertheless, with donor fatigue and competing health priorities, funding for HIV and AIDS has been shrinking. This necessitates a deep understanding of the HIV and AIDS evolving landscape to identify the key factors that will shape the response to HIV and AIDS. This extends to the need to maximise the impact of limited resources for HIV and AIDS services. As HIV and AIDS funding declines, it is important to optimise service delivery models. This understanding will assist in highlighting the efficiency imperative in HIV and AIDS resource utilisation in the face of declining global HIV funding. By focusing on the efficiency the HIV and AIDS programmes, governments and stakeholders can continue to deliver quality services and contribute to the achievement of the goal of ending HIV and AIDS as a public health issue by 2030. This chapter therefore, sets out the global and national context for the HIV and AIDS financing and its implications on the sustainability of the response in Malawi. It aims to provide the global and national health financing context to highlight the need for efficient management of HIV and AIDS programmes. More concretely, it seeks to show the importance of factors that will affect levels of HIV and AIDS funding, and how this will impose the need to minimise costs in HIV and AIDS governance in low- and medium-income countries (LMICs), including Malawi to free up resources to sustain the AIDS response. The chapter is based on a literature review focused on the changing funding landscape of global HIV and AIDS financing in the face of an enduring epidemic.

The chapter first provides an overview of the HIV and AIDS amidst funding decline, and the fiscal burden it imposes on countries heavily affected by the pandemic. This is followed by a presentation of factors that are likely to impact global HIV and AIDS

funding levels that will heighten the fiscal burden among LMICs. The HIV and AIDS fiscal burden for Malawi is further highlighted, noting the preparation of a Health Financing Strategy as a key development in the country with regard to health financing. Ultimately, the chapter argues that this changing HIV and AIDS response context and financing has heightened the need for efficiency in the implementation of the ART programme in Malawi, where transaction cost control can contribute to technical efficiency in service delivery.

3.2 An Enduring Epidemic in the Face of Declining Resources

Considerable advances have been made in the fight against HIV and AIDS globally. New HIV infections and deaths due HIV and AIDS have declined by 57% and 58%, respectively, between 2010 and 2022 (UNAIDS, 2022b). This has been largely due to treatment scale up efforts by international funders and governments in countries heavily impacted by HIV and AIDS. The gains made so far are widely acknowledged, but so is the realisation that the job is far from over as the epidemic continues to be a global public health threat. According to UNAIDS (UNAIDS, 2023), there were 1.3 million new infections and 630,000 people died of AIDS related illnesses in 2022. In addition, an estimated 39 million people were living with HIV and AIDS globally, with women and girls representing 53% of this figure. Further, 29.8 million people were accessing ART, although 9.2 million remained without access. Further, in 2022, 77% of all new infections occurred among women and girls in Sub-Saharan Africa, revealing the gendered nature of the epidemic (Cane, *et a*l., 2021). The UNAIDS (2023) reports that in 2022, adolescent girls and young women were three times more likely to be infected with HIV than their male counterparts.

Since in most African countries AIDS is spread through unprotected sex in the heterosexual population, demographic changes will continue to pose a risk for infection. A relatively youthful population in Africa means that the population of young people entering reproductive age and becoming sexually active will continue to grow, and hence, become increasingly at risk of HIV infection. In this context, the need for comprehensive sexual and reproductive health and rights (SRHR) services to address this challenge will become ever more important (Starrs *et al*, 2018). Cultural practices that encourage early marriage, low age at sexual debut (15 years or lower), unequal

power relations between men and women that make negotiation of safe sex difficult for women, as well as socio-economic conditions are likely to further fuel new infections (Chirambo, 2019). This situation is compounded by limited access to SRHR services and limited scale up of HIV and AIDS prevention programmes to prevent new infections (Poku, 2016). The persistence of the epidemic as a public health threat implies the continued need for funding to sustain the gains already made, and future prevention and treatment costs for people newly enrolled on ART (McGillen, *et al*, 2017). Further, there will be need to sufficiently expand services to reach the ambitious 95-95-95 targets championed by the UNAIDS that would help in curbing the epidemic as a public health threat by 2030.

Nevertheless, an emerging challenge is the observed flattening in international bilateral funding for HIV and AIDS since 2010, as depicted in the figure below (Figure 3).

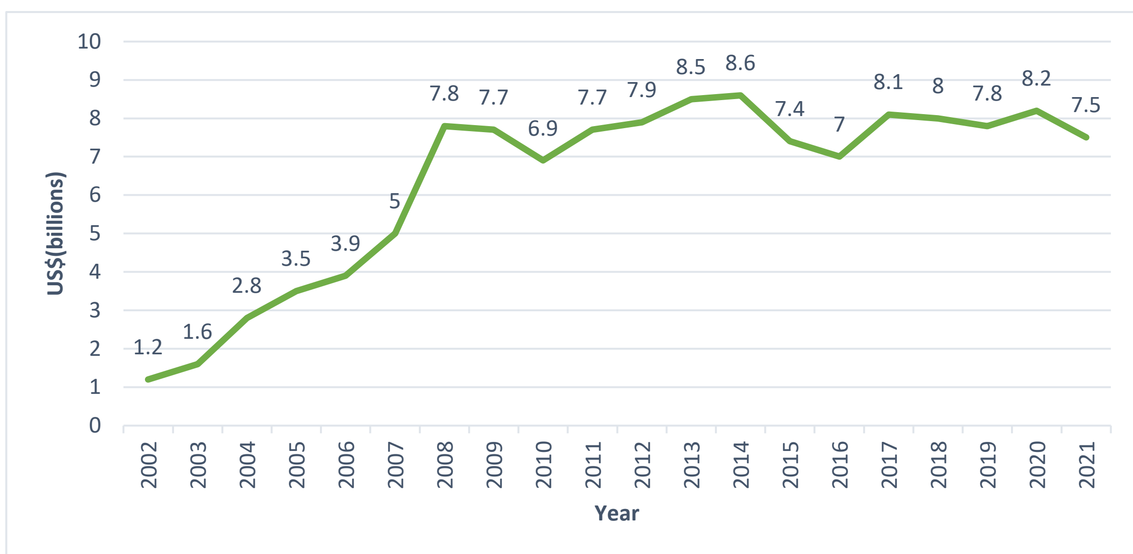


Figure 3:Trend in International Donor Country Bilateral HIV and AIDS funding-2002 to 2022

Source: Wexler, A., Kates, J., Lief, E., & UNAIDS. (2022).

As can be noted from the figure above, a rising trend in the amount of international funding since 2002 first showed a dip in 2010 from U$7.7 billion in 2009 to US$6.9 billion in 2010. This dip is observable again in 2015 from an estimated US$8.6 billion in 2014 to US$7.5 billion in 2015. This uneven trend, and the fact that not all the projected financial amounts for the fast track approach have been made available has tended to

dampen the enthusiasm of the end of AIDS (Vassal, *et a*l, 2013). This changing financial landscape implies particular challenges and vulnerabilities for low- and middle-income countries, particularly in the East and Southern Africa Region, where the HIV and AIDS burden is the highest while nearly two thirds of the financial resources for HIV and AIDS response is from international donors (Ávila *et al*, 2013; Oberth and Whiteside, 2016).

This decline in financial resources comes at a time when there is need to significantly increase the resources to fight the epidemic. For example, to be able to achieve the fast track outcomes, an estimated US$26.2 billion was required by 2020, in addition to the US$19.0 billion made available in 2015 for HIV and AIDS programmes in low and middle-income countries (UNAIDS, 2017). Given the declining availability of bilateral aid for HIV and AIDS from donor countries, the implication is that domestic spending needs to increase beyond current levels to meet the fiscal responsibilities that can sustain the HIV and AIDS response. Nonetheless, considering the large amounts of financial resources required for the response, a further implication is that the capacity of governments in low and middle-income countries to mobilise additional resources will be greatly overextended. Hence, this decline in global HIV and AIDS funding ultimately has to be understood in the context of the fiscal burden that the epidemic imposes on countries that are heavily affected, which is discussed below (UNAIDS, 2022a).

3.2.1 HIV and AIDS Fiscal Burden

The HIV and AIDS epidemic remains a key public health challenge despite the formidable efforts at resource mobilisation for the response, and discernible positive outcomes in reduced new infections and AIDS related deaths. Among the key factors likely to affect the response in the near to long term include a difficult and enduring HIV and AIDS epidemic that requires sustained efforts, and the related impact of biomedical advances through the ART revolution which is shifting the response emphasis towards medicalisation (Gitome, *et al.*, 2014). This is likely to drive costs in the battle against HIV and AIDS as more resources will be required for treatment. Further, at a time when increased commitment of resources for the global response is required, there has been an observable decline in the amount of global financial resources available for HIV and AIDS (Médecins Sans Frontières, 2019; Krishna, Ross, & Chiossi, 2023). This has implications on the financial sustainability of the response, as well as on how the

response is organised in the face of limited availability of resources. The fiscal burden of HIV and AIDS faced by low income countries heavily impacted by the epidemic, alongside constrained domestic fiscal space to raise additional resources to finance the response are further likely to constrict efforts in the fight against the HIV and AIDS epidemic. The COVID-19 pandemic and the Ukraine war, as well as other global conflict situations have tended to further cast doubt on future HIV financing certainty and sustainability (UNAIDS, 2022a; Wexler, Kates, Lief, & UNAIDS, 2022).

3.2.2 The HIV and AIDS Fiscal Burden in Eastern and Southern Africa

The HIV and AIDS epidemic has imposed huge fiscal obligations for both domestic governments and international donors. This fiscal liability is measured by the ratio of HIV and AIDS funding commitments to a country's Gross Domestic Product- the AIDS debt- to GDP ratio. A high AIDS debt- to GDP ratio implies a high fiscal liability, a mandatory requirement within the public health financing framework that may affect development planning and implementation. A low AIDS debt- to GDP ratio means that the country has a low fiscal liability, which is a preferred state for a country to be in than one with a high AIDS debt- to GDP ratio. Although the data appear somewhat dated, trends in public debt in the countries below have not significantly changed since 2015.

The table below (Table 2) depicts the extent of the AIDS fiscal liability for selected countries in the Eastern and Southern African Region

Table 2: AIDS to GDP Ratio for Selected Countries Heavily Impacted by AIDS in Eastern and Southern Africa

	#	Country	GDP/ capita	HIV Prevalence and % range (total, % of population aged 15-49 years)	Number of people living with HIV (000, range in 000s)	Public Debt to GDP ratio	HIV debt to GDP ratio	Public and HIV debt to GDP ratio
Middle Income	1	South Africa	6,483	19 (17.9-19.9)	6,800 (6,500-7,500)	49	7	56
	2	Nigeria	3,203	3 (2.9-3.4)	3,400 (3100-3,700)	13	2	15
	3	Zambia	1,722	12 (11.7-13.1)	1,200 (1,100-1,200)	38	16	54
Low Income	4	Kenya	1,358	5 (4.7-6.1)	1,400 (1,200-1,600)	53	12	65
	5	Tanzania	955	5	1,500	43	14	57

#	Country	GDP/ capita	HIV Prevalence and % range (total, % of population aged 15-49 years)	Number of people living with HIV (000, range in 000s)	Public Debt to GDP ratio	HIV debt to GDP ratio	Public and HIV debt to GDP ratio
			(4.8-5.9)	(1,300-1,900)			
6	Zimbabwe	931	17 (15.9-17.5)	1,600 (1,500-1,600)	94	29	123
7	Uganda	715	7 (6.6-8.1)	1,500 (1,400-1,800)	37	29	66
8	Ethiopia	574	1 (1.0-1.5)	730 (600-970)	44	12	56
9	Malawi	255	10 (9.3-10.8)	1,100 (990-1,100)	76	41	117

Source; Atun et al, 2016

As can be observed from the table above, except for Nigeria, Uganda and Zambia that have public debt- to- GDP ratios of less than 40 percent, the rest have unsustainable debt levels according to the International Monetary Fund (IMF) benchmarks (Vassal, *et al*, 2013). Furthermore as of 2022, the UNAIDS reports that out of US$10.0 in revenue for low income countries heavily impacted by AIDS, US$4.0 are dedicated to debt servicing, and only US$1.0 is allocated to health care (UNAIDS, 2022a). Of the nine countries included in the table above, the HIV and AIDS debt- to- GDP ratio is highest in Malawi at 41 percent. Other countries with relatively high AIDS to GDP debt ratios are Uganda and Zimbabwe at 29 percent each. When public debt and the AIDS debt is combined, all countries highlighted in the table but Nigeria reflect unsustainable levels of debt. Malawi's combined public and AIDS to GDP debt ratio of 117 percent makes it the country with the second highest combined public and AIDS debt to GDP ratio after Zimbabwe.

In general, the high AIDS-debt-to GDP ratios in countries heavily impacted by the epidemic is compounded by the limited scope for creating fiscal space[9] among most of the low and medium-income countries to finance the HIV and AIDS response (Atun, *et al.*, 2016; Remme, 2016; Chansa, *et al.*, 2018; Ithibu & Amendah, 2019; UNAIDS, 2022b). While GDP growth, increased tax revenue and innovative financing, are potential sources of generating additional funding for the response, these avenues are unlikely to generate all the required resources as international funding shrinks (Chansa,

[9] In public Finance, Fiscal space is the ability of a state to find additional resources to sustainably finance meritorious expenditure areas without jeopardizing existing commitments

et al., 2018). Further, for some countries such as Malawi, domestic and international borrowing may not be feasible options given the already high and unsustainable debt levels. An additional option is obtaining debt relief (Ithibu & Amendah, 2019), although this may take long to negotiate and achieve. Hence, achievement of efficiency in the implementation of health and HIV and AIDS programmes is emphasized (Oxford Policy Management, 2013). For Malawi in particular, prioritisation of activities that have the greatest impact (allocative efficiency), and ensuring available resources in HIV and AIDS and health programmes are deployed in a manner that reduces waste (technical efficiency) are seen as key to the creation of a financially sustainable HIV and AIDS response (Chansa, *et al.*, 2018). Indeed, the country's new Health Financing Strategy emphasizes technical and allocative efficiency as its key operative principles.

From the foregoing text, it is noteworthy that the fiscal burden for countries heavily impacted by HIV and AIDS is significant, as demonstrated by the high AIDS-debt-to-GDP ratios. The implication is that the sustainability of HIV and AIDS funding for countries with such high debt ratios remains uncertain in the face of declining international funds for HIV and AIDS. Another key implication is that increasingly, attention will turn to looking at the efficient allocation and management of available resources within the health systems and HIV and AIDS programmes by governments. (Lithman, 2014; The Global Fund, 2023). The need to achieve efficiency in the application of available resources will further be heightened by a much-needed scale up of services, whose costs are likely to be driven by ART, as discussed below.

3.3 Changing Global HIV and AIDS and factors that will affect funding Landscape

There are a number of factors that will affect the HIV and AIDS landscape that will heighten the fiscal burden of countries heavily impacted by the pandemic in the Eastern and Southern African Region. They include the promise of treatment as prevention, shifting donor development priorities, integration of HIV and AIDS programmes into the broader health sector planning, changing donor rules on HIV and AIDS funding eligibility based on country income status and HIV disease burden, and, a call for shared responsibility. These factors are summarised below.

3.3.1 The promise of Treatment as Prevention

The discovery of antiretrovirals in the late 1980s has led to the medicalisation of the HIV and AIDS response. This strategy has been informed by the effectiveness of ARVs in suppressing viral load, and the consequent effects of prolonging life for people living with AIDS. The antiretrovirals increasingly became a key element of responding to epidemic (Menzies, *et al.*, 2012; Tagar, *et al.*, 2014; Gitome, *et al.*, 2014). Countries in Sub-Saharan Africa, which account for the highest burden of HIV and AIDS were initially slow at adopting ART mainly due to cost, but use of antiretrovirals is now widespread. This has been thanks to the lowering of antiretroviral prices due to the advent of generic formulations and increased development assistance for HIV and AIDS. Indeed, the growing optimism about the end of AIDS as a public health threat is to a large extent based on the promise of ART (UNAIDS, 2014). In part, this optimism is due to the potential for using treatment as prevention as people that are on ART are unable to spread infection (Resch, *et al.*, 2015). In view of the fast track approach (UNAIDS, 2013), the need for increased resources will even be higher, and will largely be driven by the dominance of ART in responding to the epidemic. This is because the fast track approach is anchored on a biomedical platform, with ART as the main cost item.

In addition, the recognition of treatment as prevention has now prompted the test-and-treat campaign that was adopted in Malawi, which is likely to further drive the cost for HIV and AIDS programmes, at least in the short term. This drive towards universal treatment, has been buttressed by the test-and –start strategy per World Health Organisation (WHO) guidelines on ART (WHO, 2015). Although there have been concerns about the overemphasis on medicalisation of the epidemic, the impact of embracing it has led to underfunding of other critical prevention elements of the response, and dampened advocacy (Nguyena, *et al.*, 2010; Poku, 2016; Dalton, 2017). Further, while the achievements due to ART are widely acknowledged, there have also been calls to closely look at the cost effectiveness of treatment as prevention, alongside other prevention strategies as the ART programme expands (Natrass, *et al.*, 2016).

A key implication of this is that more resources than before would be needed to retain those already on ART, as well as cater for an expanding ART programme with the test-and-start approach. Consequently, and in the absence of a cure or vaccine, this focus on treatment seems to be an element that will drive costs in the HIV and AIDS response in the near to long- term. An increased fiscal commitment is, therefore, expected from

both high-income donor governments and national governments impacted by HIV and AIDS. Nevertheless, with the bulk of the ART funding coming from development aid in Malawi and other low-income countries, gaps experienced due to a decline in international HIV funding are likely to be substituted by domestic resources (Médecins Sans Frontières, 2019).

3.3.2 Changing Donor Funding Priorities and Integration of HIV and AIDS into Health Systems

The funding of HIV and AIDS has evolved over the period since AIDS was discovered and recognised as a global public health issue. In the early 1990s and 2000, development assistance for HIV and AIDS focus was the prevention and strengthening of health systems to cope with the AIDS challenge (Schnieder, *et al.*, 2016; Starrs, *et al.*, 2018). As noted above, the effect of development and success with treatment through antiretroviral medicine in the 1980s was to change the funding landscape, with more and more funding channelled to treatment and care (Menzies, *et al.*, 2012). After 2005, treatment and care has emerged as the most funded programme, accounting for 23.7 percent of development assistance for HIV and AIDS. This evolution in funding shows how funding priorities have changed over time, and is significant because it havs implications for long term allocation of funding for HIV and AIDS responses. The prioritisation of treatment over other areas of HIV and AIDS services means that countries affected by the pandemic and current donor governments have to expand and sustain funding as more people are enrolled into the ART programmes. The funding for ART once committed may not be easily reduced as people have to be retained on the drugs for the rest of their lives. Nonetheless, with donor funding levelling, this raises the issue of financial sustainability of the response. In particular, finding ways of organising the response that may yield savings that contribute to efficiency in the management of HIV and AIDS has become key. The issue of transaction costs minimisation may be considered within this broader efficiency framework.

A further factor regarding changing donor funding priorities is the focus on other non-health development priorities and other humanitarian assistance. The inclusion of only one health goal, number 3[10], with HIV and AIDS only as a target under the Sustainable Development Goals (SDGs) shows this shift in development assistance. The post 2015

[10] To ensure healthy lives and promote wellbeing for all at all ages

development agenda only includes HIV and AIDS as a target under health (Poku, 2016; Oberth and Whiteside, 2016). The implication is that allocation of international funding for HIV and AIDS in the post 2015 development agenda is likely to decline (Oberth and Whiteside, 2016). This is likely to lead to the need for increased funding allocations from domestic resources in countries in Eastern and Southern Africa that bear the heaviest burden of the epidemic. As noted above, the limited options for creating sufficient fiscal space is likely to draw attention towards greater efficiency in utilising the available resources to achieve maximum results in the fight against the epidemic.

Furthermore, there have been calls for greater integration of the HIV and AIDS programmes into the broader health sector (GoM, 2020). Consequently, the tendency has been towards financing HIV and AIDS prevention and treatment programmes as part of the overall health systems strengthening, rather than a disease specific programme. In particular, Sexual and Reproductive Health and Rights (SRHR), as well as primary health care systems strengthening appear to be favoured, with implications on how HIV and AIDS is prioritised within this scheme (Poku, 2016; Nattras, *et al.*, 2016. This shift has been informed by the perspective that financing HIV and AIDS as a disease specific intervention in a vertical manner tends to weaken health systems, as well as the need to improve efficiency. The focus on primary health reflects the spirit of the World Health Organisation (WHO) Alma Ata declaration of 1978 that emphasised the need to strengthen the primary health care system as the most cost-effective way of health service delivery (Nattras, *et al.*,2016). However, in practice, integration may imply greater costs for HIV and AIDS clients waiting in queues to access services. For instance, Natrass *et al.,* (2016) report declining quality of services in South Africa when HIV and AIDS services were integrated with basic health care services, with HIV and AIDS clients spending longer hours to access services than before the integration occurred. This South African case is an example of increased transaction costs (opportunity cost) incurred by AIDS clients. Nevertheless, at the same time, there was a dramatic decrease in the cost of ART from a provider perspective due to primary healthcare integration, in particular, due to the shift from doctor to nurse initiated management of ART (NIMART) (Meyer-Rath, *et al.*, 2017).

3.3.3 An Era of Shared Responsibility

In addition to the changing donor funding priorities, another feature with implications for the sustainability of funding for the HIV and AIDS response is shared responsibility. Governments in countries with high HIV and AIDS burdens are required to contribute towards the financing of HIV and AIDS services in a shared responsibility framework (Buse and Martin, 2012; Resch, *et al.,* 2015). Realising the great need of resources for financing HIV and AIDS in the post 2015 development agenda, the African Union has called for a shared responsibility in the financing of HIV and AIDS (African Union, 2012). This call was made to secure long term funding, realising that this financial burden could not be carried by countries heavily affected alone, nor by the high-income donor countries alone. Already, the percentage of domestic funding for the HIV and AIDS response in countries heavily affected by the epidemic is increasing (UNAIDS, 2013). For instance, UNAIDS (2017) records higher funding contribution levels for the global HIV and AIDS by governments at 57% than international donors (43%) in 2016. For the fiscal year 2018/ 2019, the Government of Malawi was the second largest funder of the HIV response at 27% of resources after the Global Fund (54%) contribution to the Malawian response. The United States Government through the President's Emergency Plan for AIDS Relief (PEPFAR) accounted for 18% of AIDS funding in fiscal year 2018/ 19 (GoM, 2021).

Furthermore, some donors have re-prioritised their funding based on disease burden and a country's ability to pay. In this reprioritisation, the transition from donor funding of HIV and AIDS to greater domestic government contribution is based on two criteria: the income level of a country and its disease burden (Burrows, *et al.*, 2016). As countries move from low income to middle income status, it is assumed that they are able to pay for their HIV and AIDS funding commitments. Burrows, *et al.,* (2016) report that under these transition arrangements, Eastern Caribbean countries, and some countries in Africa such as Botswana, South Africa and Namibia have had their funding for some areas of the HIV and AIDS response reduced, or have their total funding from PEPFAR and DFID bilateral funding withdrawn. Thus, the changing donor rules on eligibility for HIV and AIDS funding based on country income status, and the call for shared responsibility with low- and middle-income countries (LMIC) governments in HIV and AIDS funding are likely to affect the response in the near-to long-term (Médecins Sans Frontières, 2019). Furthermore, countries with a low HIV and AIDS disease burden are no longer

being prioritised in favour of those with high disease burdens. In addition, the GFTAM now requires contributions from recipient countries of between 5% and 60 % (Vassal, *et al.*, 2013; Burrows, *et al.,* 2016) depending on the country's income status, with low income countries contributing proportionately less than middle- and high-income countries. The PEPFAR also requires a 25 percent commitment of contribution from the recipient countries (Burrows, *et al.*, ibid). This implies that countries will invariably be called upon to set aside ever-increasing amounts of funds to meet HIV and AIDS service delivery needs. For countries such as Malawi, this may represent a considerable challenge in view of the relatively high dependence on external financing (Vassal, *et al.,* 2013; Resch, *et al.,* 2015; Atun*, et al.,* 2016).

With the above transition arrangements there are fears that, first, the requirement for matching contribution may also imply a crowding out of funds from other health or development priorities, and may negatively affect balanced development in a country (UNAIDS, 2022a). For instance, diversion of funds earmarked for other programmes supportive of HIV and AIDS outcomes such as food security and nutrition may mean a weakening of those aspects. Further, this transition has implications for the sustainability of HIV and AIDS financing in developing countries, and may have the effect of overburdening domestic governments when certain contribution thresholds, such as when AIDS funding is greater than 2 % of a country's GDP are exceeded (Resch, *et al.,* 2016). A speculative, yet further point, is that it may be argued that given the relatively impressive growth of some of the African economies compared with the rest of the world in the recent past, most low-income countries in Africa could transition into medium income status. Hence, they may become ineligible for HIV and AIDS financing based on their income status criteria (Burrows, *et al.*, 2016). The second transition criteria- HIV prevalence level- is further problematic, and has implications for the sustainability of the global HIV and AIDS response. This is because low national HIV and AIDS prevalence levels may hide high prevalence rates among key populations within a country, and if the weaned country is unable to fully cover aspects of the epidemic targeting its key populations, gains made in prevention may be reversed, and a rebounding of the epidemic is a likely prospect (Burrows, *et al.*, 2016).

Overall, shared responsibility based on these two transition criteria implies that countries that transition from low to higher income levels are likely to lose some of the HIV and

AIDS funding. Ultimately, some critical areas of the response may be defunded *de facto*, considering that countries may not have mechanisms to guarantee funding for players such as NGOs or community-based organisations implementing community level prevention type interventions (Burrows, *et al.*, 2016; Médecins Sans Frontières, 2019). Nonetheless, considering the severe constraints in generating additional domestic resources, this further implies the need to find ways of implementing the HIV and AIDS response at country level more efficiently.

3.3.4 Health Sector and HIV and AIDS Financing in Malawi

Health sector financing in Malawi is dominated by international development assistance for health. The 2022 National Health Accounts estimate that donor funding to the health sector accounts for 54 percent of total health expenditure (THE) (GoM, 2022). Government expenditure as a percentage of THE was 24.1%, while expenditure on HIV and AIDS accounted for 40% of THE. More importantly, donor managed expenditure on health as a percentage of THE was estimated at 54%, while that managed by Government was 39.4%. This high dependence on donor funding for health implies particular vulnerability when donor governments' priorities at home change. The case of the United Kingdom Foreign Common Wealth and Development Office (formerly DFID) that has reduced its contribution to HIV funding as part of the overall drive to reduce overseas development assistance shows that this is a distinct possibility (Wexler, Kates, Lief, & UNAIDS, 2022). It further implies that changing donor priorities and increased accountability requirements are likely to impact the availability and use of funds in the sector, including the need to use funds more efficiently. Concerns about reducing transaction costs across donors and by the Government of Malawi are also likely to feature highly in the management of DAH in Malawi (Lithman, 2014; The Global Fund, 2023; PEPFAR, 2022).

Overall, Total Health Expenditure (THE) in Malawi was estimated at 8.8 of Gross Domestic Product in 2022. Public health spending as a percentage of total Government expenditure has persistently remained below the Abuja target of 15% (Ministry of Health, 2016). In fact, public health spending as a total of public expenditure declined from 10.4% between fiscal years 2012/ 13-2014/ 15 to 8.4% in 2022 (MoH, 2022). Overall, there is allocative inefficiency in the health system as observed through low expenditure on primary health care (GoM, 2023b). The Malawian health sector is further

deemed inefficient as evidenced by a high expenditure on curative, as opposed to primary and preventive health services. Further, inefficiency is observed in the supply chain, and in the budgeting formulation process that is not linked to service delivery output. A recent study further showed that Public Finance Management in the health system is extremely weak, especially related to efficiency and accountability (Piatti-Fünfkirchen, Chansa, & Nkhoma, 2020). A relatively large share of expenditure on non-core health services such as governance suggests inefficiency in health spending, and a limited application of transaction cost economics framework in health financing. Since the bulk of HIV and AIDS ART resources are channelled through the public health system, the implication is that the observed inefficiencies in public healthcare are also reflected in the facility-based ART programme.

3.3.5 HIV and AIDS Financing and Efficiency Imperatives for Malawi

The Malawi health system is severely under-resourced with a per capita THE of US$39.9 compared with US$86.0 that is considered minimum for developing countries (GoM, 2023b). Further, and similar to general health system financing, funding for the national HIV and AIDS response is heavily dependent on international funding. External funding for HIV and AIDS accounted for 97 percent of the response (Vassal, *et al.,* 2013, GoM, 2021). The NAC (2017) finds that volumes of funding towards HIV and AIDS have been below requirement in the recent past. In the 2015/ 16 fiscal year, US$187.0 million was budgeted against projected costs of US$222.0 million. In 2017/ 18 US$184.0 million was budgeted for the national response by partners, but the projected resource requirement was US$254.0 million (NAC, 2017). Médecins Sans Frontières (2019) reports an increase in the HIV funding gap from 5% in 2017 to 22% in 2019. Further, the GFTAM is reported to have accumulated unfunded quality demand of US$111.0 million, largely due to an ARV supply shortfall (Médecins Sans Frontières, 2019).

Malawi's adoption of the National Health Financing Strategy (HFS) in January, 2023 represents an important milestone in the quest to improve the financing of the country's health system towards universal health coverage. Nevertheless, the new strategy, which includes reforms to the health financing system has to be viewed against the health financing aims, and challenges that the country faces, with particular reference to the HIV and AIDS financing. More importantly, how the HFS negotiates the changing HIV

and AIDS financing landscape will be key. With reference to the current research, the extent to which efficiency in the management of HIV and AIDS is addressed, and the role of transaction costs in efficiency aims are particularly important considerations.

A review of the new Malawi HFS shows that in many respects, it covers a broad range of the current health financing issues in Malawi, which include under resourcing, high donor dependency, system allocative and technical inefficiencies, high debt burden, weaknesses in health budgets formulation, implementation and evaluation, and, funding unpredictability (GoM, 2023b). The Health Financing Strategy's broad aim is "To set up a well-governed health financing architecture able to mobilise adequate resources, distribute the resources in an efficient and equitable way, and strategically purchase services based on a well-defined benefit package in pursuit of UHC goals." (GoM, 2023b, p22). It further has four specific objectives that cover[11]: a) adequate mobilisation of funds that are predictable and sustainable; b) improved efficiency in the pooling and management of resources in the Malawi health system; c) development and implementation of strategic purchasing measures across the health sector; and, d) establishment and strengthening institutional arrangements for health financing at all levels in the health system. Reforms to the health financing system include improvements in resource mobilisation, new pooling structures and new purchasing arrangements. Resource mobilization improvements will include capacity building for decentralised resource mobilisation and management of donor funds, while the new pooling arrangements are calculated to effect the "One Budget, One Plan and One Report" in the health system. The new purchasing arrangements will entail modifications related to tiered and better resource allocation across the service delivery levels and improvements in provider payment arrangements (GoM, 2021).

An assessment of the strategy with regard to resource mobilisation for the health system, and especially with respect to HIV and AIDS, suggests that the main sources of funding are donor agencies and the Government of Malawi (GoM, 2023b). Thus, solidarity through taxation and donor contributions remain key to the strategy. Among the specific reform strategies is consideration of community contributions and fee-paying arrangements for people in the informal sector that do not pay tax. Exceptions are the vulnerable who will access services free of charge at point of care, after vulnerability

[11] GoM (2021). The Malawi National Health Financing Strategy, Lilongwe: Ministry of Health

confirmation via the Unified Beneficiary Register (UBR). Improved donor pooling arrangements for improved funding coordination and budget support is suggested, but it needs to be negotiated, suggesting the need for the role for health diplomacy (Nimubona, 2022)[12]. A discriminating fee-paying arrangement is explicit in the HFS for essential health care services, although its application to the ART services is less clear. The strategy suggests a fee payment arrangement for non-Malawians that do not pay tax and those in the informal sector at point of care, although this is likely to pose implementation challenges and perhaps raise controversy when ART services are considered. Nevertheless, with respect to ART services, it is likely that access will continue to be free at point of care. A heavy debt burden at 59.3% for Malawi, which is extremely high by any standards, has been recognised as a binding constraint to health financing in the country. It is much higher than the 40% deemed sustainable for the low-income countries. However, the strategy does not detail the strategy to deal with this constraint. Advocacy for health debt relief via the Ministry of Finance and Economic Affairs may be one strategy to pursue.

With particular reference to the current research, the HFS identifies achievement of efficiency in the health system as a resource optimisation strategy that can be harnessed from improvements in allocative and technical efficiency. Gross allocative and technical inefficiencies in the health system stem from several sources, that include, medicine supply chain management, uncoordinated in-service training, inadequate planning and budgeting of health services and fragmented health data systems (GoM, 2023b). To the extent that the allocation is conducted by policy and programme planners through the budgeting and implementation processes, these inefficiencies reflect limited application of the Transaction Cost Economics framework in the management of health resource allocation. For instance, GoM (2023b) reports that while tools are available in HIV and AIDS modelling to guide efficient allocation of resources, these are almost never applied in the public budgeting processes. In addition, limited pooling arrangements and coordination with respect to donor funding for health is highlighted as a key challenge as it leads to allocative inefficiency, with 13% of total funds available from public and donor funds deemed to be misaligned. Technical inefficiency and the role of transaction costs is less explicit, although both can be gauged from the multiplicity of partners and

[12] Health diplomacy has been defined as "the way of gaining trust to promote multisectoral collaboration for a particular health intervention" according to Fazal TM. Health diplomacy in pandemical times. Int Organ. 2020;74(S1):E78–97. https:// doi. Org/ 10. 1017/ S0020 81832 00003 26.

their administrative set-ups in the health financing system (227 in total as of 2022) that constitute non-state institutions providing health support to households. The strategy suggests the implementation of biennial efficiency analysis, which are currently either non-existent or are not mainstreamed into the health financing governance framework. These will be focused on high expenditure programmes that include HIV and AIDS, malaria, tuberculosis, reproductive health, Expanded Programme on Immunization and COVID-19. High administration and management costs (above facility costs) that reflect transaction costs are not mentioned in the strategy, but they are implicit in the strategy's aspiration to institute technical efficiency reviews of major programmes. It is notable that HIV and AIDS has been prioritised in the efficiency analyses, but this implies the need for transaction costs to be deliberately monitored to ensure they are accounted for in the general healthcare, and specifically, the HIV and AIDS ART efficiency equation. The format of the bi-annual efficiency analyses under the health financing strategy is yet to be agreed, but the analyses offer an opportunity to integrate the tracking of transaction costs in general healthcare and HIV and AIDS service delivery, and how they impact efficiency.

Some lessons for LMICS heavily impacted by HIV and AIDS from this changing HIV funding landscape are discernible. With the decline in funding to finance the response to the pandemic, LMICS need to be proactive in instituting strategies that contribute to the generation of resources to finance HIV and AIDS programmes at the national level. Firstly, they need to put in place sustainability plans that involve increased locally generated resources into the HIV and AIDS programme, especially those that promote prevention. Secondly, efficient management through reduction of wastage in the HIV and AIDS services governance would be key. This may entail the streamlining of HIV and AIDS drugs and supplies procurement, and instituting streamlined management structures. Third, in view of the high debt burdens, LMICS should lobby for debt relief that may free up some resources for HIV and AIDS, with concrete programmes aimed at financing the HIV and AIDS response at country level. Fourth, there may be need to explore ways that are efficient at the delivery of the HIV and AIDS services, including integration of community led HIV testing that is proven to be less costly when compared with facility-based testing (Indravudh, *et al*., 2021).

3.4 Chapter Summary

This chapter has provided the context of the study in terms of the changing HIV and AIDS response landscape. It has noted the changing context of HIV and AIDS and its implications on the financing of the global response to the epidemic. In particular, the context suggests the need for increased resources to sustain the commitments already made, as well as new commitments towards ending HIV and AIDS as a health threat by 2030. A close examination of this changing financing landscape points to three key consequences. First, it is clear that the global response to HIV and AIDS needs to be maintained and expanded in view of the need to prevent new HIV infections. Second, mobilisation of adequate resources for the response is likely to be a challenge given the trends in resource availability, particularly for low- and middle-income countries where over a third of funding for the response so far has come from international donor funding. The constraints by developing countries heavily impacted by HIV and AIDS are further compounded by mounting public debt (UNAIDS, 2022a). Third, the flattening of HIV and AIDS funding by bilateral donor countries without a concomitant increase in multilateral contributions suggests increased fiscal responsibility for developing countries in responding to the HIV and AIDS pandemic. Against this background, and with a constricted fiscal space by most countries in the LMICs and Eastern and Southern African Region in particular, efficiency in the management of available resources has become a firm goal with a view to contributing to the sustenance of HIV and AIDS funding, particularly for ART. A new national health financing strategy for Malawi (GoM, 2023) seeks to consolidate resource mobilisation for healthcare, including HIV and AIDS and to improve efficiency in healthcare delivery. It holds promise for the achievement of efficiency in health financing in the country, but it needs to be implemented.

The next chapter presents the research methodology.

CHAPTER FOUR

RESEARCH METHODOLOGY

4.1 Introduction

This chapter presents the research design and the methodology. It presents the specific methods used to collect and analyse the research data. It further describes the empirical strategy that the research adopted to achieve the objectives of the study, including data collection and analysis methods. The research approach is presented first, followed by the specific methods used to collect and analyse data. The chapter further highlights the limitations of the research. A summary concludes the chapter.

4.2 Research Approach

The research approach was transdisciplinary and consistent with Global Health Research (Picard, et al., 2011,Ding, et al., 2020). It combined knowledge from the management and economics disciplines of the University of Kwazulu Natal's Health Economics and AIDS Research Division (HEARD). The research was supported by a team that consisted of researchers from various disciplines, that included political economy, psychology, anthropology and economics. The research used an exploratory concurrent mixed methods design. This consisted of a pilot study that included qualitative data gathering and assessment of the available data to determine the feasibility of completing the research as designed. The outcome of the pilot study indicated that while it was possible to get data on outputs and some inputs that were used to estimate the costs of the HIV and AIDS ART programme for the chosen year of study, it would be extremely difficult to obtain data on coordination or transaction costs under the current research. This was largely due to the limited time and financial resources available for the research. It was clear from the pilot study that these costs needed to be estimated based on previous resource and expenditure tracking studies. This was because it was not possible to obtain facility level expenditure records at public health units as only the outputs and ART activities were available at that level. It was further extremely difficult to get data on staffing, utility and other transaction costs. Further, getting the same data from the district levels would be costly and was thus not feasible given the limited resources available for the research.

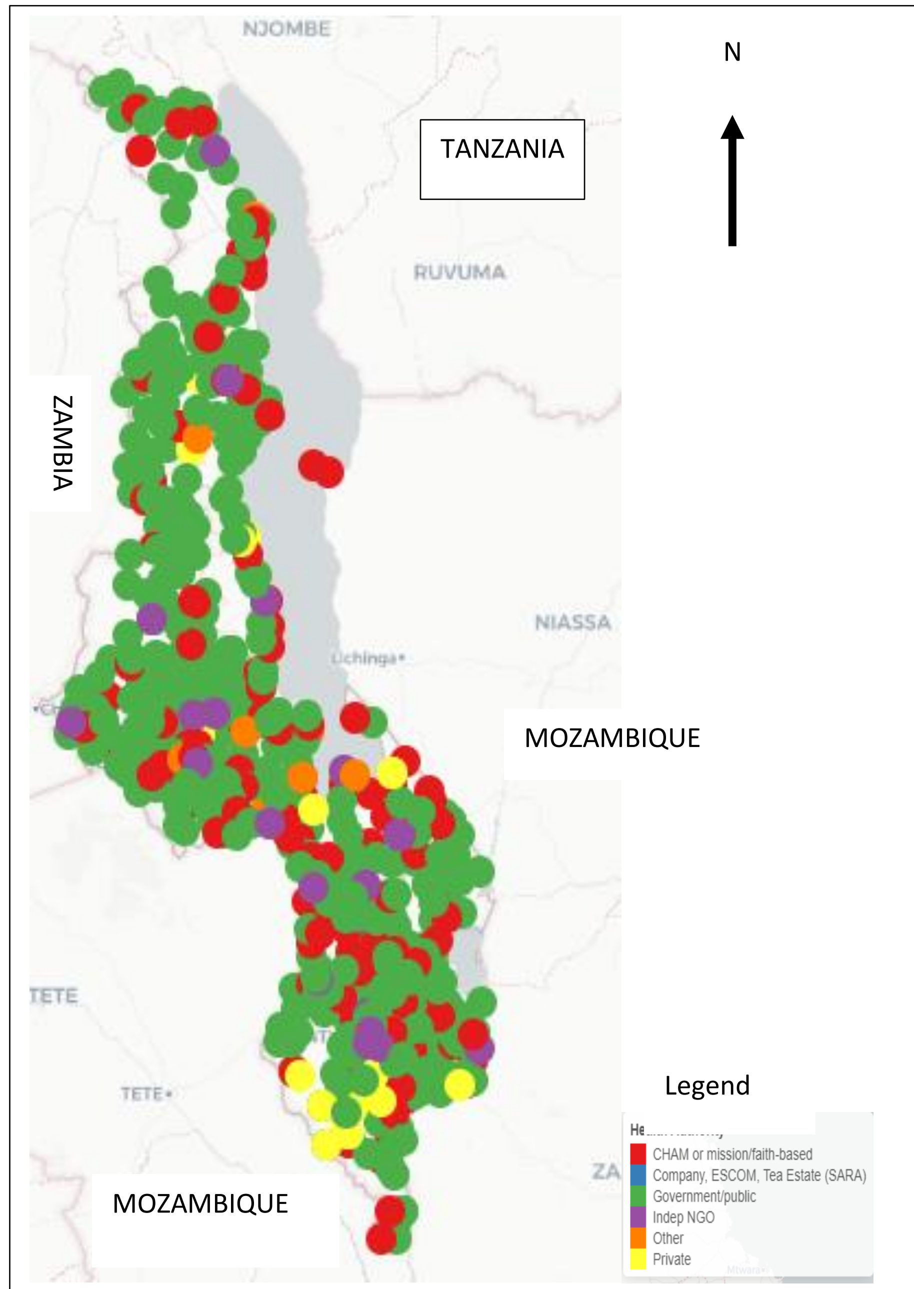
4.3 Research Setting

The research is carried out in Malawi. Malawi is located in South-East Africa and covers an area of 118,480 km^2. It is linked to its neighbours' sea ports via corridors of road transport. It is bordered by Tanzania, Zambia and Mozambique. The total population is estimated at 18 million based on the 2018 Population and Housing Census (National Statistical Office, 2018), although inter-censal population projections suggest the population is at 21 million as of 2022. The population is youthful, with 70 percent of the population aged below 30 years. This youthful population has implications on new HIV infections, since AIDS is predominantly spread through unprotected sexual intercourse in the heterosexual population. An estimated 52 percent of the population is female, while 48 percent is male (NSO and Macro, 2016). The bulk of the population is rural based, with over 80 percent estimated to live in rural areas of the country. Nearly 80 percent of the population relies on agriculture for its livelihood (National Statistical Office, 2018).

New estimates indicate that the HIV and AIDS prevalence rate is 7.7% in the general population, with nearly 1.0 million people enrolled on ART (National AIDS Commission, 2020). The HIV and AIDS prevalence is high among females 15 years and over at 9.3% when compared with males (6%) of the same age[13].The ART services are provided through a network of health facilities spread across the country. These are mainly public sector health facilities, alongside non-government health points such as Christian Health Association of Malawi (CHAM) and health units run by other religious organisations (GoM, 2023)- *see details in Chapter 6*). A select number of private health units are further included in the number of points offering ART services. Both non-government and private sector health units are extensions of the public sector HIV and AIDS service provision whose relationship with the government is through an agreement to provide the services.

[13] 2023 Spectrum, Naomi and UNAIDS Key population workbook consensus estimates for 2023

The figure below (Figure 4) depicts the map of Malawi and the location of ART service delivery points.



Source: https://mrc-ide.github.io/mwi-hiv/ART_facilities/index.html

Figure 4: Map of Malawi showing ART facilities

4.5 HIV and AIDS Areas of Focus for the Study

While there are numerous areas of the HIV and AIDS programme in Malawi, the study focused on ART and HIV testing under the HIV and AIDS Department of the Ministry of Health. This decision aimed at making the research manageable in the face of

resource limitations. The research focused on analysis of the ART programme governance and its implications for transaction costs and efficient service delivery. The ART services come under a broader category of expenditure - treatment care and support. The management and coordination expenditure area was selected as it constitutes elements that may be categorised as transaction costs. The treatment, care and support area was further selected because it accounts for the bulk of AIDS spending in the national response. It has accounted for approximately 60 percent of the AIDS spending since fiscal year 2014/ 15 (NAC, 2021). The governance costs (transaction costs) have been estimated to be between 15-30% for the period between FY 2011/12-FY 2019/20 (NAC, 2013; NAC 2021)

Thus, the ART is an area of the HIV and AIDS response in Malawi that is likely to drive costs in the near to long term in view of the new 95-95-95 targets, and is thus an appropriate area to focus on regarding efficiency. Further, procurement, distribution of drugs and actual delivery of services represent large expenditure items under ART (NAC, 2024; Songane, *et al.*, 2024).

4.6 Research Design

The research seeks to both explore the organisation for HIV and AIDS governance and to assess the factors that influence (in) efficiency in the delivery of ART services in Malawi, including transaction costs. The research assumed that the phenomena under study are complex. This includes how the delivery of ART services is organised, the interactions among key actors, behavioural attributes and the outputs they produce regarding HIV and AIDS service delivery costs and efficiency. Due to their complexity, the study assumed that they would be adequately understood through the use of multiple sources of data or perspectives. The data generated were both subjective (from a qualitative study) and objective (from quantitative study). The research is thus located within a pragmatic paradigm and deployed a concurrent exploratory mixed methods design. This design was adopted because it permitted both an in-depth exploration of how and why things are done the way they are done, as well as estimation of the influence of transaction costs using quantitative data on the efficiency of HIV and AIDS ART service delivery, which would be covered from cross-sectional or observational data. This approach is designed to respond more comprehensively to the research questions by allowing for the description of HIV and AIDS organisational phenomena,

perceptions on levels of transaction costs, estimation of transaction costs and efficiency outcomes (Bryman, 2012; Cresswell, 2012).

4.7 Research Process

The research proceeds from looking at the HIV and AIDS ART organisation in Malawi, and explores the relationship between transaction costs and efficiency in service delivery. The research problem (inefficiency in HIV and AIDS governance) is viewed from an economic governance perspective through the New Institutional Economics (NIE) lens, and in particular, through the Transaction Cost Economics (TCE) theory. The various concepts related to economic governance are reviewed by relating them to the governance of HIV and AIDS ART in Malawi. These concepts include transactions, transaction costs, efficiency, governance structure, institutions, institutional environment and governance mode or instruments. The health units offering ART services are viewed as firms whose objective is to maximise societal welfare through efficient HIV and AIDS ART service delivery. The NIE and TCE form the theoretical foundations of the research by locating HIV and AIDS ART service delivery in Malawi within the context of a theory of the firm. In particular, a key premise is that the quest for efficiency leads policy makers to design the most viable or efficient organisational arrangement in which HIV and AIDS service delivery transactions are organised. Because the HIV and AIDS services are a public good, this extended to the analysis of public sector organisation and considerations of the most equitable means to reach those in need of services. Consequently, a further key assumption is that the current organisation of HIV and AIDS services, in particular ART services, is efficient until proven otherwise (Williamson, 1999).

The research process flow is depicted in Figure 5 below.

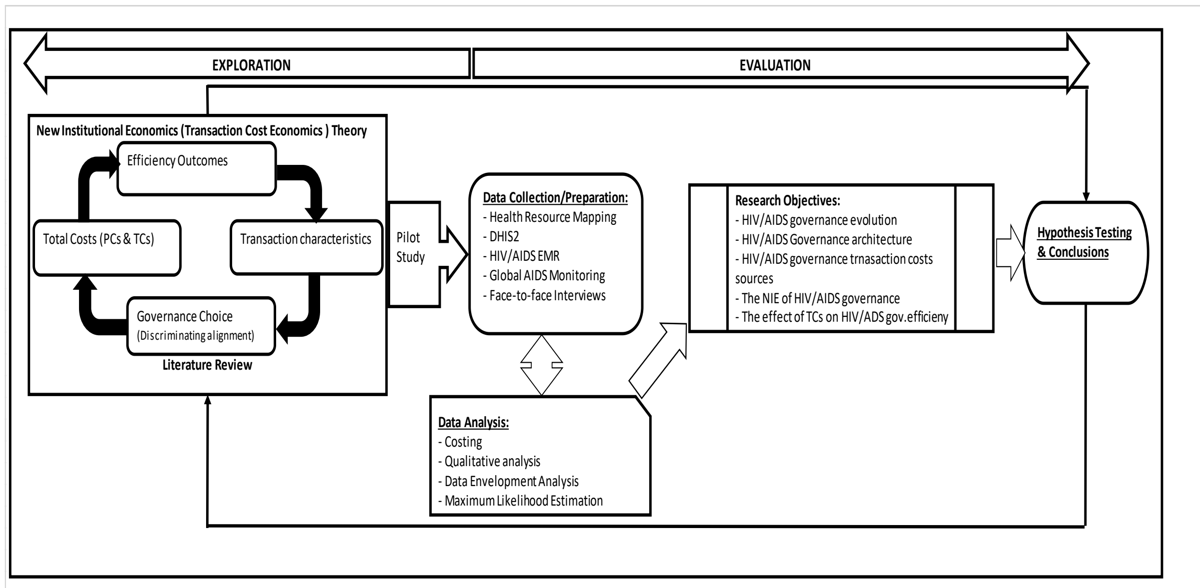


Figure 5: Research Process Flow

4.7.1 Research Operationalisation

As may be observed from Figure 5 above, the research process commenced with an exploration of transaction costs and how they impact efficient ART service provision. It explored the feasibility of implementing an evaluation of the effect of transaction costs on efficient HIV and AIDS service delivery through an exploratory pilot study. Besides literature review to understand the current knowledge and gaps related to transaction costs and the organisation of the ART programme, multiple datasets are reviewed with the aim to determine the feasibility of constructing variables to complete the efficiency analysis of HIV and AIDS ART delivery, and how transaction costs affect efficiency. The databases explored are:

- The Ministry of Health District Health Information System version 2 (DHIS2) held by the MoH Department of Policy and Planning;
- The Global AIDS Monitoring (GAM) data for Malawi for the period 2018 to 2019 and 2019-2020 held by the Department of HIV and AIDS and the National AIDS Commission;
- The Health financing Resource Mapping (RM) database for the seven mapping exercises since the resource mapping initiative started held by the Ministry of Health Policy and Planning Directorate;
- The Department of HIV and AIDS Management Information Systems electronic management records (EMR) capturing health facilities that provided ART

services for the fiscal year 2018and 2019 held by the MoH Department of HIV and AIDS and EGPAF; and,

- The Ministry of Health Community Health Systems database that captures distances from the District Health Office to each public health facility implementing ART services retained by the Community Services Department.

The theoretical concepts are operationalised through extraction of data on HIV and AIDS ART service delivery inputs (drugs, staff and infrastructure cost estimate) and outputs (number of people accessing ART and tests completed) covering the study period. For HIV testing output, the research used data capturing all tests conducted between July 2018 and June 2019, with the result that the testing numbers are higher than people on ART over the period under review. Expenditure estimation of the ART service delivery is accomplished through multiplication of the service outputs (number of people on ART and number of tests completed in the FY 2018/19 period) by per person per year unit costs based on the Global Health Cost Consortium (GHCC) database retained by the NAC. According to the GHCC database, these costs include the cost of drugs, staff, space and supplies. Where data on inputs were lacking, estimates are based on previous studies that either calculated unit costs, or proportions of costs for the particular inputs. The *efficiency* concept is operationalised by calculating efficiency scores through the deployment of Data Envelopment Analysis (DEA), a non-parametric approach to estimating the efficiency of decision-making units or production units. A description of the organisation of HIV and AIDS in Malawi provides the backdrop for understanding the current architecture of the HIV and AIDS services in the country, and determining its conformity with Transaction Cost Economics theory. It further provides insights into the rationales for the way the services are organised, and implications for transaction costs in HIV and AIDS governance.

