THE COMMON POOL PROBLEM IN FISCAL POLICY MAKING: AN EXPLORATORY EXPERIMENTAL ANALYSIS

BY

ADEDEJI PETER ADENIRAN

SUPERVISORS:

PROFESSOR OLUYELE AKIN. AKINKUGBE

PROFESSOR CHRISTOPHER MALIKANE

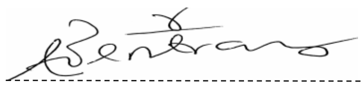
A thesis submitted to the Faculty of Commerce, Law and Management, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the Degree of Doctor of Philosophy

December, 2017

DECLARATION

I declare that this thesis is my own unaided work. It is being submitted for the Degree of the Doctor of Philosophy in Economics, University of the Witwatersrand, Johannesburg. It has not been submitted before to this university or any other university for any degree or examination.

Name of Student: Adedeji Peter Adeniran

Signed:  --

Date: December, 2017

DEDICATION

I dedicate this dissertation to the Almighty God, the invisible hand behind the success of this work; to my parents—Pastor and Deaconess Adeniran—who laid the foundation for my learning pathways and finally to my friend and partner—Idowu—the fountain of my strength and perseverance.

ACKNOWLEDGEMENTS

First and foremost, I want to profoundly appreciate my supervisor Professor Oluyele Akinkugbe for his painstaking efforts, guidance and intellectual mentorship through my PhD programme. I would also like to thank my co-supervisor, Professor Chris Malikane, for his insightful comments and advice. I am equally grateful to my friend and colleague, Emmauel Eayifah, who greatly assisted in the design of the experiment and proofreading of the instruction set. Also, I appreciate the efforts of Olanrewaju Adediran, Moses Muthijah, Denis Okugu, Jason Bell, Michelle Capazario, Dunstan Casandra, and Sedjro Alovokpinhou who assisted in the running of various experimental sessions. My sincere gratitude also goes to all members of my family for their support, understanding and encouragement. Lastly, the financial support from the University of Witwatersrand, Johannesburg, through the Postgraduate Merit Award (PMA) is graciously acknowledged.

ABSTRACT

This thesis explores implications of the common pool problem in fiscal policy making on fiscal performance. Three main research problems were investigated, using various laboratory experiments. First, we examine the effect of dynamic common pool problem on fiscal performance. Second, we test the delayed stabilization hypothesis as predicted in the dynamic common pool model. Third, we investigate the effect of multiyear budgeting on fiscal performance. Along the highlighted research problems, the thesis is structured into three core chapters that follow the much broader and general introductory chapter—chapter one.

Chapter two examines the effect of the dynamic common pool within budget institution on fiscal performance. Dynamic common pool problem occurs when fragmentation within budgeting institution creates uncertainty in the future fiscal path, thereby increasing the proclivity for budget actors to strategically incur deficit and draw down on the public-sector wealth. Given the limitation of field data in exploring a dynamic and strategic setting, the experimental method is adopted. The experimental data generated are estimated using non-parametric techniques and system GMM technique. Overall, the findings indicate that dynamic common pool problem leads to poor fiscal performance, as measured by deficit level. In addition, we observe that budget actors react to the declining public-sector wealth with more aggressive appropriation behaviour, which further exacerbates the poor fiscal outcomes. Furthermore, the effect of dynamic common problem was found to be propagated through the strategic channel. In essence, despite the absence of the conventional fundamentals for incurring fiscal deficit—the divergence between discount rate and interest rate— poor fiscal performance persists due to the dynamic common pool problem.

Chapter three empirically tests the delayed stabilization hypothesis, as predicted by the dynamic common pool model. The hypothesis suggests that there is stabilization threshold at which the presence of dynamic common pool problem no longer exerts a negative effect on fiscal outcome. In testing this hypothesis, we adopt an experimental design which builds on and extends the legislative bargaining game of "divide-the-dollar". This is implemented by applying the random stopping rule after a pre-specified period. Also, tax is imposed on the groups that exhaust their initial public-sector wealth. However, to eliminate possible identification problem, the initial public-sector wealth is restored after some periods of paying tax, although participants are not

aware of this replacement *ab initio*. Thus, it becomes possible to evaluate if the appropriation behaviours follow the pattern predicted by delayed stabilization hypothesis. Furthermore, the experimental data generated are analysed using both fixed-effect panel threshold regression model and piecewise linear regression model.

The results do not support the prediction of delayed stabilization. In fact, the deficit level is highest in the periods after which the group's initial public-sector wealth is restored; a clear contradiction of the delayed stabilization hypothesis. However, we observe a temporary stabilization during the periods when resources pool is generated by taxing the players. This evidence in part supports the crisis hypothesis, which suggests that probability of successful stabilization is enhanced when it coincides with economic crisis. Thus, this suggests a possible role for resources composition (whether taxation or natural endowment) and economic crisis in stabilization policies and plan.

In chapter four, we examine the implication of the voting equilibrium model for the relationship between multiyear budgeting and fiscal performance. Voting equilibrium model recognizes that budget actors are rational individual, who are also seeking to maximize their utility in relation to the budget proposal. Thus, voting equilibrium model predicts that voting preference of the budget actor will influence the eventual budget outcome, thereby influencing any possible effect of multiyear budgeting on fiscal performance. This prediction is tested in a laboratory setting for a five-member budgeting institution, under two different treatment conditions. In the baseline treatment, participants make their budgeting decision using multiyear budgeting procedure, while annual budgeting procedure is adopted in the secondary treamnent. The budget outcomes under two treatments are compared using various non-parametric techniques and Selten's predictive success measure.

We find no signifcant different in budget sizes under the multiyear and annual budgeting proceses. As predicted by voting equilibruim, the multiyear budgeting generates lower budget size than annual budgeting and vice versa, according to the configuration of the voters' preference. It is also observed that the predictive success of the voting equilinruim model reduces as the dimension of the planning horizon increases. In essence, increasing the dimension of planning horizon has the tendency to increase the uncertainty in the budgeting outcomes.

Chapter five concludes with a summary of the key findings and discusses the important policy implications of our findings. An important policy addition from this study points to the significant improvement in fiscal performance that could be gained by addressing the dynamic common problem within budget institution. Similarly, the study demonstrates the importance of adopting active stabilization policy such as fiscal rules or centralization of the budget institution, as against relying on budget actors to act endogenously to correct deficit bias. Finally, in designing the possible fiscal interventions, our findings suggest that the voting preference of budget actors will be crucial.