4.7.2 Pilot Study

Pilot studies are recommended in research not only for their contribution to the improvement of the reliability of research results (Gudmundsdottir & Brock-Utne, 2010), but also for their ability to offer clarity on how research can be adapted given the conditions on the ground (Malmqvist, Hellberg, Möllås, & Rose, 2019). The pilot study had four aims:

a) To use the process to gauge the appropriateness of the research data collection tools
b) To gather data for estimation of some of the variables for the main study;
c) To inform the overall research design considering the operationalisation of some of the variables such as levels of transaction costs in the HIV and AIDS programme in Malawi, and,
d) To collect qualitative data that would form the basis for deductions on the extent to which transaction costs are considered when planning and implementing the ART programme.

The pilot study did not cover data collection with clients although during the course of the field visits, how the services are organised could be observed where such site visits coincided with ART clinic days. The pilot study was useful to understanding the data collection challenges and how to circumvent them and further improve the quality of the research. This phase was largely experimental and designed to gauge what was feasible regarding data availability to meet the requirements of the research.

The qualitative exploration included a focus on modes of organisation, key considerations in the planning and implementation of the ART programme, financial and economic costs associated with search for partners, negotiation of contracts or agreements, enforcement of performance and accountability standards, in line with TCE theory. This was useful to understanding how facility level activities could be categorised into production and transaction coordination activities as envisaged by the research. The capturing of data for this level was done through an interview schedule. The pilot study yielded useful outputs. First, it led to the refinement of the data collection instruments for both the qualitative and quantitative aspects of the research. In particular, the facility level questionnaire focused only on the capture of data that were available at the health facility level. For instance, operating costs in the public sector facilities could not be captured as these costs were handled from either the district or the centre in Lilongwe.

This phase further assisted in the determination of the content and direction of the practical application of the research and methods regarding the types of data to be collected. Data collected for facilities included perceptions of service providers

regarding the effectiveness of the services, challenges faced and suggestions for improvement of services.

4.8 Data Collection Methods Description

The research relied on primary data collected from key informant interviews with representatives of agencies involved in HIV and AIDS governance. It further relied on data collected from databases in the Ministry of Health, including Global AIDS Monitoring reports and Resource Mapping. As noted above, data collected from 35 health facilities across the country formed part of the data used in the research. These data were collected through document review, semi-structured interviews, a health facility interview guide and reviews of databases.

4.8.1 Literature Review

The study reviewed existing literature on the economics of organisation and how it related to HIV and AIDS governance. In particular, it adopted a narrative review approach with a view to generating a broad evidence base on the status of application Transaction Cost Economics within the overall framework of New Institutional Economics to healthcare and HIV and AIDS governance. A narrative review is deliberately chosen to limit publication bias in the available evidence on TCE and healthcare and HIV and AIDS governance (Remme, 2018).

Approach to and Scope of the Literature Review

The literature considered reflects a deliberate selection of a broad range of sources, both printed and online, published and unpublished to limit publication bias. Within the NIE, the research makes a deliberate choice of sources to focus on the NIE and transaction costs economics as theory of the firm. The key word combinations used in the internet literature search were “Economic governance”, “Transaction costs and efficiency”, “New Institutional economics”, “transaction costs economics” and “private and public bureaucracy”. The literature review used a narrative review approach where existing literature were systematically examined and synthesised to provide a comprehensive understanding of economic governance as applied to healthcare and HIV and AIDS service delivery. The narrative review was chosen because it allows for flexibility in

interpreting findings, which is beneficial to understanding complex or evolving fields (Green et al., 2006; Baumeister & Leary, 1997). In particular, it sought to explore whether or not inefficiency in HIV and AIDS service delivery was due the limited application of Transaction Cost Economics theory. It involved conducting a broad search for relevant documents across multiple academic databases, such as PubMed, Google Scholar and Econlit to capture a wide array of perspectives on economic governance in general and HIV and AIDS service delivery in particular. In addition, policy documents and programme reports on health and HIV and AIDS design and implementation were reviewed. Furthermore, the literature review included a Cochrane review on integration of HIV and AIDS and other health service to addressed efficiency (BulstraI, *et al.*, 2021). The search yielded a number of sources which were screened to focus only on those deemed relevant to the study. This was done by skimming through abstracts as well as the main text of selected sources. The screening aimed at narrowing the sources down to those relevant to the theoretical and empirical literature related to economic organisation, transaction cost economics, the relevance of TCE to the public bureaucracy efficiency analysis, and as applied to efficiency in healthcare and HIV and AIDS service delivery.

The analysis of the data from the literature review sought to respond to questions related to the theoretical and empirical background of the research with the aim to provide the basis for analysis of the application of TCE theory to HIV and AIDS ART governance. In particular, the literature review focused on conformity or non conformity of the HIV and AIDS ART governance with TCE theory as observed in Malawi. This would enable gauging both whether the whether deviations in the application of TCE theory are contributing to inefficient service delivery, transaction cost sources and the types of inefficiency generated. The literature review focused on the following questions:

a) How is governance mechanism choice determined using a transaction cost economics framework, and its relevance to private and public sector organisation of transactions, including HIV and AIDS governance?
b) What are the sources and characteristics of transaction costs in economic transactions in general, and as they relate to HIV/AIDS governance in particular?
c) What is the status of health care and HIV and AIDS service delivery (in) efficiency, and what are the gaps in the evidence?

d) What is the relationship between transaction costs and firm or public agency HIV and AIDS service delivery efficiency?

Analysis is achieved through constant comparison of the sources around the key pre-selected themes and sub-themes of economic governance, New Institutional Economics Theory, transaction costs economics and efficiency, and, public bureaucracy and efficiency in healthcare and HIV and AIDS governance. Sub-themes such as branches of NIE, market-hierarchy comparisons, theories of the firm-market boundary and efficiency were created to summarise the data from the literature review. This became the basis for writing up the literature review and making deductions regarding economic organisation and the implications for efficient HIV and AIDS ART service delivery.

4.8.2 Semi-Structured Interviews

Semi-structured interviews were used on a face-to face basis to collect data from key informants at agencies involved in HIV and AIDS service delivery at the national level, and at health facility level. These were facilitated through checklists. Semi-structured interviews assisted in understanding in-depth organisational aspects of the HIV and AIDS ART programme in the country, their roles in the HIV and AIDS response and perceptions on key issues affecting the response related to efficiency.

Data collection was conducted between August 2020 and July 2021. National level interviews covered 8 agencies with participants that were either policy or programme managers. Some data were further collected at the facility level using a semi-structured questionnaire that formed part of a pilot phase. The semi-structured questionnaire was administered to 35 staff at the health facility level. Participants to the study at each health facility were the staff-in-charges of the health facilities, usually Medical Assistants or Clinical Officers, staff directly involved in the provision of ART service delivery. These staff are trained and certified by the Ministry of Health in the procedures for provision of ART services[14]. In some instances where the interviews were conducted through phone, interviewees were requested to estimate these areas. This was due to the safety measures imposed in 2020 at the peak of a fierce wave of the COVID-19 pandemic. This meant that the observation and confirmations related to how the services were conducted could not be made by the researcher. Nevertheless, responses to some of

[14] In a few instances, there were no staff that had been formally trained in ART provision

the critical questions, including individual assessments of service delivery and challenges appeared to be consistent with some of the sites that were physically visited.

4.8.3 Database Review and Extraction of Relevant Data

The data on testing volumes, distances and catchment population sizes were obtained from the Departments of HIV and AIDS in the Ministry of Health, and the National AIDS Commission (NAC). Some of the quantitative data were collected at facility level during the pilot study. Funding data were obtained from the National AIDS Commission databases through the Global AIDS Monitoring reports data base. Further, data on trends in resources available for HIV and AIDS programme were collected from the Ministry of Health Department of Planning as captured over the five rounds of resource mapping. The quantitative data collected as part of the research formed a cross sectional dataset from which production costs based on input prices were calculated. The main outputs are the number of people alive and on ART and the number of tests completed in the 2018/ 19 fiscal year. The per capita per year costs were multiplied directly with the number of people tested and number of people that were alive and on ART for the FY 2018/19.

4.9 Sampling

4.9.1 Inclusion and Exclusion Criteria

The study focuses on agencies involved in the HIV and AIDS response in Malawi, in particular those linked to the ART programme. The inclusion of agencies considers two broad aspects: (a) policy, funding and coordination of HIV and AIDS, and (b) direct provision of HIV and AIDS ART services.

The study did not cover HIV and AIDS prevention programmes involving behavioural change interventions. This was due to difficulties in measuring outputs and outcomes under prevention programmes. Furthermore, other forms of prevention activities such as condom distribution, Voluntary Male Circumcision (VMMC) are not included under the ART programme in the national response, due to resource limitations and to ensure the research was focused on the ART programme.

The sampling of national level agencies for the qualitative assessment was purposive. This is the most appropriate approach when undertaking qualitative research since the

aim is to understand in-depth the phenomenon being studied, and not to seek statistical representation for the generalisation of results (Bryman, 2012). The research used the 2013 NASA listing as sample frame for the qualitative assessment (GoM, 2013), the full listing is attached at *Appendix 1*. This list provides a representative list of agencies funding or implementing HIV and AIDS activities in Malawi. The study covered eight agencies through face-to-face individual interviews with representatives of these agencies. These were selected on the basis of their participation in the HIV and AIDS response at the policy or operational level such as in ART service delivery and the financing of HIV and AIDS programme. These criteria were used to ensure the different types of HIV and AIDS programmes are covered, as well as to understand the internal organisational processes and cross- agency interfaces and their implications for transaction costs and efficiency of the HIV and AIDS programme in Malawi.

The study looked at the general HIV and AIDS coordination framework to gauge policy rationales for governance choice and perceptions on efficiency in policy, programme planning and implementation. Consequently, the National AIDS Commission (NAC) as the HIV and AIDS policy coordination entity in the country was included in the sample. The Department of HIV and AIDS (DHA), Leprosy and Viral Hepatitis in the Ministry of Health was the focus of the research since it is responsible for the governance of the ART and HTS. The ART programme covers general ART and Prevention of Mother to Child Transmissions (PMTCT). The Central Medical Stores Trust was included in the sample because of its national mandate to handle public healthcare procurement, warehousing and distribution of drugs and other commodities, including those related to ART. The PEPFAR was further included in the sample given its significant contribution to HIV and AIDS funding in the country. ActionAid-Malawi is included among data sources because it coordinated nearly a third of the Global Fund resources for the period under review, focusing on enabling activities to support the HIV and AIDS programme in Malawi. Other agencies were added to the list of data sources based on recommendations from the initial agencies selected, largely to gain additional data in a snowball fashion for the qualitative assessment. These were Population Services International, EGPAF and mother2mother.

Quantitative Assessment

After the pilot phase, it became evident that collection of some of the data envisaged at design would not be feasible because of time and financial resource constraints. This was compounded by difficulties in accessing records that took longer than anticipated, particularly those related to financial transactions. This is because most of the records related to ART service delivery are considered highly confidential and managed centrally, with no records at the facility level. Variables where data would be difficult to get without substantial resources for a representative sample of the health units included numbers of trained staff to provide ART services, health facilities that received support from partners, which would have a bearing on the quality and efficiency of services, space size used to carry out ART services at each facility and staffing levels per facility to estimate facility size. Consequently, the study was narrowed down to the collection of data that enabled determination of key outputs from ART service provision, and the expenditure estimation of these elements based on the GHCC per capita per person costs. In addition, some of these costs were estimated based on literature. Facilities that had incomplete data on ART for the period covered (fiscal year 2018/ 19) were excluded as efficiency calculations via the DEA required comparison of facilities that used similar inputs and produced similar outputs. Hence, out of the 747 health facilities assessed, 725 were included in the study for the efficiency evaluation. In addition, for the second stage DEA involving regression analysis to determine the effect of transaction costs on facility technical efficiency, the sample reduced to 658 as some of the observations were removed for lack of data on some variables such as distance from headquarters which were key to the analysis.

4.10 Estimating the Financial Inputs in ART Services

The components of the ART programme consists of service provision and coordination of the logistics to deliver the services. The following broad components of the ART programme are thus identifiable: (a) diagnosis to identify people in need of ART, which is also consistent with the achievement of the first 90% target by 2020, now the first 95% for Malawi after having already achieved the 90% target; b) the drugs and commodities acquisition and distribution to facility locations to reach people that need the drugs; c) patient treatment and adherence to drugs, and, d) monitoring the patients' response to treatment. These broad programmatic activities are designed to contribute

towards the UNAIDS Global HIV and AIDS target to contain the epidemic as a public health threat by 2030. Diagnosis is consistent with the aim to achieve the first 95%, where all people that are HIV positive know their status. The acquisition of the drugs and their distribution to ensure they are available to those that need them (treatment) would contribute towards the achievement of the second 95%, where the target is to reach 95% of all people that test HIV positive with antiretroviral treatment. The programme further runs ART clinics to ensure those that are enrolled on treatment adhere to the regimens. The ART clinics are either run on specially allocated days in the week, which is in the majority of the public facilities (over 90%), or run as part of the routine facility care processes in an integrated fashion. The latter arrangement is common to private sector facilities and CSO owned facilities that are part of the ART programme. The monitoring of clients' response to treatment further contributed to the third 95% (achievement of viral suppression in at least 95% of the clients on treatment), besides monitoring toxicity levels.

4.10.1 Estimation of Input Costs

One of the tasks that the research sought to achieve was estimation of the ART programme inputs that would ultimately be used to assess the (in) efficiency of the programme. The cost estimate would include both production and transaction costs of the ART programme. Given that currently transaction costs are not explicitly tracked as a specific cost category, a further aim was to operationalise the transaction cost concept through the identification of transaction cost elements of the programme. This would extend to assessment of available data sources to estimate these costs.

The research noted that HIV testing services (HTS) are a key strategy to identify HIV infection and facilitate timely initiation on ART. The HTS is thus a key strategy towards meeting the first 95 % and contributes significantly to both the second 95% and the third 95% in the UNAIDS 95-95-95 targets by 2030. With 2015 WHO test and start guidelines, HTS represents a key piece in efforts to combat HIV and AIDS. The majority of HIV testing in Malawi is facility based, although a self-testing programme has also commenced. The current study uses facility-based testing to estimate costs and ultimately, the efficiency of HTS at the facility level based on annual testing volumes by facility. The initial test is a finger prick rapid diagnostic test (RDT), followed by a confirmatory RDT when the outcome is positive. For discordant or inconclusive tests,

where the first test is positive but the confirmatory test is negative, clients are advised to do a repeat test after four weeks. Because the initial HIV positive result has to be confirmed by a second test, it follows that an HIV positive outcome is associated with a higher cost per client.

a) ART Expenditure Estimation Approach

The study used a simple classical economic approach to estimate ART inputs and outputs to facilitate efficiency estimation for the fiscal year covered. Per capita per year service delivery unit costs for ART services and commodities were used to estimate the cost per ART site. Outputs achieved at each ART facility such as number of ART clients per facility and number of tests completed for the year at each facility were multiplied by per person per year (PPPY) unit costs. The unit PPPY cost estimates were obtained from the Global Health Cost Consortium (GHCC) cost database that was accessed through the Malawi National AIDS Commission.

The GHCC cost database (Unit Costs Programme Data) included costs for ART and HIV testing and counselling services across all facilities (747 as of 2018) in the country. The costs are reported as per person per year (PPPY) for the year 2018. They are based on full programme costs for providing free facility based ARVs and HIV testing and counselling services for one year, with 2018 as the reference year. The average cost of providing free ARVs across all facilities in Malawi in 2018 was estimated at $115.32 per person per year (average ARVs per person for all regimens). With respect to HTC services, the research used a cost estimate of $4.52 based on client-initiated facility-based testing. This cost was preferred over provider initiated for ease of calculation and the fact that there were perceptions from key informants that the provider-initiated testing and counselling appeared to have slowed down, largely due to cost in recent years.

Community outreach HTC cost estimates were excluded as the focus was on facility-based ART and HTC services. At $9.49 per person per year, the average cost for provision of HTC services for key populations was further ignored because it would have inflated the total HTC services as it was almost double that of both the client and provider initiated HTC. The high average HTC per person per year cost for key populations is attributed to the specially targeted interventions to reach the key

populations which tended to be costly. The average costs for ART and HTC were then used to estimate the total cost of providing ART and HTC services per facility, respectively. The total costs per facility were used as inputs in the calculation of efficiency scores via DEA.

Key assumptions in the GHCC costing need to be further highlighted when adopting these average cost estimates for ART and HTC services as they have implications on the interpretation of the financial cost estimates. Firstly, the GHCC estimates for both the ART and HTC services were based on a top down approach (based on total programme costs) rather than on the costing of the different cost elements at each health facility that provided ART or HTC services. Secondly, the costs do not include society costs nor those incurred by individual clients to access ART and HTC services. Thus, the average financial costs for both ART and HTC services may be underestimated. Nevertheless, given the significant amount of financial resources that are spent on ART and HTC services each year, these cost estimates are likely to represent mean per person per year costs for the programme for period for which they were estimated.

The above approach was adopted due to its practical application for the purpose of estimating the inputs for DEA efficiency analysis. Unit costs were further obtained from recent ART costing studies for elements that were not included in the GHCC cost database. As noted before, the main activities that the study was concerned with were treatment (ART) and HIV testing and counselling (HTC) services. The financial input estimation covered a total of 747 facilities, but 22 sites were removed. The reasons for removal of some of the sites were firstly, that they only provided one aspect of the ART services, such as treatment only and excluded the testing aspects. This was considering that, to complete efficiency estimation in DEA, the service provision units needed to use similar inputs and produce similar outputs to facilitate comparison. In addition, one site was reported to have discontinued provision of ART services. Ultimately, 725 sites were included in the costing database.

The ART health facility inclusion and exclusion and estimation steps are summarized in Figure 6.

Figure 6: Facility Based ART & HTC programme cost estimation process

The ART transaction costs were derived from the total costs of ART provision using Bayesian estimation in the final estimation step, which involved setting up a prior distribution of the transaction costs, and then updating the prior with the data. One of the principles for setting up a prior is the triangulation principle where at least three data points, either over time or from different sources estimate a proportion or quantum of a parameter (Stafford, 2020). Over the three year period since FY 2015/16, the NASA estimates governance or transaction costs to be 20.2% (GoM, 2021). The National Health Accounts (NHA) estimates the administration and governance costs to be 26.2% (GoM, 2018) while the Global AIDS Monitoring (GAM) reports 22% (GoM, 2022) to be programme management costs that constitute transaction costs. Based on this information, the study assumed that the proportion of transaction costs from each US$ spent on ART has a mean of 0.2, or 20 percent in a year based on data from the Malawi NASA. The rate of observing a transaction cost, λ, in ART expenditure in Malawi in a year is then estimated-*see Appendix 3a for estimation details*.

4.11 Data Management and Analysis

Qualitative Data

Qualitative data were recorded using transcripts from field notes and audios for interviews that were recorded. These transcripts were stored in a single folder, with both individual documents and folders password encrypted. These were only accessible to the researcher.

Analysis of Qualitative Data.

Qualitative data were transcribed and coded to generate themes, sub-themes and categories-the themes used to analyse the data are attached at *Appendix 1*. Analysis was guided by the theoretical framework adopted for the study, as well as the research questions for the study (Cresswell, 2012; Blanche, *et al.*, 2006). Thematic analysis was deployed to analyse the data through the software package Nvivo 12 Plus, 2018. The data were coded from transcripts in a two-stage process - an open coding through an inductive process without reference to predetermined themes to reveal the main themes emerging from the data. This was followed by reorganising the coded data further into the predetermined themes to answer the research questions in a second stage. The analysis followed an iterative process where data across different sources were compared to observe trends and peculiarities. Further, the creation of a coding framework and in-depth reading of literature and transcripts facilitated the identification of themes, causal links, similarities and differences in the emerging findings. This was the basis for summarising and interpreting the qualitative data. Main themes are presented as topics in the results, while basic codes formed the basis for writing up the paragraphs in the findings.

Qualitative description was further used to analyse and present the results of the study. In this approach, analysis is not confined to any prior theoretical position (Sandlowesski, 2000). Instead, phenomena are studied in their natural environment and as they manifest themselves (Sandlowesski, 2000; Lambert & Lambert, 2012). This applied to sections on the results on the HIV and AIDS response organisation related to treatment, care and support and conformity with TCE theory. This approach was adopted because it offered the flexibility to describe the HIV and AIDS organisation in Malawi without the restrictions of a pre-conceived epistemological framework, as it neither qualified as

phenomenology, ethnography nor grounded theory. This facilitated looking at the ART delivery framework as a system within the larger health system, and the description of the main factors that influence (in) efficiency in HIV and AIDS service delivery.

Quantitative Data

Facility level data were extracted from the databases in Microsoft Excel format from the MoH Department of HIV and AIDS MIS EMR as indicated before. To ensure they were secure, the files were also password protected and only accessible by the Researcher and a Research Assistant that assisted with the organisation of the data into a single dataset. Data cleaning was done using STATA 17 (Statacorp, 2017) using the codebook command to identify and correct unique observations. Data analysis was done in both STATA and Microsoft Excel. Data analysis involved production of descriptive statistics to identify the main patterns of the data, including minimum and maximum values and means. Further analysis involved detection of associations between the variables of interest through multivariate regression analysis.

Quantitative Data Analysis

To assess the determinants of ART efficiency across facilities, the study used a two-stage Data Envelopment Analysis (DEA), assuming variable reruns to scale (VRS). The input oriented VRS model was chosen because it is the more realistic assumption to make given that health facilities that provide ART were unlikely to operate at full efficiency, implying that the constant returns to scale (CRS) assumption would not hold. The DEA calculates an efficiency score for each decision unit, or health facility in the case of the current research. Each score takes a value of between 0 and 1. The DEA is chosen because it was designed primarily for the evaluation of efficiency in the not-for-profit sector and may also be applied to public sector programme, and, is therefore, appropriate for the current study. However, since DEA is a non-parametric technique, it does not explain the cause of variation in efficiency. To address this limitation, the study uses regression analysis through estimation of a generalised linear model using ordinary least squares model (Wooldridge, 2006; Shao and Lin, 2002; Jacobs, 2001; Charnes *et al.*, 1978) to estimate the effect of transaction costs and other control variables on HIV and AIDS ART facility efficiency.

4.12 Responding to the Research Questions

An evaluation of the effect of transaction costs on the (in) efficiency of ART service delivery is achieved via the application of various statistical and econometric analyses, including regression analysis and non-parametric analysis such as Data Envelopment Analysis, as noted above. A combination of the qualitative analyses and quantitative analyses delivers the results of the research around the inefficiency of the HIV and AIDS ART programme and the types of inefficiency observed. It further permits the assessment of how transaction costs impact the ART programme (in) efficiency. Application of the TCE discriminating alignment hypothesis to the public sector organisation to assess whether the HIV and AIDS service delivery organisation conforms to TCE theory helps to respond to the TCE conformity research question. This is achieved through a qualitative description of the organisation of HIV and AIDS services and assessing its instruments and performance attributes to answer questions related to the programme's conformity with TCE theory. Furthermore, application of the TCE model where behavioural assumptions of bounded rationality and opportunism influence the inclusion of safeguards in contractual relationships helps to uncover the sources of transaction costs in ART governance (its financing and programme implementation) and their implications on HIV and AIDS service delivery inefficiency.

4.12.1 HIV and AIDS ART Programme via Data Envelopment Analysis

As noted before, the research used the DEA to estimate facility-based ART and HTC service provision efficiency, and to determine the types of inefficiency the programme displays, and how transaction costs impact service delivery efficiency. The DEA is an application of linear programming methods to construct a piece-wise efficiency frontier over the data (Murad, Magersa, & Tewfik, 2017). All the health units are compared to the most efficient of their peers to determine their level of efficiency or inefficiency. The study adopts the variable returns to scale (VRS) model because the constant returns (CRS) model is not feasible in practice, as the latter assumes that all DMUs operate at the optimal level. The VRS model further enabled the separation of technical and scale efficiency of the health units assessed. As a consequence, while the DEA estimates the overall technical efficiency (OTE), the focusis pure technical efficiency (PTE) and scale efficiency, which are dependent on governance or management efficacy in resource allocation and control. This is consistent with the study focus of transaction cost economics analysis of the ART programme.

The DEA calculates an efficiency score for each decision unit. Each score takes a value of between 0 and 1. The DEA is chosen over parametric approaches to estimate efficiency because it allows for multiple inputs and outputs to estimate efficiency, which is difficult under the other approaches such as single stage stochastic frontier analyses. In addition, compared to other methods for estimating efficiency, it permits identification of the sources of inefficiency, and is, therefore, useful in informing policy recommendations regarding where modifications are required to make the relatively inefficient decision-making units or programmes more efficient (Jacobs, 2001; Bendoly, Rosenzweig, & Stratman, 2009). Nonetheless, since DEA is a non-parametric technique, it does not explain the cause of variation in efficiency, hence, the use of regression analysis estimation in the second stage.

The procedure of estimating efficiency scores and determining the effect of transaction costs on efficiency is described below.

a) First stage DEA

In the first stage, the efficiency scores are calculated from the following objective function:

$$\max E_r = \frac{\sum_{j=1}^{k} u_j y_{jr}}{\sum_{i=1}^{l} v_i x_{ir}} \tag{4.1}$$

Subject to:

$$\frac{\sum_{j=1}^{k} u_j y_{jr}}{\sum_{i=1}^{l} v_i x_{ir}} \leq 1 \tag{4.2}$$

$$u_j ; v_i \geq 0 \tag{4.3}$$

Where:

$i = i^{\text{th}}$ input, $i = 1, \ldots, l$;

$j = j^{\text{th}}$ output, $j = 1, \ldots, k$;

$r = r^{\text{th}}$ HIV and AIDS programme, $r = 1, \ldots, n$;

E_r = objective measure of efficiency for the rth ART DMU or facility, and r = a specific ART facility to be evaluated;

y_{jr} is the amount of output j from the ART facility r;

x_{ir} = amount of input i used by the ART facility r;

u_j =weight chosen for output j;

v_i = weight chosen for input i;

n = number of ART facilities;

l = number of inputs,

k = number of outputs.

The solution to the above problem can be solved as a minimisation problem, that is, it can be specified as a minimisation instead of a maximisation problem that converts it into a linear programme by restricting the denominator of the objective function to 1, and then adding the denominator as a constraint to the problem (Jacobs, 2001).

The linear programme form of the problem is specified as follows:

$$\max E_r = \frac{\sum_{j=1}^{k} u_j y_{jr}}{\sum_{i=1}^{l} v_i x_{ir}} \quad (4.4)$$

Subject to:

$$\sum_{i=1}^{l} v_i x_{ir} = 1 \quad (4.5)$$

$$\sum_{j=1}^{k} u_j y_{jr} - \sum_{i=1}^{l} v_i x_{ir} \leq 0 \quad (4.6)$$

$$u_j; v_i \geq 0 \quad (4.7)$$

i-1,2,…,k; j=1,2,…,l; r=1,2,…,n

$\boldsymbol{u}$ *and* $\boldsymbol{v}$ are the unknown parameters estimated by the DEA programme

The solution to the above linear programme yields the efficiency score E_r for ART facility r, and $0 \leq E_r \leq 1$. This model is specified under the constant returns to scale (CRS) assumption in line with Charnes, Cooper & Rhodes (1978). Nevertheless, the CRS model is only appropriate when all the decision-making units are operating at optimal level, which is a rare occurrence in practice. In the current study, this would mean that all the ART health units are considered to operate efficiently, which is unlikely to be the case given differing capacities, management styles and demographics that impact facility-based ART service delivery efficiency. Consequently, the model is specified under the variable returns (VRS) assumption considering the impossibility of simultaneous optimal operation for all health units that provide ART services.

Under the variable returns to scale assumption, and following (Banker, Charness, & Copper, 1984) the original CRS model is extended. The linear programme problem to be solved is:

$$Max\ h_0 = \sum_{r=1}^{k} u_r\, y_{rj_0} + u_0 \qquad (4.8)$$

Subject to:

$$\sum v_1\, x_{ij_0} = 1 \qquad (4.9)$$

$$\sum v_r\, y_{ri} - \sum v_i\, x_{ij} + u_0 \geq 0, j = 1, \ldots, n \qquad (4.10)$$

$$u_r, v_i \geq 0 \qquad (4.11)$$

$$u_0 \geq 0 \qquad (4.12)$$

With notations as given in equation (4.1), the additional term u_0 is an intercept and is unconstrained in sign (Bjurek, Hjalmarsson, & Forsund, 1990). It determines the returns to scale, where $u_0 < 0$ shows increasing returns to scale, while $u_0 > 0$ indicates decreasing returns to scale, and $u_0 = 0$ depicts constant returns to scale (efficient production).

b) Second Stage DEA

In the second stage, the aim is to estimate the determinants of pure technical efficiency (PTE) of the health facilities offering ART. A number of variables, including facility size measured by the number of ART clients per facility, catchment population, number of years that the facility has been offering ART services and other factor are set as explanatory variables. The explanatory variables include the estimated amount of transaction costs for each of the facilities. For facility size, the number of ART clients per site is chosen because it is a more stable output than either HIV tests completed or the catchment population estimate. All the variables used are for the fiscal year 2018/19.

The study used a generalized linear regression model using Ordinary Least Squares (OLS) estimation techniques. Hence, the standard assumptions regarding unbiasedness of OLS were applied, as depicted in Table 3 below.

Table 3: OLS assumptions

#	**Assumption**	**Assumption statement**
1	Linearity in parameters	$y_i = \beta_0 + \beta_1 x + \cdots \beta_k x + \mu_i$, i=1,2,…, n
2	Random sampling	(X_i, Y_i): $i = 1,2, \dots, n$
3	Sample variation in the explanatory variable	Sample outcomes on x, $\{X_i, i = 1,2, \dots, n\}$, are not all the same
4	Zero conditional mean of the error term	$E(u\|x_1, x_2, \dots, x_n) = 0$

Nevertheless, because the data were not drawn via random sampling, the study's estimated the coefficients were estimated with robust standard errors.

Considering the general case with h independent variables, the aim is to estimate the parameters, $\beta_0, \beta_1, .. \beta_h$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots, \beta_h x_h + \mu \quad (4.13)$$

The OLS are selected to minimise the sum of squared residuals as follows:

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta x_1 - \cdots \beta_h x_{ih})$$

(4.14)

In the above case , there will be *h+1* estimates.

The minimisation problem is solved with the aid of multivariate calculus, which lead to *h+1* linear equations in *h* unknowns, $\beta_0, \beta_1, .. \beta_h$ to achieve the first order conditions.

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta x_1 - \cdots \beta_h x_{ih}) = 0$$

$$\sum_{i=1}^{n}x_{i1}(y_i - \beta_0 - \beta x_1 - \cdots \beta_h x_{ih}) = 0$$

(4.15)

$$\sum_{i=1}^{n}x_{12}(y_i - \beta_0 - \beta x_1 - \cdots \beta_h x_{ih}) = 0$$

.

.

.

$$\sum_{i=1}^{n}x_{ih}(y_i - \beta_0 - \beta x_1 - \cdots \beta_h x_{ih}) = 0$$

4.12.2 Empirical Model

The empirical model to estimate health facility technical efficiency is specified as follows:

$$dmu_efficiency_i = \beta_0 - \beta_1 logtcost_i + \beta_2 sitematurity_i - \beta_3 distance_i + \beta_4 catchmentpopn_i + \beta_5 dmusize_i + \beta_6 Adminsitrationtype_i + \beta_7 locationrural_i + \beta_8 locationdistrict_i + \mu_i \quad (4.16)$$

Where:

Dmu_efficiency= estimated ART facility efficiency; *logtcost*= logarithm of estimated ART transaction cost by site; *sitematurity*=period over which facility has been providing ART; *distance*= facility distance from Lilongwe City; *catchmentpopn*= the estimated total population served by the facility; *dmu_size*=volume of ART clients per facility; *locationrural*=rural location of facility; *Adminstrationtype*=facility ownership type (public, CHAM, private, CSO); *locationdistric*t=district location of facility. The βs are the parameters to be estimated, and μ_i is a disturbance term.

4.12.3 Variables Description

The variables included in the efficiency estimation and its determinants were the efficiency scores calculated from DEA first stage, which was the dependent variable. Control variables included the length of time that the facility had been providing ART services (site maturity), facility size estimated by the number of people alive and on ART for the FY 2018/19, the catchment population of the health facility that represented a service demand factor, facility location (rural/ urban and district) and type of administration ownership (private, public, CHAM). Most of these variables are tracked in the Department of HIV and AIDS Management Information System (DHAMIS). Since not all transaction costs are consciously and adequately tracked in the ART and HTS service delivery, and as noted before, the study resorted to Bayesian estimation methods to calculate the transaction costs variable that was used to evaluate the effect of transaction costs on efficiency as presented in *Appendix 3a*.

The table below describes the variables in the empirical model and their expected signs (Table 4)

Table 4:Variables description

#	Variable name	Description	Type	Notation	Expected sign
	Dmu_efficiency	Health facility ART efficiency score calculated through	Continuous, dependent	Dmu_efficiency	

#	Variable name	Description	Type	Notation	Expected sign
		the first stage DEA analysis			
1	Log of Transaction costs	Estimated transaction costs per facility	Continuous, explanatory	Log_of_transaction_cost	(-_)
2	Site-maturity	Number of years facility has been providing	Discrete, explanatory	Site maturity	(+)
3	Distance	Facility distance from Lilongwe City and MoH headquarters	Continuous, explanatory	Dist.	(-)
4	Catchment population	The population estimated to be served by the health facility	Discrete, explanatory	Catchment popn.	(+)
5	DMU or ART facility size	The numbers of ART Clients per health facility	Discrete, explanatory	Dmu_size	(+)
6	Administration type	Set of dummy variables for public, private, CSO and CHAM ownership. Depending on ownership type, it takes a value of 1 for specific ownership, and zero otherwise	Dummy, explanatory	Public, private, CHAM-MAM, CSO	Unknown
7	Location	Ruraland urban, takes value of 1 for rural and zero otherwise	Dummy, explanatory	Rural, urban	Unknown
8	Location-district	Dummy variables for districts, takes value of 1 for specific district variable and zero otherwise	Dummy, explanatory	Name of district	Unknown

As may be observed from Table 4, the empirical model includes a number of variables. Economic literature suggests that *transaction costs* vary directly and are negatively correlated with efficiency rates of decision-making units. This implies that up to a point, they contribute to efficient production cost levels, and thereafter, additional transaction costs yield inefficiencies (Singh, 2017; Hünecke, 2019). The caveat to the inclusion of the transaction cost variable is that it is based on an estimation of the proportion of these costs out of the total costs via Bayesian statistical estimation.

Facility size, the volume of clients on ART in this research was expected to be positively correlated with efficiency. This is because larger volumes of ART clients would lead to economies of scale when compared with smaller volumes. Empirical evidence further suggests that firm size is associated with better efficiency outcomes (Martinez, 2008). Site maturity, or the period the ART site has been offering ART services is expected to have a positive effect on efficiency, and hence, should enter the efficiency function positively. This is likely to arise from the knowledge and experience that comes with learning over time. The effect of the location of the facility or ART site- that is, in in terms of rural or urban location is unknown. Nevertheless, in rural locations, the populations may be so sparse as to impact on the demand for ART services for given levels of inputs. Consequently, with urban location as the reference variable, the expected sign for rural location of ART facility was hypothesised to be negative. The district location is included to account for community and ecological factors that may have a bearing on the (in) efficiency of ART services. The expected sign is unknown for each district, although districts further away from Lilongwe City and in more remote areas should reflect a negative correlation with facility efficiency. This may be occasioned by high distribution and headquarters monitoring costs that would negatively affect facility efficiency. The type of administration-public or private- may have an effect on ART service delivery efficiency due to differing management styles, but the effect sign is unknown a *priori,* although some studies suggest private bureaucracy is likely to be efficient when compared with public bureaucracy (Etake, 2018; Freeman, Nam-Speers, & Tokac, 2020). Similarly, the health service delivery level of the facility sign is unknown a *priori*, but may similarly be impacted by demographic factors that may increase or decrease demand for services, and, hence positively or negatively affect facility efficiency. The expected sign on distance of the ART site from headquarters is expected to be negative, as the costs of distribution, monitoring and supervision are likely to increase with the remoteness of the site, and may impact efficiency negatively.

4.12.4 Data for Efficiency Analysis

The data for the variables in the efficiency analysis comes mainly from two sources, namely, the electronic management records (EMR) from the DHAMIS in the Ministry of Health, and from the Community Health Services Management Information System (MIS). The facility level efficiency scores are calculated from the first stage DEA. Data

are analysed using DEA version 2.1 (CEPA, and www.une.edu.auand econometricsand cepaw) and STATA 17 (StataCorpLLC, 1985-2021) .

4.12.5 Pre and Post Estimation Diagnostics for the second-stage DEA

A number of tests were completed to test the assumptions of the regression model and the model fit. These included the normality, heteroskedasticity, multi-collinearity and model specification tests.

Heteroskedasticity

The Breusch-Pagan and Cook-Weisberg test for heteroscedasticity was used to test the constant variance hypothesis for the residuals. Results do not support the evidence that there is heteroscedasticity (p = 0.16).

Variance inflation factor

Multicollinearity of the variables was tested using the variable inflation factor (VIF). A VIF of more than 5 suggests a variable is highly correlated with other variables (Shresha, 2020). From the VIF test, only two variables, namely, primary and secondary service delivery levels record VIF values greater than 5. This suggested that besides the two variables, multicollinearity was not a big problem in the data.

The results of the multicolinearity test via the VIF are presented in Table 5.

Table 5:Multicollinearity Test Results using VIF

	VIF	1and VIF
Primary	5.879	.17
Secondary	5.846	.171
Distance	2.722	.367
Transaction cost by site	1.976	.506
Site maturity	1.626	.615
Rural	1.596	.627
Catchment population	1.576	.635

	VIF	1and VIF
Blantyre	1.431	.699
Karonga	1.382	.724
Chitipa	1.343	.745
Mzimba	1.326	.754
Thyolo	1.259	.794
Nkhatabay	1.231	.812
Chikwawa	1.229	.814
Mulanje	1.228	.814
Rumphi	1.214	.824
Nsanje	1.213	.824
Mangochi	1.201	.833
Dedza	1.196	.836
Dowa	1.169	.856
Ntcheu	1.165	.859
Zomba	1.154	.867
Machinga	1.151	.869
Public	1.149	.87
Phalombe	1.137	.879
Chiradzulu	1.136	.881
Kasungu	1.131	.884
Mchinji	1.126	.888
Nkhotakota	1.097	.912
Neon	1.086	.921
Ntchisi	1.081	.925
Salima	1.078	.928
Balaka	1.077	.929
Likoma	1.052	.951
Mwanza	1.032	.969
Mean VIF	1.551	.

Model Specification Test

A Ramsey RESET test using powers of the fitted values of health facility ART efficiency testing the null hypothesis that the model had no omitted variables showed the model was correctly specified. The F-statistic was not statistically significant, an indication that the model was correctly specified ($p = 0.3281$).

Estimation Issues

Initially, second stage DEA was expected to be achieved via the Tobit model, and analysis through maximum likelihood estimation. Estimation via the Tobit model produced results that appeared consistent with the expected sign of variables such as transaction costs and distance regarding parameter signs. Nevertheless, the model returned a negative pseudo R squared suggesting issues with model specification and data fit. The small sample size effect was ruled out by running the model with bootstrap procedures, the first of which still produced a negative pseudo R squared. The second, a Bayesian analysis estimation produced consistent results, but convergence tests showed that the model did not converge with respect to most of the explanatory variables. Only sigma and the constant showed convergence. Trace plots showed trends, and the kernel density plots were not similar. Consequently, the Tobit model was deemed not to fit the data for modelling the variation in facility-based efficiency. A generalised linear model using Ordinary Least Squares (OLS) was fitted instead to estimate sources of variation in ART facility technical efficiency, an approach that is also adopted in other previous studies (Onyancha, 2022)-*See Appendix 4* for the regression output.

4.13 Reliability and Validity of Results

Quantitative data were cleaned to minimise errors that would affect the calculation of the various scores. This included running the codebook command in STATA to observe and correct outliers from expected variables. Further, before conducting regression analysis, the data were assessed to ensure the data assumptions were valid. This involved conducting normality tests on which basis the regression analyses were made.

Reliability is the extent to which measurements can be repeated when they are done by a different person (Drost, 2011). The research used measures such as technical efficiency that are common to organisational and economic research. For the qualitative elements of the research, internal validity of the results is enhanced firstly by triangulation of

findings from different methods and information sources. Secondly, the data were collected by the researcher who was also responsible for analysis and interpretation of both qualitative and quantitative data, and, hence, a measure of consistency in the coding and interpretation was achieved.

Drost (2011) further highlights the need to address statistical conclusion validity and internal validity in quantitative research. Statistical conclusion validity has to do with the establishment of relationships between variables, while internal validity in quantitative research addresses establishment of causal relationships and ruling out confounding factors (Drost, 2011). In the current research, statistical conclusion validity is achieved through completion of multivariate regression analysis, to establish the relationships between the dependent variable and the independent variables of interest. The current research used regression analysis to assess the effect of transaction costs on facility-based HIV and AIDS ART service delivery. In addition, internal validity is assured by controlling for factors that affect efficiency in ART service delivery. Further, to the extent that the research covered 97 percent of the health units delivering ART (725 of the 747 facilities), it was able to achieve external validity, that is, whether the results were representative of the population of health units (Lakshmi & Mohidee, 2013).

4.14 Ethical Considerations

Research ethics aim at protecting the welfare of the participants (Bryman, 2012). The study protocol was reviewed and approved by the University of Kwazulu Natal Institutional Biomedical Research Ethics Committee (BREC) (*Protocol number HSS/1743/017D*). The study protocol was further submitted and cleared by the Malawi National Research Committee in the Social Sciences and Humanities (NRCSH) at the National Science and Technology Commission (*Protocol number. P.03/19/360*). The study entailed the collection of data from individuals involved in HIV and AIDS service delivery. Participation in the study was voluntary and based on a signed written informed consent after provision of information about the study. The formal consent form was administered prior to interviewing a participant. The research further follows an anonymous reporting format of results or facts relating to individual agencies or facilities providing HIV and AIDS services. This ensures that findings are confidential and not traceable to the individual agency or facility in the report.

4.15 Research Limitations

There a number of imitations to the study. First, given the difficulties with monitoring and costing of transaction costs, the estimates for transaction costs in ART governance used in the research should be considered partial. Nevertheless, the approach adopted provides insights into how these costs can be estimated and monitored. Second, some of the costs included as part of the production costs are estimated from findings from previous studies. This includes the cost of staff at each facility. This may have led to an underestimation, although they provide an approximate value of the level of costs. Further, although a number of non-state agencies contribute towards quality service delivery for ART services, the costs related to their services were not captured, suggesting that estimates regarding service delivery and transaction costs may be an underestimation. Third, since some of the results of the study are based on perceptions and opinions of individuals, it is possible that they carry bias, and hence results need to be interpreted with caution. Nonetheless, the use of multiple sources of the data assisted in ensuring conclusions were only made after cross-checking with other sources for the qualitative data. Finally, the efficiency estimation uses a cross sectional dataset for the fiscal year 2018/ 19, and, therefore, results are confined to a single period. Hence, although the results present a baseline for ART governance (in) efficiency, longitudinal studies such as the Malmquist DEA estimation that also capture the efficiency of facility-based ART provision and total factor productivity levels over time would be more desirable.

4.14 Chapter Summary

This chapter has presented the approach and methodology adopted to collect cost data on HIV and AIDS service provision, focusing on health facility-based HIV and AIDS ART service delivery governance in Malawi. It has further highlighted how the data were analysed, emphasising the role that a pilot study played in shaping the direction and types of data collected. The study limitations have further been highlighted.

The next chapter presents the results of (in) efficiency analysis that also helps to determine the types of inefficiency that the ART and HTC programme displays. It further estimates the effect of transaction costs of facility-based ART efficiency.

CHAPTER FIVE

THE MALAWI HIV AND AIDS ART PROGRAMME (IN) EFFICIENCY AND THE EFFECT OF TRANSACTION COSTS ON PROGRAMME EFFICIENCY

5.1 Introduction

In the face of a global decline in resources available for the HIV and AIDS response, efficiency has become a key and urgent theme in HIV and AIDS governance. National governments in countries impacted by the HIV pandemic are expected to institute measures that secure efficient service delivery in the face of resource constraints. Although efficiency debates in the HIV and AIDS response point to transaction costs as a key contributing factor to the inefficiency of programmes (Zeng, Shepard, Chilingerian, & Avila-Figueroa, 2012), there is limited evidence that they are monitored to inform efficient governance mechanism choice for HIV and AIDS interventions. As a consequence, there is a limited focus on technical efficiency which arises from management processes of the response. As will be noted in chapter 7 of the thesis, public governance of the ART programme is deemed the most appropriate governance mechanism based on its competence, although it needs to be subjected to the efficiency criterion.

This chapter attempts to address that research gap by estimating the technical efficiency of facility-based ART programme in Malawi and assessing how transaction costs affect efficiency, taking into account other important covariates. With a focus on the Malawi HIV and AIDS ART programme, it seeks to answer, firstly, the research question related to whether HIV programmes in Malawi are inefficient, and if so, what types of inefficiency they display? Secondly, it aims to answer the research question regarding how transaction costs affect the technical efficiency of the HIV and AIDS ART programme in Malawi. The ART costs are estimated based on a combination of data sources, using facility level data from the HIV and AIDS electronic management records (EMR), and a GHCC database of unit costs on ART and HTC service delivery retained by the NAC. This is complemented by qualitative data gathered from ART service providers at selected health facilities. The facility level efficiency and overall

programme efficiency are estimated by means of a Data Envelopment Analysis approach as described in s*ub-section 4.12* of Chapter 4.

Overall, the results show that outputs in the Malawi ART programme can be increased by 8% with existing inputs. The results further show two types of scale inefficiency in the programme, namely, increasing returns to scale (facilities operate at too small a scale for their operations) and decreasing returns to scale (facilities operate at too big a scale for their operations).

The Chapter firstly overviews the concept of efficiency to specify the scope of the efficiency analysis for the study. This is followed by a presentation of results of the (in) efficiency estimation. The chapter further presents the results of a qualitative assessment related to ART service quality and challenges. A summary concludes the chapter.

5.2 Efficiency Concepts and Data Envelopment Analysis

In DEA estimation, efficiency is based on the ratio of weighted outputs to weighted inputs (Zeng, Rwiyereka, Amico, vila-Figueroa, & Shepard, 2014). There are three elements of efficiency that are estimated, (a) overall technical efficiency (OTE), (b) pure technical efficiency (PTE), and (c) scale efficiency (SE). Under constant returns to scale, the relevant efficiency measure is OTE. The OTE measures how a firm is able to appropriately deploy its inputs to produce maximum output in comparison to its potential production given by its production possibility frontier (PPF) (Kumar & Gulati, 2008). These inputs may be multiple, including staff, equipment and money to produce a single or multiple output (s). In the context of ART services, this may include funds to acquire drugs and test kits, personnel that directly administers diagnostic and treatment services, buildings, equipment or space for service provision. These in turn lead to the achievement of outputs such as the number of people that are alive and on ART and the number of tests completed for a given period. The OTE measure is used to determine inefficiency that arises from the input and output mix, and the size of a firm or decision-making unit size of operation.

In the DEA estimation, OTE is decomposed into two forms of efficiency, namely, pure technical efficiency and scale or allocative efficiency (Coelli, 2008). Pure technical

efficiency estimates the efficient frontier under the variable returns to scale assumption. The PTE reflects managerial performance in the organisation of inputs in the production process (Kumar & Gulati, 2008). The ratio of OTE to PTE measures the Scale Efficiency. Scale efficiency informs decisions about production levels that will yield efficient operation. This implies that if a decision-making unit or firm operates at an inappropriate scale, that is, if it is either too big or too small, technical inefficiency may arise (Zere, 2000). Constant returns to scale imply efficient operation of the firm. On the one hand, a decision-making unit is deemed scale efficient if it operates at constant returns to scale (CRS). On the other, scale efficiency is used to detect two forms of scale inefficiency. The first is decreasing returns to scale, where a decision-making unit or a firm experiences diseconomies of scale. This implies that it is too big to take advantage of scale. In relation to ART service provision, health units that experience decreasing returns to scale would reflect scale inefficiency in the provision of services. The second type of inefficiency depicted by scale efficiency is increasing returns to scale (IRS). If a firm operates at increasing returns to scale, the implication is that it is operating at too small a scale for its operations. It can thus gain efficiency by taking advantage of scale.

The amount of ART transaction costs was derived from the total costs of ART provision. It is applied directly onto the total estimated costs. At 20 percent of the total cost from the sampled facilities, the ART TCs are estimated at US$48,000,000, but this figure was rounded off to US$50,000,000 for ease of calculation. This amount was then proportionately distributed across the health units providing ART included in the study using weights based on the volume of ART clients at each facility. This means that the facilities with higher numbers of ART clients accounted for relatively higher amounts of transaction costs than those that had smaller client numbers. Clearly, and as observed before, an enumeration of transaction costs based on bottom up activity-based costing is ideal, and should improve estimation in future. Nevertheless, the contribution of the current study is to put a spotlight on transaction costs in ART service delivery by emphasizing its importance and itemising the elements that are likely to constitute transaction costs.

c) Top Down Approach

The approach used to estimate ART expenditure was largely top-down. It used two main outputs reported at each ART facility, and a database of unit costs prepared by the

GHCC in 2017 as noted before. In addition, the study used available data on cost estimation for Malawi, including Vyas et al (2020) that among others, estimated the cost of personnel input into ART service delivery. The unit cost estimates by GHCC included per person per year costs for ART, HTC and opportunistic infections treatment. These were then directly applied to facility output data to estimate the facility level cost for a particular activity. With respect to ART, the output was the number of people alive and on ART for the period 2018/19, while under the testing and counselling, the total number of tests completed in the same period was used as output. The number of people on ART was 832,165 while the number of tests completed in the same period was 8,808,897. This top down approach was informed by a pilot phase of the study that sought to establish the feasibility of deploying a bottom up approach where facility level volumes of commodities consumed and outputs produced would be estimated. The bottom up approach was found not to be feasible given the limited resources available for the research. Moreover, for the large number of ART facilities assessed, a reasonable sample would be required to complete the costing and to use the results to extrapolate to the number of ART sites in the country. Hence, a key limitation in the current expenditure estimation is that most of the cost estimates are likely to be at an aggregated level. The unit costs in the NASA are higher than in those used for the current research (i.e. from the GHCC database), which were $5.50 and $115.39 for HTC and ART, respectively. The per person per year costs estimated through the NASA was $9.75 for HTC, while that for ART was $261.00 (GoM, 2021).

5.3.3 Summary of Cost Estimates

The total ART cost for the fiscal year 2018/19 is estimated at approximately $248,300,000.00 from the study. Of this amount, the ART treatment sub-component is estimated to be $219,400,000.00, while the rest ($28,900,000.00) is estimated to be HIV testing expenditure. The weighted mean cost per site was estimated at $1,131,473.00, and ranged from $0.01 to $199,117.60. Overall, per capita per year ART costs are estimated at $298.24, and ranged from $253.64 to $1104.07. By sub-programme, the mean HTC cost was $34.59 while the mean cost for treatment (ART) was estimated at $263.65.

5.5 Results of DEA Technical Efficiency Estimation

This sub-section presents the results of efficiency estimation using DEA.

5.5.1 Summary of ART DMU Efficiency by District

Overall, of the 725 health facilities included in the research, primary health facilities accounted for 95 percent of the health units providing ART services, while secondary service facilities, mainly District hospitals accounted for 4.5 percent. The remainder of the facilities were referral facilities in the five health zones of the country. In Table 6 a number of summary statistics are reported with regard to the minimum and maximum efficiency scores for health facilities that offered ART services covered under the study. The minimum and maximum efficiency scores reported should be interpreted as being relative to health facilities that offer ART services that are operating at efficient levels. These reference health facilities may be within the same districts, or in other districts as the minimum and maximum scores reported in a particular district.

Table 6: ART efficiency summary by district

#	District	Mean	SD	Min	Max	N
1	Lilongwe	0.920	0.091	0.39	1	78
2	Zomba	0.919	0.042	0.785	1	35
3	Chiradzulu	0.901	0.037	0.846	1	15
4	Nkhotakota	0.898	0.069	0.72	1	21
5	Chikwawa	0.923	0.042	0.812	1	26
6	Thyolo	0.946	0.040	0.859	1	31
7	Mangochi	0.924	0.037	0.843	1	45
8	Karonga	0.909	0.038	0.831	0.955	17
9	Balaka	0.910	0.073	0.701	1	15
10	Blantyre	0.920	0.055	0.818	1	51
11	Dedza	0.923	0.046	0.839	1	30
12	Ntcheu	0.918	0.039	0.82	1	33
13	Rumphi	0.924	0.044	0.857	1	17
14	Mulanje	0.954	0.034	0.849	1	27
16	Dowa	0.916	0.049	0.854	1	23
17	Kasungu	0.923	0.042	0.847	1	28
18	Nkhata Bay	0.915	0.043	0.831	1	18
19	Mzimba South	0.907	0.044	0.833	1	34
20	Mzimba North	0.894	0.070	0.643	0.999	25
21	Salima	0.914	0.042	0.839	0.989	18
22	Machinga	0.926	0.035	0.852	0.977	21
23	Chitipa	0.900	0.052	0.841	0.985	10

#	District	Mean	SD	Min	Max	N
24	Nsanje	0.899	0.040	0.796	0.956	15
25	Ntchisi	0.927	0.039	0.854	0.982	11
26	Neno	0.887	0.053	0.776	1	16
27	Mchinji	0.937	0.045	0.838	1	15
28	Phalombe	0.953	0.026	0.897	1	15
29	Mzuzu City	0.929	0.058	0.804	1	8
30	Mwanza	0.882	0.049	0.831	0.937	9
31	Blantyre City	0.898	0.062	0.77	0.959	10
32	Lilongwe City	0.789	0.169	0.489	0.892	6
33	Likoma	0.912	0.079	0.856	0.968	2
	National	0.918	0.057	0.39	1	725

The mean district ART provision efficiency scores show that Lilongwe City (0.789, n=6)) has the lowest mean efficiency score. The highest mean efficiency scores are recorded for Mchinji (0.937, n=15), Thyolo (0.946, n=31), Phalombe (0.953) and Mulanje (0.954, n=27) which are above the national mean efficiency score. Other districts with relatively low efficiency scores are Mzimba North (0.894), Mwanza (0.882, n=9) and Neno (0.887, n=16). Notably, 12 of the 33 districts have mean facility efficiency scores of less than unity, suggesting overall district ART programme inefficiency.

As can further be observed from Table 6 , 21 of the 33 districts (63.6% of the districts) register at least one efficient health unit (efficiency=1) providing ART. The lowest minimum efficiency score is recorded for a health unit in Lilongwe District at 0.39 (n=78), which is likely to have dragged down the overall mean score for the district. For this particular facility, the implication is that output can be increased by 61 percent using the current inputs. The Lilongwe City Council also records at least one health unit with an efficiency score that is below 50 percent (0.498), suggesting output can be increased by 51 percent with existing inputs for this particular facility.

5.5. 2 Spatial Distribution of Health ART facility Efficiency

The spatial distribution of the ART facility efficiency shows that Rumphi, Chitipa, Karonga, Thyolo, Chiradzulu, Chikwawa and Nsanje districts have the highest number of facilities with efficiency levels of 0.93 to 1 (or 93% to 100%), where 100% efficiency scores pertain to reference ART facilities that are the most efficient among their peers. The results of the spatial distribution of facility ART efficiency are depicted in Figure 7.

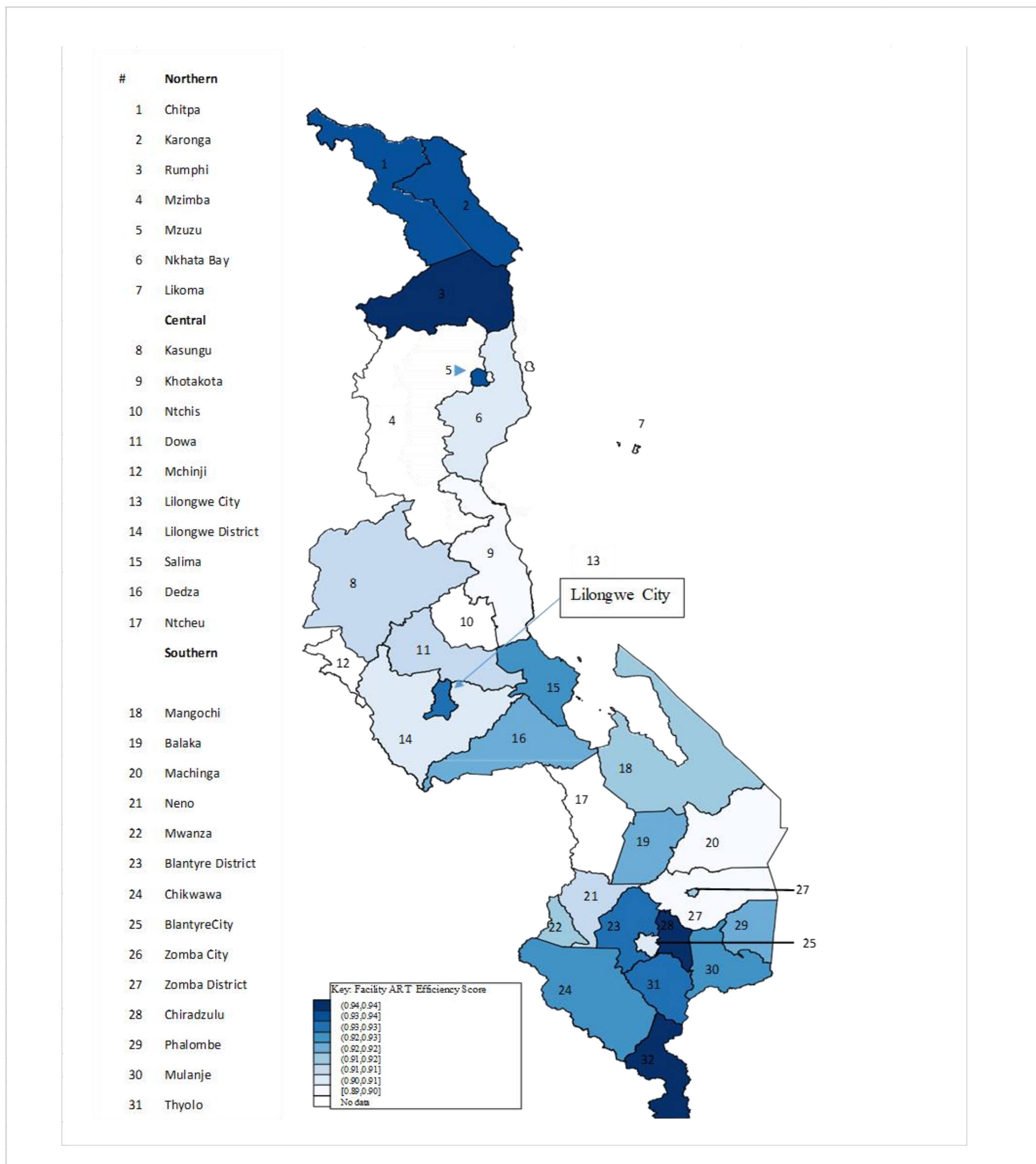


Figure 7: Spatial Distribution of ART facilities according to efficiency scores

Source: Author

As may be further noted from the figure, districts closest to Lilongwe City reflect relatively low efficiency scores of between 0.89 and 0.92, or between 89% and 92% when compared with the rest of the districts. Yet other ART facilities such as those located in Zomba, Phalombe and Blantyre City displayed low efficiency levels when compared with the most efficient ART sites.

5.5.3 ART Services Technical Efficiency

The minimum observed technical efficiency for a health facility providing ART services was 0.39, suggesting that this particular health facility can increase its efficiency by 61 percentage points if it increased output using existing resources. The maximum observed health facility efficiency is 1, suggesting some health facilities have reached optimal efficiency levels based on the DEA efficiency estimates at the current level of operations. The mean efficiency of facility-based ART service provision of the Malawi HIV and AIDS Programme under the variable returns to scale assumption is evaluated to be 0.92. This suggests that, overall, output can be increased by 8 percentage points on average with existing inputs in the programme.

The figure below (Figure 8) depicts a general picture of the facility-based ART service efficiency in the country.

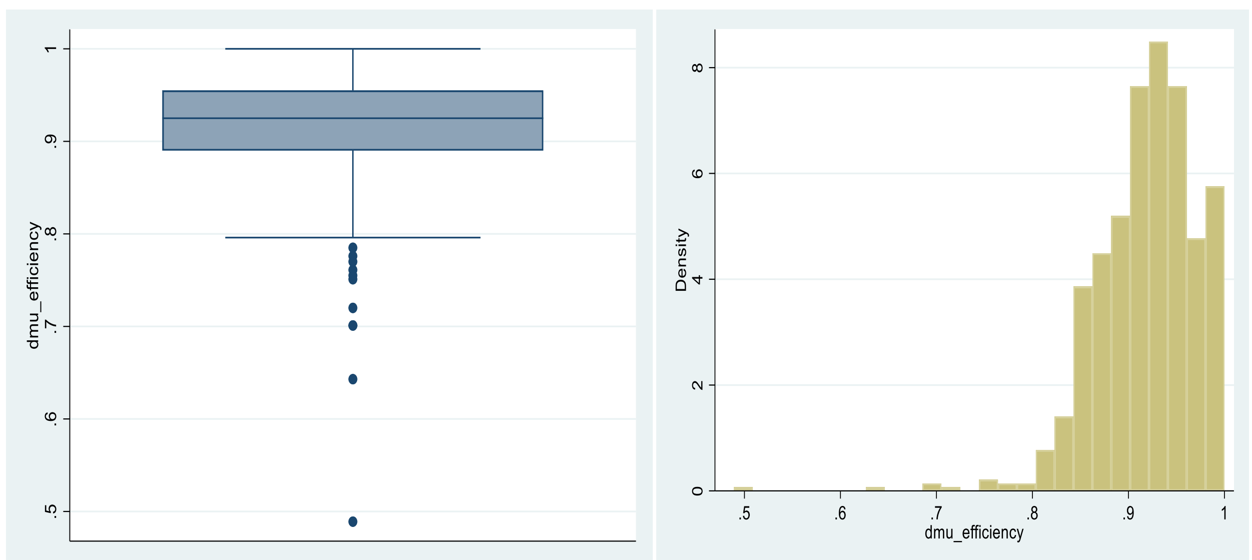


Figure 8: Summary of health facility efficiency

The figure shows that the bulk of health units that provide ART services in the country operate at efficiency levels of between 0.8 to 0.95. However, these summary figures hide individual health facility variations in terms of efficiency.

The table below (Table 7) depicts a summary of health facility efficiency by levels of efficiency attained.

Table 7: ART health facility levels of efficiency

#	DMU Efficiency Score Range	Frequency (n=725)	Share of DMUs (%)
1	0-0.59	2	0.3
2	0.60-0.69	1	0.1
3	0.70-0.79	10	1.4
4	0.80-0.89	215	29.7
5	0.90-0.99	449	61.9
6	1	48	6.6

As is observed from Table 7, 48 health units or 6.6 percent (n=725) of the facilities have their technical efficiency evaluated to be unity. This suggests that these health facilities have attained efficient service provision relative to their peers given the current input mix and outputs achieved. In addition, efficiency in nearly 62 percent of the health units is evaluated at between 0.90 and 0.99, suggesting efficiency can be improved by increasing output on a range of 1%-10% to reach comparable efficiency with their more efficient peers with existing resources. Further, nearly 30 percent (n=725) of the health facilities operate between 11%-20 % inefficiency levels. This suggests that efficiency in these health units can be improved by increasing output by margins of between 11%-20%. About 0.3 percent of the health units providing ART are evaluated to be operating at inefficiency levels of between 21%-50%, suggesting that output in these facilities can be increased by 50% and 79% with existing inputs.

5.5.4 ART DMU (In) Efficiency by Zone and District

The study further analysed health facility efficiency in relation to health zones. The Ministry of Health has divided the country geographically into health zones to facilitate technical support and coordination of various health programmes. There are currently five zones, each with at least eight districts. This division has implications on the coordination and management of the ART programme, largely with respect to monitoring and ensuring quality service delivery. Indeed, ART service providers

indicated the key role this division plays in ART programme reviews, learning and coordination. This arrangement has implications on the ART programme, firstly, because cross- district learning through reviews can impact ways in which the programme is managed. Secondly, there are transaction costs that are generated as a result of these activities.

The table below (Table 8) depicts DMU efficiency by zone from the least efficient to the relatively efficient. Health facilities with efficiency scores of less than 0.6 are categorized as least efficient while those registering 90% and above are relatively efficient.

Table 8: ART Health facilities Efficiency by Health Zone

#	Zone and District	Number of health facilities by level of efficiency category (% in facilities brackets)				
		0-0.59 (Extremely inefficient)	0.60-0.69 (Inefficient)	0.70-0.79 (Moderately inefficient)	0.80-0.89 (Fairly efficient)	0.90-1.0 (efficient)
A	**Northern Zone (=131)**					
1	Karonga	0 (0.0)	0 (0.0)	0 (0.0)	6 (2.9)	11 (2.2)
2	Rumphi	0 (0.0)	0 (0.0)	0 (0.0)	6(2.9)	11(2.2)
3	Nkhata Bay	0 (0.0)	0 (0.0)	0 (0.0)	7 (3.3)	11(2.2)
4	Mzimba South	0 (0.0)	0 (0.0)	0 (0.0)	13(6.2)	21(4.2)
5	Mzimba North	0 (0.0)	1 (50.0)	0 (0.0)	11(5.3)	13(2.6)
6	Chitipa	0 (0.0)	0 (0.0)	0 (0.0)	6 (2.9)	4(0.8)
7	Mzuzu	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.95)	6 (1.2)
8	Likoma	2 (66.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
	Sub-total	*2 (66.7)*	*1 (50.0)*	*0 (0.0)*	*51 (24.4)*	*77(36.8)*
B	**Central East Zone (n=101)**					
1	Ntchisi	0 (0.0)	0 (0.0)	0 (0.0)	3 (1.4)	8 (1.6)
2	Salima	0 (0.0)	0 (0.0)	0 (0.0)	6(2.9)	12(2.4)
3	Kasungu	0 (0.0)	0 (0.0)	0 (0.0)	8 (3.8)	20(4.0)
4	Dowa	0 (0.0)	0 (0.0)	0 (0.0)	10 (4.8)	13(2.6)
5	Nkhotakota	0 (0.0)	0 (0.0)	1(7.1)	9 (4.3)	11(2.2)
	Sub-total	0 (0.00)	0 (0.0)	*1 (7.1)*	*36(17.2)*	*64 (12.9)*
C	**Central West Zone (n=162)**					
1	Lilongwe	1(33.3)	0 (0.0)	4 (28.6)	17 (8.1)	56 (11.3)
2	Dedza	0 (0.0)	0 (0.0)	0 (0.0)	8(3.8)	22(4.4)
3	Ntcheu	0 (0.0)	0 (0.0)	0 (0.0)	8 (3.8)	25(5.0)
4	Mchinji	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.95)	13(2.6)
5	Lilongwe city	0 (0.0)	1 (50.0)	4 (28.6)	0 (0.0)	1 (0.2)
	Sub-total	*1 (33.3)*	*1 (50.0)*	*8 (57.1)*	*35(16.7)*	*117 (23.6)*

D	South Eastern Zone (n=193)					
1	Machinga	0 (0.0)	0 (0.0)	0 (0.0)	5 (2.4)	15(3.0)
2	Mulanje	0 (0.0)	0 (0.0)	0 (0.0)	2 (o.95)	25(5.0)
3	Phalombe	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)	14(2.8)
4	Balaka	0 (0.0)	0 (0.0)	1 (7.1)	5 (2.4)	9 (1.8)
5	Mangochi	0 (0.0)	0 (0.0)	0 (0.0)	11(5.3)	34(6.9)
6	Thyolo	0 (0.0)	0 (0.0)	0 (0.0)	5(2.4)	26 (5.0)
7	Zomba	0 (0.0)	0 (0.0)	1 (7.1)	9(4.3)	30(6.0)
	Sub-total	0 (0.0)	0 (0.0)	*2 (14.3)*	*38(18.2)*	*153(30.8)*
E	**South Western Zone (n=137)**					
1	Chiradzulu	0 (0.0)	0 (0.0)	0 (0.0)	6(2.9)	9 (1.8)
2	Chikwawa	0 (0.0)	0 (0.0)	0 (0.0)	5 (2.4)	21(4.2)
3	Blantyre	0 (0.0)	0 (0.0)	0 (0.0)	20 (9.6)	31(6.3)
4	Nsanje	0 (0.0)	0 (0.0)	1 (7.1)	6 (2.9)	8(1.6)
5	Neno	0 (0.0)	0 (0.0)	1(7.1)	7 (3.3)	8(1.6)
6	Mwanza	0 (0.0)	0 (0.0)	0 (0.0)0	2(0.95)	2 (0.4)
7	Blantyre City	0 (0.0)	0 (0.0)	1(7.1)	3(1.4)	6(1.2)
	Sub-total	*0 (0.00)*	*0 (0.00)*	*3(21.4)*	*49(23.4)*	*85(17.1)*
	TOTAL	**3**	**2**	**14**	**209**	**496**

When technical efficiency is considered with regard to the health sector management zones, 496 or 65.5% of the health facilities fall within the category of moderate to efficient, taking into account that these include 48 health units that record full efficiency. A further 209 facilities, representing nearly 30% of the health units were recorded as moderately efficient, with efficiency levels of between 0.8-0.89. An estimated 3 % of the health units may be categorized as inefficient, with efficiency scores of below 0-0.79%. Relative to efficient health units providing ART services in their zones, the Southern Eastern Zone records the highest number of health units that are categorized as relatively efficient at 79.3% (153 of 193 health facilities in the zone). The Central West Zone recorded 117 health units evaluated to be relatively efficient, representing 72.2% of the health units providing ART in the zone. The rest of the zones recorded below 70% in efficient ART health units in their respective zones, with the Central East Zone (63.4%), Southern East Zone (62%) and the Northern Zone (58.8%) recording the least number of relatively efficient health facilities in descending order.

5.5.5 Health Unit (In)-Efficiency Based on Returns to scale

Microeconomic theory projects the idea that the essential objective of a firm is to operate at the most productive scale. A firm's competitive advantage over competitors stems from its ability to deploy productive resources efficiently (Dresch, Collato, & Lucenda,

2018). Returns to scale relate the amount of inputs to outputs. Depending on the stage of growth they are at, in the short run, firms may operate with increasing or decreasing returns to scale, both inefficient positions to be in. Small firms are generally associated with inefficient production (Alvarez & Crespi, 2003). Nevertheless, firms will always aim to operate at constant returns to scale to maximize revenue (Kumar & Gulati, 2008). This implies that both increasing and decreasing returns to scale represent inefficient states in a firm or an agency.

Results from the evaluation show that the 69.38 percent of the health facilities experience decreasing returns to scale (n=712), while 23 percent experience increasing returns to scale. A further 7.6 percent experience constant returns to scale. The results are further depicted in Figure 9 below.

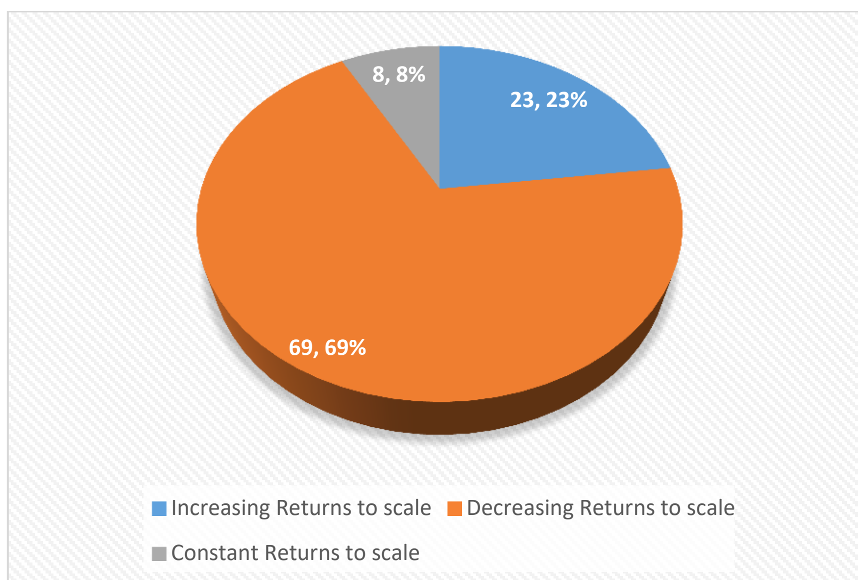


Figure 9: ART Health facility returns to scale

With constant returns to scale, a unit increase in inputs leads to a corresponding unit increase in output, and implies the DMU is operating at efficient levels. Decreasing returns to scale mean that the facility has reached a measure of efficiency, such that additional resources only yield outputs at a declining rate. For the ART health units operating with DRS, the implication is that they have become too big for the population or numbers of clients they serve (Zere E. , 2000). In ART programme terms, this suggests that the allocation of additional ART health facilities, especially when they serve in the vicinity of facilities already operating with decreasing scale, is likely to yield efficiency losses.

Facility Returns to scale by Health Service Delivery level

The different levels of health service delivery provision further displayed differing returns to scale. Nevertheless, the results show that in general, the majority of primary and secondary health facilities experienced decreasing returns to scale. Figure 10 summarises the distribution of health units providing ART services by returns to scale and health provision level.

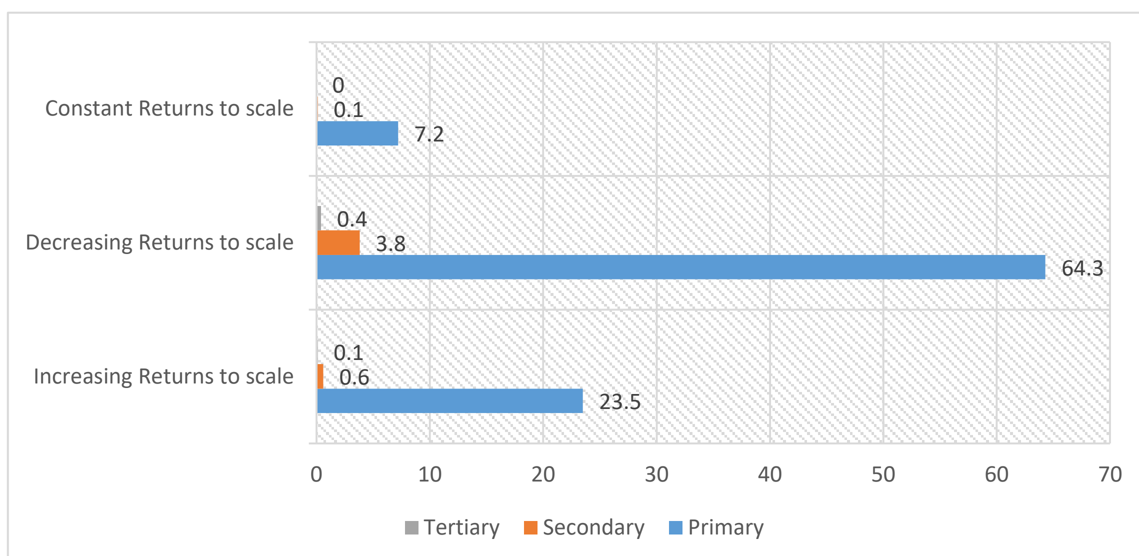


Figure 10: Percentage of ART health facilities by returns to scale and by service delivery level

An estimated 7.2 percent (n=725) of the primary health service facilities covered under the study were evaluated to operate at constant returns to scale, suggesting that they operated at efficient levels over the period covered when relative to their peers. A further 462 health facilities accounting for 93.7% (n=493) of the primary health facilities and 65.8% percent (n=725) of all facilities evaluated operated with decreasing returns to scale. Additionally, 23.6 percent (n=673) of the primary health facilities, or 21.9 percent of the 725 facilities experienced increasing returns to scale, suggesting that increasing the outputs, number of people served under the ART programme should increase efficiency levels.

Under secondary health service units, which largely consist of District Hospitals and some private hospitals, 4 of the secondary facilities (n=32) or 0.6 percent of the total facilities covered (n=725) experienced increasing returns to scale, while 27 (84.4%) of the 32 secondary facilities, or 3.7 percent of the total facilities evaluated experienced decreasing returns to scale. One secondary facility experienced constant returns to scale.

Facility Returns to scale by Rural and Urban Location

The figure below (Figure 11) depicts the distribution of ART facilities efficiency in terms of returns to scale of health units providing ART services in Malawi by geographical location (rural-urban).

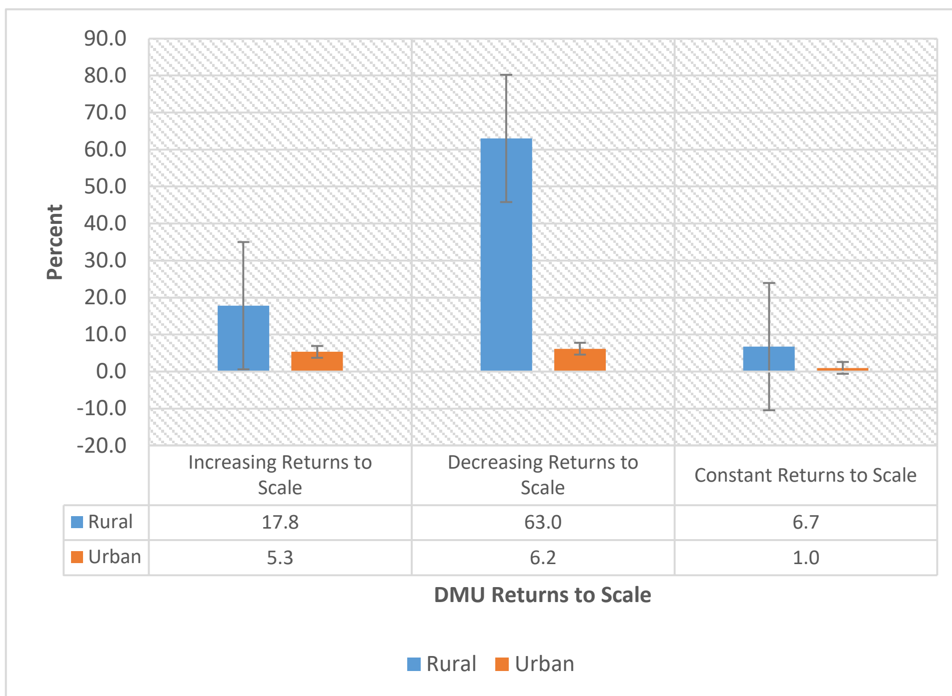


Figure 11: Health facility Returns to scale by Rural and Urban location

The figure above shows that a relatively high percentage of health units that provide ART services in rural and urban areas operate at decreasing returns to scale. The results show that 17.8% and 5.3% of the health units in the rural and urban areas, respectively, experience increasing returns to scale. With respect to decreasing returns to scale, 63% of the health units in the rural areas and 6.2% in the urban areas are evaluated to operate with decreasing returns to scale. A further 6.7% and 1 % of the health units in the rural and urban areas, respectively, operate at constant returns to scale.

Health Facility Returns to Scale by Type of Administration

The distribution of health facility in terms of returns to scale by type of management or administration is presented in Figure 12.

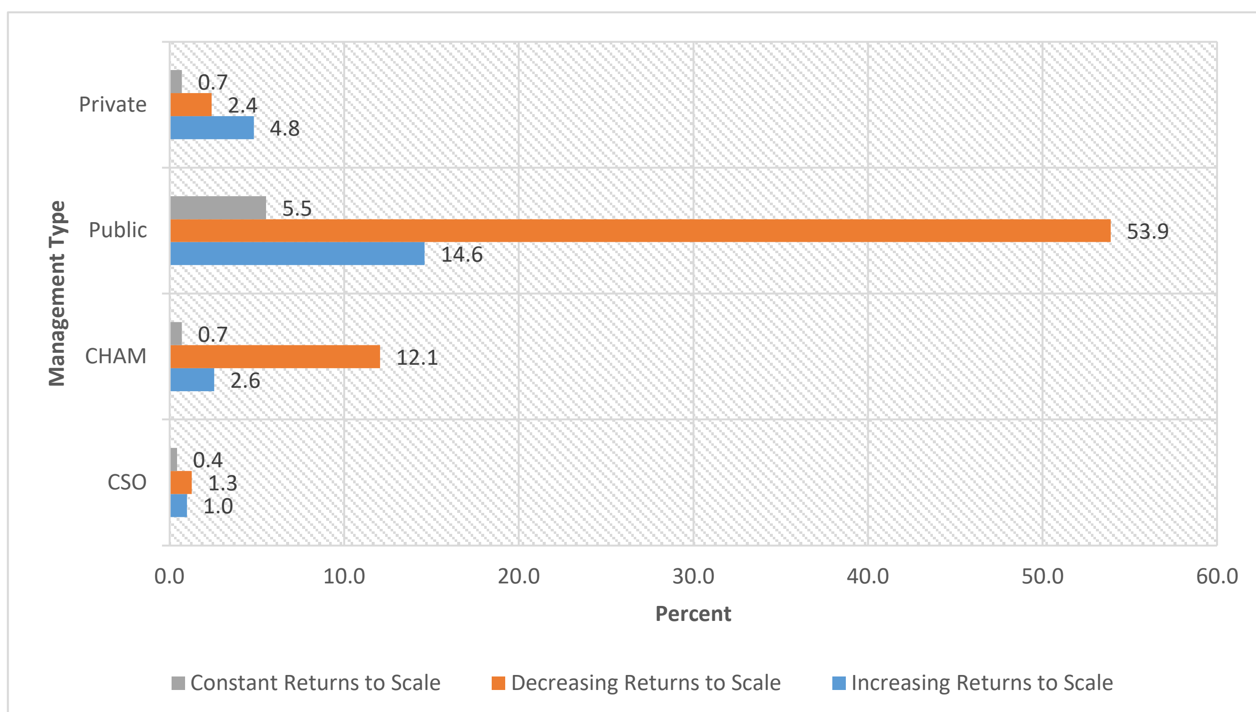


Figure 12: Facility returns to scale by type of administration

By type of administration, 53.9 % of public health units that provide ART display decreasing returns to scale, while 14.6% show increasing returns to scale. A further 5.5% are evaluated to experience constant returns to scale. A similar trend is observed for CHAM and MAM health units (n=108) where the bulk of the facilities (12.1%), experience decreasing returns to scale, while 2.6% and 0.7% experience increasing and constant returns to scale, respectively. Among the private sector health units, 4.8% of the health units display increasing returns to scale, while 2.4% operate with deceasing returns to scale, and 0.7% experience constant returns to scale. Among the CSO health units (n=19), 1.3% health units experience decreasing returns to scale while 1.0% and 0.4% show constant returns to scale.

5.6 Second Stage DEA Regression Results

This section presents the results of the second stage DEA estimated via a generalised linear model using OLS analysis. It aims to assess the determinants of facility-based

ART service delivery efficiency. It builds on the first stage DEA results as described above to assess the sources of variation in efficiency.

5.6.1 Health Units Characteristics

The following table presents a summary of the descriptive statistics for selected variables in the ART health facility efficiency regression equation (Table 9).

Table 9: ART Health Units Characteristics

#	Variable	Minimum	Mean	Maximum	Std Deviation	N
1	Distance (km) from Headquarters	5	287.9	811	157.4	725
2	Catchment population (No.)	981	22393	187764	21371	725
3	dmu size (No. of clients)	5	1148	23831	1772.3	725
4	Site maturity (years)	2	8.9	19	2.97	725

The distance from the centre is a key determinant of transaction costs and efficiency, owing to both production and supervision and monitoring costs. A minimum of 5 kilometres from the centre is observed from the data, maximum distance is recorded at 811 kilometres. The observed standard deviation for distance is 157.4 kilometres and suggests a relatively high spread of distance amongst the health facilities. The minimum and maximum catchment population (general population served by the facility) are recorded at 981 and 21,371 people, each. The mean catchment population is estimated at 22,393 people. The mean catchment population is higher than the health sector standard norm of 10,000 people, accounting for referral facilities such as district and tertiary level facilities. The ART facility size is the ART client burden per facility. The mean facility ART burden was estimated at 1148, with the lowest burden recorded at 5 and the highest at 23,831. The facility ART size has implications on the efficiency of ART service delivery as lower client burdens are likely to yield high unit costs and hence inefficient outcomes relative to high burden facilities.

The number of years that the facilities have been in operation ranged from 2 to 19 years. The mean number of years across all the facilities covered was estimated at 8.9 years. A standard deviation of 2.97 years suggests a relatively small spread in the number of years

that the facilities have been providing ART services. This may be a further indication of the rapid expansion of the ART programme over a short period of time.

5.6.2 The effect of transaction costs on HIV and AIDS Facility Based ART Efficiency

The ART facility efficiency was regressed on a number of variables as stated in Chapter 4-*see section 4.12.* A summary of the regression results in the second stage DEA are offered in Table 10, and *Appendix 4*. Health facility size was omitted from the regression model due to multicollinearity, and was thus not analysed further. It is further worth noting that the total number of facilities included in the regression analysis was reduced to 658 from 725 due to missing data. Nevertheless, at 91%, the sample remained large enough to ensure stability of the estimated parameters.

Table 10: Determinants of facility-based ART service delivery Efficiency Regression.

Dmu Efficiency	**Coefficient**	**Robust Std. Error**	**T**	**P>|t|**	**[95% conf. interval]**	
Log of transaction costs	-0.0675571	0.0325333	-2.08	0.038*	-0.1314473	-0.003667
Location (Rural/Urban)						
Rural	-0.8192774	0.4000989	-2.05	0.041*	-1.605008	-0.0335466
Facility Ownership type						
CHAM/MAM	0.0038159	0.010322	0.37	0.712	-0.0164548	0.0240866
Public	0.008361	0.0087474	0.96	0.34	-0.0088175	0.0255395
CSO	0.0036596	0.01441	0.25	0.8	-0.0246393	0.0319586
Service Level						
Secondary	0.0088343	0.018369	0.48	0.631	-0.0272394	0.044908
Primary	0.0077131	0.0178921	0.43	0.667	-0.0274242	0.0428503
Catchment population	2.84000000	1.42000000	2.01	0.045*	5.9500000000	5.63000000
Site maturity	-0.0450427	0.0216711	-2.08	0.038*	-0.0876012	-0.0024841
Facility Distance from Lilongwe	-0.001545	0.000751	-2.06	0.04*	-0.0030198	-0.0000702
District						
Zomba city	-1.615737	0.7791602	-2.07	0.039*	-3.145884	-0.08559
Thyolo	-0.6503837	0.3147269	-2.07	0.039*	-1.268458	-0.0323099
Salima	-1.103069	0.5322032	-2.07	0.039*	-2.148231	-0.0579058
Rumphi	-1.146225	0.5524795	-2.07	0.038*	-2.231208	-0.0612433
Phalombe	-0.2122788	0.1033809	-2.05	0.04*	-0.4153024	-0.0092552
Ntchisi	0.4454817	0.2187133	2.04	0.042*	0.0159635	0.8749998
Ntcheu	-0.9174926	0.4419717	-2.08	0.038*	-1.785455	-0.0495303
Nsanje	-2.055875	0.9890241	-2.08	0.038*	-3.998161	-0.1135881
Nkhotakota	1.426443	0.689219	2.07	0.039*	0.0729266	2.77996
Nkhata Bay	0.327153	0.1582591	2.07	0.039*	0.0163572	0.6379489

Dmu Efficiency	Coefficient	Robust Std. Error	T	P>|t|	[95% conf. interval]	
Neno	0.2264526	0.1105598	2.05	0.041*	0.0093306	0.4435746
Mzuzu City	-1.436862	0.6929524	-2.07	0.039*	-2.797711	-0.0760136
Mzimba South	-1.784056	0.8541324	-2.09	0.037*	-3.461437	-0.1066752
Mzimba North	-2.386309	1.144796	-2.08	0.038*	-4.634507	-0.1381108
Mwanza	0.999201	0.4816892	2.07	0.038*	0.0532398	1.945162
Mulanje	-0.4811643	0.2309656	-2.08	0.038*	-0.9347441	-0.0275846
Mchinji	-0.6155123	0.2966228	-2.08	0.038*	-1.198032	-0.0329921
Mangochi	-0.3485969	0.1683689	-2.07	0.039*	-0.6792467	-0.0179471
Machinga	0.3081936	0.1475084	2.09	0.037*	0.0185105	0.5978767
Lilongwe	0.0460453	0.025168	1.83	0.068**	-0.0033806	0.0954712
Likoma	1.606957	0.7750599	2.07	0.039*	0.0848626	3.129052
Lilongwe City	-1.911154	0.9165206	-2.09	0.037*	-3.711055	-0.1112529
Kasungu	-0.1584146	0.0784933	-2.02	0.044*	-0.3125631	-0.0042662
Karonga	-0.9341524	0.4494132	-2.08	0.038*	-1.816729	-0.0515761
Dowa	-0.0221117	0.016248	-1.36	0.174	-0.0540203	0.0097969
Dedza	-0.6556192	0.3168389	-2.07	0.039*	-1.277841	-0.0333979
Chitipa	-0.3013575	0.1472959	-2.05	0.041*	-0.5906232	-0.0120918
Chiradzulu	-1.034665	0.4982785	-2.08	0.038*	-2.013205	-0.0561249
Chikwawa	-0.8529563	0.4121149	-2.07	0.039*	-1.662285	-0.0436279
_cons	45.5682	21.18015	2.15	0.032	3.973732	87.16266

Observations 658

Rsquared 0.0804

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

As may be observed from the table above, the log of *transaction costs* enters the ART facility efficiency equation negatively as expected before estimation, and was consistent with economic theory. The result is statistically significant (p=0.038). The negative co-efficient sign suggests an inverse relationship between efficiency and transaction costs, and suggests that an increase in transaction costs would lead to a decrease in the efficiency of ART delivery.

With regard to *service delivery level* of the health facilities providing ART, results show that when compared to tertiary health facilities that provide ART, primary and secondary health facilities should be efficient.. With tertiary level as baseline, the results show that ART provision at both the primary level is efficient, as the sign on the coefficients for both primary and secondary health provision level are positive. However, the results are not statistically significant for primary level provision (p=0.667) and for secondary level provision (p=631).

The *type of ownership or administration or bureaucracy* has an influence on the efficiency of health facilities providing ART services. The types of administration recorded in the study sample were private, public, Christian Association of Malawi (CHAM) and those belonging to civil society organisations (CSO). Private and CSO facilities provided the ART based on a Service Level Agreement signed with the Ministry of Health as part of the essential health services package. Results of the regression analysis show that when compared with the private sector health facilities that provide ART, public sector, CSO and CHAM facilities are efficient. Nevertheless, the results from the model are not statistically significant for all the alternate public (p=0.340), CSO (p=0.800) and CHAM/MAM (p=0.712) health facilities providing ART services.

Both *catchment population* and facility *ART client burden* are expected to be positively correlated with efficiency. However, the facility size variable, which was based on the volume of ART clients was dropped from the regression due to multicollinearity. Nevertheless, the results with respect to the general population served by the facility show a positive correlation with health facility ART efficiency. The results are statistically significant at the α=0.05 level (p=0.045). The positive sign on the catchment population variable tended to support the notion that relatively large populations served should improve efficiency since the per capita inputs reduce with increased population using the service. The caveat is that, as noted by service providers that were participants to the research, the quality of services tends to drop with a large clientele.

It is expected that *with time*, each facility that provides ART services should improve its efficiency through learning and continuous service improvement. Studies on firm level and hospital determinants of efficiency show that the age of the firm is positively correlated with efficiency (Duong, 2016; Mahajan, Nauriyal, & Singh, 2018). However, site maturity shows a negative correlation with ART DMU efficiency under this study, and suggests that facilities that have been in operation for a long period are likely to be associated with inefficient service delivery. The result is statistically significant (p=0.038).

The *distance* of a health facility from the centre is likely to influence the efficiency of ART services. Distribution and supervision costs are likely to increase the further the

facility is from the headquarters. This is likely to reduce the efficiency of services. Moreover, where facilities are remote and the population sparse, efficiency is likely to be lower than facilities located in the more densely populated areas close to the centres of coordination. The co-efficient sign was negative for the absolute distance and the results are statistically significant(p=0.040).

The *location* of services, whether rural or urban or specific districts may impact general health and ART service delivery efficiency. The results from the study show that when compared with urban facilities, rural facilities tend to be inefficient in the provision of ART services. The result is statistically significant (p=0.041) at the $\alpha = 5\%$ critical value. The district location of the facilities was included as a covariate to control for ecological or community factors that may influence health facility efficiency in the provision of ART. The baseline district was Zomba District Council. With regard to district location, the study shows mixed results. Except for Nkhata Bay (p=0.038), Ntchisi District (p=0.042), Neno (p=0.041), Nkhotakota District (p=0.039), Mwanza(p=0.038), Machinga (p=0.037), Ntchisi District (p=0.042), Lilongwe District (p=0.068) and Likoma District (p=0.039) that were positively correlated with DMU efficiency, the rest of the districts showed a negative correlation. As can be noted, for districts displaying efficient correlation with facility efficiency, the results are statistically significant at the at the α=0.05 critical value level, except for Lilongwe District whose result is only significant at the at α=0.1 critical value. This suggests that these districts should be associated with efficient provision of ART services, holding all other things constant. Demographic factors may contribute to the district location effects on ART facility efficiency.

5.7 Qualitative Assessment of ART Services Quality

Assessment of efficiency needs to consider the quality of service provided. Hence, the research further examined the quality of services from a service providers' perspective. Among other things, participants to the interviews at service provision level were requested to provide a subjective rating of the ART services in their own facilities. On a scale of 1-5, where 1 represented extremely poor quality and 5 extremely high quality. Overall, the research found that providers tended to rate the ART service quality highly. The mean rating for the 35 facilities visited was 3.9 out of a possible 5, suggesting that the perceived quality of service by providers was relatively high. The main reason that

tended to drive the rating up was the ability of the health facility to provide the ART services as scheduled. In addition, the health facilities cited "no stock out situation of drugs and test kits, the availability of experienced staff and the ability to follow up on defaulting clients as evidence of service quality. Furthermore, the research found that where they existed, partners that supported the ART services appeared to enhance the capacity of the health facilities. These partners included Partners in Health, EGPAF, m2m, DREAM Project, Lighthouse Trust, JPIEGO, DIGNITAS, MACRO and Baylor Paediatric Foundation.

The following sentiments reflect the perception of service delivery quality by service providers:

> "*The quality of services is good. We are able to follow up missed appointments and defaulters. We implement same day initiation to all eligible clients. We are able to do counselling and conduct integrated services. A lot of work but the health centre is delivering the services" Respondent 14, Service Provider, Machinga District.*

> *"The ART services have been integrated into our programmes. Staff have been trained to provide services, we do not experience stock-outs of kits and drugs. The ART services done on daily basis, following clinical guidelines", Respondent 21, Service Provider, Zomba District*

Nevertheless, the research identified a number of challenges that constrain quality ART service delivery were further cited. As observed before, except for a few health facilities visited that have integrated ART in routine health supervision, the bulk of the facilities operate ART clinics on specific days of the week. This means that staff have to dedicate specific time to attend to the ART clinics, in addition to the normal service delivery time. The caveat is that for aspects such as HIV testing and PMTCT are integrated within the current health provision routine. The most commonly cited challenges that impacted quality ART service delivery included work overload for the few staff that provide the services, limited staff, and limited space to secure privacy in service delivery.

> *"The Health Centre uses two Community Based Distribution Assistants instead of Expert Clients. But due to staff constraints, it has deployed two Hospital Attendants-one for as ART Data Clerk and the other to conduct tests under HTS.", Respondent 11, Service Provider, Ntchisi District*

"There are no dedicated rooms for ART. There are no tracer teams for defaulters, we use phones for follow up but this poses many challenges as sometimes people cannot be reached on Phone. To alleviate the staff shortage problem, we have trained a Clinic Assistant to help out with organizing clients and counselling them". Respondent 10, Service provider, Nkhata Bay District

"Transport is a big problem, especially for tracing defaulters. We have inadequate furniture, we experience poor record keeping, and there is only one ART provider, instead of the minimum that is required by the Ministry of Health" Respondent 13, Service Provider, Phalombe District

"There is shortage of space. ART are done in the same room which is small. Absence of integrated outreach clinic means the Health Centre is unable to reach populations far from the facility." Respondent 12, Service Provider, Zomba

Other unique challenges were reported for health centres in border districts such as Nsanje and Zomba. This is with respect to high default rates, especially with some of the clients that are Mozambican residents but get initiated on ART in the Malawian health units. A further challenge reported was with respect seasonal labour in tobacco farming and in timber saw milling activities. Emergency requests for drugs tend to peak during the farming season, even where there are no proper transfer letters for the clients. Conversely, when the growing season ends, large numbers of clients drop out of care at these facilities. Due to these challenges, staff spend a lot of time following up a group of clients that was fluid, and at times provides incorrect addresses.

As may be observed from the above sentiments, the challenges highlighted confirm the issues that have long been known to impact health service delivery in Malawi. Human resource and infrastructure constraints are at the heart of the problems. Indeed, with respect to human resources that provide ART services, the research found that a provider-client ratio 1: 1870 which was high. But this ratio further hides both geographical and cadre specific disparities that have implications on the quality of ART services delivery. For instance, if only the three cadres of health facility level core providers (Nurse Midwife Technicians, Clinician and Medical Assistant are considered, that ratio is 1:7,000. Hence, although there is service provider optimism, the research found that the quality of the ART services may continue to be an issue. Moreover, even

where there are partner agencies that assist with closing some of the resource gaps, the challenges appear widespread and differ only in magnitude across the country. Moreover, participants in some health centres that provided ART indicated there are no partners that supported the facilities.

The research further found that because the support from partners is project based, some of the health centres were at a point where that support was coming to an end, or it had ended for several years before the current research. It was not possible to compare the efficiency scores for facilities that had partner support for ART and those that did not have based on the 725 facilities included in the study. This was because the EMR data supplied by DHA did not contain data on partner support to specific facilities. Nevertheless, a comparison of facility mean efficiency via a bootstrapped paired t-test (n=1000) for the facilities covered by the qualitative study, showed that facilities with partner support were more efficient (p=0.033) than those that did not, when all other things are held constant. The results may suggest there may be innovative ways that come with partner support to facilities, although a larger sample is required to verify this. In addition, there may be need to look at the various support models used by partners to gain more insights into those that promote efficiency in service delivery.

The research found, nonetheless, that infrastructure challenges have somewhat been alleviated in some health facilities, where, through the United States Government support, shelters have been erected to operate as pharmacies and ART service provision rooms. This appeared to have contributed to the quality in ART service delivery, besides the general health service provision in those particular facilities.

5.8 Chapter Summary

This chapter has presented the results of an empirical analysis of the efficiency of the ART programme in Malawi. It sought to determine whether the HIV and AIDS ART programme in Malawi is efficient and to characterise the types of inefficiency observed. It further sought to determine the effect of transaction costs on the Malawi HIV and AIDS ART governance (in) efficiency. The research finds that with an overall technical efficiency of 92%, the ART programme is inefficient as this implies that output can be increased by 8 % with existing resources. The types of inefficiency observed is scale inefficiency where 23% of the facilities operate at too small a scale (increasing returns to

scale) so as to render them inefficient. Further, 69% of the ART facilities operate with decreasing returns to scale, suggesting they are operating at too big a scale when compared to the catchment areas they served. The study shows that transaction costs affect health facility efficiency negatively, and the results are statistically significant (p=0.038). Other determinants of technical efficiency include type of administration, number of years the facility has been offering services and level of service delivery. Quality assessment of ART services by service providers shows optimism, but this has to be balanced with enduring human resources and infrastructure constraints that negatively affect the primary health care service quality.

Overall, the study results show that policy interventions that increase the use of excess capacity in the health facilities that display increasing returns to scale should lead to efficient delivery. Further, a review of the facilities that appear to be too big for their catchment areas, or volumes of clients on ART is implied given the large number of health facilities that are assessed to be experiencing decreasing returns to scale. To that end, benchmarking efficiency on the 48 ART facilities that are identified as reference peers may contribute to improved efficiency in ART service delivery in the country.

The next chapter presents an assessment of the conformity of the HIV and AIDS ART programme with TCE theory.

CHAPTER SIX

THE MALAWI HIV AND AIDS ART PROGRAMME CONFORMITY WITH TRANSACTION COST ECONOMICS GOVERNANCE THEORY

6.1 Introduction

The focus of this chapter is an assessment of the conformity of the organisational set up for the delivery of HIV and AIDS ART services in Malawi with Transaction Cost Economics (TCE) theory. It seeks to answer the research question: How does the Malawi HIV and AIDS ART programme governance conform with the governance theory (transaction costs economics theory)? It deploys a descriptive approach to the analysis of the HIV and AIDS ART programme in Malawi, and how it conforms with TCE theory. The analysis is based on a document review and perceptions from key informant interviews of national level policy makers, programme managers and service providers at the facility level. It uses the TCE discriminating hypothesis to reveal the discrete structural alternative public sector institutional arrangements (privatisation, regulation and public agency) adopted for the HIV and AIDS ART programme governance in Malawi. Conformity is assessed against the main TCE theory principles, namely, a) governance instrument transaction alignment based on transaction characteristic*s* (uncertainty, asset specificity and complexity), b) TCE efficiency goal, and, c) governance instrument competence given the public goods nature of ART services.

Overall, the research findings suggest that the chosen HIV and AIDS ART governance framework largely conforms with TCE theory, as the set up appears to have been designed to minimise transaction costs that would stem from disturbances due to asset specificity, uncertainty, complexity and probity hazards associated with ART transactions. Furthermore, the results suggest that mechanisms that can potentially be used to monitor transaction costs in HIV and AIDS ART service delivery such as the NASA exist, but there needs to be a deliberate focus to monitor and control transaction costs. Nevertheless, the governance set up does not appear to adequately deal with the wastage occasioned by a parallel supply chain set up of ART drugs and commodities.

The chapter first presents the organisation of the healthcare system within which the ART programme is implemented to characterise its governance features. The description of the organisation is important as it highlights the main features that will be used to compare and determine conformity or no-conformity with TCE theory. This is followed by an analysis of the extent to which the HIV and AIDS ART programme governance design- public agency and regulation- and implementation conforms with TCE theory. A summary section concludes the chapter.

6.2 The Malawi Health Care System

6.2.1 Key Actors in the Health Care System

The bulk of health care services in Malawi is provided by the Government of Malawi which accounts for 60 percent of health facilities (GoM, 2023). The main actors in the health care system are the Government of Malawi, the private- for- profit organisations (PFP) and the private not for profit (PNPF) organisations. The health care mandate for the public sector is derived from the Malawi Constitution that requires the Government to "provide adequate health care commensurate with the health needs of Malawian society and international standards of health care" (Government of Malawi, 1995, Cap 3., para 13 (c)) . The Constitution confers health for all citizens as a right, with the implication that every citizen is entitled to health care without discrimination. The Ministry of Health provides healthcare by the Government via a network of health units. The healthcare service delivery set up further consists of ministries and other departments and agencies (MDAs) such as the district, town and city councils, the Ministry of Defence through the army health units, the Ministry of Homeland Security that offers health services through the Malawi Police Service and Prisons health units, and parastatal agencies.

Access to health services in the public sector is free at the point of service delivery, although clients invest time and financial resources to access the services. Further, while the services are free at point of access, clients still have to spend on drugs where stock-outs of essential drugs at health facilities occur. To facilitate equitable access to essential health services, the Government has prepared and implements an essential health benefit package (HBP) that includes HIV and AIDS in its health units (GoM, 2017), which has been revised under the new HSSP III covering the period 2023-2030 (GoM, 2023). One implication of the organisation of the government health system is the equity goal which

has further ramifications on the way the ART programme is set up, and the transaction costs that would attend such a structure. Ultimately, the healthcare governance is designed in a way that ensures all people, regardless of their social, economic, political, ethnic and location access health services to meet their health needs (Dover & Belon, 2019). It extends to efforts to address avoidable inequalities such as healthcare service distribution and the elimination of health and healthcare disparities (US Department of Health and Human Services, 2010; Penman-Aguilar, *et al.*, 2018). By extension, access to ART appears to have been designed with this goal in mind, as evidenced by the rapid expansion of the programme to serve communities at sub-district level, including to the very remote parts of the country.

The second key actor in healthcare service delivery is the Not-for-profit –private (PNFP) sector bureaucracy. This consists of religious institutions, civil society organisations, statutory corporations and companies. These operate own health units, with services also provided free of charge at point of service access where there is a service level agreement with the Government, otherwise they charge user fees for access to services, albeit at subsidised rates. An administrative fee is normally levied on clients accessing these services, even for some of the services for which there is an agreement with the Government, which include those related to sexual and reproductive health and HIV and AIDS. The Christian Health Association of Malawi (CHAM) represents the largest constituency of the PNFP organisations that provide health care in the country. It accounts for 29 percent of the health care service provision (GoM, 2017; GoM, 2023a). Because it is an extension of the public healthcare service provision systems for some of the services and for particular geographical areas, this has implications on resource requirements to actively mobilise the CHAM health units network participation in ART service provision in the country.