TABLES OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AMCP	African Monetary Cooperation Program
AR (2)	Second Order Autoregressive
FRF	Fiscal Reaction Function
GDP	Gross Domestic Product
GMM	Generalized Method of Moment
HIPC	Highly Indebted Poor Countries
IMF	International Monetary Fund
IPS	Im, Pesaran and Shin
LLC	Levin-Lin-Chu
MTBF	Medium-Term Budgetary Framework
MTEF	Medium Term Expenditure Framework
MTFF	Medium Term Fiscal Framework
MTPF	Medium-Term Performance Framework
MWC	Majority Winning Coalition
OECD	Organization of Economic Cooperation and Development
OLS	Ordinary Least Squares
RHS	Right Hand Side
RSR	Random Stopping Rule
USA	United States of America
VECM	Vector Error Correction Model
VEM	Voting Equilibrium Model
Z-TREE	Zurich Toolbox for Readymade Experiments

1. Chapter One: Introduction

1.1 Background to the Study

Fiscal policy is no doubt an important economic management tool for countries to achieve macroeconomic stability and promote economic growth. Fiscal policy is defined as the use of taxation, public expenditure and/or deficit financing to influence economic activities (Horton & El-Ganainy, 2006). Despite the crucial role that fiscal policy plays in economic management, recent experiences of persistent and high levels of debt in many countries have become a concern for policy makers and researchers. In Africa, for example, high debt levels have caused a severe drag on economic development (Lopes da Veiga, Ferreira-Lopes, & Sequeira, 2014). Concerned about the sustainability of fiscal policy in developing countries and burgeoning debt levels associated with same, the International Monetary Fund (IMF) and World Bank in 1996 introduced the highly indebted poor countries (HIPC) initiative, which conditionally led to cancellation of the debt of many African countries. Despite this initiative, deficits tend to persist as data from African Economic Outlook (2015) shows that in 2014, about 64 percent of countries in Africa exceeded the deficit-GDP ratio benchmark of 3 percent recommended by African Monetary Cooperation Program (AMCP) as part of the fiscal convergence criteria. Similar experiences of poor fiscal performance and unsustainability have been documented in other regions of the world (see Corsetti et al., 2015 & Yartey, Narita, Nicholls, & Okwuokei, 2012 for the case of Caribbean and European countries respectively).

The tendency towards unsustainable fiscal policy, so termed deficit bias in the literature, as being observed in many countries is becoming increasingly difficult to reconcile with the neo-classical theory—tax smoothing hypothesis. According to the tax smoothing hypothesis, deficit is basically a temporary government response to unexpected increase in expenditure such as when an economy slides into recession (Barro, 1979). Thus, tax smoothing hypothesis predicts that over the long-run, fiscal policy should be sustainable. Given this shortcoming of the tax smoothing hypothesis, recent literature attributes the persistence of deficit bias to the common pool problem in fiscal policy making (Von Hagen, 2002; Wyplosz, 2012).

Common pool problem[1] refers to inefficiencies that arise when multiple decision makers with redistributive objective(s) have control over fiscal policy (Velasco 2000). Essentially, public revenue, once in the government coffers is a joint property of all decision makers. In a fragmented budget setting, there is incentive for individual actor to seek maximization of allocation to their district or sector. Aggregating this spending bias across budget actors leads to overspending and debt accumulation. The result of the fragmentation or common pool problem within budget institution therefore leads to poor fiscal outcomes. That is, the poor fiscal performance results from failure of individual budget actor to internalize the cost of providing public goods, while trying to maximize the benefits.

Perotti and Kontopoulos (2002) identify two forms of fragmentation that could exist within the fiscal institution as size and procedural fragmentations. Size fragmentation means the budgetary institution consists of many decision makers, each with power over the expenditure/tax items. At the executive level, the players consist of spending ministers; at the legislative level the players are constituted by parliamentary committees. The extent of size fragmentation are mostly at the prerogative of the executive and legislative arms of government. Procedural fragmentation on the other hand implies there is a weak structure governing the process of decision making among the players. This could be the case if each player can decide on allocation to its district/sector or bilateral arrangement is allowed with the finance minister. Procedural fragmentation emanates from constitutional or electoral rule regarding budgeting formulation and negotiation. In essence, the main source of size and procedural fragmentations is defined by the extant political institutions and rules. Overall, interaction of size and procedural fragmentations generate the channel through which common pool resources are overexploited.

This thesis is a collection of three essays on common pool problem in fiscal policy making. Although each essay addresses distinct research question(s), they are broadly unified within the common pool literature and by the research method adopted. Specifically, experimental approach is explored in the data generating process.

Common pool problem in fiscal making involves strategic interaction and dynamic setting, which are more suitably studied using experimental approach. This study therefore adds to the growing

[1] In the literature, common pool problem is also referred to as the government fragmentation hypothesis. Thus, the two terms are used interchangeably in this thesis.

literature adopting experimental technique in exploring the role of budgeting institution in fiscal performance. Similar studies along this lineage include, but not limited to: the strategic use of fiscal deficits (Sutter, 2003); dynamic free rider problem in fiscal policy (Battaglini, Nunnari, & Palfrey, 2012a); effect of budgeting process on fiscal performance (Ehrhart, Gardner, von Hagen, & Keser, 2007) and static common pool problem in fiscal policy making (Momeni, 2015) among others.

Economic experiment basically mimics a given market or institutional scenario in a controlled setting. For example, students take the role of decision makers, while budget institution is captured in the instruction set provided. In mimicking this controlled setting, economic experiment is required to conform to induced value conditions as postulated by Smith (1976, 1982). Smith (1982: 931) shows that experiments which meet the induced value conditions could be adjudged as being close representations of microeconomic systems. The induced value conditions are:

1) *Non-satiation condition*: This is also referred to as monotonicity condition. It implies that if subjects are faced with two equivalent alternatives which are only differentiated by the reward medium, then rationality dictates that the choice with higher reward will be preferred (Smith, 1976). Simply put, economic agent will prefer more reward to less or less to more in case of "bad". Most economic experiments use cash as incentive in order to meet non-satiation condition (see Cassar & Friedman, 2004).
2) *Saliency condition*: This condition dictates that reward accruing to subjects must be proportional to the realized experimental outcomes (Smith, 1982). To meet this condition, the experiment must tie the monetary reward offered to subjects to the cumulative payoff over the entire experiment.
3) *Dominance and privacy condition*: This condition emphasizes that the reward medium must dominate other subjective components of a subject's utility function (Cassar & Friedman, 2004). Smith (1982) separates dominance and privacy as different conditions, but Cassar and Friedman (2004) argue that privacy is a means of achieving dominance condition in a laboratory experiment. In this regard, confidentiality of the decisions made by participants during the experiment is crucial. Also, to ensure dominance, expected reward must at least be equivalent to the opportunity cost of time spent participating in the experiment.

There are a number advantages to studying common pool problem in fiscal making using laboratory experiment. First, common pool problem is as a result of strategic interactions among rational actors within budget institutions and these interactions are not well captured within field data. Field data basically captures happenstance observations from the perspective of individual economic agents, whereas the experimental data provides details on how economic agents interact and their implications for economic outcomes. In this regard, experimental approach allows for deeper and broader understanding of common pool problem as it relates to fiscal policy.

Second, laboratory experiment could help in clarifying the role of common pool problem in explaining deficit bias in relation to alternative theories in the literature. For example, the strategic deficit model of Tabellini and Alesina (1990) emphasizes a similar channel to common pool problem as the source of deficit bias. In this instance, experimental approach allows for control of decision setting, thereby ensuring the specific conditions and assumptions of a theory is tested. This ensures that a clear causal link can be established, from which important policy implication can be drawn. Relating this to the present study, in the absence of a well-designed laboratory experiment detecting the effect of dynamic common pool problem from alternative effects will be difficult.

Third, experimental approach is important for policymakers, especially for new and untested policy decisions. Essentially, laboratory experiment could be used as a testing ground, before actual implementation of policy. For example, an examination of the effectiveness of multi-year budgeting as done in this study provides a better understanding of the possible challenges that policymakers need to guide against in designing effective policy options to mitigate deficit bias.