A third key actor in the healthcare system is the private for-for-profit health sub-sector. This consists of the private hospitals and clinics that provide various health care services. These health units charge user fees, although under certain services such as sexual and reproductive health (SRH) and HIV and AIDS, they do not charge fees and cost of drugs, except for administration fees which may differ by facility depending on location (rural or urban). Under this category, the Ministry of Health further recognises the role of Traditional Healers that provide healthcare to the population. However, these

are not directly involved in the provision of ART services, although they have the potential to influence uptake of ART depending on the messages that they send out to the general public.

To summarise, there are different types of bureaucracies that deliver health services, but the dominant mode is direct public provision and regulation (combining either public and private for profit or public and private not-for- profit). The provision of ART services under agreed service level agreements (SLAs) with the Government for the private not-for-profit and the private for profit agencies represents a hybrid governance mode. This set up has implications on governance choice for healthcare and ART services in particular, which are further analysed in *sub-section 6.4*.

6.2.2 Organisation of the Health Care System and the ART Programme

In Malawi, the healthcare system is organised at four levels, viz, primary, secondary, tertiary and central level. The three levels responsible for healthcare provision are linked to each other through a referral system in the manner of a hub-and-spoke arrangement. The fourth of these levels, the central level, is responsible for policy development and coordination. Consistent with the decentralisation process, primary and secondary level health provision should ideally fall under the jurisdiction of Local Authorities or Councils administratively, that in turn fall under the Ministry of Local Government . Nevertheless, as of the time of the current research, the central level Ministry of Health Headquarters retained a certain level of functions, particularly the human resources management function (Rodriguez, 2022). The Director of Health and Social Services (DHSS) heads the district health care system, including the HIV and AIDS programme. This set up effectively reflects a dual accountability arrangement, with District hospitals and their network of health centres in a particular local authority reporting through the District Commissioner or Chief Executive Officer, while at the same time, report to the Ministry of Health and Population headquarters for policy, programmatic and other administrative matters. More Concretely, staff recruitment, discipline and health policy decisions for some health cadres are still retained by the Ministry of Health headquarters, with implications on accountability for both health resources and results at the local level. So far, indications are that the district heath care system is aligned more strongly with the centre than it is with the local level in terms of accountability. As noted above, this arrangement has implications on the level and nature of transactions and

associated costs within the health system, and the HIV and AIDS ART programme for facility-based health services. Since these costs have a bearing on the efficiency of service delivery, it is plausible to expect that the programme should track the costs and evaluate their effect on the efficiency of service provision.

Primary Level

The *primary* level of health care services consists of a cadre of Health Surveillance Assistants (HSAs) that operates at the community level around health posts, dispensaries, health centres, maternity units and community hospitals. These focus on promotive and preventive health care services through door-to-door interfaces with households, village and mobile outreach clinics. The standard norm for an HSA regarding the number of clients they should serve is 1000 people, although this is rarely achieved due to human resources limitations and inequitable distribution of resources (Government of Malawi, 2017). The HSAs are a key cadre in HIV and AIDS service provision. Their roles extend to community mobilisation for HIV testing, follow up on defaulters and encouragement of adherence to ART drugs.

The HSA works closely with community level volunteers and local community leaders on all health issues, including HIV and AIDS. The Village Health Committees (VHC), Health Centre Management Committees (HCMC), Health Action Groups and Health Centre Advisory Committees (HCAC) work closely with the HSAs on all health matters in the community and around the health units. Although HSAs are by design not directly involved in the provision of ART services, their role in the general prevention and follow up of ART defaulters imply that they generate transaction costs that should be accounted for in ART service delivery. This extends to the role that they play in the coordination of community level health volunteers related to ART service provision. In many health units, these HSAs further assist with the organisation of clients and client records during the weekly ART clinic days for facilities that have specific ART days. Further, in some health facilities, some HSAs are trained to undertake HIV testing as amateur providers. In addition, in some health units, community volunteers or Expert Clients are retained (on pay through development partners) to support the ART service provision process and follow up of defaulters.

The health centres largely deliver outpatient and maternity services, but do not usually perform any procedures. The standard norm is a catchment population of 10,000

although this norm is rarely attained. Health centres currently form the bedrock of HIV and AIDS ART service provision. A number of health staff, usually, Medical Assistants, Clinicians and Nurse Midwife Technicians (NMT) are trained as providers of ART. A cadre of HIV and AIDS Diagnostic Assistants (HDAs) supported by some partners such as MACRO, EGPAF and the Global Fund provide HIV diagnostic services at some health units that are designated as ART sites. These perform HIV testing for clients at the health centres. Nonetheless, due to human resource constraints, some health units have deployed what are called "lay HTS providers" and they include Ward Attendants and Ground Labourers to assist with HTS. Community Hospitals under the CHAM are relatively larger than Health Centres, and provide outpatient and maternity services, and also conduct minor procedures. They have a bed capacity of about 250 (GoM, 2017). The Community Hospitals further provide ART and HTS, with a number of providers trained at each hospital. Hence, the administrative costs (financial and economic) within these settings should have a bearing on the efficiency of the HIV and AIDS programme in the country. This highlights the need to track levels of both production and transaction costs for health services provision, including HIV and AIDS services such as ART because they have bearing on efficient ART service provision outcomes.

Secondary Level

The *secondary* level consists of District public and CHAM hospitals, and as of 2022, the secondary level health entities accounted for nearly 10 percent of health facilities in the country (GoM, 2023). They provide referral services for Health Centres and Community hospitals. They further provide out-patient and in-patient services within the surrounding community. The public sector District Hospitals and CHAM hospitals further provide ART services. Besides provision of ART services, they attend to HIV and AIDS referral, especially where there are complications due to ART. The District Hospitals further receive dry blood samples for clients requiring TB and viral load assays. Furthermore, District hospitals are expected to provide supervision and mentoring support to Health Centres, Community Hospitals and private units that provide ART within their areas of jurisdiction. Hence, they are both the providers and quality assurance entities for ART at the district level. An examination of their efficiency with regard to the provision of general health services, and ART services is thus relevant as it provides the basis for the determination or confirmation of governance modes that are efficient in ART transactions coordination.

Tertiary Health Care Level

The *tertiary level* is composed of the Central Hospitals that provide referral services to the District hospitals in each region. The Central Hospitals are designed to provide specialist health services. Nonetheless, in practice, much of their pre-occupation is with primary and secondary health care services, estimated at 70 percent of the services provided (GoM, 2023a). This has been attributed to a weak gate-keeping framework, where clients are not turned away for skipping the primary and secondary healthcare units. There are currently four public central hospitals, namely, Mzuzu Central Hospital that caters for the Northern Region, Kamuzu Central Hospital (KCH) that caters for the Central Region of the country, Queen Elizabeth Central Hospital for the South-Western part of the Southern Region, and the Zomba General Hospital that caters for the South-Eastern part of the Southern Region. Each of the Central Hospitals has a centre for ART services. The Central Hospitals further provide laboratory services for viral load monitoring, and in that sense, service district hospitals with this service. These activities entail both production and transaction costs in the provision of HIV and AIDS ART services.

Central Level

The central level, which consists of the MoH Headquarters with a number of Departments formatted around the various Health Programmes is responsible for policy making, setting standards, strategic planning, quality assurance, monitoring and evaluation, technical support, resource mobilisation and international representation of the health sector in the country (GoM, 2017). The central level further consists of five Zonal Offices in the Northern, Central West, Central East, South East and South-West. The zones extend the headquarters functions through provision of technical support and quality assurance monitoring to the districts and health facilities.

6.3 HIV and AIDS ART Services Governance

For HIV and AIDS services, and especially the ART services that account for the bulk of the HIV and AIDS resources, a fully-fledged Department of HIV and AIDS (DHA) is retained, and performs the headquarters roles related to HIV and AIDS. These include setting policies and standards for the biomedical response under which ART falls, as

well as their enforcement. It is further responsible for training and re-orientation of providers when policies and drug regimes change. The Department further conducts quarterly supportive supervision visits and data collection visit to every facility providing ART. The HIV and AIDS Department is housed in its own building and is complete with a procurement, M&E and accounting section, alongside the technical departments that support the various health units that provide ART and HTS.

The figure below (Figure 13) depicts the HIV and AIDS Service Delivery mechanism, showing the different actors, coordination and reporting roles.

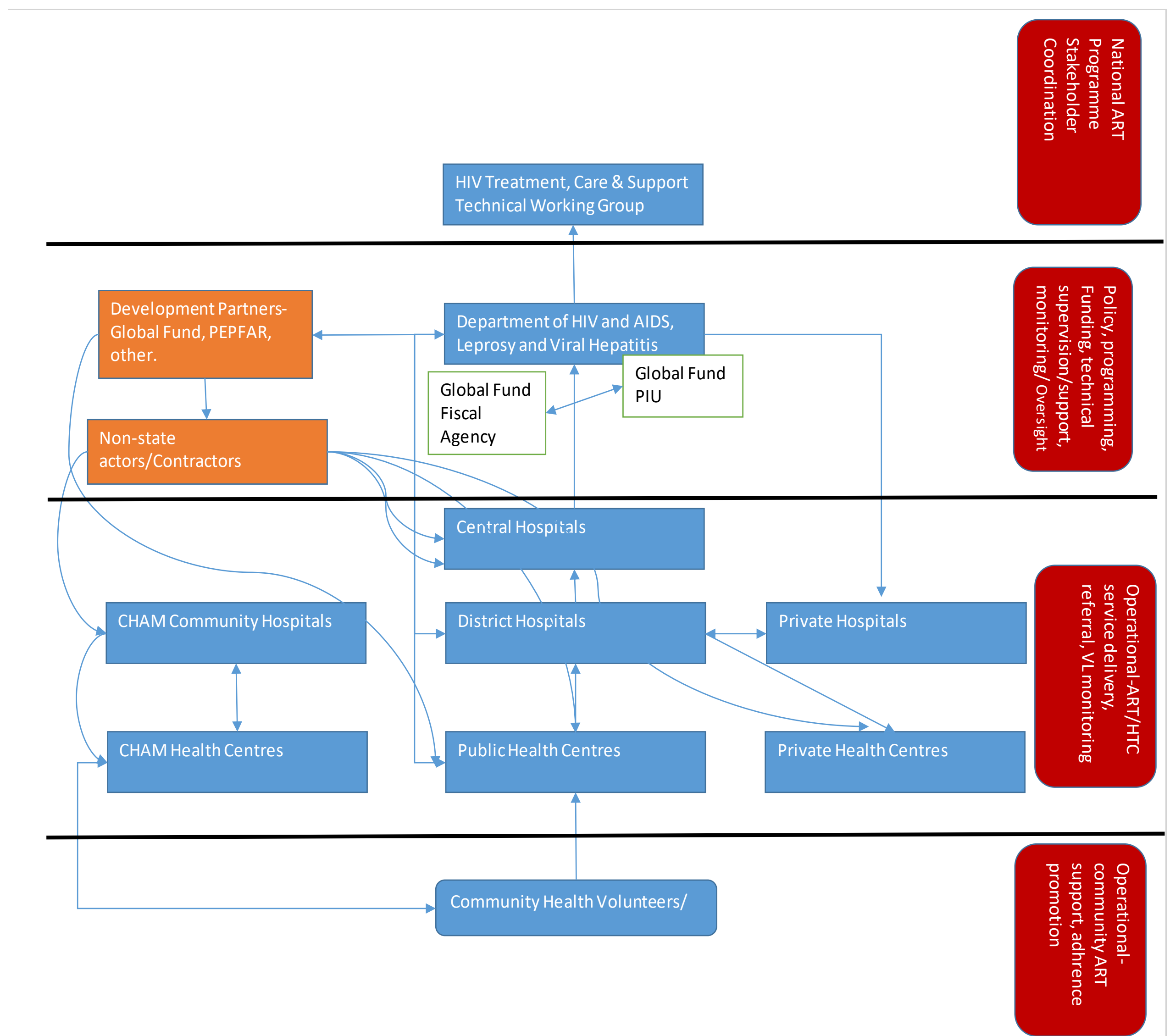


Figure 13: Malawi HIV and AIDS Service Delivery Mechanism

Note: Bottom-up arrows in the diagramme depict flow of information

Source: Author's Conception based on Literature Review and Interviews

6.3.1 National Level ART Coordination

At the national level, the ART programme is coordinated by the Department of HIV and AIDS (DHA) in the Ministry of Health. The DHA mandate extends to policy and technical guidance for all aspects of the biomedical response. A recent functional review of the Ministry of Health extended this mandate to the prevention and management of multi-morbidities linked to HIV and AIDS, including viral hepatitis[15]. The DHA is supported by technical and administrative support from the Global Fund Project Implementation Unit (PIU) housed within the MoH. The Global Fund PIU is responsible for coordinating the acquisition of HIV and AIDS ART commodities. The DHA further defines the regimen policy and ensures coordination of biomedical activities. A supply chain management firm is retained to provide supply chain management services for HIV and AIDS ART commodities. The DHA is further responsible for the overall monitoring and accountability for the biomedical aspects of service delivery. It coordinates with national level agencies such as the NAC and participates in coordination forums such the Technical Working Group on Treatment, Care and Support. Ultimately, concerning ART service delivery, the DHA represents a headquarters entity with a primary coordination role, with the different health units as service delivery nodes of the programme. The DHA undertakes quarterly monitoring and data collection visits to all the 754 ART sites (as of 2018) in the country as part of its quality assurance and monitoring and evaluation functions. This entails use of resources that indirectly link with ART, including occupancy, training and monitoring.

6.3.2 The Role of Contractors and NGOs Funded by PEPFAR and Other Agencies in ART

Funding agencies, contractors and Non-Governmental Organisations working in the HIV and AIDS sector further provide various support to the District hospitals and health facilities that entail use of resources. In some instances, this support includes provision of staff such as clinicians that support direct ART service provision, and sub-granting to NGOs that work with HIV and AIDS clients. The support from funders and CSOs financed by the USAID CDC and PEPFAR further extends to mentoring and supervision of service delivery at the facility level. This has implications on transaction costs in the overall governance of HIV and AIDS in Malawi. In other instances, this support is more indirect through training and payments to retain expert clients that work in some of the

[15] Key informant interview, MoH, Lilongwe, August 2020

health facilities, but nonetheless contributes to the cost of service delivery, as will be further analysed in Chapter 7. The expert client roles include organisation of records and follow up on clients that default on treatment. The expert clients are conceived as volunteers that support facility-based ART service delivery, although results from the research show that some of the expert clients are employed and paid. The research found non-state agencies that support selected health facilities and include expert clients with PEPFAR funding that pay them a monthly fee of MK70,000.00[16]. It is noteworthy, nonetheless, that this support is not uniform across the facilities and geographical spaces. The funders that provide this kind of support are in selected districts and facilities across the country.

6.3.3 Other Support Services to the ART Programme

The ART programme includes ancillary services to secure greater effectiveness, but which contribute to the overall costs of HIV and AIDS delivery, including transaction costs. These services include laboratory, ART products supply chain management and pharmacy. Laboratory services are provided at the secondary (District) and tertiary (Central Hospitals) health care provision levels, and are key to the viral load monitoring of clients. A National HIV and AIDS Reference Laboratory (NHRL) is in place in Lilongwe, which further assists with viral load monitoring and service quality monitoring. The NRHL has further capacity for HIV genotyping and sequencing.

The *pharmacy services* at each health facility play a critical role in the ART drug dispensation and the monitoring and ordering of ART drugs and commodities such as test kits and reagents. Although they are indirect, these services are integral to the overall provision of ART services (UNAIDS, 2011).

The HIV and AIDS products acquisition and distribution function is coordinated through the supply chain management section of the Global Fund PIU and DHA. The DHA, with support from Chemonics International and Bolloré Logistics ensure all the commodities required for the ART programme are sourced and distributed to the various health units. The commodities are ordered through the GAVI, a mechanism to aggregate procurement of drugs and other commodities by groups of countries to take advantage of efficiencies through bulk purchase.

[16] Approximately $100 in 2020

Having described the structure and implementation arrangements for the HIV and AIDS services within the context of public healthcare service delivery, the following sub-section analyse the conformity of HIV and AIDS governance with TCE theory.

6.4 HIV and AIDS ART Governance Conformity with Transaction Cost Economics Theory

Assessment of conformity with TCE by the existing ART governance instruments was evaluated against the following parameters:

- ART governance instrument suitability for dealing with transaction characteristics, namely, asset specificity, uncertainty and complexity in order to minimise transaction costs;
- Governance instrument efficiency in resource allocation
- Governance instrument competence to achieve equity

Transaction Cost Economics theory predicts that the design of economic governance instruments will be chosen considering their efficacy in adapting to disturbances for efficient production of goods and services. As with the private ordering of governance instruments, the discriminating alignment hypothesis may be applied to the public sector (Williamson, 1999). The hypothesis predicts that the choice of transaction governance instrument will be informed by the governance mode that minimises transaction costs to achieve the most efficient outcome. The current choice of public governance of the HIV and AIDS ART programme in Malawi is thus assumed to have been informed by similar calculations. Furthermore, TCE theory suggests that beyond economic efficiency considerations, remediableness considerations where competence figures prominently should be a key criterion for efficiency assessment.

The following table (Table 11), which is an extension of the private economic governance ordering (*see chapter 2*) in TCE to the public sector summarises the alternative public sector economic governance instruments and their attributes, based on the discriminating alignment hypothesis.

Table 11: Distinguishing Attributes of privatisation Regulation and Public Agency in public sector organisation

Attributes	Governance Structure		
	Privatisation	*Regulation*	*Public Agency*
I*nstruments*			
Incentive intensity	++	+	0
Bureaucratisation	0	+?	++
Performance Attributes			
Adaptive Autonomy	++	+	0
Adaptive Integrity	0	+	++
Contract Law			
Employment Relation			
Executive Autonomy	++	+	0
Staff security	0	+	++
Legalistic dispute settlement	++	+	0

Key: ++=strong; +=semi strong; 0= weak; Source: Williamson, O. E. (1999)

As observed from the table, Williamson (1999) distinguishes three types of governance instruments in a comparative public sector organisation, each with differing incentive intensity and performance attributes with respect to *ex post* adaptation to disturbances for efficient governance. These features determine each governance mode's efficacy at coordination of public sector transactions and dealing with disturbances *ex post* contracting, towards the efficiency goal. The public governance instruments are: (a) *privatisation* where public services are privatised through competitive bidding, b) *regulation*, where services are delivered using a combination of private and public sector modes. This closely resembles the hybrid governance mode under the private sector economic governance, and, (c) *public agency*, where services are provided by the

Government through a public agency. In this scheme, privatisation and public agency provision represent the two extremes of public sector organisation, with regulation representing a middle ground governance mode.

6.4.1 HIV and AIDS ART Transaction Characteristics and Governance instrument alignment

a) ART Transaction Complexity and transaction costs minimisation

As noted before, ART transactions are deemed complex and hence, the need for a hierarchical or bureaucratic governance instrument. Using the governance instrument categorisation above, the research observed that the choice of ART governance in Malawi reflected predominantly the public agency (own production through the Ministry of Health) and regulation. The main features of these governance instruments reflect that there is a high degree of bureaucratisation, or internalisation of ART transactions when compared with privatisation. Even under regulation, the extent of bureaucratisation is greater than that under privatisation. This suggests that the ART transactions are considered complex so as to warrant hierarchical or bureaucratic coordination. The chosen ART governance instruments are both weak in incentive intensity, that is, performance of health facilities would not necessarily be based on pecuniary incentives, when compared with privatisation. Furthermore, while ART service delivery through the public agency (MoH) and regulation are weak at adaptive autonomy (the freedom to innovate with minimal central control), they display strong adaptive integrity (strong coordination from the centre). Finally, with regard to the contract law that guides public agency and regulation service delivery, there is extremely low executive autonomy under public bureaucracy when compared with privatisation, and moderate under regulation. Staff employment security in ART provision is strong under public agency provision and regulation than under privatisation. Dispute resolution is largely internal under public agency provision when compared with privatisation where a third-party legalistic path is strong.

The above attributes have implications on governance efficacy for ART. Firstly, ART provision by the public sector implies a greater central coordination role with further implications to manage complexity. Secondly, the research observed that a strong adaptive integrity trait implies that innovation can only be sanctioned from the centre, even where a district or health facility has an innovative solution that is efficient. This is

important for standardisation of service delivery quality given the multiplicity of service delivery points. This finding was supported by evidence from one of the participants who emphasised the role of central policy and standards guidance, as highlighted in the following comment:

> *"We follow guidelines that are supplied by the Ministry of Health when providing the ART services. The guidelines help with ensuring that the standards are followed and maintained. They have to be followed strictly to ensure that standards are not compromised. Whenever there are changes to the guidelines by the Ministry of Health providers have to be re-oriented". Respondent 12, Facility Level ART service provider, Nkhata Bay District.*

As can be observed from the quote above, the need to maintain standards in ART provision through central level coordination is emphasised.

In addition, the research observed that internal resolution of disputes under public agency provision has further implications on transaction cost minimisation, where it is cheaper to resolve disputes than when courts or mediation entities are involved as third parties. This implies that transaction costs when managing a complex transaction are reduced. Hence, considering the governance instruments' attributes with respect to managing a complex transaction such as ART, the existing ART governance framework was assessed to be in conformity with TCE theory.

b) Efficacy to deal with Asset specificity and minimise transaction costs

The research found that this governance choice was in conformity with TCE theory which anticipates that where transaction display asset specificity, bureaucratisation (hierarchy) or own production and regulation is the most appropriate mode to coordinate the transactions. According to transaction cost economics theory, employment security is strong under public sector governance which embeds long term staff employment security. This is in contrast with the private sector where it is relatively easy to hire and fire staff, with limited staff employment security. This is an important trait as it ensures the transaction costs of staff replacement and retraining are minimised over the medium to long term. For the HIV and AIDS ART services where staff skills are highly specialised for ART provision, replacement of staff is likely to cause disruptions to

service delivery. The ease of hiring and firing staff likely to be common in the private sector would mean that each time staff that are skilled in ART provision leave the agency, human resources gaps would require the replacement and retraining of staff, and hence, lead to increased transaction costs. More specifically, staff providing ART services are recruited as civil servants with long- term employment security, where the government is less likely to fire them even where there are lapses in the quality of service delivery. Consequently, and considered with respect TCE theory, the research found that employment security of staff providing ART services served a transaction cost minimisation and efficiency role. The design and implementation of the contractual relationship in HIV and AIDS ART governance was thus assessed to be in conformity with TCE theory.

c) Dealing with Uncertainty and probity hazards to minimise transaction costs

Transaction Cost Economics theory predicts that the best governance mode is one that is able to adequately and efficiently adapt to disturbances in exchange transactions (Williamson, 1981). Regardless of the type of bureaucracy, that is, public or private, the provision of ART services is likely to face both internal and external perturbation that requires adjustment to secure efficient service provision (Mick & Shay, 2016). Internal partners consist of staff that provide the services and their principals. The staff and the organisation or firm will have contracts that define duties and responsibilities of each party in the provision of the services (Victor & Paulo, 2023) to facilitate relationships between the principal (government and staff). For Malawi, the research observed that rules governing performance of services will either be the Malawi Public Service Regulations (MPSR), the public sector guide book for public servants, or those based on contract law for private firms offering healthcare services. Due to bounded rationality, these contracts will be incomplete, and in theory, may be impacted negatively by opportunism, either on the part of employees or principals, and would hence, require mechanisms to monitor staff performance with respect to service delivery. Adverse behaviour by staff pose reputation risk to the Government, and may affect the efficiency of ART service provision. Thus, the human resources management function within firms becomes an important one, but it constitutes one of the transaction functions. The research found that for the Malawi public service, this role is fulfilled by the Department of Human Resources Management and Development (DHRMD) which has decentralised functions within the public sector Ministries, Departments and Agencies. Because the

MoH, like other MDAs, can be afflicted by the probity hazard, where staff do not provide the services as expected by the citizens, there is a reputation risk that may be politically damaging to the Government in general, and the MoH in particular.

The research found that the use of the MPSR is designed to limit the probity hazard in the performance of duties by defining guidelines and expectations of staff providing general healthcare and ART services in public health units (Spithoven, 2012). Administrative law is further deployed to limit this hazard (Banda, 2019). These requirements for safeguards and monitoring of publicly provided ART services have implications on the efficiency of service provision as they generate non-production or transaction costs that increase the total cost of service provision. Transaction Cost Economics theory predicts that these procedures will be the preferred default mechanisms for internal conflict resolution, to courts (Williamson, 2002). This is because it would be less costly to resolve the conflict internally than through third parties such as the courts. Consequently, the public agency governance mode to delivering HIV and AIDS ART services was deemed to have been in conformity with TCE theory.

External disturbances for both the public and private bureau arise from exchange interactions with partners external to the organisation (Tadelis & Williamson, 2012). In the context of ART services, the SLA arrangement poses the risk to both the MoH and private partners. For the public sector, there is a reputation risk where the private sector agency does not perform services to expected standards. The study found that as the principal partner, the MoH relied on the SLA which confers the rights for service delivery onto the Ministry. It further found that the SLA was the basis for correction or termination of the partnership where it is no longer possible to continue with the partnership. Private agencies on the other hand, face a reputation hazard from the perspective of the clients they serve. For instance, under the SLA arrangement, the research found that Government is expected to refund the financial resources used to defray the cost of service provision. If the Government fails or delays substantially to refund the expenses, the private agency will need to meet these costs to avoid service disruption. Indeed, there are documented cases of the Government delaying payments, which leads to private agencies in the SLA arrangement spending beyond what is agreed

as they have to acquire commodities on the market at relatively higher prices than agreed (Gama, 2013; Tobias, 2020).

6.4.2 Governance instrument and Efficiency in ART service delivery

a) Resource allocation

According to TCE theory, the design of the governance instruments *ex ante* should reflect the efficiency goal of economic governance, mainly through ensuring that wastage is minimised. This extends to the efficient allocation of resources. The research found that with regard to efficiency considerations in the design and implementation of the ART programme, there was deliberate administrative allocation of ART services to ensure everyone that needed it accessed it. The following sentiments tend to lend credence to this claim:

> *"We have tried to make access to ART as close as possible to the people. The current distribution of the facilities providing HIV and AIDS reflects this aim. There are still challenges in some areas, especially hard to reach areas where clients have to travel long distances, but most will access their drugs within their catchment areas"- Respondent 1, Policy and Strategy Manager*

> *"The programme could be said to be effective in reaching the vast majority of people that need the ART services. Nearly all people that are on ART can access the services now. Over the years, we have decentralised the services to the health centres. We have even included private sector health units to expand access. As long as we are able to reach those that need the services, this is a huge achievement for the HIV and AIDS programme." Respondent 1, Policy and Strategy Manager.*

From the above quotes, the design and implementation of the HIV and AIDS ART programme appears to be effective in the delivery of its mandate and thus fulfils the competence criterion of governance instruments as advanced by TCE. Accounting for the need to confirm the economic governance efficiency, the research found the HIV and AIDS governance conformed to TCE theory in as far as competence in HIV and AIDS service delivery was concerned.

Secondly, the research found that in HIV and AIDS ART programming, efficiency considerations consist of ensuring that only those activities that are effective and efficient are included in the programme. This is achieved through modelling to inform the ART programming. Modelling software such as GOALS and NAOMI are used to this end. Furthermore, the MoH conducts cost-effectiveness analyses prior to inclusion of strategic interventions and activities in the ART programme. The following quotes highlight the efforts in planning for efficiency in ART services:

> *"There's always effort, making sure that investments are made in areas, in strategic areas, areas where you're going to get a high yield, for example, for maybe if it's HIV testing...for new HIV infections.... so, there's always rigorous analysis of where do we place these machines? Where do we construct this facility? Where do we, you know, like which people, like how many more health staff do we need to allocate and where, and all those things. So, there's a lot of rigor when that is happening. ... the planning", Respondent 2, National Level Programme Manager.*

> *"The HIV and AIDS is one programme which uses a lot of research evidence before making a decision about how to implement anything." _ Respondent 2, Policy Manager, National Level*

> *"Efficiency is considered in the set-up of the ART programme but coverage to ensure everyone that needs the ART services is reached is also emphasized. We can say we have tried to reach many people that need the services" Respondent 5, Policy and Strategy Manager.*

The above quotes suggest that there are conscious policy and planning efforts through the MoH to ensure efficient allocation of health services, and by implication, ART services. Based on this result, the research found that the HIV and AIDS ART was in conformity with TCE, with the caveat that this allocation actually translates into efficiency. This is considering that empirical evidence regarding the efficiency of the overall ART programme, as well as the facility level provision is limited. In addition, perceptions on efficient management of the response show differing perspectives by policy makers, programme managers and providers, partly because there have not been consistent efficiency analyses of the HIV and AIDS ART programme in the recent past. For instance, the National Health Financing Strategy appears to suggest that even where

evidence is available to facilitate efficient resource allocation with respect to HIV and AIDS, it is rarely used in the public health sector budgeting (GoM, 2023b).

b) Regulation as a means to achieve healthcare and ART efficiency

As noted before, ART services are further provided through regulation under a public-private partnership (PPP) arrangement between the Government and non-state actors which constitutes a hybrid governance mode. This is achieved via the service level agreements (SLAs) (Gama, 2013). This hybrid governance arrangement is designed to extend the reach of ART services with a high premium placed on efficiency considerations. This arrangement obviates the need for the creation of infrastructure for service delivery by the state, and thus avoids the set up costs (which constitute transaction costs) of facility based ART provision that would have been incurred by the Government. In this sense, and from the perspective of the Government, efficiencies in service delivery are achieved via the reduction of set up costs and reduction in duplication of effort (Zeng , Mphwanthe, Huan, Nam, & A.Dutta, 2017). In addition, the government enters into these agreements where it is not efficient for the government to construct own health unit infrastructure due to proximity to a non-profit health unit, such as that of CHAM or Muslim Association of Malawi (MAM), highlighting the efficacy of the ART governance instruments in allocative efficiency of the services as predicted under TCE theory.

c) Efficiency in the acquisition of ART drugs

With regard to efficiency in procurement of ART drugs and commodities, the research found that a pooling arrangement through the GAVI, the country is able to lower the transaction costs for bidding alone as a country and also the unit costs via a negotiated price. Although it has not been evaluated, this arrangement was said to have efficiency implications as unit prices for commodities tended to be lower when purchases are pooled across a number of country requirements than when each country purchased on its own. Results from the research show that there are deliberate measures to achieve efficiency in the acquisition of ART drugs and commodities, which is consistent with TCE theory. The following quote from an HIV and AIDS policy manager demonstrates this result with respect to the efficiency contribution of a pooled arrangement to drugs procurement:

"Malawi acquires its drugs through the pool of countries. This arrangement means that we are able to get a better bargain on the prices than when we are buying as a single country" Respondent 3_Policy Manager, National Level KII

As can be observed from the above quote, pooled acquisition of commodities is a factor in the achievement of efficiency in the ART programme. However, this contribution needs to be tracked on a continuous basis to ensure it is accounted for in efficiency analyses of the HIV and AIDS service delivery. Moreover, it has to be considered in the context of the overall efficiency of the HIV and AIDS programme, especially the biomedical aspects of the programme to which it directly contributes. This assessment should include the tracking of transaction costs associated with the current arrangements for the acquisition and distribution commodities in the programme.

d) Transaction Cost reduction through a harmonised field allowance framework

Additionally, the research found the existence of efforts by donor agencies and the Malawi Government to control transaction costs. This consisted of creation of a framework for harmonised field allowance rates to be paid to Government and other staff on missions financed by the donors. These allowances applied to the ART programme, particularly in the use of Global Fund resources. Donors have harmonised subsistence, fuel and transport allowances, which are applied by the government and development partners (donors and NGOs) to curb opportunism by individuals who began to view allowances as a financial incentive to supplement their salaries. This was viewed as a perverse incentive which impacted efficiency of public sector agencies through increased transaction costs. The following extract from a circular from donors to the Government of Malawi is instructive:

"Events such as workshops, training, seminars, and alike are intended to strengthen individual capacity and improve job performance while participation in missions is considered part of one's official duties. However, participation in such events and missions often came to be viewed as one method of supplementing an individual's salary. These guidelines were

developed to minimize the financial incentive associated with participating in such events and missions."[17]

The application of administrative procedures by donors to minimise cost was further confirmed through key informant interviews with ART programme managers, as seen in the following quote:

> *"Efficiency is usually monitored through the Department of Planning. However, as a Programme Unit, we have set up procedures to ensure efficient service delivery. These include harmonization of allowances, joining the global pool in the procurement of ARVs and commodities, such as the GAVI. The setting of targets further helps the programme to ensure we maximize the use of resources available" Respondent 3, Pogramme Manager.*

The above quotes suggest that the Government of Malawi in partnership with donor agencies that support the ART programme has deployed measures for administrative cost control for otherwise high costs that arise from the apparent opportunism by staff in the public sector. In this sense, the practices and administrative control appear to be in conformity with TCE theory with regard to use of resources, particularly with respect to transaction cost control. For the HIV and AIDS programmes, this is particularly important as coordination meetings and other events such as programme monitoring missions entail financial resource needs that would substantially increase transaction costs is left unchecked. The allowance issues as a key driver of transaction costs is evidenced by recent events involving funds from the Malawi Government earmarked for the COVID-19 response that were deemed to have been "wasted" through, among other items," hefty allowances" and "unwarranted bureaucracy at Ministry of Health"[18]. This experience showed that when left unchecked, allowances and system wastages can impact efficient health service delivery, within which HIV and AIDS ART service delivery is nested.

6.4.3 Governance instrument and equity considerations to reach vulnerable populations

As observed above, the HIV and AIDS ART services are delivered via the public agency and a combination of public and private arrangements. The research found that

[17] Communication from Malawi UN Resident Coordinator to the Secretary to the President, January 2022.
[18] Centre for Investigative Journalism, online, September 26, 2023

the public governance choice aims to prevent adverse selection that would obtain if the services are provided through a privatisation governance mode due to the profit drive of agents in the market. People in the low wealth categories are likely to be excluded by virtue of their inability to pay (Muula, 2002) for the services where provision is driven by profit. More concretely, for Malawi, the public governance choice of ART is likely to have been informed by the *remediableness* principle as described in Chapter 2. The research found that the provision of the ART services through the public sector reflected the Government's choice of governance mechanisms to competently and efficiently deliver these services to reach even the vulnerable groups. The ART services in Malawi take on the characteristics of sovereign services where the public bureaucracy has the comparative efficiency advantage as governance mechanism over privatisation (Williamson, 1999; Spithoven, 2012). At its most basic, this governance choice was between using privatisation (market) and hierarchy or public agency as a governance mechanism. In theory, it should be possible to deliver the ART services through the market or privatisation, but the highly asset specific and complex nature of these services require the use of the public agency or bureaucracy to coordinate the transactions (Williamson, 2010). The relatively high levels of poverty and low incomes in a developing country such as Malawi would render the services inaccessible to the majority of the people that need them. Furthermore, the private sector's focus on cost reduction to achieve profit maximisation is likely to lead to adverse selection problems (Spithoven, 2012). Indeed, initial efforts at provision of ART with financial contribution arrangements was associated with low access by people that could not afford the payments for viral load monitoring (Muula, 2002). Consequently, the ART governance was assessed to have been in conformity with TCE theory.

Thus, the research found that the governance mode was calculated to achieve efficiency and equity in the allocation of resources with regard to ART by the MoH. The inclusion of the programme within the health systems in the country ensures that everyone that needs ART can access it. Hence, on the basis of achievement of equity and efficiency, the governance instruments were assessed to be in conformity with TCE theory. It is important to note, nonetheless, that while a general conformity of the ART governance instruments with TCE theory is observed, its effectiveness in transaction cost minimisation may be hampered by the rapid expansion and decentralisation of the programme. In addition to headquarters costs, the rapid expansion of HIV and AIDS

ART decentralisation to disparate units across the country requires investment in monitoring and standards and integrity enforcement.

6.4.4 Observed Gaps and deviation from TCE in HIV and AIDS ART governance

The forgoing text which suggests that the HIV and AIDS ART governance design is in conformity with TCE theory has to be caveated by the fact that there is limited evidence on the economic efficiency of the HIV and AIDS services in Malawi. Further, the levels of transaction costs in ART are not adequately monitored with the aim to inform efficient service delivery. In addition, there are observed structural features of the governance framework that depict wastage, and therefore, violate the first order economising principles of economic governance and Transaction Cost Economics theory. In particular, a parallel supply chain governance that is part of the ART programme and excessive safeguard tend to be in sharp contradiction with TCE theory.

a) Parallel Supply Chains/Logistics for the HIV and AIDS ART Programme

As noted before, TCE theory predicts that design of governance instruments will seek to minimise wastage. This includes the costs of acquiring inputs and their distribution. The research found that the design and implementation of the framework for the acquisition of ART drugs and supplies, was among the governance aspects that did not fully comply with TCE theory in practice. There is a parallel supply chain with duplicated roles which tended to undermine efficient supply chains for ART and related supplies and commodities in public health care commodities and supplies, including ART drugs. This parallel supply chain mainly consisted of the Central Medical Stores Trust, the Government of Malawi's official supply chain management mechanism for all health-related supplies and commodities on the one hand, and systems based on donor funded mechanisms such as the GFTAM and UNICEF. While some research participants, particularly policy and programme managers directly managing the HIV and AIDS ART programme reported that the ART supply chain is robust, others indicate that a parallel system within public sector HIV and AIDS ART governance was wasteful.

Participants to the research acknowledged that this parallel arrangement needed to be reviewed and streamlined for efficient procurement, storage and distribution of ART commodities and supplies. The following quote illustrates sentiments relating inefficiency in the current ART supply chain set up:

“Obviously, what we have now is a system where there is a parallel structure for the acquisition, warehousing and distribution of ART drugs. This amounts to wastage as the CMST capacity could have been used for this purpose. But you have to understand where we are coming from with the previous bad reputation of the CMST regarding governance, especially related to drug pilferage ”_Respondent 6, Programme Manager.

The existing arrangement was reported to be the result of limited capacity by the CMST, although the research found that the main reason was limited confidence in the CMST in the face of observed persistently reported drug pilferage and limited transparency. The pilferage that has been observed over the years reflects the existence of opportunism amongst staff in the CMST in the acquisition of supplies. This tends to increase the costs of using a public agency for procurement in the public health system in general, but also the HIV and AIDS ART programme in particular. To mitigate against this, donors have tended to favour a different set up to acquire and deliver ART drugs and supplies. Under this arrangement, the DHA, with support from Chemonics International and Bolloré Logistics (private firm) ensure all the commodities required for the ART programme are estimated, sourced and distributed to the various health units. The commodities are ordered through the GAVI, a mechanism to aggregate procurement of drugs and other commodities by groups of countries to take advantage of efficiencies through bulk purchase.

Nevertheless, because of the parallel arrangements, transaction costs from contracting different supply chain managers, storage and occupancy (office space) are likely to be high and would tend to erode the efficiency gains realised from the global pooling arrangements for the drug acquisition. Thus, from a TCE perspective, the parallel supply chain arrangement represents wastage and a failure in first order economising in economic governance. Furthermore, even when considered from the perspective of reduction of the probity and opportunism risk, the strategy adopted appears to be one that reflects a vertical arrangement that tends to significantly increase both the financial and opportunity costs of not using an already established supply chain governance structure in the CMST. Moreover, the research found that until 2019, the HIV and AIDS ART drugs and other commodities were being warehoused and distributed from CMST warehouses, where a 10% fee (warehousing cost that represented transaction costs) was levied on the cost of drugs and commodities.

The caveat to the foregoing text is that as of the time of the research, a reform agenda supported by donors such as the World Bank was in place towards resolution of the challenges that beset the CMST. Nevertheless, whether the reform means that eventually, the CMST would be used to manage the supply chain for the HIV and AIDS ART programme remains unclear.

b) Excessive fiduciary safeguards for the ART programme

As noted above, the response in Malawi is heavily donor dependent. Financing administration is thus a key part of the governance framework. The research observed that funding administration involves a number of fiduciary safeguards that appear to be excessive. As is described in the next chapter (Chapter 7) under a section on transaction cost sources, there are several layers of control that include activity approvals, audits and financial management. While the need for safeguards is acknowledged as a mechanism to deal with opportunism under the programme, a streamlining of these controls would greatly assist in freeing up resources that can be used to expand ART services. Furthermore, a multiplicity of actors supporting the ART programme in the country with own offices generate transaction costs that are likely to contribute to inefficiencies in the programme. In this connection, the research found that the wastage reduction principle was not fulfilled in the governance framework, and was thus inconsistent with TCE theory.

6.7 Chapter Summary

This chapter sought to assess whether the design and implementation of the HIV and AIDS governance in Malawi conforms with TCE theory. It has noted that these services are largely delivered through the public bureaucracy, reflected in a three-tier public health care system, comprising of primary, secondary and tertiary levels. Using the TCE discriminating alignment principle to complete the assessment, the research results suggest that the chosen governance framework largely conforms with TCE theory, except for gaps in implementation. This is demonstrated by both the design and implementation of measures that seek to deal with disturbances in transactions arising from uncertainty, probity hazards, opportunism and asset specificity. When assessed against the key TCE theory parameters that include governance instrument suitability in dealing with transaction characteristics to minimise transaction costs, efficiency in

resource allocation pertaining to the ART programme and governance competence to deliver mandate and achieve equity in access to services, the governance framework is deemed to conform with TCE theory. Use of a public agency via the Ministry of health with civil service conditions, including staff security of employment to deal with the probity hazard and asset specificity in ART services suggest compliance with TCE theory. Furthermore, deliberate measures to control costs through administrative measures such as harmonisation of allowances suggest compliance with TCE theory with regard to reduction of transaction costs towards efficiency goals.

Nevertheless, the Malawi HIV and AIDS ART governance conformity with TCE has to be caveated by the fact that in terms of the efficiency criterion, it should be viewed only in remediableness terms as will be discussed in Chapter 7. More concretely, the chosen governance mode may not actually be the best option for transaction cost minimisation, but the state mandate to distribute the services equitably renders it the most efficient. Key issues relate to indications of high transaction costs due to the multiplicity of agencies supporting the ART programme that have own set ups that increase transaction costs, including occupancy and monitoring costs. Additionally, the excessive layers of safeguards in ART programme funding are likely to increase transaction costs and reduce the efficiency of the programme. Furthermore, the governance set up does not adequately deal with uncertainty arising from reported opportunistic behaviour in the supply chain for ART drugs and commodities. A parallel supply chain structure that is set up to avoid the weaknesses of the government's health sector supply chain coordination entity, the Central Medical Stores Trust is considered wasteful and not in compliance with first order economising principles predicted by TCE theory. These gaps in the application of TCE are likely to have a negative bearing on the overall efficiency of the ART programme.

The next chapter presents results on the sources of transaction costs in the HIV and AIDS ART programme by undertaking a transaction cost economics analysis of the programme.

CHAPTER SEVEN

THE TRANSACTION COST ECONOMICS OF HIV and AIDS ART PROGRAMME IN MALAWI

7.1 Introduction

This chapter seeks to answer the research question: what are the main sources of transaction costs in the current governance of the HIV programme in Malawi, and whether these arise from noncompliance with transaction costs economics prescriptions? The aim of the chapter is to establish the rationale for the HIV and AIDS ART governance choice and highlight the consequences of that choice with regard to the sources of transaction costs. As observed in Chapter 2, TCE theory is concerned with the sources, incidence and efficient management of transaction costs (Williamson, 1999). Using TCE theory, the chapter argues that the whole ART financing and service delivery programme is a bundle of contracts involving a range of agencies that define decision making rights and responsibilities. This shapes the financing and implementation of the HIV and ART services in the country. The preparation and the management of these contracts produces transaction costs that have a bearing on total service delivery costs, and hence, should be included in the ART efficiency analysis.

The chapter is based on qualitative analysis of data collected from interviews with research participants at the health policy and programme management level, as well as literature review. Overall, the research finds that the sources of transaction costs consist of safeguards for HIV and AIDS funding, contracting, ART service delivery compliance monitoring, occupancy and supply chain management. The results suggest that safeguards structures for ART financing administration are excessive and are likely to generate substantial transaction costs, while a parallel supply chain management structure for drugs acquisition causes duplication and hence, wastage in the HIV and AIDS ART governance. These shortfalls in the HIV and AIDS ART governance are likely contributing to inefficiency.

After this introduction to the chapter, the basic TCE model is presented in section 7.2. This is followed by a summary of rationales for public agency (bureaucracy) choice as

the governance mode for the ART programme in Malawi in section 7.3. Section 7.4 presents the sources of transaction costs in the ART programme in Malawi. Section 7.5 highlights measures to reduce transaction costs in the HIV and AIDS ART programme. Section 7.6 provides highlights on the potential mechanisms to monitor HIV and AIDS ART transaction costs. Section 7.7 is a summary which concludes the chapter.

7.2 The Transaction Cost Economics Model

Transaction Cost Economics is concerned with the reduction of transaction costs using one capitalist organisational mode over another when completing exchange (Williamson,1975; Shervani, Frazier, & Challagala, 2007). These organisational modes may be the market, hierarchy or hybrid in private sector ordering of transactions. In the public sector, the choice is among privatisation, regulation (hybrid) and pure public sector agency service delivery. In the context of healthcare, HIV and AIDS services such as ART offered to clients are viewed as transactions or exchanges. The costs of completing the exchange consist of the direct production costs and pre-and post-production costs. Pre-production costs include the cost of setting up as well as market search, while post production costs may include the cost of contract monitoring and enforcement (Yousuf, 2017). The ultimate goal of transaction cost optimisation is efficient production and exchange or service delivery. Consequently, faced with different organisational forms, for services such as ART, public policy makers have the choice to provide these services throughout the continuum of care internally through the public health care system, or externally through contracting out some of services (Mick & Shay, 2016). As noted in Chapters 2 and 6, this choice is informed by a comparative institutional cost analysis through a discriminating alignment hypothesis of transaction costs at the margin. While organisational competence considerations are acknowledged, the public provision choice of health services in general, and for ART in particular has implications on the levels of transaction costs, and hence, efficiency in ART service delivery.

The TCE model is based on behavioural assumptions and transaction characteristics that inform efficient governance choice. The basic TCE model as conceived by Williamson (1975) and its logic is summarised in the following schematic (Figure 14).

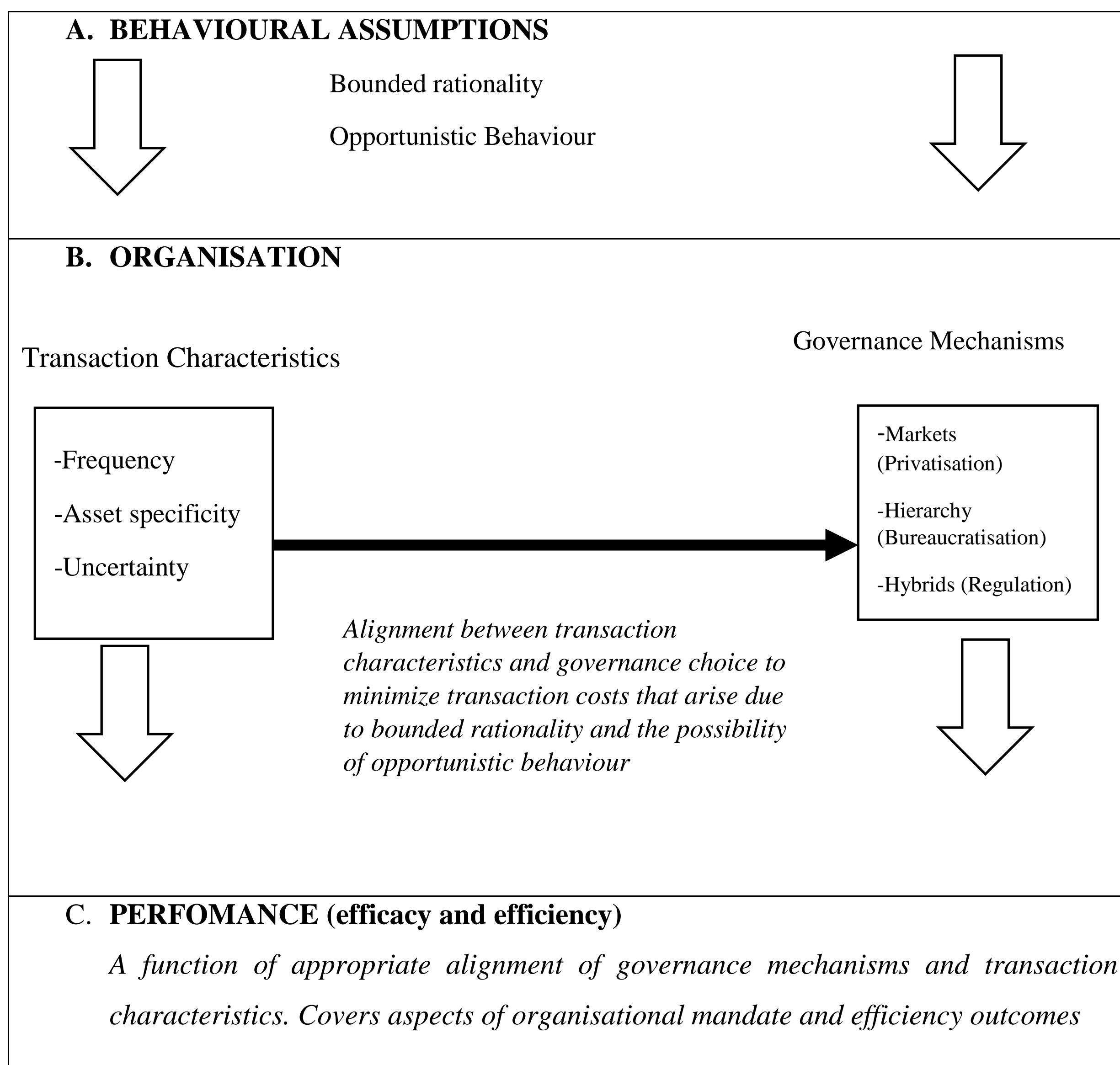


Figure 14: Transaction Cost Economics Model

Source: Adaptation from Cuypers, et al, 2021

In Block A, the TCE model highlights the role of the two behavioural assumptions, namely, *bounded rationality* and *opportunistic behaviour* and their influence on governance choice. Block B shows how exchange partners align transaction characteristics-*frequency, asset specificity and uncertainty-* with governance mechanisms-*markets (Privatisation), hierarchy (public agency) and hybrids (regulation).* For the public sector organisation, the probity hazard is a further risk that

has to be considered when designing an efficient economic governance mode for public sector organisation. This alignment determines organisational performance in Block C.

7.2.1 Governance Choice Under TCE

As noted before, TCE provides a theory for the explanation of the choice of efficient governance structures-loosely defined as the rules and the instruments that are deployed to enforce the rules (North, 1986). The governance structures define the roles, rights, duties and expectations of transacting partners. Economic governance choice is informed by efficiency considerations. The most efficient or optimal governance structure will be the governance structure with the lowest transaction costs, that is, the lowest coordination and management costs. There are different governance features and instruments. The governance features include assignment of property rights, contract law regime and reputation effects and risks (Spithoven, 2012). Governance instruments to operationalise the governance features will vary depending on the type of governance feature. The governance instruments include administration, incentives, adaptation and contract enforcement as described in Chapter 6. Administration will be common in public governance where public agencies provide services in accordance with the public interest. Public service codes of conduct regulate the behaviour of public servants when providing healthcare services (Spithoven, 2012). Incentives are associated more with the market or privatisation governance structure, with prices and other market-based instruments as key incentives, while the state maintains a regulatory role in hybrid modes. Further, in the context of the private bureau hierarchy, employee ownership schemes may act as incentives to limit shirking by staff (Wilson, Zang, & Robinson, 2003). Adaptation may also be autonomous through market forces and through hierarchy or bureaucratisation under private or public sector organisation, respectively. While the market uses the price mechanism, entrepreneur-managers in hierarchy will aim to allocate resources in ways that create value to attain business viability. Contract enforcement involves monitoring the performance and enforcement of compliance. This may be accomplished internally in hierarchy or through the courts.

Under TCE theory, governance choice involves a comparative institutional analysis where different governance modes are considered and the best mode chosen based on comparative costs as noted before.

Figure 15 depicts the process for governance choice with efficiency as an overriding objective.