Despite these advantages, issue of external validity has been raised as a major criticism of experimental approach. This relates to concerns about generalizing the experimental evidence to policy environment or field setting. However, as Cassar and Friedman (2004) noted, this criticism also affects any experimental or empirical evidence; hence there is need to rely on the principle of induction. That is, evidence observed in laboratory experiments can be expected to hold under similar condition in the real world. The key to achieve this is through replication of experimental results under different conditions to ensure robustness of evidence. In essence, moving from experimental evidence to policy formulation requires a multiple validation and replication of results. In addition, field data and other empirical approaches could complement subsisting experimental evidence, in order to clarify appropriate policy options. In summary, understanding the advantages and limitations of experimental approach is crucial to its applicability within policy space.

1.2 Statement of Problem

The effect of common pool problem or government fragmentation, from a theoretical perspective, can be explained from both the static and dynamic dimensions. In the static model pioneered by Weingast, Shepsle, and Johansen (1981), government fragmentation leads to spending bias due to failure of the competing budget actors to fully internalize the cost of providing public goods into their spending and taxation decisions. In the dynamic model introduced by Velasco (1997), deficit bias results from strategic interaction among budget actors, leading to excess spending in the current period and reduction in future public-sector wealth.

Empirical literature has found substantial support for the common pool theory as an explanation for the poor fiscal performance in many countries—developed and developing. However, this strand of the literature has largely concentrated on the static model of common pool problem. Romer (2004) and Alesina and Perotti (1999) conclude that lack of evidence on the dynamic effect of fragmentation represents a key limitation of common pool theory. This is because the static model only predicts spending bias, which is of less concern from policy perspective, compared to deficit bias that could be predicted by a dynamic model. This limitation has been addressed in the literature by assuming that revenue is exogenous, which then implies that spending and deficit behaviours are equivalent. However, recent empirical evidence has showed that effect of fragmentation differs between expenditure and deficits (Perotti & Kontopoulos, 2002), thereby refuting the assumption of exogenous revenue. In essence, the static model results on the effect of common pool problem in fiscal policy making are insufficient and weak. There is, therefore, the need to separately examine, in a dynamic framework, the effect of government fragmentation on fiscal performance.

Furthermore, the dynamic model is of interest due to its rich policy predictions that are unobserved in the static model. For example, the possibility of fiscal stabilization has been suggested in the dynamic framework (Velasco, 1997, 1998). Fiscal stabilization is a change in fiscal stance, either through tax increase or expenditure cut, to restore fiscal sustainability (Alesina, Ardagna, &

Trebbi, 2006). Velasco (1997) notes that the stabilization process is driven in large part by the dynamics of public-sector wealth; when the public-sector wealth is high, aggressive appropriation of resources occurs, leading to debt accumulation, but as public-sector wealth declines, the efficiency gain accruing from fiscal stabilization becomes attractive to each budget actor. This phenomenon has been described in the literature as "*delayed stabilization hypothesis*" (Alesina & Drazen, 1989; Velasco, 1997). Again, to the best of our knowledge, there exists to date, no research endeavour that explicitly examines this prediction[2].

Several countries have introduced multiyear budgeting procedure as a response to the common pool problem. World Bank (2013) noted that 80 percent of countries globally have adopted various versions of medium term budgeting. Multiyear budgeting is expected to have positive impact on fiscal performance based on the assumption that it corrects the short-sightedness embedded in annual budgeting which tends to generate the deficit bias (Grigoli, Mills, Verhoeven, & Vlaicu, 2012). Paradoxically, empirical evidences on the significance of multiyear budgeting on fiscal performance are inconclusive. One explanation for this paradox is provided in the voting equilibrium model developed by Ferejohn and Krehbiel (1987), which postulated that forward-looking budget actors will incorporate the effect of say voting, on expenditure in the medium-term on the time-path of budget. The implication is that budget procedure (such as top-down or multiyear budgeting) will have an ambiguous effect on fiscal performance. Existing studies have mostly examined this postulate with respect to top-down budgeting, even though it has similar consequence for multiyear budgeting. In essence, there is need to investigate the implication of voting equilibrium model on the relation between multiyear budgeting and fiscal performance.

Given the apparent gap in the literature, this study seeks to proffer answers to the following research questions?

i. What is the implication of dynamic common pool problem within budget institution on fiscal performance?

[2] The only study we are aware of in the literature that explicitly examines delayed stabilization is Alesina et al. (2006). However, their study is based on the war of attrition model, and not the common pool problem model which the present study focuses on. The implication is that source of deficit bias is basically imposed rather than being endogenously determined, as will have been necessary in testing delayed stabilization in relation to common pool problem model.

ii. Can public-sector wealth dynamics create a coordinating mechanism leading to delayed stabilization?
iii. What is the effect of multi-year budgeting on fiscal performance?

1.3 Research Objectives

The broad objective of this study is to explore the implication of common pool problem of fiscal policy making on fiscal performance. In line with this broad objective, the sub-objectives are to:

i. examine the effect of dynamic common pool problem on fiscal performance;
ii. examine the responsiveness of budget actors through delayed stabilization to public-sector wealth dynamics;
iii. evaluate the effect of multiyear budgeting on fiscal performance.

1.4 Study Rationale and contribution

Chapter two contributes to empirical studies on common pool problem in fiscal policy making. Specifically, the chapter extends the existing studies on static common pool problem of budget institution to a dynamic setting. Given that both static and dynamic common pool problems have different channels through which they affect fiscal performance, it becomes imperative for policymakers to understand how these different channels operate, in order to design appropriate policy measures. The chapter also contributes to the strand of literature that investigates the importance of strategic effect channel. A number of political economic models on fiscal performance, such as the strategic deficit model of Tabellini and Alesina (1990), have underscored the crucial role of this channel. In a similar vein, this study examines the role of strategic effect channel. However, while previous studies have mostly dwelt on strategic effect resulting from polarization and re-election uncertainty, the present study focuses on its importance in relation to the common pool problem.

Chapter three empirically tests the delayed stabilization hypothesis. Establishing the validity of delayed stabilization hypothesis or otherwise is crucial to designing an appropriate timing and strategies to address deficit bias. For example, while fiscal rules is often times advocated to address the problem of deficit bias, the possibility of delayed stabilization implies that budget actors could endogenously take action that correct the deficit bias. Thus, this study is a step in the direction of facilitating deeper understanding of areas of potency for fiscal rules as well as measures that could better complement its effectiveness. Implicitly, the chapter also helps to illuminate on the possible role of economic crisis for successful stabilization. In essence, the experimental design for the chapter separately examines the subjects' appropriation behaviour with and without taxation (equivalent to having fiscal crisis). This represents an empirical contribution to the literature on "crisis hypothesis".

The fourth chapter contributes to policy discourse on the effectiveness of multiyear budgeting. Given the mixed finding on the performance of multiyear budgeting, this study serves to deepen policymakers' understanding of institutional enablers (such as voting preference, decision sequence, and dimension of planning horizon) that could be crucial to its effectiveness. Furthermore, in deriving the testable predictions for empirical analysis, the study extends the voting equilibrium model of Ferejohn and Krehbiel (1987). Specifically, the voting equilibrium model is extended to a setting in which budgetary decision involves intertemporal allocation. This provides a framework to compare the adequacy of alternative budgeting processes across periods.