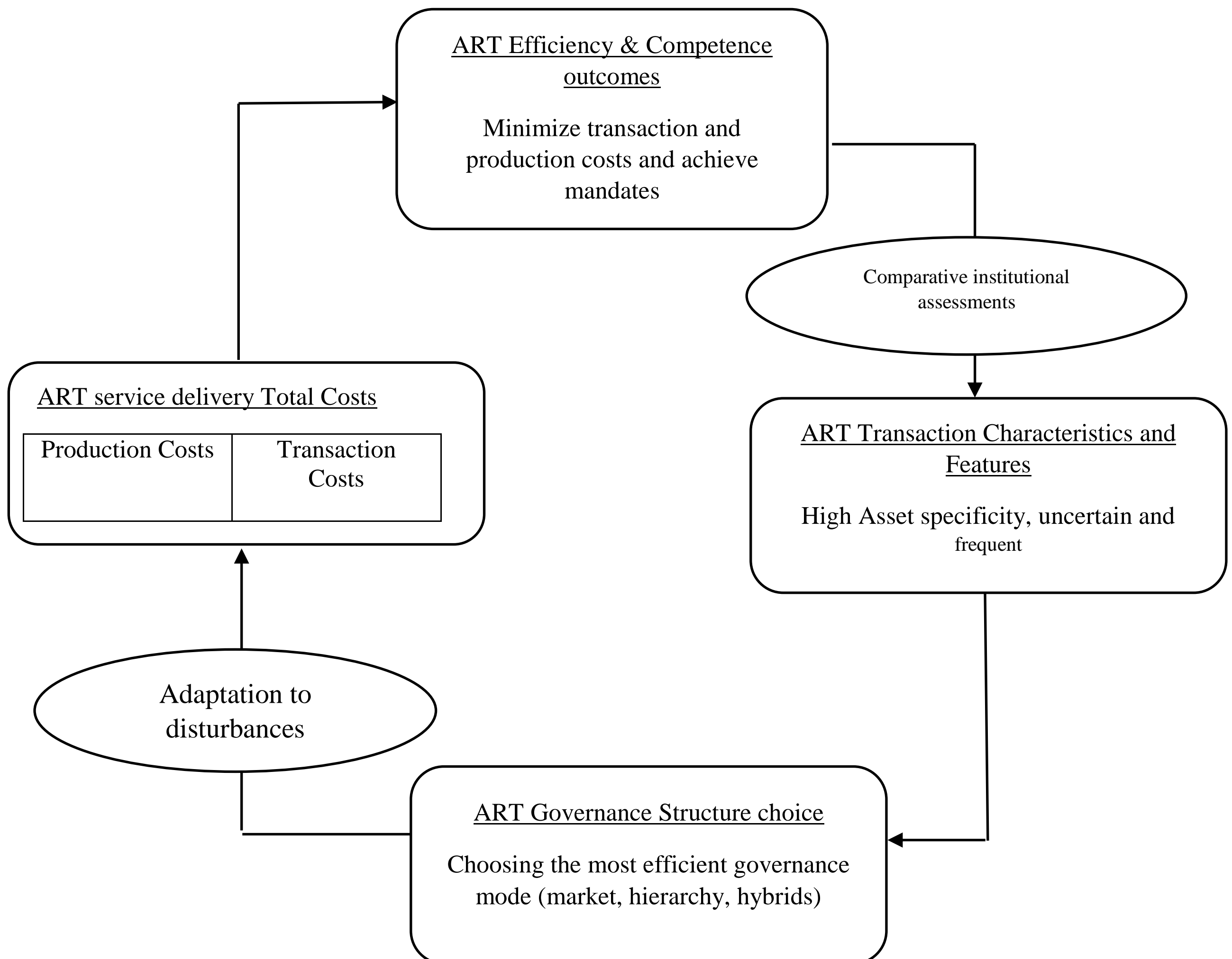


Figure 15: TCE institutional choice model

Source: Adaptation from Sophia (2010)

As may be observed from Figure 15, considering the behavioural assumptions and transaction characteristics, the comparative institutional analysis process starts with the consideration of transaction characteristics, namely, asset specificity, uncertainty, complexity and frequency of ART transactions. Depending on the degree of asset specificity and uncertainty, market (privatisation) or hierarchy (bureaucratisation) governance structures are chosen under private (public) sector organisation. The chosen governance structure is then deployed as a mechanism for adaptation to temporal disturbances of service delivery, and in the case of the current research, HIV and AIDS

ART service delivery. The efficiency outcomes from cost minimisation is the basis for further assessment and re-assessment in a discriminating alignment fashion ((Ketokivi & Mahoney, 2017). The starting point, or primordial governance mode may either be the market (privatisation) or hierarchy (bureaucratisation) under private (public) sector organisation. This management of adaptations generates costs that consist of transaction costs as described above. With a focus on the reduction of transaction costs for efficient production or service delivery, this cycle repeats, in a continuous fashion. This cycle provides a useful basis for unveiling the decision-making process leading to governance choice and the sources of transaction costs in the delivery of public services such as HIV and AIDS ART. More formally, and following Plunket & Saussier (2003), the HIV and AIDS ART service governance choice may be modelled as follows:

$$Z^* = \begin{cases} Z_p, if\ S_p > S_a \\ Z_a, if\ S_p \geq S_a \end{cases} \qquad (7.1)$$

Where Z^* is the public governance mechanism chosen for ART, Z_p and Z_a are public and alternate governance mechanisms,such as privatisation and regulation, respectively. The terms S_p and S_a are the perceived surplus or value by decision makers or exchange partners.

Since it is unlikely that decision makers will calculate or observe accurately (bounded rationality) the expected value realised from using one mechanism of ART service delivery governance over another, there is need to relate the benefits and costs of alternative governance mechanisms to the observable features of the transaction itself. This facilitates the testability of economic governance hypotheses (Plunket & Saussier, 2003). Hence, the following relations are incorporated in the model alongside the above stated arguments:

$$\mathrm{S}^{\mathrm{P}} = \mathrm{S}^{\mathrm{P}}(\mathrm{X}, \mu_{\mathrm{P}})$$

(7.2)

and

$$\mathrm{S}^{\mathrm{A}} = \mathrm{S}^{\mathrm{A}}(\mathrm{X}, \mu_{\mathrm{A}})$$

(7.3)

Where X is a vector of attributes that affect the gains from using the relevant governance mechanisms for ART services, while μ_P and μ_A are error terms that capture omitted variables or errors by decision makers or the analyst about the true values of S_p and S_a , largely due to bounded rationality and information asymmetry. If the above relationships are assumed to be linear, they can be restated linearly as follows:

$$S^P = \beta X + \mu_P \qquad (7.3a)$$

and

$$S^A = \alpha X + \mu_A \qquad (7.3b)$$

Hence, the probability of a particular form of governance mechanism emerging, such as public bureaucracy, over alternative governance mechanisms can be stated as follows:

$$\Pr(Z^* = Z^P) = \Pr(S^P > S^A) = \Pr(\mu_P - \mu_A) < (\beta - \alpha) \qquad (7.4)$$

This probability shows that the effect of a parameter of X , β, on the efficiency or competence gains of the governance mechanism that is chosen is greater than its effect on an alternate form of governance (e.g. public bureau versus market, private bureau, hybrid or CSO bureau), α. In the context of the HIV and AIDS ART provision, the implication is a high likelihood of the public bureau emerging as the preferred governance mode increases.

Using the market (privatisation) and hierarchy (public bureaucratisation) as the two extremes of transaction governance modes under private (public) sector ordering, and considering transaction costs only, it is further possible to demonstrate the comparative cost assessment and governance choice based on the following decision rules (Dietric *et al.*, 2015), holding all other things constant.

$$C_{tm} - C_{tf} > 0 \qquad (7.5a)$$

$$C_{tm} - C_{tf} < 0 \qquad (7.5b)$$

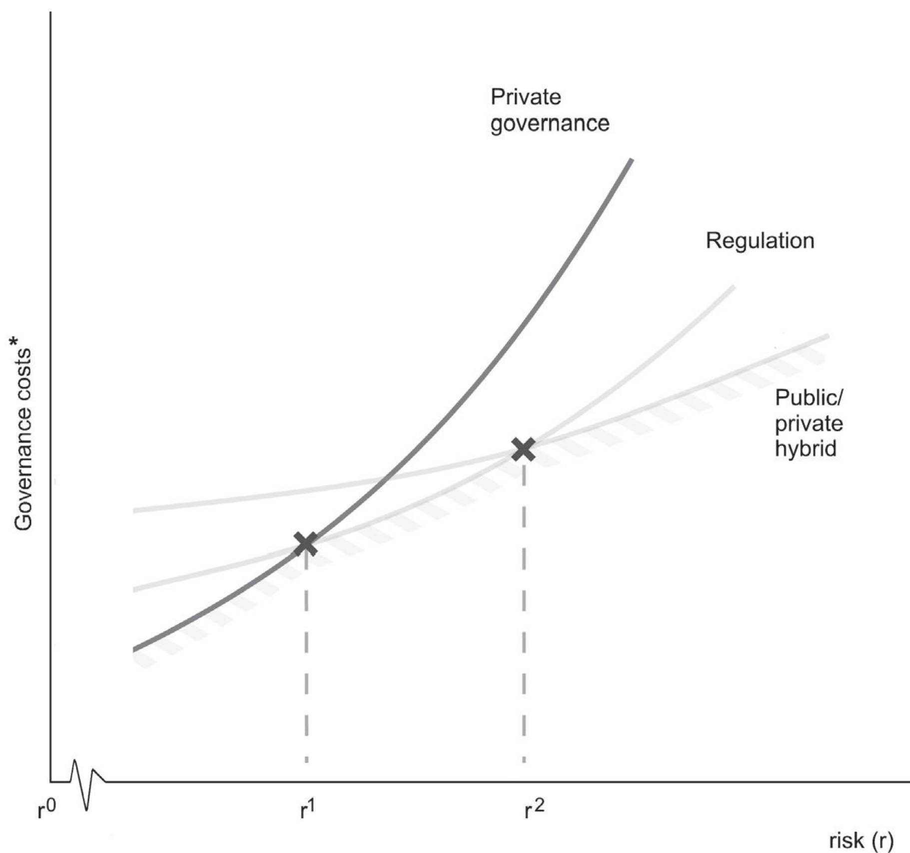
$$C_{tm} - C_{tf} = 0 \qquad (7.5c)$$

Where C_{tm} and C_{tf} are market (privatisation) and firm (bureaucratisation) transaction costs, respectively. The condition in equation 7.5a predicts that transactions will be brought under hierarchical coordination as the cost of internal coordination is less than that through the market. In equation 7.5b, transactions will be outsourced through the market governance or coordination as the cost of service provision or production by the firm is higher than that of production through the market. Condition 7.5c is a point of indifference in choice between hierarchy and market organisation of transactions.

Transaction Cost Economics theory projects the idea that it is important to look at how transactions in the production of a product or service are organised and how they impact the total cost of a firm. This perspective allows an assessment of efficiency in the production and distribution of the product or service. Related to HIV and AIDS ART services, TCE implies the need for an explicit focus on transaction costs besides the cost of drugs and commodities and human resources involved in the direct provision and distribution of products or the services. The total cost of producing and distributing a product or service consists of the *activity costs* on the one hand and *transaction costs* on the other (Reuster, 2010). The activity costs comprise of production costs or the direct costs of inputs and distribution of the product or service (North, 1986). Activity or production costs further include any gains lost as a result of not using the resources to the next best alternative use, that is, their opportunity costs. Transaction costs consist of infrastructure costs such as the occupation costs or investments in the buildings. The transaction costs further include coordination costs, which are both internal and external to the organisation. Internal transaction costs include those related to coordination of work across different units, for recruitment, contracting and monitoring of staff performance as well as the enforcement of contracts. In ART services, these include the central and district level monitoring costs of service delivery. External transaction costs are the costs of dealing with actors external to the firm, including information search, partner search, contracting, monitoring and contract enforcement.

The ART service delivery through public governance will thus reflect a choice based on the perceived differing governance costs among privatisation, regulation (hybrid) and pure public bureaucracy (Spithoven, 2012). The cost function differs depending on the institutional environment, as well as hierarchy type, whether private, public or hybrid

bureau. Following from the consideration of costs, decision makers will undertake a comparative institutional analysis to determine levels of transaction costs under the available governance modes, as noted before. This will reflect in differing transaction curves across the different governance modes, as depicted in the figure below (Figure 16).



* Measured in remediableness terms.

r^0 = low risk r^1 = medium risk r^2 = high risk

Figure 16: Transaction Cost Curves with different Governance modes

Source: Spithoven, 2012

In the graph above, the governance costs are depicted with a remediableness frame in mind, as observed in the forgoing text. When considered in terms of uncertainty terms only, that is, when transaction asset specificity and frequency characteristics are ignored, the different governance mechanisms carry differing levels of risk. The resulting

remediableness related transaction costs will differ according to the type of governance mechanism adopted. Governance costs increase with the level of risk. At each level of risk, the intersection point represents the point of indifference where governance mechanisms can be switched taking into account the comparative governance costs under each governance mode. In Figure 16, r^0 is a low risk point, r^1 is medium risk point while r^2 is a high-risk point. With respect to HIV and AIDS ART governance in Malawi, private provision (privatisation) could represent a governance mode with the highest governance costs in Figure 16, while provision through a public agency (bureaucratisation or own Government provision) represents the transaction cost curve for the most efficient governance mode in remediableness terms. Regulation represents an alternate governance mode, but the relatively high risk means that it is not a preferred option due to the uncertainty that comes with market and other modes other than public provision. More concretely, if exclusively left to private firms and civil society organisations, with the state assuming only a regulatory role, there is the risk that some sections of society may not access ART. This may be due to cost containment measures, or the lack of capacity by civil society organisations, as well as differing policies and aims by donors, such as a focus on a particular population targets only. The state would incur a reputation risk as it would be in breach of the constitutional provisions for the right to healthcare- universal access to health care principles. This in large part explains why the public provision of ART services is the preferred governance mode for ART in Malawi. With public provision, the state is able to secure and allocate resources through the primary health care system to deliver ART services free of charge at service delivery point. Hybrid governance modes, such as public-private partnerships reflected in Service Level Agreements are a further alternate mode. As will further be noted below, the state uses both private and civil society organisations to deliver ART services, while retaining the ultimate responsibility for service provision.

7.3 Distinguishing Production and Transaction Costs in HIV and AIDS ART Governance

In considering the sources of transaction costs, there is need to distinguish between the transaction costs and production costs. The following algorithm (Figure 17 below-next page) helps to distinguish the costs.

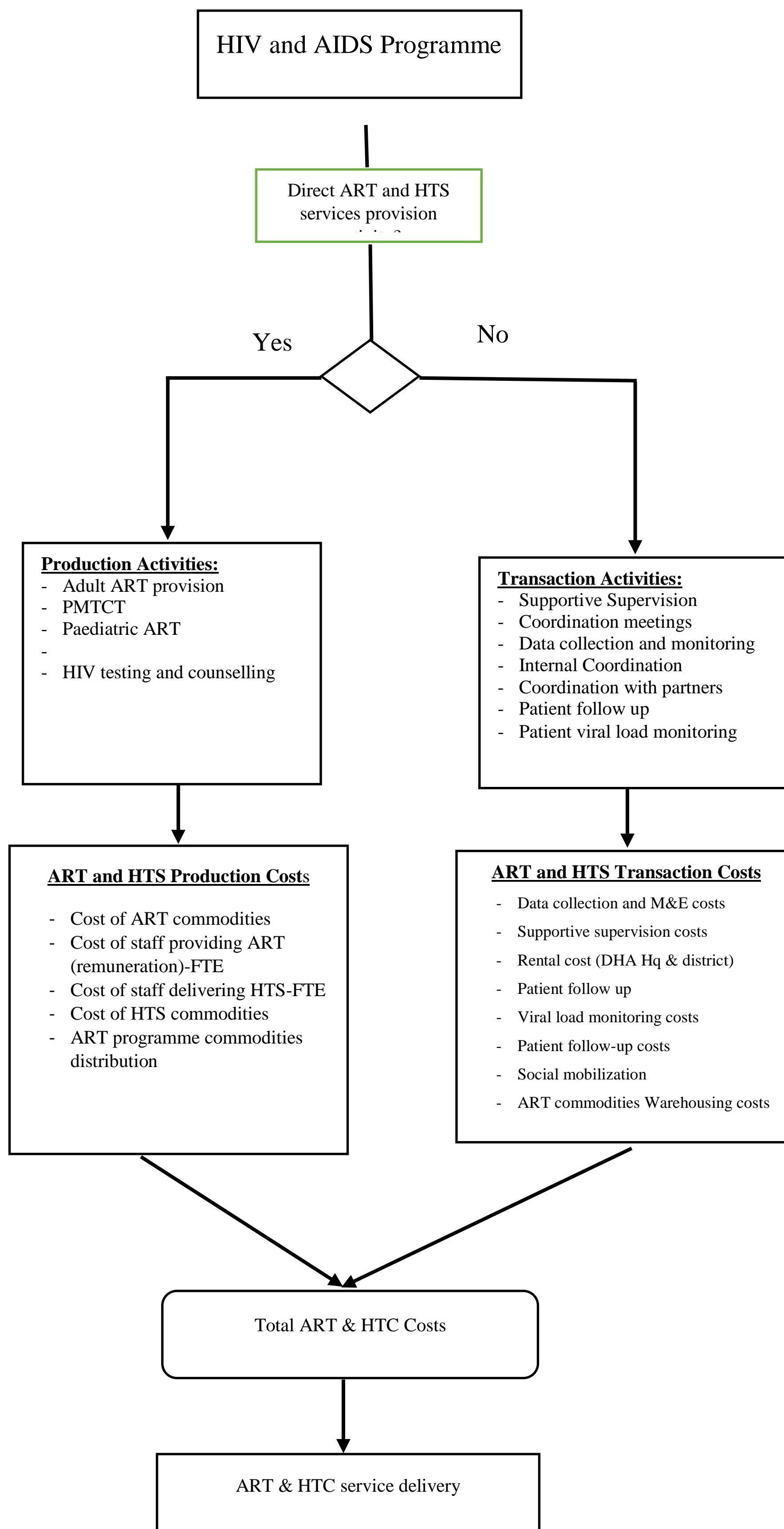


Figure 17: Delineating Production and Transaction Costs in ART
Source (Author)

In Figure 17, a focus on tracking transaction costs views the ART Programme as the firm. This includes the Department of HIV and AIDS and the service delivery points at the national, sub-national (central hospitals), district (district hospitals) and sub-district levels (health centres). For this research, the focus is ART services offered at the different levels of service delivery. To delineate the transaction and production costs, the activities are categorised into direct service provision (production activities) and indirect (transaction activities). The direct service provision activities include ART provision-ART clinics, PMTCT, paediatric ART, PrEP, testing and counselling. The production costs arising from direct service provision (direct costs) include the cost of ART commodities, cost of staff providing ART (remuneration)- full time equivalents (FTE), cost of staff delivering HTS-FTE, cost of HTS commodities. The transaction costs associated with ART service delivery include service delivery supervision costs, data collection and M&E costs, rental cost (DHA headquarters and district), patient follow up, viral load monitoring costs, health facility space and patient follow-up costs. The transaction costs are necessary and contribute to ensuring the set targets and quality standards of the ART programme are met. When viewed from a TCE perspective, the transaction costs represent the costs incurred in coordination, enforcement of contracts (standards) and searching for clients such as finding new ART clients and follow up of defaulters. The search costs extend to education and outreach activities designed to encourage communities and individuals to test for HIV and patients on ART to adhere to the ART.

7.4 Sources of Transaction Costs in HIV and AIDS ART Programme in Malawi

Having considered the basis for governance choice, this section answers the following specific research question:

- *What are the main sources of transaction costs in the current governance of the HIV programme in Malawi? Are these due to non-compliance with the prescription of transaction cost economics theory?*

As noted from the foregoing text, the characteristics of transactions are a key determinant of the types of institutional arrangements that emerge to coordinate transactions. With particular reference to the ART service provision, it is important to analyse these characteristics by looking at the ART service delivery process. Within this analysis, the application of TCE is facilitated by a description of the ART services and

the implications of completing these activities on transaction costs. In the context of the fight against HIV and AIDS, strategic choices are made on how to govern service delivery to meet the aims of access to ART services. Hence, the public provision of the ART programme in Malawi should be viewed in the context of sovereign or public goods problem, for which public governance through the public bureau is deemed the most efficient means over a range of alternatives. The public goods problem pertains to under production, over-use and degradation, where, if left to the spontaneous actions of individuals or private actors, it will not be produced (Chen, 2021).

The following text describes the ART services in the context of the key transaction characteristics of frequency, asset specificity and uncertainty, and how these contribute to the generation of transaction costs. In addition, the financing of ART, as well as supply chain management are analysed through the TCE lens to unveil the underlying transaction characteristics that generate transaction costs. It is shown that both ART financing institutional arrangements can be viewed from a contract lens, with the governance of the services characterised by rights and responsibilities for decision making. It is further shown that the governance of ART services is shaped by the characteristics of the transactions and their alignment with organisational modes that secure efficient delivery of the services. A description of the HIV and AIDS ART programme funding architecture and activities will help to characterise the nature of service and funding transactions.

7.4.1 Transaction Costs from Aid Administration for the Malawi ART Programme

The ART programme is the largest component of the HIV and AIDS response in Malawi. It accounts for over 70 percent of the financial resources earmarked for the response. The main financiers of the response in Malawi are the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFTAM) and the President's Preparedness for the Fight against AIDS Response (PEPFAR) who jointly accounted for more than 85% of the financial resources in the period between 2018 and 2022 (GoM, 2021). As noted previously, for the fiscal year 2018 and 2019, the GFTAM provided the largest funding share for HIV and AIDS at 54%, while the PEPFAR contributed 18% of the AIDS funding in the same year. An analysis of the funding structure and the fiduciary mechanisms shows that the TCE theory can be applied to both the funding

and the programmatic elements of the ART programme. The ART programme and other aspects of funding are governed by the rules set up by the financing entities. In this context, the governance of ART programme financing is viewed as a choice of contractual relationships between the financiers of the HIV and AIDS response in development assistance for health (DAH) on HIV and AIDS and the Government of Malawi. This contractual relationship embeds and enforces a collection of decision rights and responsibilities that govern ART programme financing. More concretely, the financing agreements or memoranda of cooperation between the funding agencies such as the GFTAM, PEPFAR and other agencies that support the ART programme represent the governance mechanisms that facilitate and constrain the actions of the partners in the receipt and utilisation of HIV and AIDS DAH. These rules are mainly designed to protect the funding products of the financing agency, although the ostensible objective is reaching clients with HIV and AIDS services.

From the perspective of the funders, there are probity and reputation hazards that can potentially afflict the funding agencies when resources are transferred to the recipients of the grants earmarked for HIV and AIDS ART programmes. These hazards stem from the behavioural assumptions under the TCE theory, namely, the opportunism and bounded rationality of the actors. For instance, due to bounded rationality, the policy and programme planners may not be able to fully integrate in the initial ART programme planning the sources of future disruptions, or may indeed underestimate certain factors that cause inefficiencies in service delivery. Probity hazards may stem from the recipient government's representatives behaviour and actions in the application of the resources and delivery of ART services, including waste through corruption. To mitigate against the effects of these hazards, the funding agencies deploy safeguards, including audits, establishment of a Financing Agency, channelling funds through private sector agencies and requirements to have a dedicated unit within the recipient agency with fiduciary and programme coordination roles. These measures entail costs in support of proper use of funds for the ART programme implementation, but add to the cost of delivering aid for ART programme. The following sections discuss funding and administration and transaction cost implications for two key HIV and AIDS funding agencies in Malawi to show the sources of transaction costs.

a) *GFTAM Funding*

The GFTAM is explicit in its funding approach in insisting that programmes funded are state-led and that the programming reflects the participation of all key country stakeholders (Triponel, 2009). Hence, for the GFTAM funding, proposals are prepared by recipient countries following the GFTAM guidelines. Once approved, the funds have to be managed using financing arrangements agreed between the GFTAM and the Malawi Government. Further, a number of safeguards to guarantee fiduciary compliance are reflected in the institutional and implementation arrangements for administering the funds. As noted in Chapter 6, there is a requirement to set up a Project Implementation Unit (PIU) under the Principal Recipient of the GFTAM grant. As of 2018/ 19 through 2023, the GFTAM Principal Recipient is the Ministry of Health, which is further responsible for the planning and management of the ART programme. A PIU was thus set up in the MoH, complete with a staff contingent comprising of a Programme Planning and Coordination and Finance and Administration Department. The PIU is housed within the MoH Department of HIV and AIDS (DHA). The purpose of the PIU is to ensure oversight of ART programme for GFTAM funding. Its functions include receipt and verification of activity funding requests from MoH programme in the implementation of ART activities, including drug and commodities acquisition. Its functions extend to the monitoring of the ART activities. It further offers technical assistance to the MoH programme units implementing ART. As PR for GFTAM, the MoH through PIU further provides support to World Vision International, a non-governmental organisation, that coordinates the non-biomedical aspects of the GFTAM funding. However, the focus of the current research is the ART programme, and hence, discussion will be confined to the ART programme. In TCE terms, the PIU is a safeguard instrument to secure ART programming is consistent with GFTAM and MoH standards and fiduciary requirements. To the extent that the PIU attracts resources to enable it function, these resources are part of the overall transaction and management costs that are likely to affect levels of transaction costs and the efficiency of ART service delivery.

In addition, an independent Financing Agency (FA) is retained to vet all activity funding requests submitted to it by the PIU before they are funded. This is another fiduciary compliance mechanism under GFTAM, and is an investment in safeguards to secure

fiduciary compliance with GFTAM funding procedures. Additionally, annual financial audits of the application of the GFTAM Grant are conducted by independent auditors. This further attracts resources and has implications on the total cost of ART programme.

As may be observed from the foregoing text, the deployment of safeguards adds to the costs of delivering DAH for the ART programme in Malawi. While the current study focuses on the facility based ART service delivery to analyse transaction costs and efficiency, it is argued that the financing of the programme should not be divorced from the programme itself as the safeguards attract resources that constitute transaction costs and add to the unit costs of ART delivery. The implication is that arrangements where the coordination unit with funding responsibilities is part of the routine government bureaucracy, transaction cost reduction may be facilitated. Moreover, an independent Financial Management Agency adds to the layers of safeguards and hence transaction costs in ART service delivery. Thus, although the inclusion of a standalone PIU and independent FMA is calculated to reduce some of the hazards related to aid delivery, it may also be construed as a gap in the pursuit of cost minimisation via the application of TCE theory.

b) *PEPFAR Funding Arrangements*

The PEPFAR follows a different funding arrangement, with a major feature of the funding support to ART being the presence of private contractors. It is instructive that the channelling of PEPFAR funding is not largely through the public bureaucracy, but rather a different type of bureaucracy that runs parallel to the public bureaucracy, although ultimately, the focus for both is reduced HIV and AIDS infections and AIDS related deaths in the population. The caveat is that although funding is not directly through government departments, PEPFAR and the Centre for Disease Control (CDC) finance staffing, equipment and research that benefits the public sector HIV and AIDS service delivery. The contractors set up own offices in the country, with most located in Lilongwe. Agencies such as Family Health Services (formerly Population Services International) and Elizabeth Glazer Paediatric Foundation (EGPAF) are cases in point. The EGPAF in particular sub-grants to mother2 mother (m2m) to implement a range of activities that include the financing of "volunteers" or "Expert Clients" that assist with ART records, support for people on ART and follow up of ART defaulting clients.

The above PEPFAR support or financing model implies that there is a component of transaction costs that should be apportioned to the ART programme through central level monitoring and technical assistance costs incurred by the implementing contractors. Additionally, the ART ancillary costs such as payment for Expert Clients that has a bearing on the success of the programme constitute transaction costs. These technical assistance and headquarters monitoring costs likely add to coordination and management costs, with implications on the overall technical efficiency of ART programme. Nevertheless, the limited tracking of these costs suggests that the magnitude of transaction costs is not adequately monitored to trigger ways of containing them using the TCE theory. This has implication on the efficiency of the ART programme as high transaction cost are likely to impinge on cost minimisation, and hence, efficient ART service delivery.

7.4.2 Transaction Costs from HIV and AIDS Programme activities

To further understand the transaction costs associated with the ART programme, it is important to describe its nature and content with respect to the transactions involved. The programme largely consists of a continuum of care reflected in diagnosis and treatment services.

a) HIV Testing Services

The HIV testing services (HTS) are an entry point into the ART programme for clients. It is thus a key aspect of the continuum of care under ART. The primary point of access to HIV Testing Services is the health facility, although the services are further provided outside of the health facility during community outreach clinics under community health services, or other promotional activities. As of 2019, trained nurses and mid-wife technicians, as well as amateur HTC providers carried out HTS. The amateur HTS providers included Health Surveillance Assistants, and in some instances, ground labourers trained in HTS. To qualify for training as an amateur provider, the staff had to possess a Malawi School Certificate of Education, the equivalent of Ordinary Level qualification. A further cadre of HTS providers called Health Diagnostic Assistants (HDAs) were deployed in some health units by the Ministry of Health with Global Fund support, although this was a non-established or temporary position in the Ministry of Health. Providers deliver HTS on a daily basis based on demand by clients. Furthermore, for women attending ante-natal care (ANC) services, an HIV test is mandatory. These

tests are conducted on the first ANC visit if they have not had a recent test. Because the HTS are facility based, they attract resources such as space which form part of transaction costs.

The HTS include a counselling component before and after undergoing the HIV test. Two tests are usually administered. For a first test positive result, there is a follow up confirmatory test. Once a positive test is confirmed after the second test, the client is referred to and immediately initiated on ARVs. There is the further possibility of an inconclusive positive outcome where the two tests produce contradicting results. For a negative test result, the client is informed of the outcome and no further tests are conducted to reconfirm the negative status. The services are provided in dedicated rooms, although in some instances, the same rooms are used for other health care services. The research found that a key challenge in the provision of HTS services is inadequate space due to limited infrastructure within the health units. In some instances, this challenge affected the extent to which privacy in the provision of services could be achieved and maintained. With regard to transaction costs, the research found that both the staff time and space (occupancy) for these services attract economic costs that should be accounted for when monitoring transaction costs. Further, the creation of demand for testing in outreach clinics and other community initiatives contributed to transaction costs in HTS. The research did not find evidence that these transaction costs were monitored, especially at the facility level, so as to inform cost control per TCE theory. This has a bearing on service provision efficiency, and implies the need for cost monitoring arising from activities that support HTS.

b) HIV and AIDS Treatment Services

The ART services have been decentralised from the District Hospitals (secondary care) to Health Centres (primary care), and hence, from doctor-based to technician-based provision since 2006. This decentralisation implies the need for strong coordination to maintain standards. The providers at the health centres include clinicians and Medical Assistants. While most will have general knowledge with respect to care from their in-service training, they require specific training and skills with certification in ART service delivery (Harries, *et al*., 2016). The training covers the HIV clinical guidelines that the DHA in the MoH prepares and updates periodically. These guidelines in turn follow the WHO Clinical Guidelines on HIV prevention and treatment. As observed

before, there are expert clients that assist with the organisation of clients as well as sorting out patient records and preparation of the patients before accessing the ART services in some health units. In TCE terms, the bespoke skills and experience of providers in the provision of ART become asset specific and therefore need to be secured in the long term. This is because disturbances such as attrition will require hiring and retraining of ART providers. Further, the involvement of volunteers such as Expert Clients (paid or not paid) has implications on transaction levels.

Except for a few health facilities (largely private or not profit) that have integrated ART services with other general primary health care, the ART services are offered on dedicated ART clinic days in a week. Space used for other services are the same ones used for ART, while in a few sites there are structures dedicated for ART clinics, and thus imply transaction costs in terms of occupancy. The services are usually conducted once a week for different cohorts of clients. The cohorts are categorised into two broad categories that consist of the newly initiated who have shorter drug replenishment periods (three months), and the more stable clients who have relatively longer replenishment periods. The newly initiated clients are on shorter replenishment periods because of the need to closely monitor their adherence, and the uncertainty of whether or not they will stay on treatment. The more stable clients are unlikely to default on the medication having experienced its benefits. The research found that the cohort framework permitted efficiency in the provision of the services given the extreme human and infrastructure resource constraints in primary healthcare in the country. Further, participants to the research at service provision level indicated that the clients do not need to make repeated visits to health care points, thereby reducing transport and other costs associated with accessing ART services. Nonetheless, the research found that there were challenges related to clients' long waiting time before they are able to access a service. In addition, long distances to the nearest service point for some clients mean high costs of accessing the ART services due to transportation and food costs.

7.4.2.1 Types and Sources of Transaction Costs in the ART Programme

The ART services play a vital role in the fight against HIV and AIDS, particularly with regard to prevention and management of the epidemic. Yet, the research found that their management and coordination attract significant resource that may negatively impact their efficiency. These costs relate to information search, contracting, monitoring and

compliance, coordination of actors at the national and sub-national level and uncertainty and risk costs. These costs are outlined below to highlight how they are generated and why they should be taken into account and monitored in ART service delivery. This monitoring will aid the programme to gauge their levels and minimize them to contribute towards efficient service delivery, particularly when resources are constrained. Hence, a reasonable expectation is that the transaction costs would be periodically evaluated to ensure they remain within levels that contribute to efficient service delivery. Although the NASA is currently the main framework used to track expenditure, it does not explicitly focus on transaction cost monitoring to inform efficiency considerations in ART in the planning and delivery of ART services.

Noting the core ART service delivery as part of the overall transactions (production transactions), the research identified a number of areas as sources of governance or transaction costs, as will be describe below (Table 12).

Table 12: Transaction Sources in ART Governance

#	Transaction	Transaction Cost elements	Location	Monitoring Responsibility
1	Search & Information	Preparation of ART funding proposals Assessment of health facility eligibility to provide ART Follow up on ART non-adherent clients	National, district, community	MoH DHA National AIDS Commission
2	Contracting- including the cost of negotiation with new providers/suppliers	Staff recruitment Procurement Agent recruitment Contract monitoring	National	MoH DHA
3	Monitoring and evaluation costs	Data collection and analysis Management Information System costs (DHAMIS)	National District	MoH DHA National AIDS Commission District Health Office (DHO)
4	ART coordination	Headquarters costs (office, personnel, financial audits) Technical Working Groups (Treatment, Care & Support) at national and district levels	National District	National AIDS Commission MoH DHA MoH Planning Department
5	Supportive supervision and quality monitoring	Quarterly supportive supervision missions logistics Staff time Coordination of supervision missions across agencies	National District	MoH DHA
6	ART Supply chain	Third party Procurement Agent recruitment Warehousing Procurement Agent Contract monitoring Parallel ART supply chain governance	National	National AIDS Commission MoH DHA

#	Transaction	Transaction Cost elements	Location	Monitoring Responsibility
7	ART financing fiduciary risk safeguards	Project implementation units, financial management agencies, external audits	Principal Recipients, External units set up and retained by funders	Donors and development partners
8	Occupancy costs	Headquarters rentals, facility ART service provision spaces	National, sub-national (referral heath units), district	Government (MoH DHA, NAC, health facilities), Private health units, NGOs (donor contractor with national or district offices supporting ART service delivery)

Source: Key Informant Interviews and literature review

a) Search and Information Costs in ART and HTS

One source of transaction costs as explained under TCE are search and information costs (Tadelis &Williamson, 2012). The research found that there are activities that constitute search and information that are essential to the provision of these services. These search and information costs are generated in a number of ways. To begin with, and with specific reference to ART clients, the Ministry of Health and partners spend extensive time and financial resources to promote HIV testing and getting people onto the ART if found sero-positive. The promotion and advocacy costs represent transaction costs that may affect the efficiency of ART as they add to the service provision input costs. To this end, the research found that the various campaigns and community mobilisation efforts for HIV testing and promotion of ART as both means for addressing the individual health of clients and as a prevention strategy constitute essential activities, but generate transaction costs in the ART programme. These costs may be viewed as internal transaction costs to the ART programme (Victor & Paulo, 2023).

The research further found that search costs are incurred when looking for partners to implement the ART services. These costs are due to the information asymmetry problem where the MoH, which has the primary responsibility for the provision of health and ART services, but does not have the full information about the agencies that would be engaged to provide the services. Due to this information asymmetry, the MoH through DHOs invest in staff time and financial resources to gather information that becomes the basis for decisions about contracting out the ART services to third party agencies such as CHAM and other civil society organisations. This point is best illustrated by the processes that go into the confirmation of non-state health units to provide ART services even when these already provide essential health services under

the SLA framework. More concretely, the Ministry of Health via the District Health Offices (DHOs) will engage in due diligence exercises aimed at confirmation of the suitability of health units to provide ART services. These activities generate transaction costs.

In the context of financing ART, the proposal development process and completion of agreements between the Government of Malawi and partners further represent transaction activities. As noted above, these costs include staff time and financial resources deployed to prepare funding proposals by the recipient governments, technical assistance provided by funding agencies in the preparation of these proposals and the coordination efforts at country level to complete the preparation of the proposals in a participatory fashion (Triponel, 2009). An example is the preparation of country support strategies by funding partners that involves search for information through consultations to enable realistic assessments and prioritisation of activities. Further, at recipient country level, the research found that broad based consultations that include the hiring of technical assistance, coordination across a broad range of stakeholders are undertaken when completing the preparation of funding proposals. These activities are not directly linked to ART service provision, but add to the cost of service provision. Moreover, periodic technical support by funding partners entail costs that are linked to the ART programme. The research found that these transaction costs are incurred prior to the implementation of the ART services but should be accounted for in the final efficiency analysis of the ART service delivery. In addition, Research and Development activities in ART are important. The research found that while these activities are undertaken to improve both the drugs and models of service delivery, they also constitute elements that generate transaction costs. The search for effective interventions and monitoring attracts resources that have a bearing on the total costs of the ART programme. This has further implications on the efficiency of service delivery. Thus, while the costs are beneficial to the ART programme, limited monitoring of their magnitude in service delivery suggests inadequacies in TCE application as the incidence of overall transactions costs may not be fully appreciated to inform programme design and resource allocation.

b) *Contracting Costs*

Contracting costs are the costs incurred when concluding contracts with various agencies. The process of development and maintenance of contracts with a number of agencies may be the source of significant administrative costs that have a bearing on the level of transaction costs in ART (Williamson, 1985). These include time and financial resources used negotiating contracts before the commencement of the activity. In the context of the Malawi ART programme, the research found that contracts have to be concluded between the Ministry of Finance and a range of partners. First, agreements with funding agencies are concluded between the Government of Malawi through the Ministry of Finance and the funding partners. Second, Service Level Agreements with agencies are prepared and concluded between the Government of Malawi via the Ministry of Health and non-state actors such as CHAM and private hospitals to deliver primary healthcare and ART services. These entail costs of negotiation involving time and staff in the drafting and negotiation of the agreements. The costs extend to those for facility assessments towards certification for ART service delivery. Third, and in the context of the logistics to ensure ART commodities reach the intended beneficiaries, third party logistics agencies such as Bolloré Logistics are engaged to acquire, ship, store and distribute the drugs and test kits in-country. This yields storage or warehousing costs that should be accounted for as transaction costs in the ART programme. Fourth, at the firm or agency level, recruitment of staff and the conclusion of contracts with them further requires resources that have a bearing on the level of transaction costs in ART. Although healthcare workers under MoH are not recruited with the sole purpose of ART service provision, they are relevant for the analysis of transaction costs since ART services are embedded within the broader health care establishment costs. For instance, the MoH lists various cadres of staff that are involved in the provision of ART services, whose recruitment, contracting and orientation require time and resources. Fifth, although funding is available through agencies such as the GFTAM, the costs of negotiating contracts with pharmaceutical companies for the supply of drugs are non-trivial and need to be accounted for in the analysis of ART service delivery.

c) *Monitoring and Compliance Costs*

The provision of ART services involves intensive monitoring to ensure compliance with quality of care standards, and to account for the progress registered in the programme. The monitoring activities include quarterly data collection visits, quarterly supportive

supervision, mentoring, and training on new guidelines when modifications have been effected. Support for ART is further provided by non-state agencies that assist service providers provide quality services in some districts, as observed below by a research participant:

> *"And then, mostly in the south we are in Blantyre, Zomba, Thyolo, Chiradzulu, Mwanza and Neno. And so, in all of these districts, we are supporting the district teams to achieve 95 95 95 targets. So, to support HIV testing, we work with them to make sure everybody's trained on the new revised guidelines. We support them to implement some of the new interventions like index testing, self-testing, the contact tracing that we have been doing for index testing. and we do this in two ways. One is actually doing training and mentorship. The second is we work with… Lighthouse and Macro and have actually employed HIV Diagnostic Assistants who actually provide these services.* Respondent 5, National Programme Manager, CSO

As illustrated by the above quote, the research found that there are likely to be differential levels of transaction costs by facility owing to the additional support that some districts and facilities receive from partners, with implications on both efficiency and quality of ART provision. With regard to efficiency, the differing support levels may further imply that while some health facilities that provide ART can demonstrate efficiency due to the support from partners, others are likely to demonstrate inefficiency.

Furthermore, using Malawi as an example, the research found that the costs of monitoring and compliance appear to be significant. Quarterly data collection visits to each health unit that provides ART services are undertaken. These visits are organised centrally and require significant financial and human resources. In addition, a dedicated unit for data and monitoring and evaluation is part of the monitoring framework, including the existence of an HIV and AIDS Management Information System (DHAMIS) managed by the Department of HIV and AIDS in the MoH (GoM, 2022). Further, electronic data management is contracted out to a third-party agency which implies further transaction costs. As of the time of the study, this role was performed by Boabab Health, and subsequently by EGPAF. Since the development of information systems, data management protocols, and data quality assurance incur transaction costs,

it is important to monitor these costs so as to inform measures to optimise them and contribute to more efficient use of the resources.

Although actual estimates are not available to gauge the magnitude of quarterly data collection and supervision visits costs by MoH and partners that focuses on ART services, there is a perception among both primary health care policy and programme managers that the amount is non-trivial. The investment in resources for monitoring shows in the relatively high-quality data that has come to be associated with the HIV and AIDS programme in the country. The monitoring of these compliance costs is captured as governance and sustainability cost in some monitoring reports[19], although the monitoring costs are not clearly delineated. Nevertheless, in the face of global declining resources, these costs need to be closely monitored so that they can inform the crafting of measures that support efficiency of data gathering and analysis, including the need to consider harnessing advances in technology such as distributed ledger technology (DLT), including block chain technology in information management (Kamau, et al., 2018; Seed, *et al.*, 2022).

d) *Supportive Supervision for Service Quality Monitoring*

In addition to the quarterly data collection, there are quarterly supportive supervision missions for ART services to each ART facility. The compliance monitoring that is undertaken is aimed at mitigating opportunism and enforcement of probity in service delivery. In essence, the supervision visits are quality compliance enforcement mechanisms that involve onsite observation of service delivery, mentoring support to ART service providers and recognition of performance by service providers through issuance of certificates of excellence. The supervision visits are thus critical to the maintenance of minimum ART service quality standards in the country. The standards are defined in the Rapid Testing Guidelines and comprehensive HIV Clinical Guidelines (GoM, 2018). In addition, various partners that include EGPAF, mother2mother, Médecins Sans Frontières (MSF) deploy teams of staff and technical assistants to mentor and monitor compliance to standards.

[19] For instance the annual Global AIDS Monitoring Report prepared by the Malawi National AIDS Commission

Another aspect of monitoring and compliance that occurs in ART services is adherence to treatment promotion, or what is popularly called follow up on clients that have defaulted on treatment. Defaulters are reported to be clients that have missed two months of their appointment without providing information. Because non-adherence to treatment regimens can lead to treatment failure, drug resistance, and increased healthcare costs (Shukla, et al., 2016), arrangements are made at facility level to ensure defaulters are followed up and brought back to care. Teams of staff are deployed to follow up and bring the concerned clients back to treatment. The follow up team usually consist frontline staff and community volunteers variously called mentor mothers and expert clients. The research found that there are costs (both financial and economic) attached to the follow ups. This is evidenced by the following quotes:

> *Interviewer: "Okay....you talked about, ... expert clients… are these paid staff or they are volunteers?"*
>
> *Respondent: "No, they're fully paid. So, we work with mothers to mothers, an organisation with headquarters from South Africa, who work with mentor mothers and expert clients, who are fully paid, contracted staff. I think they are paid around 70,000 Kwacha per month" Respondent 1, National Programme Manager*

As observed from the quote above, these follow up activities entail resources in human resources, time and funds, and thus have implications on the efficiency of service provision. These costs need to be monitored and accounted for in the ART efficiency. Nonetheless, there did not appear to be a systematic monitoring framework for these costs within the national programme, and the need for instituting a cost tracking mechanism to capture these costs is implied.

Monitoring patient adherence, clinical outcomes, and programme performance across decentralised facilities further require data collection, analysis, and reporting. To monitor toxicity of drugs, laboratory tests are periodically carried out on clients. At the facility level, blood samples are taken and sent for laboratory testing before results can be delivered to the client. This process involves time and financial resources that add to the total ART cost, and hence increase unit costs of the services. The implication is that

such costs should be actively brought into the ART service efficiency analysis framework.

e) *Coordination Costs*

Provision of ART services involves high- level coordination of partners supporting the programme. The research observed some coordination mechanisms in ART provision aimed at mitigating some of the challenges that come with complexity of service delivery, although they have transaction cost implications. Interviews and observations of ART clinics at facility level revealed that there were a number of activities to coordinate service delivery. Coordinating appointments, laboratory testing, counselling sessions, and results delivery involves administrative costs. These coordination activities are essential to ART, although their costs are not routinely tracked. Front line staff, nurses, providers and volunteers spend significant amounts of time organising client files. Furthermore, the coordination of client appointments requires a lot of time. Based on the observed facility level processes during ART clinics, staff further spend time counselling clients, first as groups of individuals, especially focused on promotion of adherence to treatment regimens, and secondly, as individual clients.

Coordinating testing and counselling services across different healthcare providers or public health agencies can also lead to coordination challenges. Consequently, the research found that there were financial resources and time are invested to secure effective coordination of the services. The costs of such coordination are non-zero and need to be adequately monitored and addressed. As noted before, these services are provided by public, private- not- for profit and private- for- profit agencies. They are further provided across geographically disparate sites across the country and, therefore, render the provision of these services complex. This requires deliberate mechanisms for coordination to ensure a measure of uniformity in the application of minimum testing and care standards. The quarterly supervision missions observed above, and district level coordination through the District Health Offices are important for the mitigation of challenges that come with ART service complexity, but they have resource implications that should be tracked and accounted for in ART service delivery, and how they impact efficiency.

At the national and district level, a number of committees have been set up as part of the HIV and AIDS programme coordination. More particular to ART, there is a Treatment, Care and Support Technical Working Group (TWG) that brings together key players at the policy and programme management level to coordinate the activities. The TWG is expected to meet once in every quarter and is designed to facilitate cross sector and agency collaboration. This coordination is important for ART service delivery as the group reviews coordination challenges, emerging issues and crafts strategies to improve delivery of the services. However, the convening of the meetings requires resources for meeting venues and associated costs. Furthermore, staff time incurred through participation in the TWGs represents an economic cost that is rarely tracked and acknowledged in resource expenditure tracking. This implies that these costs should be documented and accounted for in the overall ART service provision for a more complete understanding of the magnitude of transaction costs and how they impact efficiency in service delivery.

f) *Supply Chain Transaction Costs*

Transaction costs related to the supply chain in ART arise from the need to acquire, store and distribute antiretroviral drugs and testing supplies to decentralized facilities. Inventory management and coordination among multiple facilities thus increase transaction costs within the supply chain. While distribution costs are counted as activity or production costs (North, 1986), those related to acquisition of drugs, testing supplies and storage constitute transaction costs. Hence, the governance of the ART services in Malawi needs to be further considered in the context of the logistics that make commodities available where they are needed. This concerns the acquisition, warehousing and distribution of the commodities. These processes produce both production and non-production costs. The warehousing and management fees for the procurement agencies involved contribute to transaction costs. These costs may have a huge bearing on efficient service delivery. As further noted in Chapter 6 the programme has deployed measures to secure efficient procurement of drugs through pooled regional bids, but the efficiency gains from these measures need to be evaluated. Nevertheless, the research observed a parallel procurement management arrangement which represents a good case note in the less than efficient transaction cost management in ART drugs and that constitute the bulk of the resources in the HIV and AIDS response (Wild & Cammack, 2013)-*see next section for details*. The main entities involved in the parallel

supply chain management of the ART programme include the Government of Malawi/GFTAM through Bolloré Logistics, the Central Medical Stores Trust (mainly for antibiotics for treating opportunistic infections), the International Development Association (IDA) and UNICEF.

The parallel ART supply chain arrangements are illustrated in Figure 18 below:

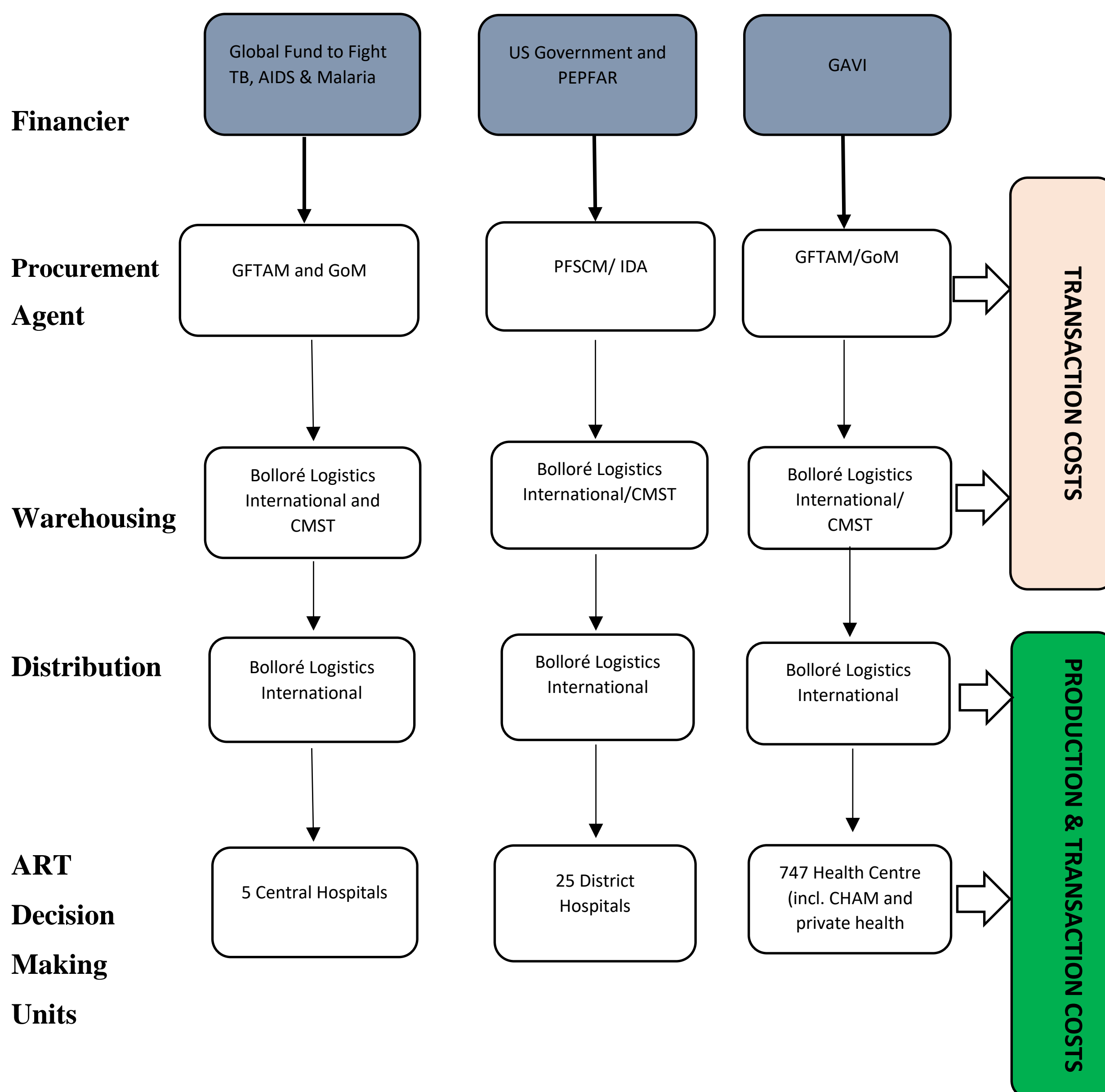


Figure 18: Illustration of Transaction costs in ART Supply Chain process

Source: Adaptation from Central Medical Stores Trust (2014)

As may be observed from the above schematic (Figure18), the research found that the ART commodity supply chain management has several implications for transaction costs and efficiency. First, and as noted above, the conclusion and management of contracts with third party agencies for supply chain management generates search and information costs. These costs contribute to the level of transaction costs in ART services. With

regard to efficiency, the long tenured nature of the contracts with third party private procurement agencies is likely to bring about the small numbers and lock-in effects associated with such arrangements. Thus, while the initial search for suitable logistics and supply chain agencies was based on competition, retention of one provider over a long period may introduce lock-in costs and hence inefficiency. The relationship eventually becomes such that a monopoly situation develops with one service provider negotiating further extensions of the contract.