The final contribution of the study is from a methodological perspective; the three essays adopt experimental method. In relation to dynamic setting, this method is yet to gain any significant attention in the context of common pool problem in fiscal policy. However, experimental analysis has been used to explore related research objective on the common pool problem in natural resource exploration. For example, Mason and Phillips (1997) examine how repeated interactions and number of firms in an oligopoly market could affect resources exploitation in the commons. Similarly, Osés-Eraso, Udina, and Viladrich-Grau (2008) investigate, in a dynamic setting, the effect of reduction in stock of natural resources on appropriation behaviour. While these studies address related issue, the experiments conducted in the current study include features of fiscal institutions that are not present in natural resources setting. For example, budget constraint is binding on actors in a fiscal setting, unlike natural resource exploitation experiment in which resource preservation is not binding or its extinction costless[3].

[3] The possibility of debt default also implies that budget constraint may in reality not be binding in fiscal policy. However, debt default is not costless and countries will rather avoid it. Rose (2005) estimated that debt default reduced bilateral trade between debtor and creditor countries by approximately 8 percent and persists for 15 years. In addition, there is mild effect on reputation measured by downgrade in credit rating, reduction in domestic output and shorter duration of government (Borensztein & Panizza, 2009).

The rationale for adopting experimental analysis is due to difficulties posed when using field data to address the aforementioned research questions. For example, the dynamic analysis of common pool problem requires the existence of a state variable such as the public-sector wealth or dynamic externality that links the present action to the expected future outcome. However, using field data, these state variables cannot be accurately observed or measured. Moreover, experimental data have higher identification advantage in studying a structured and strategic dynamic environment (Battaglini et al., 2012a). Also, the prediction of the model relies on behavioural assumptions of rational economic agents, infinite horizon and the use of trigger strategy, all of which are most suitably studied in a control condition (Cullis, Jones, & Jones, 2009; Mason & Phillips, 2002).

In general, it must be emphasized that experimental data can only complement rather than substitute the field data. In fact, field data remain crucial in designing an effective institution and policy framework to address deficit bias; experimental data however serves as a starting point in testing the analytical power and validity of the theory, as well as the conditions and institutions under which the theory operates.

1.5 Organization of the Thesis

The rest of the thesis, after this general introductory chapter, is organized into four parts. Chapter two examines the implication of dynamic common problem for fiscal performance. Chapter three investigates the prediction of the dynamic common pool problem regarding delayed stabilization. Chapter four analyses the effect of multiyear budgeting on fiscal performance. Finally, we conclude in chapter five with summary of the major findings, policy recommendations and suggestions for future research.

2. Chapter Two: Dynamic Common Pool Problem in Fiscal Policy Making: Experimental Evidence

2.1 Introduction

The vast empirical literature on the common pool problem in the fiscal policy processes build its underlying theoretical foundation on the static model of Weingast et al. (1981), which presupposes that spending bias is due to limited internalization of the cost of fiscal policy by budget actors. Alesina and Perotti (1999) note that since high level of spending does not necessarily translate to high deficit level, the static model becomes of limited policy relevance. In essence, it is crucial to investigate the mechanism through which dynamic common pool problem generates deficit bias. The dynamic dimension of the common pool problem has received little attention. Dynamic common pool problem occurs when fragmentation within fiscal institution makes the budget actors to strategically incur deficit, thereby reducing future public-sector wealth. Romer (2004) concludes that the lack of empirical evidence and rather limited empirical literature on the dynamic common pool problem represents a major drawback of the deficit bias theory based on common pool model.

This paper explores the dynamic dimension of the common pool problem in fiscal policy. While several scholars have emphasized the need for empirical evidence on this, the scope of field data has constituted a major challenge. For example, Buiter (1985) argues that the standard measure of fiscal performance, cyclically-adjusted deficit, is a poor indicator of government fiscal stance as it tends to exclude changes in net worth of the public sector. The problem with cyclical-adjusted deficit becomes even more pronounced in a dynamic setting, making the use of field data difficult. Furthermore, common pool problem as generally measured by the degree of fragmentation within budget institutions has been observed to be very sensitive to classification of government, thereby creating a comparability problem for cross-country analysis (Padovano & Venturi, 2001).

To address these challenges, this study uses experimental data. A simplified experiment, which to a large extent captures the key features of a dynamic common pool problem faced in fiscal settings, was designed and implemented. The fiscal institution is made up of $n > 1$ legislators representing various interest groups. The decision setting employed follows a legislative bargaining game of "*divide-the-dollar*", played over an infinite horizon. Also, the discount rate and interest rate used are equalized to eliminate conventional intertemporal reason to incur deficit or surplus. Thus, deficit only occurs due to strategic effect as predicted by the dynamic common pool model. Based on this set up, the experimental data generated were analysed to determine the extent to which the appropriation behaviour of players matches the theoretical predictions of dynamic common pool models.

This paper contributes to existing literature on the subject in three distinct areas. First, it extends the literature on static common pool problem in fiscal policy making to the dynamic setting. From policy standpoint, understanding the dynamic common pool problem is crucial to designing appropriate intervention to mitigate the deficit bias. Second, the study contributes to experimental studies on the dynamic common pool problem. Notable earlier studies in this strand of the literature include Herr, Gardner, and Walker (1997), Mason and Phillips (1997) and Osés-Eraso et al. (2008). However the central focus of most of these studies is on the dynamic common pool problem in the natural resource context. This study, on the other hand, relates to a fiscal policy setting.

Third, this paper builds on and extends the experimental studies on fiscal policy by Sutter (2003) and Fréchette, Kagel, and Morelli (2005, 2012). Sutter (2003) developed an experimental approach to test a strategic debt accumulation model. Strategic debt accumulation and dynamic common pool models are related, because the strategic effect is the main driver of deficit bias in the two models. However, while Sutter (2003) analysed the strategic effect arising from political polarization and re-election uncertainty in a two-period dynamic model, this study extends it to an infinite time horizon and focuses more on the common pool problem in fiscal policy making. Fréchette et al. (2005, 2012) examined the implication of different budget procedures—universalism, demand bargaining and alternating-offer with close and open rules—for public goods provision. Again, these works differ from the present study as their focus was on the free riding problem in public goods provision.

The rest of this paper is structured as follows. Section two reviews literature on the static common pool problem in fiscal policy making. In addition, a brief review of experimental studies on the dynamic common pool problem, in the resource exploitation setting is undertaken. Section three describes the dynamic common pool model, which motivates the design of our experiment. In section four, we detail the experimental procedure and implementation strategy for the study. Section five presents the experimental results from both the non-parametric and econometric analyses. Finally, policy implications of the results are discussed in section six.

2.2 Literature Review

2.2.1 Studies on Common Pool Problem in Fiscal Policy

Extensive literature abounds on the common pool problem in fiscal policy making. Many of these studies examine the effect of government fragmentation on fiscal performance across countries in different regions of the world and many of them are based on the static model which predicts that government fragmentation leads to excess spending. To extend the analysis in explaining deficit or debt, it is generally assumed that revenue is exogenous, which implies that spending and deficit behaviours are equivalent (Eslava, 2011). One of the early evidence on static common pool problem is provided by Von Hagen (1992) which investigates the effect of budgeting procedures in the Euro zone on fiscal discipline as measured by gross debt and net lending ratio. The study finds that a country with processes in which the prime minister or finance minister has strategic dominance over other cabinet members is conducive to improved fiscal discipline. De Haan, Moessen, and Volkerink (1999) extend the analysis more broadly to OECD countries and also report a negative impact of government fragmentation on fiscal performance.