Second, warehousing costs are incurred as part of the commodity and drugs management for ART. The HIV and AIDS commodity storage and distribution services are offered by Bolloré Logistics, although until June 2019, Bolloré Logistics used the Central Medical Stores Trust (CMST) warehouses. A 10 percent levy was applied to the cost of the drugs and other commodities for ART services. Hence, all the warehousing costs contribute to total transaction costs and may thus impact efficiency in ART service delivery.

Third, the CMST is the official Government agency for procurement, storage and distribution of drugs. Nevertheless, it does not participate in the acquisition of the ART commodities, bar antibiotics for opportunistic infections that are also used for general treatment. This role implies transaction costs as the cost of the service is charged to District Health Offices, and thus contributes to the overall costs for the services. Nevertheless, with respect to ART commodities, the research found that this role has been left with the MoH DHA, although until July 2019, the ART commodity supply chain relied on the CMST storage facilities. This was due to limited capacity for storage elsewhere in the country.

Fourth, an important aspect of supply chain related to the generation of transaction costs is that the architecture currently consists of parallel supply chain arrangements to the CMST, as reported in Chapter 6. A number of agencies, including UNICEF, Chemonics International (PEPFAR and CDC) are involved in the quantification and procurement of some drugs and commodities used in ART services. Participants to the study indicated that this set up reflects a duplication of structures and, hence, increases transaction costs related to the administration, storage, staffing and occupancy (offices). The reasons for this set up are historical, and have to do with the reduced confidence in the CSMT by the funding agencies.

From the foregoing text, the supply chain management and storage of commodities implies transaction costs in the ART service delivery. Further, because the CMST already has infrastructure to manage the entire health commodities supply chain system, the costs of using the parallel structure imply some transaction costs may be saved through a more harmonised ART supply chain framework. The need for improvement of the image and capacity of the CMST has been recognised and a reform programme has been under implementation since 2015. Clear areas of improvement have been outlined with clear performance indicators that are periodically reviewed with funding agencies. Interviews with the CMST showed that the performance indicators were improving, although a formal assessment had not been undertaken as of the time of the research. The cost of managing contracts (time, monitoring costs), with Bolloré Logistics also contribute to transaction costs.

g) Transaction costs from Donor Agencies and NGO operations

When the role of various participants in the ART programme is considered, the transaction costs due to administration, coordination and monitoring may actually be higher than what is reported. For instance, all agencies that sub-granted or technically supported NGOs that carried out activities related to ART were potential candidates as transaction entities. In this connection, participants to the research noted that there were numerous agencies that have set up own offices in the country with implications on transaction costs when the totality of the effort was accounted for. The following sentiment captures this perception:

> *"As you know, there are many agencies involved in HIV and AIDS in the country focusing on one aspect or the other. So, there are likely to be administrative costs associated with office space and other costs. Perhaps it may be useful to identify local agencies that are already on the ground instead of every organisation coming to set up their own offices. But we understand that there may be issues of capacity, as well as mistrust regarding whether funds would be used for intended purposes". Respondent 9, Programme Manager*

As can be noted from the above quote, although not explicit, one would discern elements related to assumptions of how actors decide the type of organisation to effect

transactions with, in this case HIV and AIDS. For instance, decisions to set up offices in country rather than work with local agencies may be driven by perceptions of a specialised set of skills, culture and attitudes to effectively deliver the services which may be construed as asset specific, a key assumption under TCE. Second, perceptions of mistrust reflect notions of opportunism, and the possibility of misuse of funds by local agencies due to interests other than those originally intended by the financing agencies.

Furthermore, non-state agencies that support selected health facilities incur transaction costs that ought to be accounted for as they provide funding and technical support. These costs are incurred for data collection and support for follow up of defaulting clients. The research found that in some districts and health facilities, there were dedicated teams whose role was to follow up defaulting clients. For instance, in Nsanje District, a follow up team of professional staff and volunteers, complete with resources that include transport is part of the framework for ART service delivery. In TCE terms, this is akin to looking for exchange partners, and, therefore, constitutes transaction costs. In yet other districts where these follow up teams are not present, Expert Clients are used and a range of other volunteers in committees at the community level. In some instances, they have been provided with bicycles and a lunch allowance for each follow up visit that they undertake. Nevertheless, the research found that this was not consistent across all districts and facilities which made estimation of these costs somewhat difficult. It was not readily evident that these costs are tracked and included as part of the ART service provision costs in the NASA. Even where they are added or mentioned, it is at a highly aggregated level, and also not explicitly labelled as transaction costs.

In summary, the transaction costs in the Malawi HIV and AIDS ART programme stem from two broad sources, a) development aid administration for the ART programme, and, b) ART programme operations, including healthcare and ART supply chain governance. Development aid administration involves due diligence and the establishment of safeguards such are establishment of an independent Project Implementation Unit, Financial Management Agency, and a requirement to conduct annual financial audits by the Government and the funding agencies. Furthermore, to avoid the fiduciary risks associated with financing public agencies in Malawi, some funding agencies channel funding via NGOs and at times private sector agencies. Transaction costs from the ART programme coordination are occasioned by a number of sources. First, promotion of aspects such as HTC and adherence to drugs by patients that

require substantial resources, including for follow up of defaulting clients. Second, contracting costs for different actors involved in ART service provision besides the MoH. This involves due diligence, for example, when determining the appropriateness of ART service provision by a private health unit, there is due diligence before the signing of a service level agreement. Third, the MoH and partners undertake compliance and monitoring costs that involves human and financial resources. Fourth, the MoH undertakes quarterly supportive supervision visits to each facility to mentor staff providing ART services, as well as to collect data. This further involves significant amounts of resources. Fifth, at both the central and district levels, there are coordination forums designed to share progress and resolution of policy as well as operational challenges and learning. They further entail human and financial resources. Finally, under ART programme transaction costs, there are costs associated with supply chain management for ART. These include the contracting and management costs of the third party agent that handles the acquisition and distribution of drugs and commodities. The transaction costs extend to warehousing costs.

7.4.3 HIV and AIDS ART transactions and TCE theory prescriptions

From the forgoing text, it is clear that there are a number of sources of transaction costs in the HIV and AIDS ART governance in Malawi. As observed in Chapter 6, the design and implementation of the ART governance largely conforms with TCE theory. Nevertheless, the research also found some deviations from the prescription of transaction cost economics theory. Firstly, the HIV and AIDS ART aid management or finance appears to consist of an excessive design of safeguards. While these safeguards are predicted by TCE theory (enforcement of post contract performance and compliance), they have led to the setting up of a number of fiduciary and administrative structures that drive transaction costs in ART. The setting up of PIU and independent financial management agency by the GTFAM, the largest financier of the ART programme is in evidence. In the same connection, rather than channel resources to support the ART programme through the government systems, a multiplicity of non-state actors operate within the ART governance space and contribute to transaction costs generation. Both these aspects appear to be inconsistent with the first order economising principles of TCE- wastage reduction.

Secondly, the ART programme management framework is designed to reduce the risks of differing standards and quality through a centralised coordination system. This system is consistent with TCE theory, although there did not appear to be a framework for transaction costs monitoring for the purpose of efficiency calculations. This limited monitoring of transaction costs appears to constrict efforts at looking for better ways to minimise transaction costs, for instance in the area of data and information management. In addition, as with ART finance administration, wastage is observed in supply chain management where a parallel drug procurement system has been retained despite the existence of an existing structure for medicine procurement in the CMST.

7.5 Monitoring Transactions Costs and their Incidence in the ART Programme

The application of the TCE to improve efficiency in ART management implies the need for an adequate mechanism to track ART service delivery transaction costs. There are currently four frameworks that track resources that go into the HIV and AIDS programme that can provide indications of levels of transaction costs in HIV and AIDS service delivery. These approaches are: (a) the National AIDS Spending Assessments (NASA), (b) Health Sector Resource Mapping, c) Global AIDS Monitoring (GAM) report, d) National Health Accounts (NHA). Table 13 summarises expenditure or budget categories that may be used as proxy for transaction costs monitoring.

Table 13:Existing frameworks to proxy HIV and AIDS ART transaction costs

#	Expenditure tracking framework	Explicitly tracks transaction costs (Y/N)	Category proxying transaction costs	Transaction costs estimate (proxy)	Recent Update
1	NASA (National AIDS Commission)	N	Health systems strengthening with “management administration costs” sub-category	20%	2021
2	Health Sector Resource Mapping (HSRM) (Ministry of Health)	N	Community outreach, technical assistance, administration & management – Other, Research, M&E and supervision, planning & policy	30%	2020

#	Expenditure tracking framework	Explicitly tracks transaction costs (Y/N)	Category proxying transaction costs	Transaction costs estimate (proxy)	Recent Update
			activities, communication costs (print, TV, radio), and resource mobilization.		
3	Global AIDS Monitoring (GAM) report (National AIDS Commission)	N	Programme Management	22%	2022
4	National Health Accounts (Ministry of Health)	N	Health systems administration and financing administration	26.2%	2018

Source: Literature review

The research found that while there exist initiatives to track resources and expenditure for the health sector and the HIV and AIDS sub-sector, an explicit framework to track transaction costs for efficient ART management was not evident. Nevertheless, the NHA laments the high expenditure on governance and finance administration as reflective of inefficiencies in health expenditure in Malawi (GoM, 2018). Consequently, the need for an explicit focus on transaction costs monitoring in healthcare and HIV and AIDS ART governance is implied. Data from the NASA, NHA, Health resource mapping can be used to monitor transaction costs with a view to integrating them into the efficiency calculations. Additionally, the HIV and AIDS EMR and data from other HIV and AIDS input and output monitoring, including numbers and effort by staff providing ART and HTC service can meaningfully be used to track transaction costs and efficiency of the ART programme.

7.6 Measures to Reduce Transaction Costs in ART Service Delivery

A key issue with respect to ART governance is the monitoring of transaction costs to inform programme efficiency. This is important to the tracking of the effectiveness of measures aimed at optimization of transaction costs in ART service delivery. Within the constraints of public governance of ART service delivery, achievement of efficiency will largely depend on the types of instruments deployed to reduce the transaction costs.

The research revealed a number of measures deemed to contribute to transaction cost reduction for quality ART service delivery. They include standardization and guidelines, the tenure of contractual arrangements with suppliers and sub-contracted entities, and the drive towards integration of service delivery within the public primary health care system.

First, and drawing on the WHO clinical guidelines, the Government seeks to standardize service delivery to reduce the probity hazard in ART service delivery. The Government has developed HIV Clinical guidelines as well as clear protocols for HIV testing and counselling. This helps to reduce the information asymmetry and search costs by providers and patients. At each ART service delivery point, guidelines are in place to guide the provision of the services. This further contributes to the quality of the services delivered. As noted before, the extent to which these guidelines are implemented and adhered to at each ART site is monitored via quarterly supportive supervision visits to each health facility. Nevertheless, it is important to observe that while monitoring compliance with the guidelines presents the Ministry of Health with information to facilitate improvements, the enforcement of the guidelines also represent a source of transaction costs, and highlight the need for tracking them for cost control.

Second, current contractual arrangements with service providers are reported to contribute to limiting disturbances in service provision, and hence, reduce transaction costs. The current agreements and contracts with partners in the ART programme, including the SLA framework of service provision have performance measures and targets. This helps to align the interests of the MoH and the different stakeholders to the aspirations of improved health of the citizens through provision of services that meet the minimum quality standards. Furthermore, noting that the conclusion and enforcement of the contracts generates transaction costs, clear and relatively long-term contracts appear to be one instrument that can help to reduce transaction costs. The long tenured drugs and commodity distribution contracts with Bolloré Logistics, as well as the drug acquisition partnership through a regional pooled arrangement appear to lower transaction costs, lowers drug prices and uncertainty and stabilizes drug availability across service delivery points. The caveat is that long tender terms do not automatically reduce costs- if input costs, such as those of an active pharmaceutical ingredient, fall during the term of the contract. It is unlikely that this price reduction is carried through

into savings. Further, whether or not savings can be realised depends on a) the frequency, and b) direction of such input price fluctuations. Further, this perceived effectiveness of the contract management framework and its associated coordination costs was not evaluated, although there were indications service providers and suppliers generally remained within prescribed targets. Except for the supply chain disruptions during the COVID-19 pandemic reported to have caused delays and an increase in the cost of drugs and commodities, there appeared to be a contracts management framework in place that was functional. Nevertheless, periodic assessments of the framework in terms of costs and technical efficiency should contribute towards the monitoring and generation of information of where efficiencies can be realized through a reduction in wastage in the face of dwindling HIV and AIDS aid.

Third, innovations in electronic records management (EMR) is an area that can potentially contribute to the reduction in transaction costs in ART. Policy managers indicate a movement away from the current module that appears to be "heavy" on equipment and financial resources should contribute to efficient data collection and management, and hence, overall ART governance efficiency. A lightweight module of EMR is being considered for the ART programme, and the following quotes exemplify the contribution of innovation in reducing transaction costs in ART records management:

> *"So, now, there is a lightweight version of the EMR being tested, with the small investments. Lightweight in the sense that it doesn't take a lot of heavy duty equipment, like heavy investment to run the EMR"_ Respondent 2, National Policy Manager.*

> *"So, another thing that has facilitated this has always been the systematic monitoring and evaluation that has happened. And even that has changed over time from being entirely paper based to now..., supported with electronic solutions, which donor X has also facilitated this last year. In fact, this quarter, you know, with COVID-19, they didn't do their usual supervision, but they still have data for the quarter. So, this is the first time ... a significant chunk of the HIV ART programme data is coming completely from an electronic solution". Respondent 5, National Programme Manager.*

As may be observed from the above quotes, it is envisaged that investment in efficient e-solutions can greatly reduce the cost of data collection and processing. To this may be added the use of decentralized ledger technology to improve data validation and fidelity. This would insure against opportunistic tendencies by staff that manage data, and thus represent an enhancement of measures to curb opportunism in ART data management. For instance, with the current practice of incentives such as the award of certificates of excellence, or other facility-based awards related to quality of care, there is the possibility that staff may manipulate the data where reporting is done remotely. Finally, the research found that the Government through the MoH has embarked on a drive towards integrated primary health service delivery within which the ART programme is embedded (GoM, 2023c). The third generation National Health Sector Strategic Plan (NHSSP III) has focused on integration of primary health services towards achievement of efficiency and quality of care. Since ART is offered within this primary health care framework, it is expected that efficiencies to be realized from the integration will further benefit the ART programme.

To an extent, the provision of ART through the public health units' network has contributed to the reduction in the coordination costs and improvement of the continuity of care. This co-location of services and sharing of patient information among providers such as between the antenatal clinics and ART services appears to have streamlined the care delivery, and is thus an example of how TCE can be applied to public agency service provision to improve efficiency. It is noteworthy, nonetheless, that within the bulk of government health units, ART clinics are conducted as separate clinics and on specific days. The research found that the clinics handle the clients in cohorts, with each cohort attending the clinic every three or six months depending on whether they are assessed to have become stable clients or not. At times this scheduling of clinics depends on the quantities of drugs available. This arrangement is aimed at the optimisation of human and physical resources given the large numbers of people on ART per clinic, as well as the reduction of patient-level costs (travel cost, days lost from work). Handling clients in cohorts is considered the most efficient way to the provision of the ART services given the current staffing and facility constraints. It further reduces the time the clients need to visit the health facilities for supplies and review. However, major challenges that constrict effective integration include inadequate skilled staff handle co-

morbidities, space limitations and perceived personality problems that seek to retain ART as a vertical programme. Nevertheless, the costs and efficiency of integrated health services delivery is not adequately tracked. Only anecdotal evidence is available, and evidence elsewhere suggests that integration of ART services with other general services become inefficient when viewed from the perspective of clients.

7.7 Chapter Summary

This chapter sought to answer the research question related to the sources of transaction costs in HIV and AIDS ART governance in Malawi, including whether the transaction costs are due to lack of conformity with TCE theory. The research has found that the main sources of transaction costs are the HIV and AIDS ART financing administration where what appear to be excessive fiduciary safeguards feature prominently. Another source of transaction costs is the ART programme management where contracts management, service quality monitoring and compliance activities likely generate transaction costs. The presence of substantial transaction costs in HIV and AIDS ART service provision means that they affect the overall unit costs of service provision. Nevertheless, since transaction costs are essential to the provision and access of ART, the implication is that they affect levels of efficiency of ART services as they add to the total costs of service provision. Transaction Cost Economics suggests that the continued increase of such costs may mean that beyond a certain threshold, they are likely to contribute to inefficiency in the production and distribution of a product or service.

Furthermore, the research noted that the transaction costs may not necessarily be arising from a lack of governance design consistency with TCE theory, as also noted in the previous chapter. Furthermore, an overdesigned safeguards framework and wastage reflected in a parallel procurement management scheme appear to infringe on the TCE first order economising principles-reduction of waste towards efficient economic governance.

The next chapter discusses the findings of the research.

CHAPTER EIGHT

DISCUSSION OF RESULTS

8.1 Introduction

The research sought to assess whether inadequacies in the design and implementation of the HIV and AIDS ART programme as compared with Transaction Economics theory was resulting in inefficiencies in Malawi. The research hypothesised that (in) efficiency of the HIV and AIDS ART programme in Malawi was a consequence of inadequate monitoring and control of internal transaction costs, which in turn arises from the limited application of a transaction cost economics (TCE) theory to inform (in) efficiency considerations in ART policy, programme planning and implementation. The sub-research questions were the following:

a) Are HIV programmes in Malawi inefficient, and what type of inefficiency do they display, and what type of inefficiency do they display?

b) Are transaction costs affecting the overall efficiency of the HIV and AIDs programme?

c) How does the Malawi HIV and AIDS ART programme governance comply with the governance theory (transaction costs economics theory?

d) What are the main sources of transaction costs in the current governance of the HIV programme in Malawi? Are these due to non-compliance with the prescription of transaction cost economics theory?

This chapter summarises the results of the research as presented from Chapters 5 to 7, highlighting the key research findings and how the research questions have been addressed.

8.2 Results Discussion

Overall, the research found that the design of the HIV and AIDS ART programme governance conformed with TCE theory as it has all the features to deal with disturbances that arise from asset specificity, uncertainty and complexity. Additionally, the results suggest that the governance design is adequate to deal with probity hazards that are likely to obtain with public provision of services. Nevertheless, the results show

that there are inadequacies in monitoring and internal control, particularly with respect to supply chain management that tends to generate wastage in service provision. Furthermore, while mechanisms such as the NASA are available and capture data related to transaction costs, the monitoring of transaction costs for efficiency decisions is not evident, implying scope for improvement. The programme shows inefficiencies, with transaction costs impacting efficient provision negatively. An excessive safeguards framework in HIV and AIDS financing administration further tends to generate transaction costs and increases the costs of ART service delivery, and hence, likely impacts efficient provision negatively.

The following sub-sections reflect on the research sub-questions.

8.2.1 HIV and AIDS ART Technical (In) Efficiency and how Transaction Cost affect Efficiency

The research sought to answer the question as to whether the HIV and AIDS programme in Malawi was inefficient, and if it was inefficient, to determine the types of inefficiency it displayed. It further sought to assess how transaction costs effect the efficiency of the HIV and AIDS programmes in the country. With a focus on the HIV and AIDS ART programme, the research finds that with an overall technical efficiency of 92%, the ART programme is inefficient as this implies that output can be increased by 8 % with existing resources. The results are comparable with evidence from Rwanda, a similar setting in the Sub-Saharan region, where average efficiency using DEA estimation in HIV and AIDS ART services was 78 percent (Zeng, Rwiyereka, Amico, vila-Figueroa, & Shepard, 2014). Nonetheless the ART service delivery efficiency levels reported in the current study appear to be higher than those reported for Rwanda. Further, the research identifies 48 efficient health facilities that are operating at efficient levels-they displayed unity efficiency scores. For these facilities, the implication is that inputs cannot be changed (increased or decreased) without impacting on efficient operation. A further implication is that these “efficient” facilities can be used as a benchmark for their peer health facilities towards efficiency improvement in the ART programme.

Two types of scale inefficiency in the facility-based HIV and AIDS ART programme in Malawi are revealed by the research- *increasing* and *decreasing* returns to scale. Increasing returns to scale imply that available inputs are less than fully utilised given

current operations of the facilities, and hence, the need to fully deploy available inputs to achieve technical efficiency in ART service delivery. In other words, the results imply that facilities experiencing increasing returns to scale are operating at too small a scale to be efficient (Kumar & Gulati, 2008). Further, the results appear to indicate that some of the facilities are too big for the catchments or the ART client volumes they serve. The results support recent empirical evidence from a systematic review that use a similar approach to the current study in estimating technical efficiency of healthcare facilities, the DEA, that suggests that an estimated 60 percent of health facilities in Sub-Saharan Africa are inefficient (Babalola & Modley, 2020; Ibrahim, 2023). For ART programming, this implies that efficiency would arise from an increase in the outputs, which may be achieved via a streamlining of the number of health facilities that provide ART services in a particular geographical area. Another means might be intensification of testing services to ensure more people are brought onto the ART care stream, which tends to support findings in Malawi by (Indravudh, *et al*., 2021) who study the effectiveness of community-based HIV with regard to utilisation of community–led HIV testing. They found that community-based HIV testing could be delivered at low cost when compared with facility-based testing. Furthermore, the results tend to support findings by Carlson, Chirwa, Hall, & Cammack, (2015) who report inefficiency in the health sector in Malawi that is occasioned by resource mismanagement and limited budget control by budget holders. The results are important for Malawi and other resource constrained countries given the dwindling global resources dedicated to HIV and AIDS response.

The study further shows that the transaction costs affect health facility efficiency negatively. The log of transaction costs was negatively correlated with facility or decision-making unit efficiency as hypothesised, and the results are statistically significant. The result was consistent with transaction cost economics theory that affirms that beyond a certain level, transaction costs will negatively impact firm efficiency. For instance, in a theoretical analysis of transaction costs in aid delivery, Lawson (2009) and Lithman (2014) suggest that while a measure of transaction costs is required to coordinate and achieve efficient aid delivery, as the levels of transaction costs rises, it impacts efficiency negatively. Furthermore, the results support empirical literature on the adverse effects of high levels of transaction costs on firm efficiency. The results support findings by Todorova (2011) among Eastern Europen countries who use a

production function model and find that at certain levels of transaction costs and community isoprofit line[20], firm output is reduced. More particular to access to healthcare, the current research results support the findings of Drake, Anderson, Cai, & Sacks (2023) who use regression discontinuity analysis and argue that eliminating transaction costs in healthcare insurance in the USA increased benefit take-up of health insurance, and hence, access to healthcare services by low income groups in the USA. The results are consistent with findings by Eze, Idemili, & Lawani, (2024) who analyse health systems efficiency using data envelopment analysis and find that 78.5% of the countries studied had inefficient health systems. This suggests that the inefficiency problem in health systems needs to be prioritised and ways found to increase efficiency in the face of declining development aid for health. Furthermore, the results presented in the study tend to support findings by Khanakwa and Monigaba (2022) who used a qualitative analysis of different institutional arrangements for the delivery of HIV and AIDS prevention services in Uganda. They found that governance modes with low transaction costs tended to be efficient when compared with those with high transaction costs for the same service and outcome. The results imply that the Malawi ART governance framework may need to consider other governance modes for service delivery, including community-led as well as ICT based teleclinics and digital health that lower transaction costs (Theodore,*et al.*, 2015).

While a number of studies have assessed health systems or HIV and AIDS delivery efficiency, they have not specifically quantified transaction costs and assessed their effects on efficiency. In addition, while some studies highlight the role of TCE in HIV and AIDS governance, they have remained qualitative. Hence, one of the main contributions of the current study is the quantification and linking transaction costs to technical efficiency of HIV and AIDS programmes. This may have practical policy relevance in the general health systems governance, as well as HIV and AIDS policy in the prioritisation of transaction costs minimisation.

Other factors that affect HIV and AIDS ART service delivery efficiency

Other factors that affect facility-based HIV and AIDS ART programme efficiency include distance from Lilongwe (headquarters), type of bureaucracy or ownership,

[20] Refers to the graph of a firm profit function that represents an infinite number of solutions all of which yield the same profit (toppr.com)

location, service delivery level, the number of years the facility has been in operation and the population served.

The coefficient for *distance* squared to reflect the limit of distance was negative as hypothesised, but the result was not statistically significant. Results support empirical evidence that indicates that provision of services to remote populations is associated with inefficient provision, for instance, Abate, Dereje, Hirvonen, & Minten (2020), who use econometric analysis to explore the relationship between public service delivery and remoteness in Ethiopia. They report inefficient public service provision to remote communities due to high transaction costs, suggesting that distance from centre increases transaction costs in public service delivery that may impact efficiency. Further, other studies such as that by Berg, Radicke, Stentzel, Hoffman, & Flessa (2019) who used linear programming to examine hospital planning based on economic efficiency criteria in rural areas in Germany show that it would lead to closure of some hospitals due to inefficiecy. They further found that application of economic efficiency criteria in the allocation of the facilities would reduce access to paediatric and obstetric care for 8% of the population, reflecting the conflict between the apparent opposing goals of efficiency and equity or universal access.

The type of *organisation ownership or bureaucracy* is a further determinant of efficiency in ART provision. With private ownership as baseline, public, CHAM/MAM and CSO facilities appear to be more efficient. This result is consistent with Tiemann & Schreyögg (2009) who used bootstrapped DEA to assess technical efficiency of hospitals in Germany, and found that public sector hospitals performed better than privately owned hospitals. However, using a quantitative meta- analysis of the technical efficiency of hospitals in the USA, Freeman, Nam-Speers, & Tokac (2020) find that the evidence has been mixed regarding the influence of ownership type on technical efficiency of health facilities. They report that privately owned health care facilities are not universally technically more efficient, and that the effect of ownership is non-liear, mediated by time and facility types. Furthermore, using a top-down approach to estimate the cost of ART delivery in Malawi, GoM & Options (2012) find varying costs between MoH and CHAM run health facilities providing ART services, with CHAM run facilities displaying less variation in costs than public or private facilities. This may have implicatons on efficient ART service dleivery. The type of ownership is reported to

be a key determinant that affects efficiency, as reported by Mahajan, Nauuriyal & Sing (2018) who applied DEA to anaylse efficiency in the Indian pharmaceutical industry. Nevertheless, the results from the current study are not statistically significant, and hence, it was not possible to conclude that the other forms of bureaucracy were more efficient than private run health facilities that provided ART services.

Results on *location* by rural and urban show that facilities in the rural areas are less efficient when compared with urban facilities, and the result is statistically significant. The reasons for this outcome are not obvious, although demographic factors such as population density and numbers of people accessing ART services may have a bearing on the inefficiency of ART service provision in rural facilities. A study that used DEA to assess the determinants of technical efficiency in Kenyan manufacturing firms found that location was a key a determinant of technical efficiency (Cheruiyot, 2017). Results from the current study further suggest that facility location in Nkhata Bay, Ntchisi, Neno, Mzimba, Mwanza, Machinga, Lilongwe and Likoma Districts should be associated with efficient ART service provision. The rest of the districts were negatively correlated with DMU efficiency, suggesting that they should be associated with inefficient provision. Some studies observe the influence of district characteristics, including demography and management style that are likely to influence efficiency of service delivery. For example, Bautista-Aredondo, *et al*., (2014) who used DEA to assess the technical efficiency of ART service provision in Rwanda report these factors to be important determinants of technical efficiency.

Furthermore, *spatial (geographical) distribution of* ART facility efficiency showed that particular districts in the northern and southern parts of the country had high efficiency scores relative to their peers. Notably, these districts appear to display higher ART facility efficiency scores when compared with districts close to Lilongwe City, where central coordination of the ART programme occurs from. The reasons for this outcome are not readily apparent, although 6 of these 7 remote districts with health facilities displaying efficiency are border districts. It therefore, appears that the effect of both high client load and catchment population per facility is mediated by their geo-spatial location. This is likely occasioned by extension of ART services to the Zambian and Mozambican population along the border areas in these districts, which tends to increase demand and hence decrease unit costs. The extension of ART services to citizens of the

two countries was revealed through interviews with service providers in the qualitative component of the current research. Hence, although the facilities along the borders are likely to reflect high technical efficiency for ART provision, there are likely to be challenges to track ART clients, as was revealed through qualitative component of the study. In addition, increased numbers from the immediate vicinities of the communities in the neighbouring countries is likely to put pressure on the available ART resources, with policy implications on resource allocation for ART.

The results from the research further show that *healthcare level of service delivery* is an important determinant of efficiency. With the tertiary level as baseline, ART provision through both primary and secondary level provision should be efficient. Nonetheless, the results are not statistically significant and, hence, inconclusive. One reason for this may be the fact that while tertiary health units are expected to largely provide referral services from lower levels, they also provide some of the primary health services, including ART. Hence, despite the tertiary and secondary facilities being relatively well resourced than primary facilities, patient loads may be similar. The Government of Malawi (2023b) notes that because of this, there are inefficiencies in the healthcare governance. Thus, the need for an effective referral system is implied across health service provision levels towards efficient ART management.

The number of years that the facility has been offering ART services (*Site maturity*), did not display the expected coefficient sign before estimation. That is, while it was expected that with longer time periods of operation health facilities that provide ART should become efficient due to accumulated knowledge and experience with service provision routines, the results appear to indicate that for the sites included in the model, site maturity was associated with less efficient ART service provision. The result was contrary to other studies such as Berruti and Blandford (2012), that find the period that a health facility has been operating is positively correlated with efficiency. Similar results were reported by Zeng *et al.*,(2014) for Rwanda. The reasons for this inconsistency with previous data are not readily available, although staff turn-over and transfers away from site, leading to use of less experienced staff might be one reason. Another reason may be use of outdated ways of organising patient cohorts, despite quality monitoring by the MoH.

Catchment population is positively correlated with ART health facility efficiency, and the result is statistically significant. The result tends to support the hypothesis that the higher the demand in service delivery, the greater the efficiency in service provision of a decision-making unit. Other studies also find that catchment population is positively correlated with health facility efficiency. For instance, Berruti and Blandford (2012) who analysed the determinants of HIV treatment costs in resource constrained countries covering Botswana, Ethiopia, Mozambique, Nigeria, Uganda, and Vietnam found that patient volume and the number of months that a facility provided treatment services were strong predictors of average treatment costs. Sites with large volumes of clients were positively correlated with efficiency. This implies that ART programmes need to consider improved ART service delivery efficiency in the context of population levels that can potentially access the ART services.

8.2.2 HIV and AIDS ART Programme Service Delivery Governance Conformity with TCE Theory

The research further sought to answer the question on whether the HIV and AIDS programme in Malawi conforms with the governance theory according to Transaction Cost Economics theory. Overall, the research findings suggest that the HIV and AIDS ART programme design and implementation in Malawi conforms with TCE theory. Using the discriminating hypothesis principle to analyse conformity, the governance structure is assessed to have been designed and implemented in ways that are calculated to reduce transaction costs. These costs arise from disturbances due to asset specificity, uncertainty and complexity of transactions. Furthermore, the governance instruments are designed to contain probity hazards common with public agency economic governance. The research found that the dominant ART governance mode, the state mandate and its sub-attributes such as long-term contract relationship with staff seem suited and were in conformity with TCE theory. The results are consistent with findings by Spithoven (2012) who reports that delivery of health services via alternative arrangements to public provision such as privatisation are likely to lead to the exclusion of some groups due to the cost control drive by the private sector. More to the point, efficiency calculations for ART using the public agency (MoH) have to be further considered in terms of the broad objectives of service delivery beyond technical efficiency to include equity in service delivery.

Furthermore, the research results suggest that there are attempts to achieve efficiency through allocation of HIV and AIDS resources based on the HIV and AIDS disease burden through public agency provision of HIV and AIDS ART services. Distribution of the ART facilities is achieved by public fiat, and thus helps to allocate service delivery to even the most disadvantaged communities and households. Hence, the research found the governance of the programme was in conformity with TCE, assuming the actual transaction costs are lower than alternate governance arrangements. However, other studies report less than efficient allocation of resources in the general healthcare sector in Malawi (World Bank Group & Global Financing Facility, 2021). The caveat is that studies that report ineffectiveness in resource allocation consider a broader range of HIV and AIDS services, when compared with the current research that focused only on the ART programme.

The research further finds examples of cost control measures, such as unification of donor per diems and allowances for the public sector missions, although the effectiveness of these measures in contributing to the reduction of transaction costs has not been evaluated. Nevertheless, these represent conscious administrative cost control measures that assist with transaction cost reduction. Other studies applying TCE theory in HIV and AIDS service delivery find that public provision is one of the dominant governance modes of HIV and AIDS service delivery (Khanakwa and Mbonigaba 2022; 2024). They report that pure public provision appeared not to be the most efficient governance mode at transaction cost minimisation. Instead, partnership institutional arrangements with strong community oversight appear to be more effective at transaction cost minimisation (Khanakwa and Mbonigaba, 2022). Similarly, in a study comparing HIV prevention service provision in India through public and non-state provision using transaction costs economics theory, Guinness (2011) found that public provision did not display transaction cost minimising tendencies due to corruption and limited transparency in service provision.

The research findings further suggest that there are aspects of the Malawi HIV and AIDS ART governance that showed less than full compliance with TCE theory. These are mainly with regard to inadequacies in reducing wastage, or first order economising in TCE terms. For instance, the existence of a parallel supply chain governance for drugs and commodities suggests there is wastage in the broader healthcare supply chain, which

also affects ART service delivery efficiency. This represents a deviation from the TCE prediction that managers or policy makers in the public sector are likely to pursue governance modes that reduce wastage (first order economising). While it may be argued that the aim is to avoid the costs occasioned by the much-highlighted drug pilferage at the CMST, the parallel arrangement likely increases the transaction costs of the public healthcare supply chain and induce further wastage due to the duplication effect. This result supports findings by Kaupa & Naude (2021) who highlight the parallel supply chain as a binding constraint to efficient public healthcare supply chains. The need to streamline the supply chain process, and introduction of other more efficient approaches with transaction costs reducing tendencies is thus implied (Omari, *et al*., 2021; Roecka, Sternberg, & Hofmanna, 2020).

Because the ART programme governance has been embedded into the network of healthcare facilities that are geographically spread, the research observed that the rapid expansion and decentralisation of the programme may further hamper its efficiency. Thus, while general conformity of with TCE of the HIV and AIDS ART programme is observed, the rapid expansion and geographic spread of the programme may be associated with high transaction costs that affect its efficiency. The results from the research are consistent with findings by Oliveira, Santinha, & Marques (2024) who completed a systematic review of empirical evidence on the effect of decentralisation on healthcare effectiveness, equity and efficiency and found that decentralisation fell short of delivering healthcare results. This research, therefore, supports recent evidence from the Organisation for Economic Development (OECD) countries that public expenditure in highly decentralised health care systems tends to increase Dougherty, Lorezoni, Marino, & Murtin (2022), and likely yields additional transaction costs, with implications on efficient healthcare service delivery. However, the research results contradict Dick-Sagoe *et al*., (2021) who used a qualitative design to assess the effect of decentralisation on healthcare services in Lesotho. They report that decentralisation of healthcare increases efficiency mainly due to the proximity of management boards to the health facilities in the rural areas. Nevertheless, their study ignores the significant resources that are spent by the centre to enforce standards.

The current study expands the literature on the application of TCE to HIV and AIDS and healthcare management by highlighting the extent to which HIV and AIDS governance

conforms with TCE. It specifically points out areas of conformity and gaps that may assist in directing action to improve service delivery efficiency.

8.2.3 Sources of Transaction costs in the Malawi HIV and AIDS ART programme and compliance with TCE prescriptions

The research further sought to answer the questions as to what are the sources of transaction costs in the HIV and AIDS ART service delivery governance. Additionally, it sought to assess whether the transaction costs in the HIV and AIDS programme obtain because of noncompliance with TCE theory. The results of the analysis of the ART programme using Transaction Cost Economics confirm that a deliberate choice is made by the Malawi Government to provide services through the public agency (public bureaucracy) in the Ministry of Health healthcare provision system. The results are consistent with TCE theory, and suggest that this choice is designed to address adverse selection in service provision and reputation risk. The results are consistent with literature on general economic governance choice (Williamson,1975; Shervani, Frazier, & Challagala, 2007). The results further support emprical literature which indicate that public governance of healthcare services is designed to avoid adverse selection risks and state failure reputation risk. For instance, a study by Spithoven (2012) found that the choice of public goverance of the Affordable Care services in the United States of America was designed to obviate similar risks. This implies that ART service provision is likely to continue to be provided through the public agency in the country, until a feasible alternative can be defined, implemented and will have demonstrate greater efficiency when compared with the public agency. Nevertheless, there appears to be a limited consideration and application of TCE in ART governance. Intervention specific cost efficiency assessments appear to receive a high premium in ART efficiency considerations when compared with governance aspects. The existence of other mechanisms for HIV and AIDS service delivery such as community-led HIV testing suggests that the current framework can be modified in ways that improve efficiency. Thus, the study contributes to the need to specifically and deliberately consider modifications in ART organisational service delivery mode to improve efficiency, and is in support of findings by Indravudh, *et al* (2021) who report evidence of the effectiveness and effeciency of community led HIV self-testing in Malawi.

The results further highlight the sources of transaction costs in the ART programme in the country. The main sources of transaction costs are finance administration for ART, stemming mainly from safeguards and compliance monitoring. The creation of an independent Project Management Unit in the MoH, as well as a Finance Management Agency and various financial audits under the GFTAM contribute to the generation of transaction costs. A multiplicity of private sector and not for profit agencies through which funding support for ART is channelled by donors, including PEPFAR, further contribute to transaction costs in ART governance in Malawi. Furthermore, centralised coordination to maintain quality standards and service compliance through the programme and supply chain management generate transaction costs. On the ART programme financing end, the safeguards appear excessive while duplication in the healthcare supply chain management violate TCE economising principles. The results imply the need to review the safeguard mechanisms when financing the ART programme and to streamline the healthcare supply chain management.

A key limitation in the HIV and AIDS ART governance framework is that transaction costs appear not to be an explicit focus of efficiency calculations. Thus, while the National Health Accounts has raised concern regarding the high levels of governance costs, there does not appear to be a conscious effort at monitoring them and using them as a specific cost category. This implies that it is less likely to prioritise these costs in efficiency calculations. There is an opportunity to use data collected via the NASA and NHA, as well as Health Resource Mapping to track transaction costs, as well as assess efficiency in ART provision. These data can be used in combination with the DHA electronic management records which track patient and HIV testing volumes, as well as staff effort in the programme.

The current research highlights the sources of transaction costs in HIV and AIDS governance and hence expands knowledge of the sources of transaction costs in service delivery. In addition, by characterising production and transaction costs in HIV and AIDS ART service delivery, the research has pointed towards how transaction costs can be monitored.

8.3 Research Limitations and Further Research

This study has drawn attention to the TCE application as a possible means towards efficient HIV and AIDS ART management in Malawi. It has shown that the facility-based HIV and AIDS ART programme displays some inefficiencies, with some of the facilities operating at too small or too big a scale to be efficient. In estimating the determinants of technical efficiency of facility-based ART, it has further shown the feasibility of monitoring efficiency, with transaction costs in mind, and the means through which they can be minimised. The study has further attempted to quantify transaction costs in ART service provision, achieved through Bayesian analysis in a context where these costs are not deliberately monitored to inform efficient ART service provision in the country. The study further covers the bulk of ART sites in the fiscal year 2018/ 19, and, therefore, provides a national baseline regarding facility-based ART service delivery efficiency.

Nevertheless, a number of limitations are also noted in the research which have to be taken into account when interpreting the results. First, some of the cost elements for the ART have had to be assumed and percentages from other studies used for those particular elements. These include personnel, distribution, occupation costs and warehousing costs. Further, the study uses cross-sectional data that does not permit analysis of changes in total factor productivity in ART that can only be achieved with longitudinal data. Further, due to the cross-sectional nature of the dataset, causality between explanatory variables and the dependent variable may be difficult to infer. Hence, although the study provides a baseline for technical efficiency analysis of the ART programme in Malawi, and the influence of transaction costs on ART delivery efficiency, further research is required using longitudinal or panel data to track intertemporal changes in both transaction costs levels, total factor productivity and efficiency. Applying a Malmquist DEA that takes into account total factor productivity over time is therefore, recommended in future research. Furthermore, periodic monitoring of transactions and their effect on efficiency would be important.

CHAPTER NINE

CONCLUSION

9.1 Research Overview and Aim

The growing efficiency imperative in HIV and AIDS governance is occasioned by a dwindling trend in resources available to fight the AIDS pandemic. At a time when there are calls to increase resources to contain the HIV and AIDS pandemic as a public health threat by 2030, the initial optimism by donor countries to finance the pandemic response appears to have waned, casting doubt on the sustainability of the response, and raising the spectre of a rebound in the epidemic. This jeopardy is exacerbated by the low-income countries' high dependency on global health aid to finance their ART programmes, the largest component by any measure of the AIDS response at country level. While new HIV infections show a downward trend in Malawi, the reduction in aid for the HIV response implies particular vulnerabilities, considering the need to sustain people that are already enrolled on ART, and those that will need be brought onto the ART stream. Both the GFTAM and PEPFAR, the two largest funders of HIV and AIDS in Malawi, are likely to reduce funding over the next years under the expectation that LMICs like Malawi increase the domestic share of HIV funding. While policy discussions are currently under way on how to sustain the response in Malawi, including use of innovative financing to finance the response, they appear to be inconclusive and alternatives to increase domestic funding for HIV and AIDS appear limited.

Against this backdrop, the motivation for the research was that efficient economic governance of the AIDS response to minimise costs is an area that receives limited attention, but which can potentially assist in the minimisation of HIV and AIDS ART costs. Consequently, the research focused on a Transaction Cost Economics analysis of the ART programme in Malawi, with a view to determining the sources of transaction costs and their efficiency implications in ART service delivery. In particular, given the chosen mode of provision of ART via the public agency, the study applied a Transaction Cost Economics lens to assess the efficiency of the ART programme in Malawi and gauge its governance compliance with TCE theory.

9.2 Methodology

Specific methods included an extensive literature review focused on economic governance and transaction cost economics to foreground the research, policy and service provision level key informant interviews, database reviews for ART related data on current client volumes by ART site, prices for ART drugs and commodities. The collected data were analysed via thematic analysis applied to qualitative data from interview transcripts. Qualitative data were coded and subsequently organised into themes to detect trends and unique traits in the data. Quantitative data were analysed through descriptive statistics and regression analysis. Facility level ART delivery efficiency was completed via a two stage Data Envelopment Analysis (DEA) to first estimate a non-parametric facility (in) efficiency score and, second, to estimate the sources of variation in the facility level ART (in) efficiency scores.

The study reviewed concepts related to transaction costs to facilitate their application to HIV and AIDS governance, particularly to ART service delivery. The study notes that transaction Cost Economics Theory informs economic governance of transactions that apply to the public and private sectors in the ART programme in Malawi. The need to look at the key factors internal (to the Ministry of Health) and external (interaction with other actors in ART provision) that influence ART (in) efficiency which need to be explored in the evaluation of the efficiency of HIV and AIDS programme is implied from the research. Transaction characteristics, governance choice, costs and efficiency, institutional environment, technology, individual factors and time and geography, are key in this respect.

9.3 Results

Overall, the research results show that the HIV and AIDS ART programme in Malawi is inefficient as output can be increased by a magnitude of up to 8 percent with current inputs. An estimated 8 % of the health facilities offering ART services are assessed to be efficient, while the rest appear to operate at inefficient levels. The results further show that the HIV and AIDS ART programme displays two types of scale inefficiency, namely, increasing returns to scale and decreasing returns to scale. An estimated 23% of the facilities display increasing returns to scale. Increasing returns to scale imply that the facilities are operating at too small a scale so as to be inefficient. A further 69% are

assessed to operate at decreasing returns to scale implying that they are operating at too big a scale for their current operations. With specific reference to efficient service delivery, the research policy implication is that policy interventions that increase the use of excess capacity in the ART health facilities that display increasing returns to scale should lead to more efficient service delivery. Further, a review of the facilities that appear to be too big for their catchment areas, or volumes of clients on ART is implied given the large number of health facilities that are assessed to be experiencing decreasing returns to scale. To that end, benchmarking efficiency on the 48 facilities that are identified in the current research as reference peers may contribute to improved efficiency in ART service delivery in the country.

The research further found that, consistent with economic theory, transaction costs impact negatively on efficient ART service delivery. Results from the regression analysis show that transaction costs enter the efficiency equation negatively, implying the need to find means to reduce these costs and thus contribute to efficient service delivery. To the extent that a reduction in transaction costs can contribute to freeing up resources, and hence, act as a resource mobilisation strategy, the implication is the need to closely monitor them and identify areas where efficiency can be gained. While the initial costs may be significant, investment into decentralised ledger technology to assist in monitoring and adherence to standards in a highly decentralised ART service delivery framework may be one area to consider in ART governance. This is particularly important given the declining global resources available for HIV and AIDS that is likely to impact the sustainability of the ART programme.

The research further finds that in general, ART governance choice and implementation conforms with TCE theory. Delivery through a public agency and a partnership arrangement ensures that there is no adverse selection and that services are equitably distributed to all people that need the service. The long term, relatively secure employment relationship between the state and employees that provide healthcare services that include ART ensures there are minimal disruptions that may arise from asset specificity in human resources. Furthermore, probity hazards are minimised via a strong monitoring framework and an insistence on centralized standards. In addition, the research observes that mechanisms to track HIV and AIDS expenditure and transaction costs exist. Nonetheless, it is not clear whether there is an explicit focus on

the tracking of transaction to inform efficient service delivery. Furthermore, there is evidence that there are attempts to improve allocative efficiency and reduce wastage in HIV and AIDS services. This extends to measures such as equitable distribution of ART facilities, as well as reduction in transaction costs arising from excessive allowances in the public sector. Nonetheless, there are gaps in in implementation that show non-conformity with TCE theory. These include a parallel public healthcare medical and supply chain that generates inefficiencies from duplication. Further, an excessive safeguards framework for HIV and AIDS financing tends to generate transaction costs that likely impact efficient service delivery.

The study observed that HIV and AIDS commitments regarding treatment and care have to be sustained over the long term as people already enrolled onto the ART have to continue receiving the life sustaining drugs. Yet, the decline in global development aid for HIV and AIDS has heightened the need for efficient governance of the response. Notions of shared responsibility by both domestic governments in low- and medium income countries and donor countries are being promoted. Nevertheless, as the quest to secure a sustainable response over the long haul continues, efficient ART governance is one avenue towards that goal. The application of TCE thus finds relevance in ART governance. The implication is that ART policy needs to continuously interrogate the existing ART governance mechanisms and the levels of transaction costs entailed towards optimisation. In particular, a review of the geographical location and population levels, and use of service delivery modes such as community-based HIV testing that are shown to be efficient may need to be explored.

9.4 Policy Recommendations

Application of a transaction cost economics approach to ART governance in Malawi would help to identify and address inefficiencies, uncertainties and areas that generate and drive transaction costs related to the delivery of ART services. This would not only assist in the optimisation of available resources to greatly improve the ART outputs and outcomes. The existing standardised clinical guidelines for ART are a useful tool towards efficient delivery as they minimise errors and prevent uneven care standards, but need to be complemented with other interventions designed to reduce transaction costs in ART services. More importantly, given the efficiency imperative that is largely

occasioned by dwindling international funding for HIV and AIDS, efficient implementation of the ART programme would contribute towards ART programme sustainability. Consequently, the results of the study have several policy implications for efficient ART service delivery with respect to the application of a transaction cost economics framework to ART governance. Data from the research suggests that optimal ART facility allocation should consider catchment population for efficient service delivery. The researcher intends to engage policy makers through dissemination of the findings to generate discussion around transaction costs and efficiency in the HIV and AIDS programme. This may include presentation in review meetings, such as the Joint Annual Review meetings, specifically around Technical Working Groups, particularly the Treatment, Care and Support Technical Working Group. Nevertheless, the following policy recommendations require the attention of the Malawian Government policy makers.

1. Review the location of ART healthcare facilities in view of the revealed allocative inefficiencies in sites with increasing and decreasing returns to scale. Patient load and catchment population are specific areas that require attention in decisions about service location and delivery. Accounting for universal and equitable access to ART, this requires the Ministry of Health to ensure that the location of ART facilities takes into account both patient load and catchment population;
2. Efficiency analysis of the ART Programme and broader HIV/AIDS response needs to consider the transaction costs that emanate from safeguards instituted to manage risk by financiers, including financial management and procurement agencies set up by donors. This is because in accounting terms, these resources are earmarked to ensure delivery of the HIV/AIDS response. There is, therefore, need to streamline the safeguards framework to ensure cost savings, as currently, these seem to be over-elaborate at the moment;
3. Consider institution of a harmonised and centralised procurement drug and supplies system for ART in view of the parallel arrangement that currently exists that consists of the Central Medical Stores and private procurement agencies. The current arrangement is viewed to be inefficient, and inconsistent with cost minimisation as championed by TCE theory as it increases supply chain transaction costs which have an impact on broader healthcare service delivery that include ART. A harmonised and centralised procurement system would

assist in the reduction of multiple negotiations, information search costs and would facilitate economies of scale in the acquisition and delivery of ART drugs and commodities;

4. The Government of Malawi and its partners should consider investing in adequate information systems and ICT and decentralised ledger technology, including block chain technology to reduce information asymmetries, improve communication and validation of information to reduce uncertainty in ART service delivery;
5. Explore and implement mechanisms to deliberately track transaction costs to inform efficiency considerations. This means that the existing AIDS resource tracking mechanisms have a specific category that tracks transaction costs in ART service delivery (health systems strengthening and administration costs, mainly above facility costs linked to ART services). This should aim at optimising the resources, including costs related to quarterly service quality monitoring, central level occupancy costs, facility level occupancy and utility costs and administration and management costs. Expenditure data from the NASA provides a starting point, which can be used with data from the DHA electronic management records. The staff inputs for those directly and indirectly involved in ART and HTS provision need to be adequately and consistently monitored, which implies the need for consistent data on this input in the NASA;
6. Institute allocative and technical efficiency monitoring in ART, as these can be influenced directly by management decisions in the health planning and financing system. Above facility costs would require decomposition to identify their drivers and where savings can be made- including administration, management costs, central coordination, capacity building, utility and occupancy costs;
7. Consider benchmarking efficiency of ART on the 48 facilities assessed to be efficient to improve efficiency in ART service delivery in the country, and,
8. In view of other studies in Malawi indicating the efficacy and efficiency of HIV testing, consider the active integration of community-led delivery of testing and counselling elements of the ART programme as an option to increase ART numbers.

10. REFERENCES

Abate, G. T., Dereje, M., Hirvonen, K., & Minten, B. (2019). *Geography of public service delivery in rural Ethiopia.* Adis Abbaba: IFPRI and The Federal Democratic Republic of Ethiopia Policy Studies Institute.

Abimbola, S., Ukwaja, K. N., Onyedum, C. C., Negin, J., Jan, S., & Martiniuk, A. L. (2015). Transaction costs of access to health care:Implications of the care-seeking pathways of tuberculosis patients for health system governance in Nigeria. *Global Public Health*, 1060-1077.

Accorsi, S., Dhaba, S., Vella, V., Tadesse, L., S. B., Hirpa, W., Bilal, N. K. (2010). Monitoring Hospital Performance: the challenges of efficiency in service delivery in Ethipopian Hospitals. *Quarterly Health Bulletin*, 28-46.

Adesina, A., Couture, T., & McGinn, E. (2015). *Estimating the Costs of Nonintegrated and Integrated Family Planning and HIV Facility Services in Malawi.* Washingto DC: Futures Group, Health Policy Project.

Agafonow, A. (2017). Transaction Cost and Business Ethics. In A. M. D.C. Poff, *Encyclopedia of Business and Professional Ethics* (pp. 1-4). online,DOI 10.1007and 978-3-319-23514-1_175-1, viewed February 9,2021: Springer International Publishing AG 2017 .

African Union. (2012). Establishing Shared Responsibility and Global Solidarity as a vision for global health in the Post-2015 development agenda, at https:and and apf.francophonie.org, viewed February 15, 2018

Agboola, A. (2015). Neoclassical Economics and New Institutional Economics. *Property Management. 33.*, 412-429.

Alvarez, R., & Crespi, G. (2003). Determinants of Technical Efficiency in Small Firms. *Small Business Economics*, 233-244.

Ankarloo, D., & Palermo, G. (2004). Anti-Williamson: a Marxian Critique of New Institutional Economics. *Cambridge Journal of Economics, 28*, 413-429.

Atun R, Chang AY, Ogbuoji O, et al (2016). Long-term financing needs for HIV control in sub-Saharan Africa in 2015–2050: a modelling study. BMJ Open 2016;6

Arrow, K. J. (1974). *The limits of organization.* New York: WWNorton & Company.

Babalola, T. K., & Modley, I. (2020). Assessing the Efficiency of Healthcare facilities in Sub-Saharna Africa: A systematic review. *PubMed Central*, Viewed online, December 15, 2023.

Banda, C. (2019). Adminsitrative Justice, Environmental Governance and the Rule of Law in Malawi. *Maryland Journal of International Law, 34*, 20-72.

Banker, R. D., Charness, A., & Copper, W. W. (1984). Some models for estimating technical and scale efficiencies in data envelopment analysis. *Management Sciences, 30(1)*, 1078-1092.

Bautista-Aredondo, S., Gosa-Rubi, S., Opuni, M., Kwan, A., Chaumont, C., & Coetzee, J. (2014). Assessing cost and technical efficiency of HIV prevention interventions in sub-Saharan Africa: the ORPHEA study design and methods. *BMC Health Services Research, 14*(599), 1-11.

Bautista-Arredondo, S., Hera-Fuentes, G. L., Contreras-Loya, D., Kwan, A., Van Buren, S. J., Amanze, O. O., . . . Sosa-Rubõ, S. G. (2018). Efficiency of HIV services in Nigeria:Determinants of unit cost variation of HIV counseling and testing and prevention ofmother-to-child transmission interventions. *PLoS ONE 13(9): e0201706. https:and and doi.organd 10.1371and journal.pone.0201706*, 1-19.

Bendoly, E., Rosenzweig, E. D., & Stratman, J. K. (2009). The efficient use of enterprise information for strategic advantage: A data envelopment analysis. *Journal of Operations Management , 27*, 310-323.

Berg, N., Radicke, F., Stentzel, U., Hoffman, W., & Flessa, F. (2019). Economic efficiency versus accessibility:Planning of the hospital landscape in rural regions using a linear model on the example of paediatric and obstetric wards in the northeast of Germany. *BMC Health Services Research, https:and and doi.organd 10.1186and s12913-019-4016-2*, 1-13.

Besley, T., & Ghatak, M. (2009). Property Rights and Economic Development. In D. Rodrik, & M. Resenzweig, *Handbook of Development Economics* (pp. 3-92). London: London School of Economics.

Bjurek, H., Hjalmarsson, L., & Forsund, F. (1990). Determinsitic parametric and non-parametric estimation of efficiency in service production: A Comparison. *Journal of Econometrics, 46(1and 2),*, 213-227.

Briere, M., Lehalle, C.-A., Nefedova, T., & Raboun, A. (2020, June 30). *Modelling Transaction Costs when Trades May Be Crowded: A Bayesian Network Using Partially Observable Orders Imbalance.* Retrieved from onlinelibrary.wiley.com, viewed October 30, 2023: https:and and onlinelibrary.wiley.com

BulstraI, C. A., Hontelez, J. A., Otto, M., Stepanova, A., Lamontagne, E., Yakusik, A., . . . Ba¨rnighausen, T. (2021). Integrating HIV services and other health services: A systematic review and meta-analysis. *PLoS Medicine*, 1-31.

Buitelaar, E. (2004). A Transaction Costs Analysis of Land Development Process. *Urban Studies, Volume 41, Number 13*, 2539-2553.

Burrows, G, Oberth, G, Parsons, D and McCallum, L. (2016). Transitions from Donor funding to domestic reliance for HIV countries: recommendations for transitioning countries, AMG Global & Aidspan

Buse, K and Martin, G. (2012). AIDS: Ushering in a new era of shared responsibility for global health, Globalization and Health 2012; at http:and and www.globalizationandhealth.comand content and 8 and 1and 26, viewed March 12, 2018

Cane, R. M., Melesse, D. Y., Kayeyi, N., Manu, A., Wado, Y. D., Barros, A., & Boerma, T. (2021). HIV trends and disparities by gender and urban–rural residence among adolescents in sub-Saharan Africa. Reprod Health 2, 1-10.

Coase, R.H. (1937). The Nature of the Firm, *Economica*, New Series, Vol. 4, No. 16. (Nov., 1937), pp. 386-405.

Carlson, C., Chirwa, W. C., Hall, N., & Cammack, D. (2015). *Study of Health Sector Efficiency in Malawi.* Lilongwe: Government of Malawi and Options.

Cordella, A., & Simon, K. A. (1997). “The Impact of Information Technology on Transaction and Coordination Cost”, Proceeding of Scandinavian Conference on Information System, . IRIS20.

Chadad, F. (2009). *Both Market and Hierarchy: Understanding the Hybrid Nature of Cooperative.* Columbia: Univeristy of Missouri and Insper.

Chansa, C., Mwase, T., Matsebula, T. C., Kandoole, P., Revill, P., Makumba, J. B., & Lindelow, M. (2018). Fresh Money for Health? The (False?) Promise of “Innovative Financing” for Health in Malawi. *Health Systems and Reform*, 324-335.

Chassagnon, V. (2014). Between Markets and Hierarchy: an economic analysis of quasi-vertical integration. *revue de philosophie économique, Vol 15, No. 1*, 135-165.

Chen, J. (2021). *Public Goods:Re-examination, related concepts and Disticntions.* Retrieved November 3, 2023, from https:and and doi.org

Chibambo, T. L. (2009). *Cultural practices among the Sena in Nsanje: HIV and AIDS and the African Evangelical Church.* Durban: University of Kwazulu Natal.

Claude, C. M. (2018). Research frontiers of new institutional economics. *RAUSP Management Journal, vol. 53, no. 1*, 3-10.

Coelli, T. (2008). *A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program, CEPA Working Paper 8and 96.* Armidale: University of New England, http:and and www.une.edu.auand econometricsand cepawpand htm.

Contreras, Ó. F., & Zanarone, G. (2018). Managing Social Comparison costs in organizations. *Working Papers, No. 2018-25, Banco de México, Ciudad de*.

Cordella, A. (2006). Transaction costs and information systems: does IT add up? *Journal of Information Technology 21*, 195-2002.

Cheruiyot, K. J. (2017). *Determinants of Technical Efficiency in Kenyan Manufacturing firms.* Washingtn DC: World Bank.

Cuenca, J. S. (2020). Efficiency of local governments in health service delivery: A stochastic frontier analysis. *PIDS Discussion Paper Series, No. 2020-06, Philippine Institute for Development Studies (PIDS), Quezon City*, 1-22.

Cuypers, I. R., Hennart, J.-P., Silverman, B. S., & Ertug, G. (2021). Transaction Cost Theory: past progress, current challenges, and suggestions for the future. *Academy of Management Annals, 15*(1), 111-150.

Dalton, A. (2017). 'Just take a tablet and you'll be okay': medicalisation, the growth of stigma and the silencing of HIV, HIV Nursing 2017; 17: 63–68

David, S. (2008). Increasing Firm Efficiency usign Trasnaction Cost Economics, *feaa.ucv.ro.* Retrieved October 30, 2023, from https:and and feaa.ucv.ro

Deitrich, M., Krafft, J., & McHardy, J. (2015). Real Firms, Transaction Costs and Development. *SERPS No 2015004,www.sheffield.ac.ukand economics*, 1-28.

Ding, Y., Pulford, J., & Bates, I. (2020). . Practical actions for fostering cross-disciplinary global health research: lessons from a narrative literature review. *BMJ Global Health 2020;5:e002293. doi:10.1136and bmjgh-2020-002293*, 1-13.

Dick-Sagoe, C., Asare-Nuamah, P., & Dick-Sagoe, A. D. (2021). Public choice and decentralised healthcare service delivery in Lesotho: Assessing improvement and efficiency in service delivery. *Cogent Social Sciences*, 1-16.

Dougherty, S., Lorezoni, L., Marino, A., & Murtin, F. (2022). The impact of decentralisation on the performance of health care systems: a non-linear relationship. *European Journal of Health Economics*, 705-715.

Dover, D. C., & Belon, A. (2019). The health equity measurement framework: a comprehensive model to measure social inequities in health. *International Journal for Equity in Health*, 1-12.

Drake, C., Anderson, D., Cai, S.-T., & Sacks, D. W. (2023). Financial Transaction costs reduce benefit take-up-evidence from zero premium health insurance plans in in Colorado. *Journal of Health Economics*, Online, viewed December 14, 2023.

Dresch, A., Collato, D., & Lucenda, D. (2018). Theoretical Understanding between competeitiveness and producitivity: firm level. *Inginiera y competitividad, vol.20, no.2*, 1-20.

Drost, E. A. (2011). Validity and Reliability in Social Science Research . *Education Research and Perspectives,https:and and www3.nd.eduand ~ggoertzand sgamethand Drost2011.pdf, 38*(21), 105-124.

Duong, V. H. (2016). Technical efficiency of FDI firms in the Vietnamese manufacturing sector, . *REVIEW OF ECONOMIC PERSPECTIVES – NÁRODOHOSPODÁŘSKÝ OBZOR, 16*(3), 205-230.

Edwards, E. C., & Guilfoos, T. (2020). *The Economics of Groundwater Goverance Instititions around the Globe (CenREP Working Paper No 20-001.* Raleigh: Center for Enironmental and Resource Economic Policy.

Etake, E. H. (2018). *An Analysis of the Technical Efficiency of Public Hospitals in Togo: a non-parametric approach.* Nairobi: The African Economic Research Forum.

Eze, P., Idemili, C. J., & Lawani, L. O. (2024). Evaluating health systems' efficiency towards universal health coverage: A data envelopment analysis. *INQUIRY: The Journal of Health Care*, 1-12.

Fernandez, J. (2016). Douglas North's Theory of Institutions: Lessons for Law and Development. *Hague Rule of Law*, 373-419.

Freeman, A., Nam-Speers, J., & Tokac, U. (2020). A quantitative meta analysis of organizational ownership and technical efficiency: non-linear influence by facility types and time for nonprofit and for-profit healthcare providers. *International Review of Public Administration,*, 106-128.

Friese, M., Heimeshoff, U., & Klein, G. (2018). *Property Rights and Transaction Costs - The Role of Ownership and Organization in German Public Service.* Düsseldorf: Düsseldorf University Press .

Galbács, P. (2017). Some methodological aspects of the controversy between mainstream economics and institutionalism. *CHOPE Working Paper, No. 2017-09,Duke University, Center for the History of Political Economy (CHOPE), Durham, NC*, 1-27.

Gama, E. S. (2013). *The Implications of Contracting Out Health Care Provision to Private Not-for-Profit Health Care Providers: the case of service level agreements in Malawi, PhD Thesis.* online at eresearch.qmu.ac.uk, viewed April 20, 2023.

Garbus, L. (2003). *HIV and AIDS in Malawi.* California: AIDS Policy Research Center, University of California San Francisco.

Government of Malawi.(1995). The Republic of Malawi Constitution, Lilongwe: Ministry of Justice

GoM, & Options. (2012). *Multi-Country Analysis of Treatment Costs for HIV and AIDS (MATCH): MALAWI.* Lilongwe: Ministry of Health.

GoM. (2013). *The National AIDS Spending Assessment-FY 2012and 2013.* Lilongwe: National AIDS Commission.

GoM. (2017). *Health Sector Strategic Plan II.* Lilongwe: Minsitry of Health.

GoM. (2017). *Integated HIV Programme Report-April to June 2017.* Lilongwe: Ministry of Health, Department of HIV and AIDS.

GoM, (2017). Health Sector Strategic Plan II, Lilongwe: Minsitry of Health.

GoM. (2017). *The National Community Health Strategy 2017-2022.* Lilongwe: Ministry of Health.

GoM. (2018). *Malawi Clinical HIVGuidelines.* Lilongwe: Ministry of Health.

GoM. (2020). Malawi Biological and Behavioral Surveillance Survey 2019-2020, Zomba: National Statistical Office.

GoM. (2020). *Health Sector Resource Mapping*. Lilongwe: Ministry of Health.

GoM .(2020). Malawi National HIV and AIDS Strategic Plan 2020-2025, Lilongwe: National AIDS Commission.

GoM .(2020). Malawi Poverty Report 2020, Zomba: National Statistical Office.

GoM. (2021). *National AIDS Spending Assessment (NASA) Report: FY 2015and 16, FY 2016and 17, FY 2017and 18 and FY 2018and 19.* Lilongwe: National AIDS Commission.

GoM. (2022). *Malawi Global AIDS Monitoring Report.* Lilongwe: National AIDS Commission.

GoM. (2022). *National Health Accounts.* Lilongwe: Ministry of Health.

GoM. (2023a). *Government of the Republic of Malawi Health Sector Strategic Plan III: 2023-2030-reforming for universal health coverage, First Edition.* Lilongwe: Ministry of Health.

GoM. (2023). *The Extended National HIV and AIDS Strategic Plan 2023-2027.* Lilongwe: National AIDS Commission.

GoM. (2023). The National Health Financing Strategy: 2023-2030. Lilongwe: Ministry of Health.

GoM. (2023a). *Government of the Republic of Malawi Health Sector Strategic Plan III: 2023-2030-reforming for universal health coverage, First Edition.* Lilongwe: Ministry of Health.

GoM .(2023a). Government of the Republic of Malawi Health Sector Strategic Plan-2023-2030: Reforming for Universal Health Coverage, Lilongwe: Ministry of Health.

Gräbner, C., & Ghorbani, A. (2019). *Defning institutions – A review and a synthesis, ICAE Working Paper Series, No. 89.* Linz: Johannes Kepler University Linz,Institute for Comprehensive Analysis of the Economy (ICAE).

Gulati, R., & Nickerson, J. (2008). Interorganizational Trust, Governance Choice, and Exchange Performance. *Organization Science 19(5)*, 688-708.

Hart, O., & Moore, J. (1990). Property Rights and the Nature of the Firm. *Journal of Political Economy*, 1119-1158.

Harries, N. D., Ford, N., Jahn, A., Scouten, E. J., Libamba, E., Chimbwandira, F., & Maher, D. (2016). Act Local, Think Global how the Malawi experience of

scaling up antiretroviral treatment has informed global policy. *BMC Public Health, 16*(938), 2-9.

Heiman, B., & Nickerson, J. A. (2001). *Towards Reconciling Transaction Cost Economics and the Knowledge-based View of the Firm: The Context of Interfirm Collaborations.* Berkeley: University of Calfifornia at Berkeley.

Holstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership and job design. *Journal of Law Economics and Orgnaization, 7*, 24-52.

Ibrahim, M. D. (2023). Efficiency and productivity analysis of maternal and infant healthcare services in Sub-Saharan Africa. *The International Journal of Health Planning and Management*.

Indravudh, P. P., Fielding, K., Kumwenda, M. K., Rabecca Nzawa, R. C., Desmond, N., Nyirenda, R., . . . Corbett, E. L. (2021). Effect of community-led delivery of HIV self-testing on HIV testing and antiretroviral therapy initiation in Malawi: a cluster randmised trial. *PLOS Medicine*, 1-18.

Institute of Health Metrics & Evaluation. (2022). *healthdata.org*. Retrieved from Institute of Health Metrics and Evaluation: https:and and www.healthdata.org

Ithibu, A., & Amendah, D. (2019). Domestic Financial Contributions to HIV, TB, and Malaria. Nairobi: Aidspan.

Khanakwa, S., & Mbonigaba, J. (2022). Institutional Arrangements for providing HIV and AIDS Services in Uganda: A Transaction Cost Economics Analysis. *Health Services Insights*, 1-13.

Krishna, K., Ross, E., & Chiossi, S. (2023). Securing political and financial support for the global HIV response: A case study of the post-2000 surge in donor funding. London: Chatham House.

Kamau, G., Boore, C., Maina, E., & Njenga, S. (2018). Block Chain Technology: is this the solution to EMR innteroprability and security issues in Developing Countries? IST-Africa 2018 Conference Proceedings. www.IST-Africa.org.

Kano, L., & Verbeke, A. (2015). The three faces of bounded realibility: Alfred Chandler and the microfoundations of management theory. *California Management Review, 58*(1), 97-122.

Kaupa, F., & Naude, M. J. (2021). Central Medical S. *African Journal of Governance and Development*, 33-59.

Ketokivi, M. (2020). Transaction Cost Economics as a Theory of Supply Chain Efficiency. *Production and Operations Management, Vol.29,No.4*, 1011-1031.

Ketokivi, M., & Mahoney, J. T. (2017). Transaction Cost Economics as a Theory of the Firm, Management, and Governance. *https:and and doi.organd 10.1093and acreforeand 9780190224851.013.6*, online, viewed Janury 31, 2021.

Kharsany, A. B., & Karim, Q. A. (2016). HIV infection and AIDS in sub-Saharan Africa: current status, challenges and opportunities. *Open AIDS Journal*, 34-48.

Kumar, S., & Gulati, R. (2008). An Examination of Technical, Pure Technical,and Scale Efficiencies in Indian Public Sector Banks using Data Envelopment Analysis . *Eurasian Journal of Business and Economics 2008, 1 (2),* , 33-69.

Kurup, K. K., Ross, E., & Chiossi, S. (2023). *Securing political and financial support for the global HIV response:A case study of the post-2000 surge in donor funding*. London: Chatham House.

Langlois, R. N. (1992). Transaction Cost Economics in Real Time. *Industrial and Corporate Change, Volume 1, Number 1*, 99-127.

Lawson, A. (2009). *Evaluating the Transaction Costs of Implementing the Paris Declaration.* Oxford: Fiscus Public Finance Consultants Ltd.

Leonard, K. (2002). Why both states and markets fail: assymentric information and the role of NGOs in African health care . *International Review of Law and Economics 22*, 60-81.

Lépine, A., Vassall, A., & Chandrashek, S. (2015). The determinants of technical efficiency of a large scale HIV prevention project: application of the DEA double bootstrap using panel data from the Indian Avahan. *Cost Effectiveness and Resource Allocation*, 2-11.

Lieberherr, E. (2009). *Policy Relevance of New Institutional Economics?Assessing Efficiency, Legitimacy and Effectiveness.* Delft University of Technology.

Lithman, E. (2014). *Transaction Costs and Development Aid through SIDA.* Stockholm: SIDA.

Locke, A. (2013). Property Rights and Economic Growth. *https:and and www.refworld.organd pdfidand 523ab6dc4.pdf, viewed January 30,2021*, 1-32.

Loppo, L., & Zenger, T. (1998). Testing Alternative Theories of the Firm: transaction cost, knowledge-based and measurements formake or break decisions in information services. *Strategic Management Journal*, 853-877.

Lye, A., Cicirello, A., & Patelli, E. (2021). Sampling methods for solving Bayesian model updating problems: A tutorial. Mechanical Systems and Signal Processing, 159, 1-43. Article 107760. https://doi.org/10.1016/j.ymssp.2021.107760

Luciana, M. (2008). The application costs to vertical integration decision: evidences from a Brazilian beef processor. *Organizações Rurais & Agroindustriais, 10*(3), 317-327.

Mahajan, V., Nauriyal, D. K., & Singh, S. P. (2018). Efficiency and Its Determinants:Panel Data Evidence from the Indian Pharmaceutical Industry, Margin. *The Journal of Applied Economic Research, 12*(1), 19-40.

Makowski, D., Wallach, D., & Trembley, M. (2002). Using a Bayesian approach to parameter estimation:comparison of the GLUE and MCMC methods. *Agronomie*, 191-203.

Marinescu, C. (2012). Transaction Costs and Institutions' Efficiency: A Critical Approach, American Journal of Economics and Sociology, Vol. 71, No. 2 (April, 2012).

Martinez, M., Espiter, E., Marta, F., & OlomoJorge, R. (2008). An Empirical Test of Transaction Cost Theory: validating the analysis of discrete sturctural alternatives. *University of Rioja Working Papers*, 1-30.

McGillen, J B, Sharp, A, Honermann, G M, Collins C and Hallett T B. (2017). Consequences of a changing US strategy in the global HIV investment landscape, AIDS 2017, 31:F19–F23, Wolters Kluwer Health, Inc.

Meyer-Rath, G., Johnson, L. F., YoganPillay, MarkBlecher, Brennan, A. T., LawrenceLong, . . . SydneyRosen. (2017). Changing the South African national antiretroviral therapy guidelines: The role of cost modelling. *PLoS One*, 1-15.

Munthali, S, M. (2010). An Institutional Analysis of Community and Home Based Care and Support for HIV and AIDS Sufferers in Rural Households in Malawi, Doctoral Thesis in Economics presented to the University Rhodes, South Africa.

Murad, A., Magersa, D., & Tewfik, B. (2017). Technical Efficiency of Selected Hospitals in Eastern Ethiopia. *Health Economics Review,Health Economics Review, ISSN 2191-1991, Springer,Heidelberg, Vol. 7, Iss. 24,http:and and dx.doi.organd 10.1186and s13561-017-0161-7*, 1-13.

MédecinsSansFrontières. (2019). Burden Sharing or Burden Shifting? How the HIV and TB Respeonse is being derailed. Brussels and Geneva: Médecins Sans Frontières.

Menzies NA, Berruti A A, Blandford JM .(2012). The Determinants of HIV Treatment Costs in Resource Limited Settings. PLoS ONE 7(11): e48726. doi:10.1371and journal.pone.0048726

Mick, E., & Shay, P. (2016). Accountable Care Organizations and Transaction Cost Economics. *Medical Care Research and Review, 73*(6), 649-659.

Meramveliotakis, G. (2018). New Institutional Economics: A Critique of Fundamentals & Broad Strokes Towards an Alternative Theoretical Framework for the Analysis of Institutions. *Asian Journal of Social Science Studies Vol. 3, No. 2;* , 50-64.

Messac, L. (2013). *pulitzer center*. Retrieved September 12and 09and 2023, 2023, from http:and and www.pulitzercenter.org

Mugwagwa, I., Bijman, J., & Trienekens, J. (2020). Typology of contract farming arrangements: a transaction cost perspective. *Agrekon, 59:2, DOI:10.1080and 03031853.2020.1731561*, 169-187.

Mulaga, A. N., Kamndaya, M. S., & Masangwi, S. J. (2022). Decomposing Scoioeconimc inequality in catastrophic out-of-pocket health expenditures in Malawi. *PLOS Glob Public Health, 2*(2), 1-13.

Murad, A., Magersa, D., & Tewfik, B. (2017). Technical Efficiency of Selected Hospitals in Eastern Ethiopia. *Health Economics Review,Health Economics Review, ISSN 2191-1991, Springer,Heidelberg, Vol. 7, Iss. 24,http:and and dx.doi.organd 10.1186and s13561-017-0161-7*, 1-13.

Muula, A. S. (2002). HIV and AIDS Prevention Efforts in Malawi. *Malawi Medical Journal*(Review Article), 26-27.

Mwenge, L., Sande, L., Mangenah, C., Ahmed, N. K., Sibanda, E., Kalua, T., . . . Terris-Prethold, F. (2017). Costs of facility-based HIV testing in Malawi,Zambia and Zimbabwe, PLoS ONE 12(10): e0185740. https:and and doi.organd 10.1371and journal.pone.0185740. *PLoS ONE*, 1-16.

Nagle, F., Seamans, R., & Tadelis, S. (2020). Transaction Cost Economics in the Digital Economy: A Research Agenda . *Working Paper 21-009, Harvard Business School, online, viewed January 31, 2021*.

Nickerson, J., & Yen, C. J. (2018). Transaction Cost Economics. In M. T. Augier, *The Pelgrave Encyclopedia of Strategic Management.* (pp. https:and and doi.organd 10.1057and 978-1-137-00772-8_569). London,: Pelgrave Macmillan,.

Nickerson, J., & Zenger, T. R. (2008). Envy, Comparison Costs and the Economic Theory of the Firm. *Strategic Management Journal*, 1-21.

Nimubona, A. (2022). Health diplomacy to promote multisectoral participation in fighting against fragmentation and increasing budget for internalization of the health financing progress matrix in Burundi. *Health Econoics Review*, 12-31.

Nguyena V, Nathalie Bajosb, et al. (2010). Vinh-Kim Nguyena, , Franc¸oise Dubois-Arber , F, O’Malleye, J and Pirkle. CM (2010). Remedicalizing an epidemic:

from HIV treatment as prevention to HIV treatment is prevention , AIDS 2010, Vol 24 No 00

North, D. C. (1971). Institutional Change and Economic Growth, The Tasks of Economic History. *The Journal of Economic History, Vol. 31, No. 1*, 118-125.

North D.C. (1993). The New Institutional Economics and Development. *Economic History 9309002, University Library of Munich, Germany*.

Oberth, G and Whiteside, A .(2016). What does sustainability mean in the HIV and AIDS response?, African Journal of AIDS Research, 15: 1–9

Obińska-Wajda, E. (2016). The Institutional Economics-main theories. *Financial Internet Quarterly, eFinanse, Vol 12and nr 1,s.*, 78-85.

Obloj, T., & Zenger, T. (2015). Incentives, Social Comaprison and the Proximity of Envy's Object. *HEC Paris Research Paper*, 1-35.

Oderkirk, J., & Slawomirski, L. (2020). *Policy Brief: Opportunities and Challnges of Blockchain Technologies in health care.* OECD.

Oliveira, R., Santinha, G., & Marques, T. S. (2024). The Impacts of Health Decentralization on Equity, Efficiency,and Effectiveness: A Scoping Review. *Sustainability 2024*, 1-35.

Penman-Aguilar, A., Talih, M., Huang, D., Moonesinghe, R., Bouye, K., & Beckles, G. (2018). Measurement of Health Disparities, Health Inequities, and Social Determinants of Health to Support the Advancement of Health Equity. *J Public Health Manag Practice*, 1-17.

Pérard, E. (2009). Water Supply: Public or Private?, Policy and Society, 27:3. *An approach based on cost of funds, transaction costs, efficiency and political costs*, 193-219.

Plunket, A., & Saussier, S. (2003). Theories of the firm: How to rule out competing views? *Economie & Institutions, 3*(2), 1-15.

Piatti-Fünfkirchen, M., Chansa, C., & Nkhoma, D. (2020). Public Financial Management in the Health Sector: An assessment at the local government level in Malawi. Washington DC: The World Bank.

Pindyck, R. S., & Rubinfeld, D. L. (2018). *Microeconomics* (9th ed.). Pearson Education.

Polishchuk&GeorgiySyunyaev, R. (2011). New Institutional Economics: A State-of-the-Art Review for Economic Sociologists. *economic sociology_the european electronic newsletter*, 12-21.

Poniatowicz, M. (2017). Transaction Costs in the Context of the Requirements of Effective Finance Management for Local Government Unit. *Ekonomia i Prawo. Ecomomics and Law, 16(3):,doi:10.12775and EiP.2017.023.*, 325-341.

Reimers, K., Guo, X., & Li, M. (2018). Beyond Markets, Hierarchies and Hybrids: an institutional perspective on IT-enabled two-sided markets. *Springer, Electronic Markets, published online,https:and and doi.organd 10.1007and s12525-018-0319-0*, 1-19.

Resch, S, Ryckman, T, and Hecht, R. (2015). Funding AIDS programmes in the era of shared responsibility: an analysis of domestic spending in 12 low-income and middle-income countries, *Lancet Glob Health* 2015; 3: e52–61

Reuster, S. (2010, January 14). *Recent Developments in Transaction Cost Economics, Resource Market Working Paper No. RM- 18.* Retrieved August 12, 2023, from http:and and dx.doi.organd 10.2139and ssrn.1535903

Rindfleisch, A. (2019). Transaction cost theory: past, present and future. *Academy of Marketing Science,https:and and doi.organd 10.1007and s13162-019-00151-x*, 1-14.

Ritchter, R. (2005). The New Institutioal Economics-Its Start, Its Meaning, Its Prospects. *The European Business Organization Law Review (EBOR)*, 161-200.

Rudasingwa, M., Muula, A., Yeboah, E., Bonnet, E., Ridde, V., & Allegri, M. d. (2021). *Inequlaities on health care spending on curative services: Evidnece from Malawi.* Paris: Agency Francoise de Development (AFD).

Sautter, C. M. (2021). Transaction Cost Economics & MAES: The Deal Maker's Crystal Ball. *Foldham Law Review Onine*, 41-55.

Seed, H., Malik, H., Bashir, U., Ahmed, A., Riyaz, S., llyas, M., Khan, M. I. (2022). Blockchain Technology in Healthcare: a systematic review. *PLoS ONE, e0266462, https:and and doi.organd 10.1371and journal.pone.0266462, 17*(4), 1-31.

Shervani, T. A., Frazier, G., & Challagalla, G. (2007). The Moderating Influence of Firm Market Power on the Transaction Cost Economics Model: an emprical test in a forward market channel integration context. *Strategic Management Journal, No.28*, 635-662.

Songane, M., Mukherjee, S., Chamanga, R., Maphosa, R., Longwe, T., Namathanga, B., . . . Tiam, A. (2024). Cost of Providing Advanced HIV Disease Treatment Services through Malawi’s Hub-and-Spoke Model. *The American Journal of Tropical Medicine and Hygiene.*, 897-903.

Stafford, R. D. (2020). *"Analysis of Beta Distribution for Subjective Uncertainty Analysis in Cost Models".* Ohio: Air University.

Starrs, A M, et al. (2018). Accelerate progress—sexual and reproductive health and rights for all: report of the Guttmacher–Lancet Commission. Lancet 2018, 2642–92.

Shukla, M., Argawal, M., Singh, J. V., Tripathi, A. K., Srivastava, A. K., & Singh, V. K. (2016). Nonadherence to Antiretrovia=ral Therpay Among People Living with HIV and AIDS attending Tow Tertuary Care Hospitals in District of India. *Indian Journal of Community Medicine , 41*(1), 55-61.

Spithoven, A. (2012). Public governance of healthcare in the United States: a transaction costs economics (TCE) analysis of the 2010 reform. *Journal of Institutional Economics, 8*(4), 429-457.

Siberberg, E., & Suen, W. (2001). *The Structure of Economics: A Mathematical Analysis, Third Edition.* New York: McGraw Hill.

Singh, S. (2017). Do transaction costs affect technical efficiency? Case of Indian microfinance institutions. *International Journal of Management and Development Studies*, 16-25.

Shresha, N. (2020). Detecting multicollinearity in Regression Analysis. Amrican Jornal of Appplied Mathematics and Statistics, 39-42.

StataCorpLLC. (1985-2021). Stata Statistical Software: Release 17. College Station, TX.

Stover J, Bollinger L, Izazola JA, Loures L, DeLay P, Ghys PD, et al. (2016) What Is Required to End the AIDS Epidemic as a Public Health Threat by 2030? The Cost and Impact of the Fast-Track Approach. PLoS ONE 11(5): e0154893. doi:10.1371and journal. pone.0154893

Steinecker, M., & Czirfuszová, M. (2021). Examining the three types of medical care centres in relation to transaction cost theory. *International Journal of Health*, 64-70.

Tadelis, A., & Williamson, O. (2012). Transaction cost Economics. In R. Gibbons, & J. (. Roberts, *Handbook of Organizational Economics* (pp. 1-55). New Jersey: Princeton Unveristy Press.

Tagar E, Sundaram M, Condliffe K, Matatiyo B, Chimbwandira F, et al. (2014) Multi-Country Analysis of Treatment Costs for HIV and AIDS (MATCH): Facility-Level ART Unit Cost Analysis in Ethiopia, Malawi, Rwanda, South Africa and Zambia. PLoS ONE 9(11): e108304. oi:10.1371and journal.pone.0108304

Todorova, T. (2011). Adverse effects of transaction costs in East European Economies. *Organizations and Markets in Emerging Economies*, 34-50.

Tobias, M. (2020). *A case study of Public Private Partnerships in the Health Sector in Malawi.* Harare: EQUINET.

Todorova, T. (2016). Transaction Costs, Market Failure and Economic Development. *Journal of Advanced Research in Law and Economics, ISSN 2068-696X, ASERS Publishing, s.l., Vol. 7, Iss. 3(17), Summer*, 1-18.

Theodore, B. R., Whittington, J., Towle, C., Tauben, D. J., Endicott-Popovsky, B., Cahana, A., & Doorenbos, A. Z. (2015). Transaction Cost Analysis of In-Clinic Versus Telehealth Consultations for Chronic Pain: Preliminary Evidence for Rapid and Affordable Access to Interdisciplinary Collaborative Consultation. *Pain Medicine*, 1045–1056.

Trachtman, J. (1997). The Theory of the Firm and the Theory of International Economic Organization: Toward Comparative Institutional Analysis. *17NW, Journal of International Law and Business,470*, 470-555.

Triponel, A. (2009). Global Fund to Fight AIDS, Tuberculosis and Malaria: A New Legal and Conceptual Framework for Providing International Development Aid. *North Carolina Journal of International Law, 35*(1), 174-232.

Twea, P., Manthalu, G., & Mohan, S. (2020). Allocating resources to support universal health coverage: policy processes and implementation in Malawi. *BMJ Global Health*, 1-5.

UNAIDS. (2021). *End AIDS, End Inequalities.* Geneva: UNAIDS.

UNAIDS. (2021). *End Inequalities. End AIDS. Global AIDS Strategy 2021-2026.* Geneva: UNAIDS.

UNAIDS. (2022a). *IN DANGER: UNAIDS Global AIDS Update 2022.* . Geneva: Joint United Nations Programme on HIV and AIDS, 2022. Licence: CC BY-NC-SA 3.0 IGO.

UNAIDS. (2022b). A Pandemic Triad: HIV, COVID-19 and Debt in developing countries. Geneva: UNAIDS.

UNAIDS. (2022c). *UNAIDS DATA 2022.* Geneva: UNAIDS.,

UNAIDS. (2023). World AIDS Day 2023 Global AIDS Statistics. Geneva: United Nations Joint Programme on HIV and AIDS.

UNAIDS. (2023). *World AIDS Day 2023 Factsheet.* Geneva: UNAIDS.

US Department of Health and Human Services, O. o. (2010). *The National Plan for Action Draft as of February 17, 2010.* Washington: US Department of Health and Human Services, Office of Minority Health.

Vassal A, Remme M, Watts C, Hallett T, Siapka M, et al. (2013). Financing Essential HIV Services: A New Economic Agenda. PLoS Med 10(12): e1001567. doi:10.1371and journal.pmed.1001567

Victor, J., & Paulo, A. (2023). Transaction costs in healthcare: empirical evidence from Portuguese hospitals. *Cogent Business & Management, 10*(2), 1-26.

Vitola, A., & Senfelde, M. (2015). *Business: Theory and Practice, 16(3),doi:10.3846and btp.2015.498*, 271-279.

Vyas, S., Songo, J., Guinness, L., Dube, A., Gies, S., Kalua, T., . . . Wringe, A. (2020). Assessing the costs and efficiency of HIV testing and treatment services in rural Malawi: implications for future "test andstart" strategies. *BMC Health Services Research,2020,20:740, https:and and doi.organd 10.1186and s12913-020-05446-5*, 1-11.

Wahab, S. A., Rose, R., & Osman, S. I. (2012). The Theoretical Perspectives Underlying Technology Transfer: A Literature Review . *International Journal of Business and Management*, 277-288.

Wang, N. (2003, February 21). *Measuring Transaction Costs: An incomplete survey.* Retrieved July 2019, from http:and and www.coase.organd workingpapersand wp-2.pdf.

Weinstein, O. (2012). " Firm, Property and Governance: From Berle and Means to the Agency Theory, and Beyond. *Accounting, Economics and Law, Vol. 2, Iss. 2, Article 2,DOI: 10.1515and 2152-2820.1061*, 1-55.

Wallis, J and North D. (1986). Measuring the Transaction Sector in the American Economy,1870-1970, http:and and www.nber.organd chaptersand c9679, viewed May 25, 2016

Wild, L., & Cammack, D. (2013). *The supply and distribution of essential medicines in Malawi.* London: Overseas Development Institute.

Wooldridge, J. M. (2006). *Introductory Econometrics: A modern approach* (3 ed.). Mason: Thomson South Western.

Wilson, N., Zang, H., & Robinson, A. (2003). Specific human capital and partial employee ownership: a trasaction cost analysis. *Advances in the Economic Analysis of Participatory and Labour Manged Firms*, 211-226.

Williamson, O. (1975). *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: Free Press.

Williamson O. E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations, Journal of Law and Economics, Vol. 22, No. 2 (Oct., 1979), pp. 233-261

Williamson, O. E. (1981). The Economics of Organization: a Transaction Cost Approach. *The Amercian Journal of Sociology, Vol. 87, No. 3*, 548-577.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism.* New Yorkand London: The Free Pressand Collier Macmillan Publishers.

Williamson, O. E. (1991). Comperative Economic Organization: The Analysis of Discrete Structural Alternatives. *Adminstrative Science Quarterly, (36)*, 269-296.

Williamson, O. (1993). Transaction Cost Economics and Organization Theory. *Industrial and Corporate Change,2. 10.1093and iccand 2.1.107.*, 77-107.

Williamson, O. (1995). Markets, Hierarchies and Power in the Economy. *https:and and www.reseachgate.netand publicationsand 5212637, viewed December 3, 2022*, 21-49.

Williamson, O.E. (1999). Public and private bureaucracies: a transaction costs economics perspective. Journal of Law, Economics, and Organisation 15, 306–342.

Williamson O,E.(2000). The New Institutional Economics: Taking Stock, looking ahead, Journal of economic literature, Vol. 38, No. 3, 595-613

Williamson, O. (2002). The Lens of Contract: Private Ordering. *The American Economic Review, 92*(2), 438-443.

Williamson, O.E. (2007). Transaction Costs: An Introduction, at www.economics-ejournal.organd economicsand discussionpapers, accessed, March 28, 2018

Williamson, O. (2008). Comparative Economic Organization: the analysis of discrete structural alternatives. *Adminstrative Science Quarterly 36*, 269-296.

Williamson, O. (2010). Transaction Cost Economics: The Natural Progression. *The American Economic Review, Vol 100, No.3*, 673-690.

Williamson, O., & Ghani, T. (2012). Transaction cost economics and its uses in marketing. *Journal of the Academic and Marketing Sciences, 40, DOI 10.1007and s11747-011-0268-z*, 74-85.

Williamson, O. (2014). The Transaction Cost Economics Project. *Montenegrin Journal of Economics, Vol. 10, No.1 (July 2014)*, 7-14.

Wexler, A., Kates, J., Lief, E., & UNAIDS. (2022). Donor Government Funding for HIV in Low- and Middle-Income Countries in 2021. San Francisco: KFF.

World Bank and Global Financing Facility.(2021). Spending for Health in Malawi: current trends and strategies to improve efficiency and equity in health financing, Washington DC: World Bank.

Yousuf, A. (2017). Trasaction Costs: a conceptual framework. *International Journal of Engineering and Management Sciences, 2*(3), 131-139.

Zimbauer, D. (2001). From Neo-classical Economics to New Institutional Economics and Beyond-prosepcts for an interdiscsiplinary research programme? *LSE Development Stuides Institute Working Paper Series, No.01-12*, 1-19.

Zeng, W., Mphwanthe, H., Huan, T., Nam, J., & A.Dutta. (2017). *Cost-Effectiveness of Malawi's Service Level Agreement with the Christian Health Association of Malawi.* Washington DC: Palladium, Health Policy Plus.

Zeng, W., Rwiyereka, A. K., Amico, P. R., vila-Figueroa, C. A., & Shepard, D. S. (2014). Efficiency of HIV and AIDS Health Centers and Effect of Community-Based Health Insurance and Performance-Based Financing on HIV and AIDS Service Delivery in Rwanda. *American Journal of Tropical Medicine and Hygiene, 90*(4), 740-746.

Zeng, W., Shepard, D. S., Chilingerian, J., & Avila-Figueroa, C. (2012). How much can we gain from improved efficiency? An examination of performance of national HIV and AIDS programs and its determinants in low- and middle-income countries. *BMC Health Services Research, 12*(74), 1-9.

Zogning, F. (2017). Agency Theory: A Critical Review. *European Journal of Business and Management, Vol 9, No. 2*, 1-8.

11. APPENDIXES

Appendix 1: Sampling frame for qualitative study

#	Institution
	External/International Institutions
1	AFRICARE
2	AFRICAV DEVELOPMENT BANK
3	AMERICAN INSTITUTE FOR HEALTH
4	ASSOCIATION SCHOOL OF PUBLIC HEALTH
5	ACADEMY FOR EDUCATIONAL DEVELOPMENT
6	BAYLOR COLLEGE OF MEDICINE CHILDRENS FOUNDATION (BCMCF)
7	BAOBAB HEALTH PARTNERSHIP
8	CARE INTERNATIONAL
9	CATHOLIC RELIEF SERVICES (CSR)
10	CENTRE FOR COMMUNITY ORGANISATION & DEVELOPMENT (CCOD)
11	CHRISTIAN AID
12	CPAR
13	DIGNITAS
14	DANISH CHURCH AID
15	Dev Tech
16	DFID
17	EUROPEAN UNION
18	THE ELIZABETH GLAZER PAEDIATRIC AIDS FOUNDATION (EGPAF)
19	FAMILY HEALTH INTERNATIONAL (FHI)
20	FAO
21	GTZ
22	GH Tech
23	GOAL MALAWI
24	HOWARD UNIVERSITY
25	INTERNATIONAL LABOUR ORGANISATION (ILO)

26	IRISH AID
27	INTRAHEALTH INTERNATIONAL Inc
28	JOHN HOPKINS UNIVERSITY CENTRE FOR COMMUNICATIONS PROGRAMME
29	JOHN SNOW Inc
30	JHPIEGO
31	KNVC TB FOUNDATION
32	LAND O'LAKES
33	MSF-BELGIUM
34	MACRO INTERNATIONAL
35	NORWAY (THE KINGDOM OF NORWAY)
36	NORWEGIAN CHURCH AID
37	OXFARM
38	PACT Inc
39	PACT MALAWI
40	PARTNERS IN HOPE
41	PARTNERSHIP FOR CHILDHEALTHCARE Inc
42	PLAN MALAWI
43	PROGRESSION
44	PROJECT MALAWI
45	PROJECT CONCERN INTERNATIONAL
46	PROJECT PEANUT BUTTER
47	POPULATION SERVICES INTERNATIONAL
48	PROFESSIONAL AND SCIENTIFIC ASSOCIATION
49	SAVE THE CHILDREN
50	SWAM
51	UNAIDS
52	UNICEF
53	UNFPA
54	UNDP

55	UNIVERSITY OF WASHINGTON
56	UNIVERSITY OF NORTH CAROLINA
57	THE USG (USAID, PEPFAR, CDC)
58	WORLD FOOD PROGRAMME (WFP)
59	WORLD BANK
60	WORLD HEALTH ORGANISATION (WHO)
61	WORLD VISION
	Public Institutions
62	MALAWI POLICE
63	MALAWI PRISON
64	MALAWI DEFENCE FORCE
65	MALAWI REVENUE AUTHORITY
66	M INISTRY OF LOCAL GOVERNMENT
67	MINISTRY OF AGTRICULTURE
68	MINISTRY OF EDUCATION SCIENCE AND TECHNOLOGY
69	MINISTRY OF GENDER
70	MINISTRY OF HEALTH
71	MINISTRY OF YOUTH DEVELOPMENT AND SPORTS
72	MINISTRY OF CULTURE AND TOURISM
73	DEPARTMENT OF NUTRITION, HIV AND AIDS
74	MINISTRY OF WOMEN, CHILDREN AND COMMUNITY DEVELOPMENT
75	MINISTRY OF INFORMATION
76	MINISTRY OF FINANCE
77	BLANTYRE WATER BOARD
78	LILONGWE WATER BOARD
79	CENTRAL REGION WATER BOARD
80	NATIONAL AIDS COMMISSION
81	UNIVERSITY OF MALAWI
82	UNIVERSITY OF MALAWI COLLEGE OF MEDICINE
83	NATIONAL BANK OF MALAWI

	Private Institutions
84	ALLIANCE ONE TOBACCO (MW)Ltd
85	LIMBE LEAF TOBACCO
86	MALAWI BUSINESS COALITION AGAINST AIDS
87	MALAWI BLOOD TRANSFUSION
88	SUNBIRD TOURISM
89	MALAWI RURAL FINANCE COMPANY
90	MALAWI CONGRESS OF TRADE UNION
91	UNILEVER SOUTH EAST Ltd
92	MALAWI RED CROSS
93	SOUTHERN BOTTLERS
94	CATHOLIC UNIVERSITY OF MALAWI
95	LIGHTHOUSE
96	MUSLIM ASSOCIATION OF MALAWI
97	ACTIVE YOUTH INTERVENTION FOR SOCIAL ENRICHMENT
98	MACRO
99	PAKACHERE HEALTH PROMOTION & COMMUNICATION
100	CONSOL HOMES
101	FAMILY PLANNING ASSOCIATION
102	CHISOMO CHILDRENS CLUB
103	MALAWI INTERFAITH AIDS ASSOCIATION
104	JOURNALIST ASSOCIATION AGAINST AIDS
105	GUIDANCE, COUNSELLING & YOUTH DEVELOPMENT FOR AFRICA
106	FEDOMA
107	BANJA LA MTSOGOLO
108	ADRA
109	MANASO
110	NAPHAM
111	EPISCOPAL CONFERENCE OF MALAWI
112	CHRISTIAN HEALTH ASSOCIATION OF MALAWI

113	QUADRIA MUSLIM ASSOCIATION
114	COALITION OF WOMEN LIVING WITH HIV
115	MATINDI YOUTH ORGANISATION
116	LIFELINE MALAWI
117	MANET++
118	BLANTYRE SYNOD HEALTH & DEVELOPMENT COMMISSION
119	WORD ALIVE COMMISSION FOR RELIEF DEVELOPMENT
120	MANERALA+
121	WILSA

Appendix 2: Qualitative Data Analysis Themes

#	*Theme*	*Sub-theme*	*Relevant thesis Chapter*
1	HIV and AIDS ART service delivery organisation	• Governance instrument choice • Service delivery structure • Key actors and relationships • Healthcare/ART supply chain governance	HIV and AIDS ART governance conformity with Transaction Cost Economics theory (Chapter 6) and the TCE of HIV and AIDS ART (chapter 7)
2	HIV and AIDS ART Planning and resource allocation	• Efficiency considerations • ART access for vulnerable population • Efficiency monitoring • Monitoring transaction costs	Malawi HIV and AIDS programme (in) efficiency and the effect of transaction costs on efficiency (Chapter 5)
3	HIV and AIDS ART facility level service delivery	• Capacity to deliver services-staff, buildings • Service delivery mode-separate ART clinic days or integrated into daily routines • Service quality • Challenges with services	Determinants of HIV and AIDS ART efficiency and TCE conformity (Chapter 5 & 6)
4	HIV and AIDS ART financing	• HIV and AIDS ART financing administration • Fiduciary compliance and Safeguards	Transaction cost economics of the HIV and AIDS ART programme (Chapter 7)

Appendix 3a: Estimating Transaction Costs in the Malawi ART Programme

1.0 Introduction

The primary objective of this Appendix is to contribute to the clarification and hence operationalization of the transaction costs concept in ART service delivery. The secondary objective is to estimate transaction costs levels based on existing literature, expenditure and resource tracking mechanisms. The Appendix starts with a review of the transaction costs in general and how they apply to HIV and AIDS services. This is followed by a review of current mechanisms for tracking HIV and AIDS resources in the country and the extent to which transaction costs are an exclusive focus to aid in the management of the ART programme towards technical efficiency. In the third sub-section of the Appendix, results related to approximation of facility-based ART expenditure are presented, alongside the estimation of transaction cost levels for the ART programme in 2018and 2019. Transaction costs are estimated as a component of the total costs through a Bayesian scheme. While the analysis does not cover all elements of the HIV and AIDS programme, it provides an indication of where the broad areas for attention are with regard to the monitoring of transaction costs in ART and its application in HIV and AIDS service delivery efficiency calculations.

1.1 Revisiting the Transaction Costs Concept

Transaction Cost Economics projects the idea that it is important to look at how transactions in the production of a product or service are organised, and how they impact the total cost, and, hence, the efficiency of a firm (Williamson & Tadelis, 2012; Briere, Lehalle, Nefedova, & Raboun, 2020)). This perspective allows an assessment of efficiency in the production and distribution of the product or service. Related to HIV and AIDS services, TCE implies the need for an explicit focus on transaction costs in the way services are delivered. As observed in Chapter 2 transaction costs are the costs of partner search, drawing up, conclusion and enforcement of contracts in exchange that involves complex transactions. Transaction Cost Economics requires that decision units or firms that produce goods or provide services need to look at each transaction and the implications for the choice of particular organisational modes to manage the transaction. This aids in the choice of the mode that offers the most efficient outcomes. Ultimately, attainment of efficiency in the coordination of the transaction is the goal of the firm. In this connection, it is important to remember that transaction costs form part of the total

cost function of the firm. With specific reference to the ART programme, this implies that policy makers should be concerned with the levels of governance costs, since high levels imply high unit costs which may impact efficient service delivery.

The difficulties in the measurement of transaction costs are widely conceded (Wallis and North, 1986; Lawson, 2009; Vandeninden and Paul, 2012). This is partly due to problems with achievement of consistent monitoring and measurement of the performance of transactions, as well as limited agreement on what should be counted as transaction costs. However, the literature provides a number of pointers towards identification and quantification of transaction costs. They include the information costs of searching for potential transaction partners, costs of definition and negotiating contracts, as well as enacting and enforcing contracts once they are concluded (Wallis and North, 1986; North 1994; Williamson, 2000; Marinescu, 2012; Gërdoçi *et. al*, 2016). They further include the costs of adaptation to changes in the contractual arrangements or due to changing contexts.

The literature further distinguishes between *production* and *transaction costs* (Wallis and North, 1986). While production costs refer to the direct cost of actual inputs used to produce a product or a service, transaction or governance costs refer to those incurred by the producer of the product or service to facilitate exchange (North and Wallis, 1986).

The following diagramme (Figure A3. 1) summarises the firm decomposition of total costs between transaction and production costs.

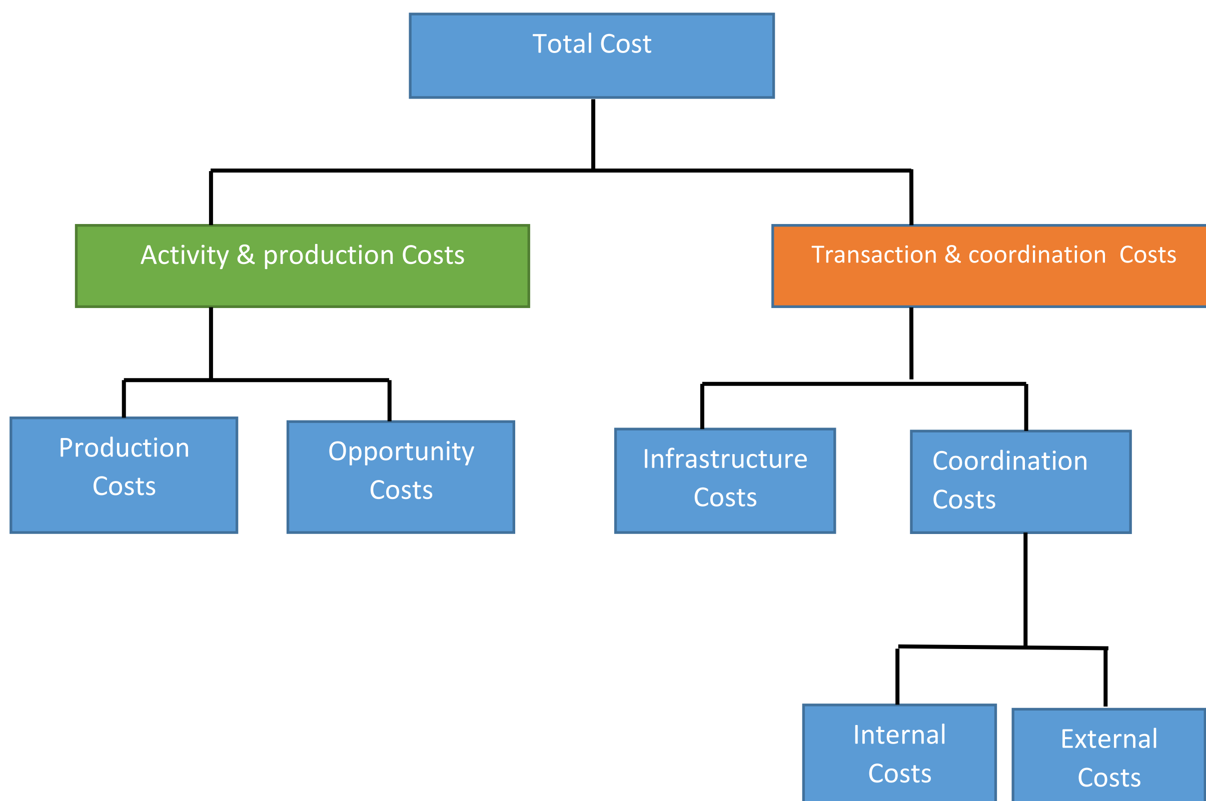


Figure A3.1: Production and Transaction Costs in a firm

Source: Cordella & Simon, 1997 cited in David, 2008

As can be observed from Figure A3.1, total firm costs in the production and distribution of a product or service consist of transaction costs and production costs. The production costs are the direct costs incurred to produce the product, as well as the opportunity cost of using the resources for other uses. Transaction costs consist of two further sub-categories, namely; infrastructure costs and coordination costs. Infrastructure costs refer to the costs of physical buildings and setting up of an organisation to deliver the services or products. Coordination costs within transaction costs consist of internal and external costs. Internal coordination transaction costs consist of costs incurred to manage and coordinate staff and other resources within the organisation. Thus, the setting up of control and coordination structures such as audit, human resources management, monitoring and evaluation and finance, among others, constitute the internal coordination transaction costs. On the other hand, external coordination transaction costs include the costs of relating with agencies or partners external to the firm or organisation

(Svetlov, 2009). In the context of ART service delivery, these may include the processes towards concluding agreements and enforcing agreements through various processes with the partners. Coordination of the Treatment, Care and Support Technical Working Group convened by the Ministry of Health in the ART coordination is a case in point. With respect to transaction governance choice, this implies that a careful balance has to be struck between choosing the levels of transaction and production costs that yield efficiency. A focus on transaction costs in HIV and AIDS service delivery would entail delineating the transaction costs in the HIV and AIDS service delivery to inform resource allocation and control decisions aimed at achievement of efficiency.