Earlier literature has also identified the common pool problem as partly responsible for the poor fiscal performance in developing countries. For example, Alesina, Hausmann, Hommes, and Stein (1999) replicate Von Hagen's (1992) study in Latin America by constructing an index of budget institutions based on the extent of procedural rules and transparency. Again, countries with fragmented budget institutions are observed to have high primary deficit. Gollwitzer (2011) also reaches a similar conclusion on the negative effect of government fragmentation on fiscal performance in African countries. In essence, there is broad consensus in the literature on the negative influence of the common pool problem on fiscal performance.

However, Perotti and Kontopoulos (2002) criticize the existing literature on the common pool problem in fiscal policymaking in two significant respects. First, they observe that existing literature has focused on procedural fragmentation, whereas the theoretical literature also suggests the notion of size fragmentation. Second, they argue that the underlying static models on which existing studies are based predict spending rather than deficit bias. These observations were tested on a panel of 19 OECD countries and the results indicate that size fragmentation measured by cabinet size has a more significant impact than procedural fragmentation on fiscal performance. More importantly, their study finds that government fragmentation has a higher magnitude and statistically significant effect on expenditure than deficit as predicted in the static model. Similar results have been reported in subsequent studies by Volkerink and De Haan (2001) for OECD countries and Elgie and McMenamin (2008) for a panel of developed and developing countries.

One of the implications of these findings is that the static model is more directly related to spending behaviour. A more profound implication according to Alesina and Perotti (1999) is that policy relevance of evidence provided on the basis of the static model is limited because high government spending level does not necessarily translate to high deficit. Alesina and Perotti (1999) further highlight one of the technical difficulties encountered when exploring the dynamic dimension of common pool problem as being the dilemma faced by budget actors between choice of expenditure in the present period and the different path of public-sector wealth. While at the theoretical level, some measure of progress has been attained in overcoming these difficulties as Velasco (1997) demonstrates, the process through which the dynamic process generates a deficit bias—the empirical evidence—remains elusive. This is mainly due to the limitation of field data in analysing the strategic interactions involved in a dynamic environment. Thus, this study fills an obvious gap in the literature through the use of experimental data. As Battaglini, Nunnari, and Palfrey (2012b) noted "if one is going to take model to data, there are clear identification advantages for experimental studies when studying structured and dynamic environment" (Ibid, pp. 408).

2.2.2 Experimental Studies on Dynamic Common Pool Problem

Several laboratory experiments have been conducted to examine the dynamic common pool problem. However, these experiments are designed to evaluate the problem in the context of natural resource exploitation. For example, Herr et al. (1997) compare the appropriation externality that is present in the use of common-pool resources such as groundwater aquifer in both static and dynamic settings. They find that in the dynamic setting, the common pool problem is worsened, with appropriation levels significantly above the theoretical prediction by the non-cooperative equilibrium. Specifically, the players in the experiment are observed to adopt a myopic strategy by neglecting the cost that the current extraction rate has on the sustainability of resources.

Mason and Phillips (1997) also investigate the welfare effect of increasing the number of firms extracting resources in the commons. The experiment finds that an expansion in the industry size leads to a significant welfare gain in the static setting, but has the opposite effect in the dynamic environment. A similar result is reported by Fischer, Irlenbusch, and Sadrieh (2004) in their experimental analysis of intergenerational altruism hypothesis. Subjects in the experiment display altruistic motive for future generations in their demand for resources, but the motive is offset by financial incentive gained from high appropriation in the current period. In addition, they observe an increasing tendency to free ride on future generations, given the uncertainty regarding the extent to which other players are concern about sustainability of the resources.

Generalizing these results to the fiscal setting has some obvious limitations. For example, resources in the fiscal setting are reversible, which means private goods can be transformed back to public goods through taxation (Velasco, 1997). This feature could amplify the effect of uncertainty, leading to more aggressive appropriation when resources are abundant and more cooperative behaviour during scarcity. In the natural resource context, on the other hand, it is widely observed that cooperation is less likely either when resources is scarce or in abundance, which may suggest a Laffer curve relationship between resource level and cooperation (De Janvry, McCarthy, & Sadoulet, 1998).

Nevertheless, there are still useful inferences that could be drawn. Specifically, the dynamic common pool problem differs from the static problem. Thus, choosing the policy measures to address poor fiscal performance will depend on which dimension of common pool problem is manifesting within the fiscal institution. In essence, there is no one-size-fits-all measure to address both static and dynamic common pool problem. Even though the dynamic setting does not seem to ameliorate all the difficulties associated with the common pool problem, it provides a more useful method of approach by introducing repeated interactions in the fiscal policy making process.

2.3 Theoretical Model

2.3.1 The Dynamic Common Pool Model

The theoretical literature on the common pool problem in fiscal policy making was first proposed by Weingast et al. (1981). According to the authors, the existence of common pool problem in fiscal policy process generates spending bias as interest groups only internalize the benefit of public goods while neglecting the cost imposed on collective tax resources. However, the model is static in nature and can only explain the persistence of high government spending rather than deficit or debt. As Perotti and Kontopoulos (2002) argue, spending bias is less of a concern for fiscal sustainability in as much as revenue rises proportionately. Thus, the dynamic common pool model has been proposed as a remedy to this shortcoming. Originating with seminal work of Chari and Kehoe (1993), the dynamic common pool model was further extended by Velasco (1997, 1998, 2000). The conceptualization and formulation that follows are based on Velasco (1998). However, as a motivation for our experimental design, the basic model in Velasco (1998) is extended to "*n*" interest groups.

Consider an economy consisting of n symmetric interest groups (indexed $i = 2,3,...n$), which are equally represented within the budget institution, say by legislators from different interest groups or constituencies. For expositional convenience, the model is restricted to the domain of distributive policy; the extension to non-distributive policy complicates the model with little additional insight. The main assumption of the model is that appropriation decision takes place under a fragmented fiscal structure. This implies that a legislator from interest group i could directly decide the level of spending g_i accruing only to it. This assumption can be interpreted as budget procedure following "universalism criteria". "Universalism criteria" ensures that every group get a share of the budget, resulting into a near-unanimous approval of budget proposal (Shepsle & Weingast, 1981). From the perspective of the individual budget actor(s), universalism approach is superior and more acceptable than majority rule in deciding appropriation (see Koh, 2015; Shepsle & Weingast, 1981). In line with the literature, the assumption of $n > 1$ implies the existence of size fragmentation, while the assumption of weak fiscal structure implies procedural fragmentation.