1.2 Monitoring Transaction Costs in HIV and AIDS ART Programmes in Malawi

The application of the TCE to improve efficiency in ART management implies the need for an adequate mechanism to track ART service delivery transaction costs. While there are initiatives to track resources and expenditure for the health sector and the HIV and AIDS sub-sector, an explicit framework to track transaction costs for efficient ART management appears inadequate. There are currently four frameworks that track resources that go into the HIV and AIDS programme that can provide indications of levels of transaction costs in HIV and AIDS service delivery. These approaches are: (a) the National AIDS Spending Assessments (NASA), (b) Health Sector Resource Mapping, c) Global AIDS Monitoring (GAM) report, d) National Health Accounts (NHA).

Table A3.1 summarises expenditure or budget categories that may be used as proxy for transaction costs monitoring.

Table A3.2: Potential Mechanisms to Track Transaction Costs in ART

#	Expenditure tracking framework	Explicitly tracks transaction costs (Y/N)	Category proxying transaction costs	Estimated proxied transaction costs	Recent Update
1	NASA	N	Health systems strengthening with	20%	2021

			"management administration costs" sub-category		
2	Health Sector Resource Mapping (HSRM)	N	Community outreach, technical assistance, administration & management – Other, Research, M&E and supervision, planning & policy activities, communication costs (print, TV, radio), and resource mobilization.	30%	2020
3	Global AIDS Monitoring (GAM) report	N	Programme Management	22%	2022
4	National Health Accounts	N	Health systems administration and financing administration	26.2%	2018

Source: Literature review

a) National AIDS Spending Assessment

The NASA is coordinated by the NAC and follows the government of Malawi fiscal year, and is considered to be the more comprehensive framework for HIV and AIDS expenditure tracking in the country of the four frameworks. The FY2018/19 NASA sought to reconcile the different frameworks with respect to AIDS budget and

expenditure tracking (GoM,2021). The NASA was designed to track HIV and AIDS resources and expenditure covering various sources, including bilateral and multilateral donor agencies, public and private sector; and non-governmental organisations. It tracks historical expenditure and thus represents the framework that is considered likely to provide more accurate indications about the AIDS expenditure and transaction costs patterns[21]. It guides the tracking of expenditure trends and used to identify areas requiring additional funding and thus informs the budgeting process by both the Government and stakeholders. The first NASA was conducted in 2010 covering fiscal years 2007/8 and 2008/ 9. It was initially intended to be an annual exercise to ensure HIV and AIDS planned resources and consequent expenditure by function or key priority area as specified in the policy and planning documents were tracked. The annual schedule has not been followed strictly, as the next NASA after that in 2010 was conducted in 2013, covering the fiscal years 2010/11 and 2011/12. With respect to transaction costs, the NASAs do not explicitly specify a category labelled “transaction costs”, although it tracks “management and administration costs” from which a picture of transaction costs may be gauged. For instance, the management and administration costs accounted for approximately 30 percent of the HIV and AIDS expenditure for the period FY 2008/9, FY 2011/12 and FY2012/ 2013. The management and administration costs are sub-categorized into planning, administration, coordination and management, monitoring and evaluation operations research, serological surveillance, HIV drug resistance surveillance, drug supply systems, information technology, upgrading and construction of infrastructure, programme management and administration not disaggregated by category, and programme management and administration. In particular, transaction costs related to the administration and disbursement of funds, planning, programme management and coordination, monitoring and evaluation, operations research and maintenance and construction of buildings may paint a general picture of the levels of transaction costs in HIV and AIDS programme in the country. More particular to facility-based ART and HTS, the NASA further tracks costs related to the management of the programme which is labelled “above facility costs” that can be used as a proxy for transaction costs. The most recent estimate averaging transaction costs from FY 2015 to FY 2019 is 20.2% (GoM, 2021). Nevertheless, to the extent that transaction costs monitoring is not explicit in NASA, it may be surmised that there is limited consideration of transaction cost minimisation towards efficient service delivery.

[21] Communication from a National AIDS Spending Assessment Expert via a virtual meeting

b) Health Sector Resource Mapping (HSRM)

The Health Sector Resource Mapping (HSRM) exercise collects broad health sector funding data. The HSRM collects data on budgeted funds rather than expenditure, but is useful in gauging what key financiers are allocating the funding towards, and in what magnitudes. This assists the MoH to plan and advocate for resource allocation to areas that are likely to receive limited or no funding. The HIV and AIDS component is one of the key themes in the RM as it accounts for nearly 30 percent of health funding. The bulk of these funds are allocated towards treatment and care where antiretroviral drugs are one of the drivers of expenditure. Within the HIV and AIDS funding, resources are tracked by programme or strategic pillar. Under the RM, the aspects that relate to transaction costs in HIV and AIDS are discernible, and include community outreach, technical assistance, administration & management – Other, Research, M&E and supervision, planning & policy activities, communication costs (print, TV, radio), and resource mobilization. Most recent estimate is for 2020 at 30%[22]. There is a category labelled administration & management – salaries, but this should be viewed as management and investment costs as the cost is not broken down into staff salary costs for direct service providers and those that are supporting staff. However, the lack of clarity of transaction costs appears to confirm the limited attention these costs receive as a basis for efficiency decisions.

c) Global AIDS Monitoring (GAM) report

A further mechanism that includes elements of AIDS expenditure tracking is the preparation of a voluntary Global AIDS Monitoring (GAM) report by the Government of Malawi. The preparation of the Global AIDS Monitoring Report is coordinated by the National AIDS Commission, with the participation of the Department of HIV and AIDS in the MoH, government departments, donor agencies and non-governmental organisations. The expenditure item in the GAM that may be used to track transaction costs is programme management, estimated at 22% in 2021 (National AIDS Commission, 2022). Nonetheless, a shortfall of the GAM reporting framework is that expenditure estimates include only a limited number of financiers. Hence, the GAM has a limited scope for general HIV and AIDS expenditure tracking as the focus is only on

[22] Ministry of Health Resource Mapping database, 2020

the public sector and large donor support such as the PEPFAR and the Global Fund. This further implies that it may not be adequate to track transaction costs in HIV and AIDS services for efficiency decisions.

d) National Health Accounts

The National Health Accounts are prepared by the Ministry of Health. Data are collected from a range of agencies such as government agencies (including District Councils), donors, non-governmental organisations and insurance companies. Household health spending data are obtained from national household surveys, in particular, the integrated household survey (IHS) in the most recent past. The expenditure categorization shows five main expenditure functions, namely, preventive care, medical goods, governance, health systems administration and financing administration, and ancillary services. From these categories, governance and administration costs stand out as transaction costs, although ancillary services may also be added to that category. The most recent NHA data estimates these "transaction costs" at 26.2% (GoM, 2020). Nevertheless, the NHA covers expenditure for the entire health sector, and thus does not particularly focus on HIV and AIDS expenditure, and may thus be ill suited to track transaction costs in HIV and AIDS ART service delivery.

The forgoing text does not specifically show that there is a specific expenditure category labelled "transaction costs", although it is possible to gauge them from governance and administration costs. As such it was not particularly clear as to whether the MoH is focused on the reduction of transaction costs as an area where potential efficiencies can be gained.

1.4 The Significance of Transaction costs in healthcare and ART services

When transaction costs are considered as the administrative and financial costs not directly incurred due to ART service provision, perceptions are that the levels may actually be high. Participants looked at the roles that various actors in the ART programme played and linked them to activities that could be characterised as transaction activities. Further, all agencies that sub-granted or technically supported agencies that carried out activities related to ART were potential candidates as transaction entities. In this connection, participants noted that there were numerous

agencies that have set up own offices in the country with implications on transaction costs when the totality of the effort was accounted for. The following sentiment captures this perception:

> *"As you know, there are many agencies involved in HIV and AIDS in the country focusing on one aspect or the other. Some support efforts are useful for the ART programme. So, there are likely to be administrative costs associated with office space and other costs. Perhaps it may be useful to identify local agencies that are already on the ground instead of every organisation setting up their own offices. But we understand that there may be issues of capacity, as well as mistrust regarding whether funds would be used for intended purposes".*
> *Respondent 9, Programme Manager*

As can be noted from the above quote, although not explicit, one would discern elements related to assumptions of how actors decide the type of organisation to effect transactions, in this case HIV and AIDS. For instance, decisions to set up offices in country rather than work with local agencies may be driven by perceptions of a specialised set skills, culture and attitudes to effectively deliver the services which may be construed as asset specificity, a key assumption under TCE. Second, perceptions of mistrust reflect notions of opportunism, and the possibility of misuse of funds by local agencies due to interests other than those originally intended by the financing agencies.

Furthermore, non-state agencies that support selected health facilities incur transaction costs that ought to be accounted for as they provide funding and technical support. These costs are incurred for data collection and support for follow up of defaulting clients. The study found that in some districts and health facilities, there were dedicated teams whose role was to follow up defaulting clients. For instance, in Nsanje District, a follow up team with resources that include transport is part of the framework for ART service delivery. In TCE terms, this is akin to looking for exchange partners, and, therefore, constitutes transaction costs. In yet other districts where these follow up teams are not present, Expert Clients are used and a range of other volunteers in committees at the community level. In some instances, they have been provided with bicycles and a lunch allowance for each follow up visit that they undertake. Nevertheless, the study found that this was not consistent across districts and facilities which made estimation of these

costs somewhat difficult. It was not readily evident that these costs are tracked and included as part of the ART service provision costs in the NASA. Even where they are added or mentioned, it is at a highly aggregated level, and also not explicitly labelled as transaction costs.

1.4 Opportunities and Constraints to Tracking ART Transaction costs

The study has noted a number of mechanisms used to track resources in the health sector, and for the HIV and AIDS sub-sector. More particular to the current study, the NASA appears to be the main frame for AIDS spending monitoring, and therefore, offers the opportunity to explicitly track transaction costs to inform decisions regarding achievement of technical efficiency of the ART programme in Malawi. Nevertheless, there appear to be binding constraints that limit the estimation of transaction costs of the ART programme. First, a limitation of the NASA is that not all agencies are covered and hence, expenditure figures remain estimates. Second, some of the elements that may constitute transaction costs require further breakdown. For instance, and as noted before, some items combine management and transaction costs in their categorization of costs under management and administration. Without that breakdown, the estimates are likely to be crude. Third and critically, related to the current study, the NASA does not track facility level coordination costs which include utilities, space, supporting services that constitute transaction costs. This, nonetheless, is understandable, as the NASA does not have as its deliberate focus the monitoring of transaction costs.

Nevertheless, despite the above weaknesses, the NASA represents the closest framework that may be used to monitor transaction costs in ART service delivery. With respect to facility-based ART service provision, one NASA expenditure category or sub-category that is most relevant to the analysis, and may be a proxy for transaction costs is the "above facility costs". This carries the connotation that these costs are directly focused on supporting ART services at the facility level. In financial accounting terms, these costs may be applied to ART delivery as indirect costs. With the above caveats, the study adopts the NASA framework to estimate and track the transaction costs of ART service delivery.

1.5 Estimating Financial Inputs in ART Services

The components of the ART programme consist of service provision and coordination of the logistics to coordinate the services. The following broad components are thus identifiable: (a) diagnosis to identify people in need of ART, which is also consistent with the achievement of the first 90% target by 2020, now the first 95% after having already achieved the 90% target; b) the drugs and commodities acquisition and distribution to facility locations to reach people that need the drugs; c) patient treatment and adherence to drugs, and, d) monitoring the patients' response to treatment. These broad programmatic activities are designed to contribute towards the UNAIDS Global HIV and AIDS to contain the epidemic as a public health threat by 2030. Diagnosis is consistent with the aim to achieve the first 95%, where all people that are HIV positive know their status. The acquisition of the drugs and their distribution to ensure they are available to those that need them (treatment) would contribute towards the achievement of the second 95%, where the target is to reach 95% of all people that test HIV antiretroviral treatment. The programme further run ART clinics to ensure those that are enrolled on treatment adhere to the regimens. The ART clinics are either run on specially allocated days in the week, which is in the majority of the facilities (over 90%), or run as part of the routine facility care processes in an integrated fashion. The latter arrangement is common to private sector facilities and CSO owned facilities that are part of the ART programme. The monitoring of clients' response to treatment further contributed to the third 95% (achievement of viral suppression in at least 95% of the clients on treatment), besides monitoring toxicity levels.

1.5.1 Expenditure Estimation Methodology

One of the tasks that the study sought to achieve was estimation of the ART programme inputs that would ultimately be used to assess the efficiency of the programme. The cost estimate would include both production and transaction costs of the ART programme. Given that currently transaction costs are not explicitly tracked as a specific cost category, a further aim was to operationalize the transaction cost concept through the identification of transaction cost elements of the programme. This would extend to assessment of available data sources to estimate these costs. The study used facility level inputs, mainly drugs and test kits of the ART programme, as well as results of cost elements from recent studies to estimate both production and governance or transaction costs.

HIV testing services are a key strategy to identify HIV infection and facilitate timely initiation on ART. The HTS is thus a key strategy towards meeting the first 95 and contributes significantly to both the second 95 and the third in the UNAIDS 95-95-95 targets by 2030. With 2015 WHO test and start guidelines, HTS represents a key piece in efforts to combat HIV and AIDS. The majority of HIV testing in Malawi is facility based, although a self-testing programme has also commenced. The current study uses facility-based testing to estimate costs and ultimately, the efficiency of HTS at the facility level. Initial test is a finger prick rapid diagnostic test (RTD), followed by a confirmatory RTD when the outcome is positive. For discordant or inconclusive tests, where the first test is positive but the confirmatory test is negative, clients are advised to do a repeat test after four weeks. Because the initial HIV positive result has to be confirmed by a second test, it follows that an HIV positive outcome is associated with a higher cost per client.

1.5.2 ART Production and Transaction Costs

Expenditure on ART reflects both production and transactions costs as noted in sub-section 8.1 above. Consequently, in the estimation of costs, these two broad categories were kept in view.

The following algorithm (Figure A3.2 next page) was deployed as an organizing framework for the ART cost classification between production and transaction costs.

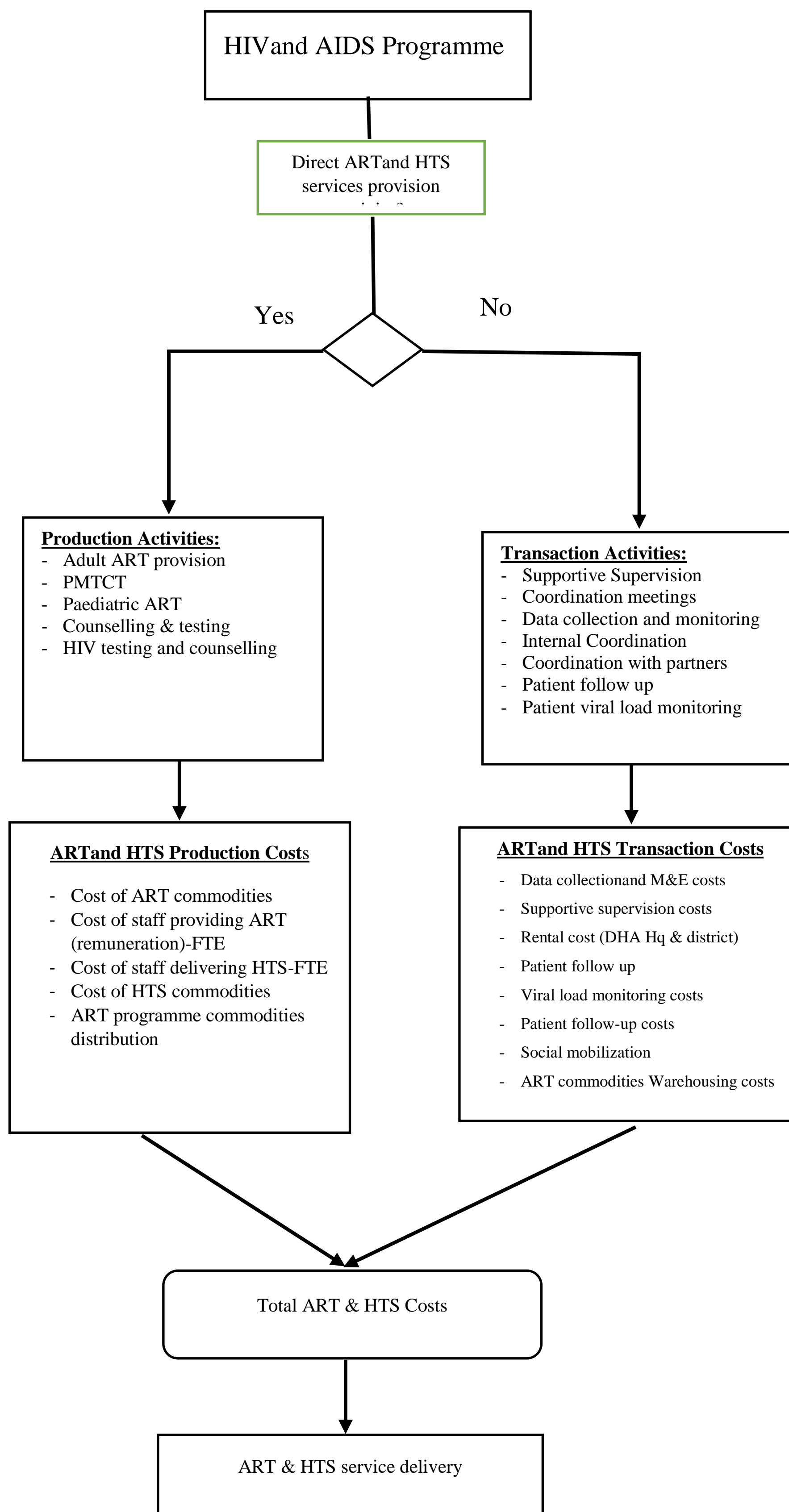


Figure A3.2: Delineating Production and Transaction Costs in ART

Source (Author)

As can be observed from Figure A3.2, a focus on tracking transaction costs views the ART Programme as the firm. This includes the Department of HIV and AIDS and the service delivery points at the national, sub-national (central hospitals), district (district hospitals) and sub-district levels (health units). For this study, the focus is ART services offered at the different levels of service delivery. To delineate the transaction and production costs, the activities are categorized into direct service provision (production activities) and indirect (transaction activities). The direct service provision activities include ART provision-ART clinics, PMTCT, paediatric ART, PEP, testing and counselling. Transaction activities include supportive supervision, internal coordination (human resource management, infrastructure), monitoring and evaluation, audit, research, coordination meetings, patient follow-up and viral load monitoring. These activities drive the cost of ART service delivery. The production costs arising from direct service provision (direct costs) include the cost of ART commodities, cost of staff providing ART (remuneration)- full time equivalents (FTE), cost of staff delivering HTS-FTE, cost of HTS commodities. As will be noted below, the ART costs are the main drivers of ART and HTS, and the HIV and AIDS programme in general. The transaction costs associated with ART service delivery include service delivery supervision costs, data collection and M&E costs, rental cost (DHA headquarters & district), patient follow up, viral load monitoring costs and patient follow-up costs. The transaction costs are necessary and contribute to ensuring the set targets and quality standards are met. When viewed from a TCE perspective, the transaction costs represent the costs incurred in coordination, enforcement of contracts (standards) and searching for clients (follow up of defaulters). The search costs extend to education and outreach activities designed to encourage communities to test for HIV and patients on ART to adhere to the ART.

d) ART Expenditure Estimation Approach

The study used a simple classical economic approach to estimate ART inputs and outputs to facilitate efficiency estimation for the year covered. Service delivery unit costs for ART services and commodities were used to estimate the cost per ART site. Outputs achieved at each ART facility such as number of ART clients per facility for the year were multiplied by per person per year unit costs. The Global Health Cost Consortium (GHCC) cost database was accessed through the National AIDS Commission. This framework was adopted due to its practical application for the

purpose of estimating the inputs for DEA efficiency analysis. Unit costs were further obtained from recent ART costing studies. As noted before, the main activities that the study was concerned with were treatment (ART) and HIV testing (HTC) services. The financial input estimation covered a total of 747 facilities, but 22 sites were removed. The reasons for removal of some sites were firstly, that they only provided one aspect of the ART services, such as only the treatment aspect and excluded the testing aspects. This was considering that to complete efficiency estimation in DEA, the units need to use similar inputs and produce similar outputs to facilitate comparison. In addition, one site was reported to have discontinued provision of ART services. Ultimately, 725 sites were included in the costing database.

The health facility inclusion and exclusion and estimation steps are summarized in Figure A3.3.

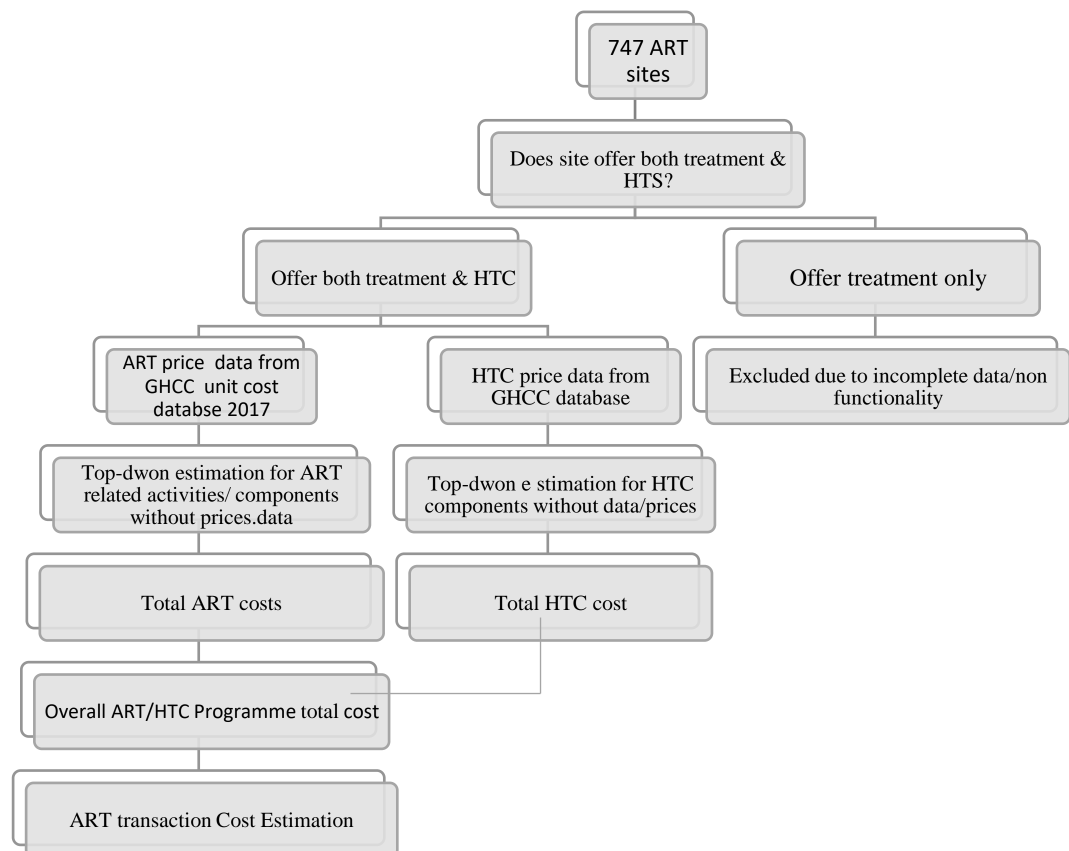


Figure A3. 3: Facility Based ART programme cost estimation process

e) Top Down Approach

The approach used to estimate ART expenditure was largely top-down. It used two main outputs reported at each ART facility, and a database of unit costs prepared by the GHCC in 2017. The research used the GHCC cost database (Unit Costs Programme Data) that included costs for ART and HIV testing and counselling services. The costs are reported as per person per year (PPPY) for the year 2018. They are based on total costs for providing free facility based ARVs and HIV testing and counselling services for one year, with 2018 as the reference year. The average cost of providing free ARVs across all facilities in Malawi in 2018 was estimated at $115.32 per person per year (average ARVs per person for all regimens). With respect to HTC services, the research used a cost estimate of $4.52 based on client-initiated facility-based testing. This cost was preferred over provider initiated for ease of calculation and the fact that there were perceptions from key informants that the provider-initiated testing and counselling appeared to have slowed down, largely due to cost in recent years.

Community outreach HTC cost estimates were excluded as the focus was on facility-based ART and HTC services. At $9.49 per person per year, the average cost for provision of HTC services for key populations was further ignored because it would have inflated the total HTC services as it was almost double that of both the client and provider-initiated HTC. The high average HTC per person per year cost for key populations is attributed to the specially targeted interventions to reach the key populations which tended to be costly. Key assumptions in the GHCC costing need to be further highlighted in adopting these average cost estimates for ART and HTC services as they have implications on the interpretation of the financial cost estimates. Firstly, the GHCC estimates for both the ART and HTC services appear to have been based on a top down approach (based on total programme costs) rather than on the costing of the different cost elements at each health facility that provided ART or HTC services. Secondly, the costs do not include society costs nor those incurred by individual clients to access ART and HTC services. Thus, the average costs for both ART and HTC services may be underestimated. Nevertheless, given the significant amount of financial resources that are spent on ART and HTC services each year, these cost estimates are likely to represent mean per person per year costs for the programme for period for which they were estimated.

In addition, the study used available data on cost estimation for Malawi, including Vyas et al (2020). The unit cost estimates by GHCC included per person per year costs for ART, HTC and opportunistic infections treatment. These were then directly applied to facility output data to estimate the facility level cost for a particular activity. This approach was informed by a pilot phase of the study that sought to establish the feasibility of deploying a bottom up approach where facility level volumes of commodities consumed and outputs produced would be estimated. The bottom up approach was found not to be feasible given the limited resources available for the study. Moreover, for the large number of ART facilities assessed, a reasonable sample would be required to complete the costing and to use the results to extrapolate to the number of ART sites in the country. Hence, a key limitation in the current expenditure estimation is that most of the cost estimates are likely to be at a very aggregate level. The per person per year costs used in the current study was estimated through the NASA was $9.75 for HTC, while that for ART was $261.00 (GoM, 2021). The unit costs in the NASA are higher than in those used for the current research, which were $5.50 and $115.39 for HTC and ART, respectively.

.

1.5.3 Summary of Cost Estimates

The total ART cost for the fiscal year 2018/ 19 is estimated at approximately $248, 300.00. Of this amount, the ART treatment sub-component is estimated to be $219,400,00.00, while the rest ($28,900,000.00 is estimated to be HIV testing expenditure. The weighted mean cost per facility was estimated at $1,131,473.00, and ranged from $0.01 to $199,117.60. Overall, per capita per year ART costs are estimated at $298.24, and ranged from $253.64 to $1104.07. By sub-programme, the mean HTS cost was $34.59 while the mean cost for treatment (ART) was estimated at $263.65. The per capita costs from the current study appear high for HTC when compared with those reported by the NASA ($9.57), while the cost for ART is comparable with the NASA estimate ($261.00).

1.5.4 Approximating Levels of Transactions Costs in the ART Programme

Tracking transaction costs in any setting is difficult because of observation challenges. Moreover, for the HIV and AIDS programme, and particularly the ART programme, transaction costs tracking for the purpose of informing efficiency considerations appears not to be a deliberate policy currently. While there are frameworks that would make it

possible to an extent to track transaction costs with a view to actively include them in the ART efficiency decisions, they do not exclusively track theme for this reason. Among the key frameworks for expenditure tracking are the National AIDS Spending Assessment (NASA), Resource Mapping (RM), and the National Health Accounts. The Global AIDS Monitoring initiative, a voluntary reporting commitment by the Government of Malawi on the progress of the HIV and AIDS response further reports expenditure estimates, although the sample of agencies considered to be relatively small (GoM, 2021).

For the current study, a Bayesian updating approximation approach is deployed using Binomial distribution modelling and Markov Chain Monte Carlo Simulation. The Bayesian approach is suitable for estimation of transaction costs because it allows for estimation of probability without knowing the prior distribution of a variable. It further allows for theoretically updating our knowledge with past experience, as well as apply probability directly without requiring the long run proportions of a single event. It is a subjective probability estimation approach that permits consideration of events that are not repeated. This is in contrast to the objective uncertainty estimation that uses observed data. Bayes updating is a statistical method that applies the Bayes Theorem to update the probability of distribution as more information becomes available. The subjective element in Bayesian estimation may be viewed as a weakness, although it is increasingly used in in empirical estimation, including in cost estimation (Stafford, 2020).

Uncertainty about parameters is expressed via a prior distribution, $\mathrm{P(A)}$. This is followed by updating the priors by using a dataset B into posterior distributions by deploying the following Bayes rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad \text{(A2.1)}$$

Where:

$\mathrm{P(P|A)} = \text{likelihood}$

$\mathrm{P(A)} = \text{prior probability.}$

$\mathrm{P(B)} = \text{a scaling factor}$, which is the marginal density of the data, and

$$P(B) = \int_A P(B|A)P(A) \quad \text{(A2.2)}$$

Since $P(B)$ is a normalising scaling factor, the Bayes Theorem can be written proportionally:

$$P(A|B) \propto P(B|A) * P(A) \tag{A2.3}$$

The study uses the Beta distribution as a conjugate prior for Bayesian updating. The parameter of interest is θ of the total expenditure on ART. The total expenditure may be considered as the total sample from a distribution X from which transaction costs as a specific category of ART service delivery costs is drawn. This parameter of interest, $\theta \epsilon (0,1)$ is the proportion of transaction costs of the total ART expenditure in a particular year. Outcome y records the absolute volume of transaction costs in the total expenditure. It is observed that part of the total expenditure does not constitute transaction costs, but consists of production costs or other costs such as investment costs. This suggests that if a sample was drawn from each USD spent on HIV and AIDS service delivery, the non-transaction cost expenses would have the outcome $y = 0$, while the transaction cost outcome would be $y = 1$. The occurrence of transactions costs is considered as an event in the course of expenditure on the ART programme. The binomial distribution is, therefore, defined as follows:

$$\mathrm{y}|\theta \sim \text{binomial}\ (\mathrm{X}, \theta) \tag{A2.4}$$

The most recent NASA (GoM, 2021) estimates administrative costs to be approximately 15 percent of the total HIV and AIDS expenditure. This expenditure consists of non-ART related expenditure, although it also excludes the facility level administration and management costs, as well as the cost of infrastructure and running of the facility level. It further excludes utility costs such as electricity, water and follow up of clients that drop out of treatment. Thus, there is a lot of uncertainty as to the exact measure of transaction costs associated with ART service delivery in Malawi. Nevertheless, the coordination cost estimates from the NASA provide a general idea of the overall transaction costs. The study used these current estimates as the starting point, that is, in the setting of the prior. Moreover, for the fiscal years covered under the 2021 NASA report (GoM, 2021), these costs are estimated to lie between 15 percent and 30 percent. For FY 2018/ 19, total facility-based ART are estimated at US$244,500,000.00. The

estimated costs for transaction costs from the study is US$50,000,000[23]. Based on this information, the study assumes that the proportion of transaction costs from each US$ spent on ART has a mean of 0.2, or 20 percent in a year based on data from the Malawi NASA. The rate of observing a transaction cost λ in ART expenditure in Malawi in a year is then estimated. The study uses this conservative expected value to represent the incidence or the rate at which transaction costs occur in ART for FY2018and 19, that is, θ roughly follows a Beta distribution, $\theta \sim Beta\ (50000000{,}244500000)$.

The goal is to estimate a given distribution with a beta distribution. The expenditure distribution of ART is assumed to be known and to follow a Beta distribution. The study thus estimates the defining parameters of a beta distribution, α and β, to achieve the best model for transaction costs via the ART expenditure distribution. Considering that the Bernoulli distribution is a single trial of the binomial distribution $(n = 1)$, the equation depicting the Binomial distribution (Equation A2.4) reduces to:

$\alpha_{t+1} = \alpha_t + 1$ $\beta_{t+1} = \beta$

(A2.4a)

And

$\alpha_{t+1} = \alpha\beta_{t+1} = \beta_t + 1$ (A2.4b)

for a success $(x = 1)$, and a failure in the case of a success $(x = 0)$, respectively.

The study assumes that the prior and posterior distribution are in the same probability distribution family as the prior probability distribution. Hence, the prior and the posterior are conjugate. The study further assumes that the ART expenditure follows a Beta distribution. and, hence, uses a Beta distribution as prior to the likelihood function of the ART expenditure. The prior is specified as follows:

$p(\lambda) \propto \ for\ 0 < \lambda < 1$

(A2.5)

The Metropolis-Hastings algorithm is executed through the *bayesmh* command in STATA 17 (StataCorpLLC, 1985-2021), to estimate the probability of transaction costs. This is one of the most commonly used sampling strategies in Bayesian estimation.

[23] These figures are rounded off to the nearest whole number for ease of calculation

The standardization factor requires complex calculations, made possible by computer simulations. The study thus used the Hastings Metropolis Markov Chain Monte Carlo (HM MCMC) algorithm simulations to estimate the standardization factor in the Bayes formula. Since the normalizing concept is unknown, the study starts from a proportionality concept. It considers that the distribution of transaction costs in ART spending is a proportion of the expenditure distribution. The study assumes that the transaction costs distribution in ART spending follows a random walk HM MCMC algorithm (S*ee Appendix 2b*).

1.6 Transaction Cost Estimation Results

The research used a beta distribution as a prior distribution, with a prior probability of 0.2 to estimate transaction costs in the HIV and AIDS ART programme in Malawi. This prior was set considering available data from the NASA covering a three year period since fiscal year 2015/16 on transaction costs. Transaction costs are proxied on programme enablers and systems strengthening costs related to HIV and AIDS service delivery expenditure as reported by the most recent NASA report (GoM, 2021), which estimates the costs to be around 20%, from real-world ART programme operations. The model was estimated using a sample of 10,000 observations, with 12,500 iterations performed to ensure robust results.

The results of modelling the transaction costs show that the hypothesized mean value of the transaction costs levels in the ART programme is consistent with those estimated by the NASA. The results are presented in Table A3.2.

Table A3.2: ART Programme Transaction Costs estimation

Model summary:

Likelihood:		
Y	~	binomial({theta},1)
Prior:		
{theta}	~	beta (50000,250000)

Bayesian	Binomial	model MCMC	iterations=	12,500
Random-walk	Metropolis–Hastings	sampling Burn-in	=	2,500

MCMC	Sample	Size	=	10,000		
Number of Obs			=	1		
Acceptance rate			=	0.4396		
Log	marginal-likelihood	=-.17582133	Efficiency	= 0.2162		
					Equal-tailed cred.	
	Mean	Std. dev.	MCSE Median	[95%	cred.	interval]
Theta	0.166663	0.000685	.000015 .1666756	0.165313	0.168037	

Source: Author estimations

The table above presents estimates the posterior probability of transaction costs in the ART programme, reflecting a refined probability after incorporating empirical data. In summary, the Bayesian beta model with a prior of 0.2 yielded a posterior mean of 0.166663. The estimated posterior mean is somewhat lower than, but close to 0.2, suggesting that the proportion of transaction costs in the ART programme is likely to be close to 0.2, or 20%. The reported 95% credible interval for the posterior probability is (0.165, 0.168). This interval offers a range within which the actual transaction cost likely falls, given the data and the prior information from the NASA report. The narrow credible interval suggests relatively high precision in the estimate, which strengthens confidence in the updated 0.16663 probability as a reflection of transaction costs in the ART programme.

The NASA estimates transaction costs, proxied by programme enablers and systems strengthening over the four fiscal years since 2015/16 up to 2018/19. The programme enablers consist of costs related to programme administration, strategic information, M&E, public system strengthening and community system strengthening. The highest transaction costs are recorded in 2017/18 and 2015/16 estimated at 24.0% of total costs. It is noteworthy that the NASA reports that after systems strengthening, programme administration and management costs were the second highest, averaging 26.25% of the programme enablers and systems strengthening costs over the four years, suggesting the importance of transaction costs in HIV and AIDS service delivery. These transaction costs are largely financed by the United States Government and the GFTAM. This suggests that optimization of transaction costs is likely to contribute to efficiency gain in

the ART and HTC services given the size of the transaction costs in the overall HIV and AIDS service delivery.

1.6.1 Model Robustness

The estimation shows a high level of precision as indicated by the standard deviation and the Monte Carlo Standard Error (MCSE). The standard deviation of 0.000685 implies a narrow posterior spread, indicating high certainty around the mean estimate. The Monte Carlo Standard Error (MCSE) for the Median (0.1666756) is estimated at 0.000015, indicating the precision of the median estimate. The small MCSE from the results suggests that the median is a reliable representation of the central tendency in the posterior distribution.

The sampling process was efficient and reliable, as evidenced by the acceptance rate and efficiency metrics, making the results a robust estimation of the underlying parameter, *theta*. The results of the model estimation show an acceptance rate for the sampling process of 0.4396, indicating a reasonably efficient sampling process, as values between 0.2 and 0.5 are typically considered acceptable in Bayesian sampling(Lye, Adolphus; Cicirello, Alice; Patelli, Edoardo, 2021) . An acceptance rate of 0.4396 suggests a balance between exploration of the parameter space and the efficiency of the sampler. The efficiency of the model was 0.2162. Although not extremely high, an efficiency of 0.2162 is acceptable for most Bayesian analyses (*ibid*), suggesting that the samples are informative.

Model robustness was further confirmed through post estimation model convergence diagnostic plots, as shown below -see Figure A3. 4.

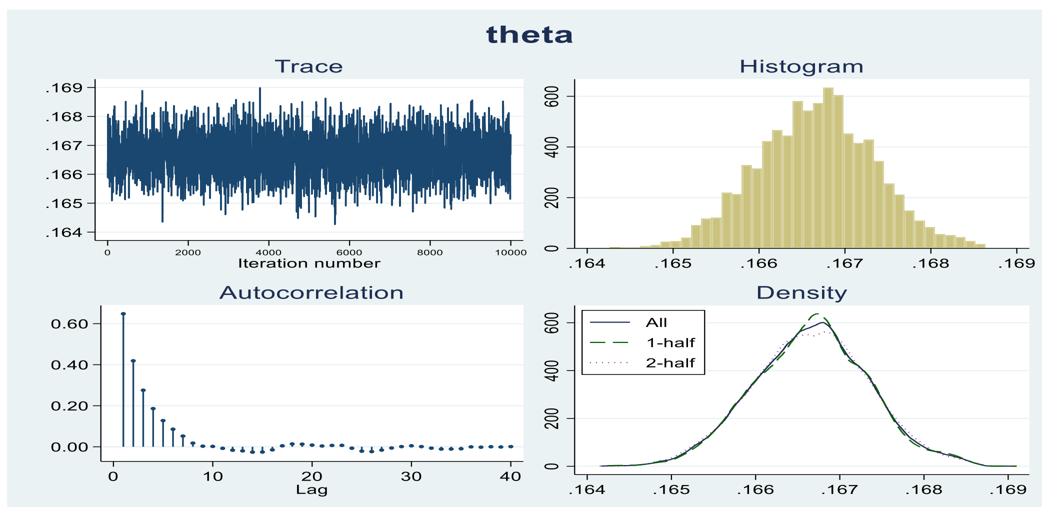


Figure A3.4: Transaction cost estimation model diagnostics

As may be observed from the figure, the trace plot shows that there is convergence and good mix in the model. It further looks plausible as it does not show trends and appears to traverse the parameter range reasonably well. The autocorrelation plot further indicated good convergence, reaching zero after 10 lags. Further, the density plot showed good convergence as the density for the first and second half are similar. In addition, the histogram appears to show an observable distribution, as it is well shaped. The implication is that MCMC chain converged well for the parameter of interest, *theta* that was used to estimate the proportion of transaction costs in the HIV and AIDS ART programme

1.7 Summary

This Appendix has argued that despite the existence of mechanisms to track HIV and AIDS expenditure, there is limited attention towards the tracking of transaction costs to inform efficiency decisions in the governance of the pandemic. Accounting for the limitations of the estimation method followed, the Appendix notes that the estimation facilitates the modelling of transaction costs in ART. The modelling results show that the estimation is largely consistent with observed levels of governance costs. The estimated proportion of transaction costs is close to the observed data from the NASA.

Appendix 3b: Hastings-Metropolis MCMC Algorithm

The Hastings Metropolis algorithm is adopted as it provides a way to sample from an unknown distribution. Let f be a function that is proportional to the desired probability distribution, called the target distribution. Further, let $q() = q(.|\text{x})$ be the proposal distribution to be simulated via the Markov Chain Monte Carlo (MCMC) process. The algorithm constructs a series of variables $x_1, x_2, \ldots x_n$ for a given x_n.

1. Generate $y_n \sim (qIx_n)$,
2. Generate $u \sim \mathcal{U}[0.1]$ a uniform distribution
3. Compute the acceptance rate, $\alpha = min\{\frac{f(y_n)(qx_n|y_n)}{f(x_n)q(y_n|x_n)}, 1\}$
4. Accept the new candidate y_n with probability α if $\mu \leq \alpha$, otherwise reject.

$$X_n \begin{cases} y_n \; if \; \mu \leq \alpha \\ x_n, \text{ otherwise} \end{cases}$$

Appendix 4:Regression Output for transaction costs on health facility ART efficiency

```
. tobit dmu_efficiency log_of_transaction_costs i. location_type2 i.Ownership_type_ i.Level_
> of_service dist2 catchpon2 zombacity thyolo site_maturity rural salima rumphi phalombe nt
> chisi ntcheu nsanje nkhotak nkhatab neno mzuzucity mzimba_south mzimba_north mzimba_dist
> mwanza mulanje mchinji mangoc machinga lilong likoma llcity kasungu karong dowa distance
> dedza chitipa chiradz chikwaw catchmentpopulation lambda, ll(0) ul(1) vce(robust)
```

Refining starting values:

Grid node 0: log likelihood = 869.591

Fitting full model:

```
Iteration 0:   log pseudolikelihood =   869.591
Iteration 1:   log pseudolikelihood = 872.92003
Iteration 2:   log pseudolikelihood = 872.95818
Iteration 3:   log pseudolikelihood = 872.95821
```

```
Tobit regression                        Number of obs =            658
                                          Uncensored =            610
Limits: Lower = 0                      Left-censored =      0
        Upper = 1                     Right-censored =             48

                                        F(43, 615) =        .
                                        Prob > F     =        .
Log pseudolikelihood = 872.95821        Pseudo R2    =  -0.0336
```

dmu_efficiency	Coefficient	Robust std. err.	T	P>|t|	[95% conf.	interval]
log_of_transaction_costs	-.0705635	.0352692	-2.00	0.046	-.1398262	-.0013007
location_type2						
rural	-.8467592	.4328386	-1.96	0.051	-1.69678	.0032617
Ownership_type_						
chammam	.0047343	.0105346	0.45	0.653	-.0159538	.0254224
public	.0097377	.0090031	1.08	0.280	-.0079429	.0274182
CSO	.006444	.0157072	0.41	0.682	-.0244023	.0372902
Level_of_service						
secondary	.0124711	.0194132	0.64	0.521	-.0256531	.0505952
primary	.0106605	.0190819	0.56	0.577	-.0268132	.0481341
dist2	7.27e-07	3.78e-07	1.92	0.055	-1.64e-08	1.47e-06
catchpon2	-1.27e-10	6.39e-11	-1.99	0.047	-2.52e-10	-1.58e-12
zombacity	-1.686861	.8441411	-2.00	0.046	-3.34461	-.0291124
thyolo	-.6798858	.3404827	-2.00	0.046	-1.348536	-.0112361
site_maturity	-.0470777	.0235175	-2.00	0.046	-.0932621	-.0008934
rural	.017509	.0295693	0.59	0.554	-.04056	.0755779
salima	-1.152147	.5763136	-2.00	0.046	-2.283929	-.0203659
rumphi	-1.193509	.5975091	-2.00	0.046	-2.366914	-.0201035
phalombe	-.2201741	.1116511	-1.97	0.049	-.4394378	-.0009104
ntchisi	.4669026	.2367158	1.97	0.049	.0020333	.9317719
ntcheu	-.9570887	.4786454	-2.00	0.046	-1.897066	-.0171112
nsanje	-2.138842	1.070527	-2.00	0.046	-4.241174	-.0365096
nkhotak	1.490771	.746165	2.00	0.046	.0254303	2.956111
nkhatab	.3413344	.1714377	1.99	0.047	.00466	.6780087
neno	.2356176	.1194771	1.97	0.049	.0009851	.4702502
mzuzucity	-1.499997	.7512533	-2.00	0.046	-2.97533	-.0246643
mzimba_south	-1.854483	.9237585	-2.01	0.045	-3.668587	-.0403799
mzimba_north	-2.48629	1.239269	-2.01	0.045	-4.920003	-.0525775
mzimba_dist	2.29073	1.142358	2.01	0.045	.0473332	4.534126
mwanza	1.044	.5218258	2.00	0.046	.0192237	2.068777
mulanje	-.500493	.2500629	-2.00	0.046	-.9915737	-.0094123
mchinji	-.6413608	.3214347	-2.00	0.046	-1.272604	-.010118
mangoc	-.3628457	.1821348	-1.99	0.047	-.7205274	-.0051641
machinga	.3219205	.1600683	2.01	0.045	.0075737	.6362673
lilong	.0484287	.0269613	1.80	0.073	-.0045188	.1013762
likoma	1.678879	.8392342	2.00	0.046	.0307672	3.326992
llcity	-1.994176	.9937695	-2.01	0.045	-3.945769	-.0425833
kasungu	-.1645282	.0850437	-1.93	0.053	-.3315394	.002483
karong	-.973435	.4860343	-2.00	0.046	-1.927923	-.0189468

dowa	-.0240881	.0169264	-1.42	0.155	-.0573286	.0091524
distance	-.0016219	.0008146	-1.99	0.047	-.0032217	-.0000222
dedza	-.6860577	.3432715	-2.00	0.046	-1.360184	-.0119313
chitipa	-.3174032	.1587781	-2.00	0.046	-.6292162	-.0055902
chiradz	-1.077663	.5393569	-2.00	0.046	-2.136868	-.0184587
chikwaw	-.8925789	.446174	-2.00	0.046	-1.768788	-.0163696
catchmentpopulation	2.93e-06	1.53e-06	1.91	0.056	-7.54e-08	5.94e-06
lambda	-138.4378	67.9663	-2.04	0.042	-271.912	-4.963628
_cons	47.5658	22.93712	2.07	0.039	2.521214	92.61038
var(e.dmu_efficiency)	.0027801	.0002285			.0023657	.0032669

```
. reg dmu_efficiency log_of_transaction_costs  location_type Ownership_type_ Level_of
    _service  zombacity thyolo site_maturity rural salima rumphi phalombe ntch
    isi ntcheu nsanje nkhotak nkhatab neno mzuzucity mzimba_south mzimba_north mzimba_dist mw
    anza mulanje mchinji mangoc machinga lilong likoma llcity kasungu karong dowa distance de
    dza chitipa chiradz chikwaw catchmentpopulation , vce(robust)
```

Linear regression

```
                Number of obs   =        658
                F(43, 613)      =          .
                Prob > F        =          .
                R-squared       =     0.0804
                Root MSE        =     .05126
```

dmu_efficiency	Coefficient	Robust std. err.	t	P>|t|	[95% conf.	interval]
log_of_transaction_costs	-.0675571	.0325333	-2.08	0.038	-.1314473	-.003667
location_type2						
Rural	-.8192774	.4000989	-2.05	0.041	-1.605008	-.0335466
Ownership_type_						
chammam	.0038159	.010322	0.37	0.712	-.0164548	.0240866
Public	.008361	.0087474	0.96	0.340	-.0088175	.0255395
CSO	.0036596	.01441	0.25	0.800	-.0246393	.0319586
Level_of_service						
secondary	.0088343	.018369	0.48	0.631	-.0272394	.044908
primary	.0077131	.0178921	0.43	0.667	-.0274242	.0428503
Catchmentpopulation	2.84e-06	1.42e-06	2.01	0.045	5.95e-08	5.63e-06
site_maturity	-.0450427	.0216711	-2.08	0.038	-.0876012	-.0024841
Distance	-.001545	.000751	-2.06	0.040	-.0030198	-.0000702
Zombacity	-1.615737	.7791602	-2.07	0.039	-3.145884	-.08559
Thyolo	-.6503837	.3147269	-2.07	0.039	-1.268458	-.0323099
Salima	-1.103069	.5322032	-2.07	0.039	-2.148231	-.0579058
Rumphi	-1.146225	.5524795	-2.07	0.038	-2.231208	-.0612433
Phalombe	-.2122788	.1033809	-2.05	0.040	-.4153024	-.0092552
Ntchisi	.4454817	.2187133	2.04	0.042	.0159635	.8749998
Ntcheu	-.9174926	.4419717	-2.08	0.038	-1.785455	-.0495303
Nsanje	-2.055875	.9890241	-2.08	0.038	-3.998161	-.1135881
Nkhotak	1.426443	.689219	2.07	0.039	.0729266	2.77996
Nkhatab	.327153	.1582591	2.07	0.039	.0163572	.6379489
Neno	.2264526	.1105598	2.05	0.041	.0093306	.4435746
Mzuzucity	-1.436862	.6929524	-2.07	0.039	-2.797711	-.0760136
mzimba_south	-1.784056	.8541324	-2.09	0.037	-3.461437	-.1066752
mzimba_north	-2.386309	1.144796	-2.08	0.038	-4.634507	-.1381108
Mwanza	.999201	.4816892	2.07	0.038	.0532398	1.945162
Mulanje	-.4811643	.2309656	-2.08	0.038	-.9347441	-.0275846
Mchinji	-.6155123	.2966228	-2.08	0.038	-1.198032	-.0329921
Mangoc	-.3485969	.1683689	-2.07	0.039	-.6792467	-.0179471
Machinga	.3081936	.1475084	2.09	0.037	.0185105	.5978767
Lilong	.0460453	.025168	1.83	0.068	-.0033806	.0954712
Likoma	1.606957	.7750599	2.07	0.039	.0848626	3.129052
Llcity	-1.911154	.9165206	-2.09	0.037	-3.711055	-.1112529
Kasungu	-.1584146	.0784933	-2.02	0.044	-.3125631	-.0042662
Karong	-.9341524	.4494132	-2.08	0.038	-1.816729	-.0515761
Dowa	-.0221117	.016248	-1.36	0.174	-.0540203	.0097969
Dedza	-.6556192	.3168389	-2.07	0.039	-1.277841	-.0333979
Chitipa	-.3013575	.1472959	-2.05	0.041	-.5906232	-.0120918
Chiradz	-1.034665	.4982785	-2.08	0.038	-2.013205	-.0561249
Chikwaw	-.8529563	.4121149	-2.07	0.039	-1.662285	-.0436279
_cons	45.5682	21.18015	2.15	0.032	3.973732	87.16266