The state variable of the model is defined by the public-sector wealth that relates behaviour across periods. The public-sector wealth represents the income side of government budget and its flow is assumed to be constant overtime[4]. Public-sector wealth is a common property of the various interest groups as well as debt burden incurred. Every period, each legislator chooses spending/appropriation level (g_i) condition on stock of public-sector wealth. The government inter-temporal budget constraint over infinite-horizon is given as follows:

$$W_t = (1+r)W_{t-1} - \sum_{i=1}^{n} g_{it} \qquad (1)$$

[4] This assumption allows us to focus on the expenditure side of government budget. This is not to ignore the issue around government revenue, but in the interest of realism. In practice, expenditure and revenue are separated, with bureaucrats involved in revenue generation, while politicians deciding the expenditure items. It is this separation of revenue and expenditure decisions that give rise to common pool problem.

where W is the public-sector wealth, defined as net financial assets including present value of all revenue, r is the constant real interest and subscript t denotes timing of the event. The initial stock, W_0, is taken as positive and is exogenously given. Equation (1) implies, *ceteris paribus*, that changes in public-sector wealth overtime could only occur through fiscal deficit or surplus. To rule out incidences of default, the solvency condition is imposed on the budget constraint as:

$$\lim_{t\to\infty} W_t (1+r)^{-t} \geq 0 \tag{2}$$

Furthermore, since the spending decision across interest groups is not coordinated, it is important to impose a spending limit to prevent violation of the solvency condition. Thus, the maximum spending for each group is set at $g_{it} \leq (1+r)W_t / n$ or the group gets zero if violated. This condition implies that:

$$\sum_{i=1}^{n} g_{it} \leq (1+r)W_t \quad \forall t \tag{3}$$

Lastly, the interest groups are assumed to have identical preference over g_{it} represented by the utility function (index):

$$U_i = \sum_{s=t}^{\infty} Log(g_{is})(1+r)^{-(s-t)} \tag{4}$$

A logarithm utility function is used due to its simplicity, but the result is robust to alternative functional specification[5]. It should be noted that equation (4) takes the discount rate as equal to interest rate across periods, thereby excluding the standard inter-temporal explanation for fiscal deficit or surplus. With this set up, the group objective is to maximize equation (4) conditional upon equation (1).

2.3.2 Social Planner versus Fragmented Government Solution

2.3.2.1 Social Planner Solution

In order to benchmark the outcome of the fragmented government decision, it is important to first determine the level of g_{it} that a benevolent social planner will choose, denoted as the optimal

[5] Velasco (2000) uses a quadratic loss function and derives similar result.

spending, g_{it}^*. The social planner jointly maximizes the utility of the interest groups, subject to the budget constraint. By imposing symmetry across the interest groups, the Lagrangian function to the social planner problem is given as:

$$L = \sum_{s=t}^{\infty} Log(g_{is})(1+r)^{-(s-t)} + \lambda(W_t - (1+r)W_{t-1} - \sum_{i=1}^{n} g_{it} \qquad (5)$$

The first order condition is derived as:

$$g_{i,t+1} = g_{it} \qquad (6)$$

Substituting equation (6) into the equation (1) gives the spending level that will be chosen by the social planner as:

$$g_{it}^* = \frac{rW_{t-1}}{n} \qquad (7)$$

Equation (7) shows that the social planner equalizes spending across groups and periods, thereby balancing the budget at all times. A benevolent social planner will incur fiscal deficit or surplus only when interest rate is not equal to discount rate, as social welfare can be improved as a result[6]. Since this is ruled out in our analysis, there is no incentive to borrow or save.

2.3.2.2 Fragmented Government

In the absence of central coordination, each legislator decides the level of spending based on equation (4). The players are involved in a non-cooperative game and their actions can be characterized by the symmetric Markov perfect equilibrium. With the Markov perfect equilibrium, each player action can be summarized by a policy rule which is a function of the state variable only. Thus, following on Velasco (1998), we define a policy rule that is a simple linear function of the public-sector wealth as:

$$g_{it} = \varphi W_{t-1} \qquad (8)$$

Based on this policy rule, the budget constraint facing each group becomes:

[6] Another reason for deficit highlighted by Krogstrup and Wyplosz (2010) is the existence of productive spending that generates revenue for future periods. Again, we exclude this case from the analysis in order to focus on deficit resulting mainly from government fragmentation. In fact, it easy to show that if the effect of the productive spending is non-durable, then the optimal level provided by either the social planner or fragmented government is zero.

$$W_t = [1 + r - (n-1)\varphi]W_{t-1} - g_{it} \tag{9}$$

The new budget constraint in equation (9) shows that public-sector evolves over time based on the interest rate and the rate of spending "φ". The Bellman equation (or valuation function) corresponding to this problem is:

$$V(W_t) = [Log(g_{it}) + (1+r)^{-1}V(W_{t+1})] \tag{10}$$

The legislator representing group i therefore maximizes equation (10) subject to equation (9). The Euler equation associated with the set-up yields:

$$\frac{g_{i,t+1}}{g_{it}} = \frac{(1+r)-(n-1)\varphi}{1+r} \tag{11}$$

In addition, if all groups follow the policy rule given by Equation (8), then Equation (9) becomes:

$$W_t = W_{t-1}(1 + r - n\varphi) \tag{12}$$

Combining equations (11) and (12) gives:

$$\frac{W_t}{W_{t-1}} = \frac{g_{i,t+1}}{g_{it}} \tag{13}$$

Based on equation (13), we derive the optimal value of φ as:

$$\varphi = \frac{r(1+r)}{1+nr} \tag{14}$$

Finally, we can substitute for φ into equations (8) to derive the optimal spending level in the non-cooperative game as:

$$g_{it}^{nc} = \frac{r(1+r)W_{t-1}}{1+nr} > g_{it}^{*} \tag{15}$$

Similarly, substituting φ into equation (9) yields:

$$\frac{W_t}{W_{t-1}} = \frac{1+r}{1+nr} < 1 \tag{16}$$

Equation (15) shows the spending level derived from the interaction among fragmented budget actors using Nash/non-cooperative strategy. The spending level chosen, g_{it}^{nc} , is greater than g_{it}^{*} , which implies that deficit is incurred. Existence of deficit in this model is mainly due to strategic effect. Specifically, the interest groups are in a prison dilemma game, where maximizing their respective utility constitutes the Nash equilibrium of the game, but a sub-optimal outcome compared to the social planner choice. This dilemma is identical to the tragedy of commons faced in natural resources exploitation. In both cases, the coordination problem among interest groups, leads to overexploitation of resources or budget deficit. In fact, the mechanism driving the two outcomes is similar—lack of property right over the resources. Similarly, equation (16) shows the path of public-sector wealth over time. Evidently, the public-sector wealth declines and the rate at which this occurs correspond to the number of interest groups (equivalently, the degree of fragmentation in fiscal policy making). Compared to the case where $n = 1$, which reduces to social planner solution, the public-sector wealth level remains constant over time.

This result shows that deficit is endogenously generated as a result of fragmentation of budget institution. Substituting equations (7) and (15) into group i utility function (equation 4) reveals that the instantaneous utility derived from non-cooperative strategy is higher compared to cooperative strategy. However, the present value of discounted utility obtained along each path, presented in equations (17), (18) and (19) that follow indicates that cumulative utility from cooperative/stabilization strategy is strictly higher than that of non-cooperative strategy.

$$U_i^{nc} = \frac{1+r}{r}[Log(\frac{r(1+r)}{1+nr}W_{t-1}) + \frac{1}{r}Log(\frac{1+r}{1+nr})] \tag{17}$$

$$U_i^{*} = \frac{1+r}{r}[Log\frac{rW_{t-1}}{n}] \tag{18}$$

and

$$U_i^{nc} < U^{*} \tag{19}$$

In equation (17), the first term in the bracket on the right hand side (RHS) is the static gain accruing to interest groups; this moreover comes at a cost of lower wealth in the future as captured by the second term. The second term is the shadow cost that current action imposes on future resources; it represents the dynamic externality[7] and negative if $W_t > W_{t-1}$. Only when the interest groups choose on average g^*_{it} is the shadow cost equal to zero. The dilemma faced by each interest group is reflected by the shadow cost, which is determined by the aggregate groups' decision. In essence, there is weak incentive for players to choose the optimal spending that sustains the common resources, as choices of other players are unpredictable at the point of decision making. This makes non-cooperative strategy the equilibrium of this game.

Proposition: *Deficit increases (equivalently, public-sector wealth declines), the higher the level of government fragmentation.*

2.4 Experimental Procedure

This section details the experimental procedure for testing the proposition derived above. We first describe a laboratory setting that captures the basic structure highlighted in equations (1) to (4). The theoretical predictions based on the specified parameters of the model are then derived. These predictions serve as the benchmark with which to compare the results of the experimental design in this study.

2.4.1 Decision Setting

The experiment combines key features of demand bargaining and dynamic common pool experiments used by Fréchette et al. (2005) and Mason and Phillips (1997), respectively. In the demand bargaining experiment, players engage in the distributive politics game of "divide-the-dollar", in which they make sequential demand on a fixed sum of resources. The game continues till a majority winning coalition (MWC) that satisfies the budget constraint is formed. In our set-up, the participants taking the role of legislators make demand (g_i) on the stock of public-sector wealth (W_t), represented by experimental currency (points), which are exchanged for cash at the

---

[7] In general, the dynamic externality is given by $Log(\frac{W_t}{W_{t-1}})$ ………. $(17')$

end of the experiment. However, the players' move simultaneously in our case and the fixed sum is shared over an infinite horizon. Once decisions conform to equation (3), there is no need to meet the requirement of MWC, as this constitutes a unanimous decision. The payoff to participant is defined by equation (4), aggregated over the entire rounds of the game. The stock of wealth evolves over time according to equation (1). That is, at the end of each round, the remaining points are multiplied by $1+r$ and they constitute the total resources for the subsequent rounds.

Similar to Mason and Phillips (1997), the dynamic common pool problem is reflected by the shadow cost which captures the dynamic externality a player action imposes on other players. High level of spending in the current period implies low spending level in the future. This is introduced into the experiment through revealing to the participants the shadow cost associated with different levels of spending. Essentially, the shadow cost gives the rate of resource depletion and allows participants to assess the effect of their current action on the sustainability of the resources. The shadow cost is given by equation $(17')$, which becomes zero if the Nash strategy is followed. The participants are informed about how the shadow cost is derived, but this is automatically generated by the computer program managing the experiment.

We mimic the infinite horizon game by introducing a random stopping rule. The rule is defined by the discount factor which is set at $1-r$. This is calculated based on the model assumption that discount rate equals interest rate and it gives the probability of continuing the game to the next round at $1-r$. A similar procedure has been used by Mason and Phillips (1997) and Battaglini et al. (2012a). A random number between 0 and 1 will be generated after each period by the computer and the game proceeds if the number generated is below $1-r$. The game ends either through the random stopping rule or when the points (public-sector wealth) are exhausted by the players. The public-sector is taken as exhausted if the number of points in the group account is below the group size (n).

2.4.2 Design Parameters and Treatments

The key parameters of the experiment are defined by W_t, r, and n. We implement two treatments: primary treatment (low W_t) and secondary treatment (high W_t). Only the value of $r = 10$ percent is fixed across treatment. In the primary treatment, the legislative size[8] is set at $n = 3$ and $n = 5$, combined with initial wealth per person normalized to real numbers within [0, 20]. Hence, the initial wealth for $n = 3$ is 60 points and 100 points for $n = 5$. The initial wealth is normalized across groups to control for possible wealth level effect. The budget constraint, for this treatment, is tight and the possibility of reaching debt ceiling is higher, even if players make judicious appropriation early in the game. As Table 2.1 indicates, the social planner solution that satisfies the parameters is for each participant to appropriate 2 points in each period.

In the second treatment, the legislative size remains the same but the economy is much wealthier, with the initial wealth per person normalized to real numbers in the range [0, 100]. Hence, groups with $n = 3$ have initial wealth as 300 points, while that of $n = 5$ is 500 points. Public-sector wealth decline is not explicitly controlled for as treatment, because we are dealing with behaviourally induced changes. Thus, the secondary treatment majorly compares appropriation behaviour between abundant and low wealth level. Also, the initial wealth is normalized in both primary and secondary treatments to ensure that the results are not driven by wealth level. Overall, there are four treatment categories: low/scarcity wealth and small group, low/scarcity wealth and large group, high/abundant wealth and small group, high/abundant wealth and large group.

Table 2.1 summarizes the theoretical predictions given the specified parameters. The last two columns show the predicted deficit for the first two periods. The deficit is computed as the difference between appropriation levels chosen by the social planner (g_{it}^{*}) and budget actors in a fragmented government (g_{it}^{nc}). The results reveal that by using Nash strategy, the deficit is incurred irrespective of group size and initial wealth. Moreover, despite initial wealth being normalized across groups, predicted deficit is higher within larger groups. This validates proposition above that deficit increase with higher level of government fragmentation, as measured by the group size.

Table 2.1: Theoretical Predictions

Treatment		Strategy		Deficit	
Initial Wealth	Group size	Nash	Social Planner	Period 1	Period 2
60	3	$0.085W_{t-1}$	$0.033W_{t-1}$	9.3	6.9

[8] The choice of legislative size matches a related experiment by Battaglini et al. (2012a) that tested group size effect in a dynamic free rider problem.

100	5	$0.073W_{t-1}$	$0.02W_{t-1}$	26.5	16.8
300	3	$0.085W_{t-1}$	$0.033W_{t-1}$	46.8	34.9
500	5	$0.073W_{t-1}$	$0.02W_{t-1}$	132.5	84.1

Source: Author's Computation

2.5 Sampling Procedure

2.5.1 Optimal Sample Selection

The determination of the optimal sample depends on the type of treatment. There are two types of treatment—dichotomous and varying. The dichotomous treatments are restricted to only two contrasting level of measurement, while varying treatments are allowed to take different levels at a given point in time. Another determinant of the optimal sample is the type of outcome—continuous or binomial. Continuous outcome occurs when the treatment variable are measured on a numerical scale, while binomial outcome relates to ordinal scale. The experiment designed for this study has a varying treatment levels and continuous outcome. Thus, following List, Sadoff, and Wagner (2011), the optimal sample size is given by[9]:

$$M = mk = 2k(t_{\alpha/2} + t_{\beta})^2 (\frac{\sigma}{\delta})^2 \qquad (20)$$

where M is the total sample size; k is the number of treatment groups, which is four in our design; m is the optimal size of each treatment; $t_{\alpha/2}$ and t_{β} are the lower and upper critical values of the t distribution; σ is the variance of the treatment effect; δ is the minimum average effect. The ratio $(\frac{\sigma}{\delta})$ gives the standard deviation of the change in the outcome variable. In line with the literature, we take $t_{\alpha/2}$ and t_{β} as 1.96 and 0.84 respectively (see Gujarati, 2003, p. 122). Therefore, to detect one standard deviation change in the outcome variable, the optimal sample size (M) required is 64 participants. In essence, the experiment retains this optimal size of 64 participants.

[9] The formula given by Equation (18) is adopted on the assumption that the treatment effect is linear. Prior studies strongly suggest this assumption is valid in an experiment dealing with variation in resource stock (see Osés-Eraso et al., 2008; Osés-Eraso & Viladrich-Grau, 2007).

2.5.2 Sampling Technique

Participants in the experiment are recruited from first and second year undergraduate students of University of the Witwatersrand, Johannesburg. First and second year students are preferred due to lower opportunity cost[10] and high learning curve both of which increase the cost-effectiveness of the experiment (see Cassar & Friedman, 2004; Frechette, 2011). To select the participants, targeted emails were sent to students that meet the study's inclusion criteria. The criteria are: (i) the participant must have attained 18 years of age; (ii) the participant must be computer literate; and (iii) the participant must be in the first or second year undergraduate degree program.

A total number of 446 students volunteered to participate in the experiment. Using a simple random sampling procedure, the required 64 participants for the experiment are selected from this initial sampling frame. Table 2.2 shows the demographic distribution of the respondents by sex, year of study and faculty. The distribution reveals that the participants are well spread across each of demographic variable. The only exception is distribution along racial line which is skewed to the blacks. This is as a result of low volunteers from the other races. However, the sample distribution is relatively proportional to size of each race in the total student population. The proportionality of the sampling distribution has two advantages that reduce sampling error. First, the standard deviation is lower, as it ensures that the actual student composition are better reflected. Second, it yields improved statistical precision at lower sample size. In addition, given that subjects are randomly assigned to treatment condition, this eliminates selection bias. In essence, this indicates that sample selection bias is less of a concern for this study.

Table 2.2: Demographic Characteristics of Participants

Variable	Frequency	Percentage (%)
Sex		
Male	30	46.9
Female	34	53.1
Year of Study		
1st	31	48.4
2nd	35	54.7
Race		
Blacks	59	92.2
Indian	2	3.1

[10] For example, Croson (2005) recommended that participants in the experiment must earn the opportunity cost of the time in an equivalent on-campus job. Within the university setting, first and second year students have lowest opportunity cost in this spectrum.

White	3	4.7
Faculty		
Commerce, Law and Management	24	37.5
Health Sciences	9	14.15
Humanities	18	28.1
Science	13	20.3

2.5.3 Experimental Implementation

The experiments for this study were conducted at the PhD Computer laboratory of the University of Witwatersrand in the month of September, 2016. Four separate sessions were run, based on different combinations of treatment. Also, between-subject design was used, such that each participant was only exposed to a treatment. Overall, no session exceeded an hour and the average earnings per participant was R65, in line with induced value theory. This is higher than earning level proposed by Crosson (2005), which recommends that participants' earnings should commensurate with the opportunity cost of engaging in any other on-campus duty over the equivalent period. An equivalent hour spent in part-time student on-campus employment currently pays about R20 and R30 for first and second year students. The sessions were computerized using Z-TREE (Zurich toolbox for readymade experiments) developed by Fischbacher (2007).

The experiment was conducted as follows. Participants, upon arrival, were given the instruction. The instruction set details the group size, random stopping rule, the initial wealth in the group account and the spending rule, together with implications if violated (see appendix 2A). The participants were intimated with the fact that decisions are made simultaneously; hence no coordination with other group members is allowed. However, after each period, the total group demand, the remaining wealth in the group account and the average wealth for the next period were displayed to all group members. Lastly, participants are informed on how the shadow cost will be computed. Throughout the duration of the experiment, the shadow cost and public-sector wealth are referred to as endowment dynamics and endowments respectively, in order to prevent possible framing effect. After fielding questions from participants on the clarity of instructions, there was a pilot-test session before the main experiment.

2.6 Experimental Results

This section reports the findings based on analyses of the experimental data. First, we present a non-parametric analysis of the data. This is followed by the discussion of the empirical strategy and estimation technique for the specified model. Furthermore, we examine the extent to which the results match the theoretical prediction derived from the dynamic common pool model.

2.6.1 Non-parametric analysis

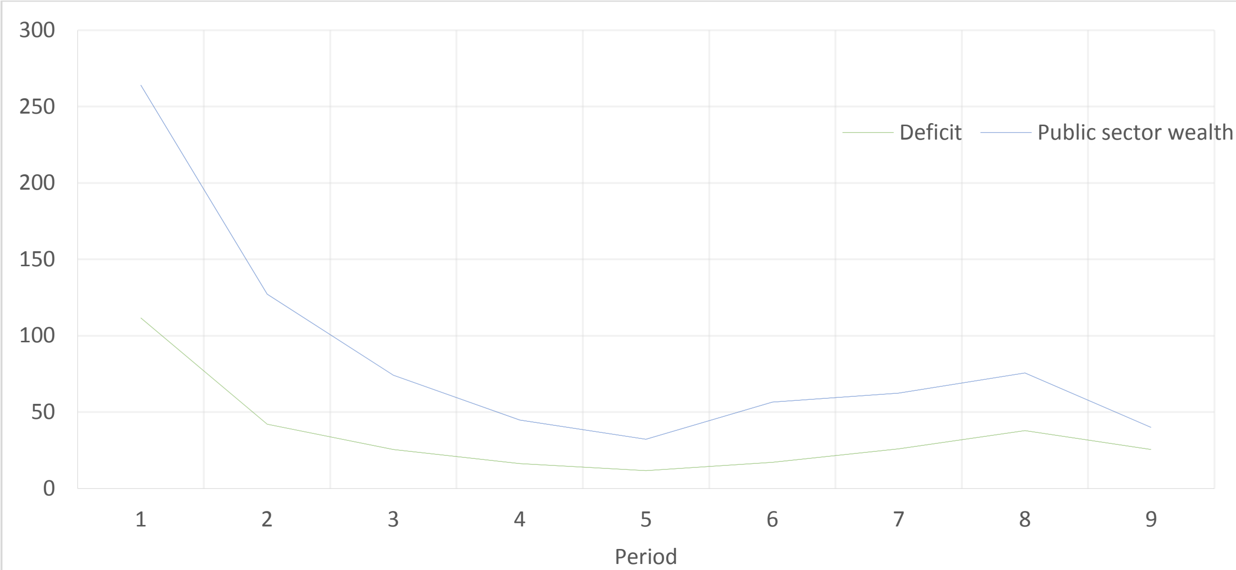
2.6.1.1 Dynamics of public-sector wealth and deficit

Figure 2.1 presents the trends in average public-sector wealth and deficit across groups over the periods. The result shows that public-sector wealth declines consistently over the period, except between periods 5 and 8, when it experiences a slight increase. Given that socially optimal appropriation implies that the public-sector wealth is constant overtime, the result indicates that deficit is mostly incurred. This concurs with the theoretical prediction that deficit is generated within a fragmented government. Equally confirmed from the experimental results in figure 2.2 is the prediction that deficit will decrease over time. Again, the only exception to this observed trend in deficit occurs between periods 5 and 8.

A further investigation of this trend from the data shows that only one of the sixteen groups in the experiment continue the game beyond period 6. Thus, upward trend in public-sector wealth between periods 5 and 8 is accounted for by the optimal appropriation of this group. Out of the remaining fifteen groups, only one reached the termination period due to random stopping, while the rest exhausted their initial public-sector wealth. In fact, the average period it takes the group to exhaust their initial wealth is 5 periods. This is comparable to average periods reported in previous common pool games (see Ostrom, Gardner, & Walker, 1994). Essentially, despite controlling for dynamic externality, the majority of the groups tend to accumulate deficit, which reduces public-sector wealth over time.

Observation 1: *Public-sector wealth declines and deficit is generated within a fragmented government.*

Figure 2.1: Evolution of public-sector wealth and deficit

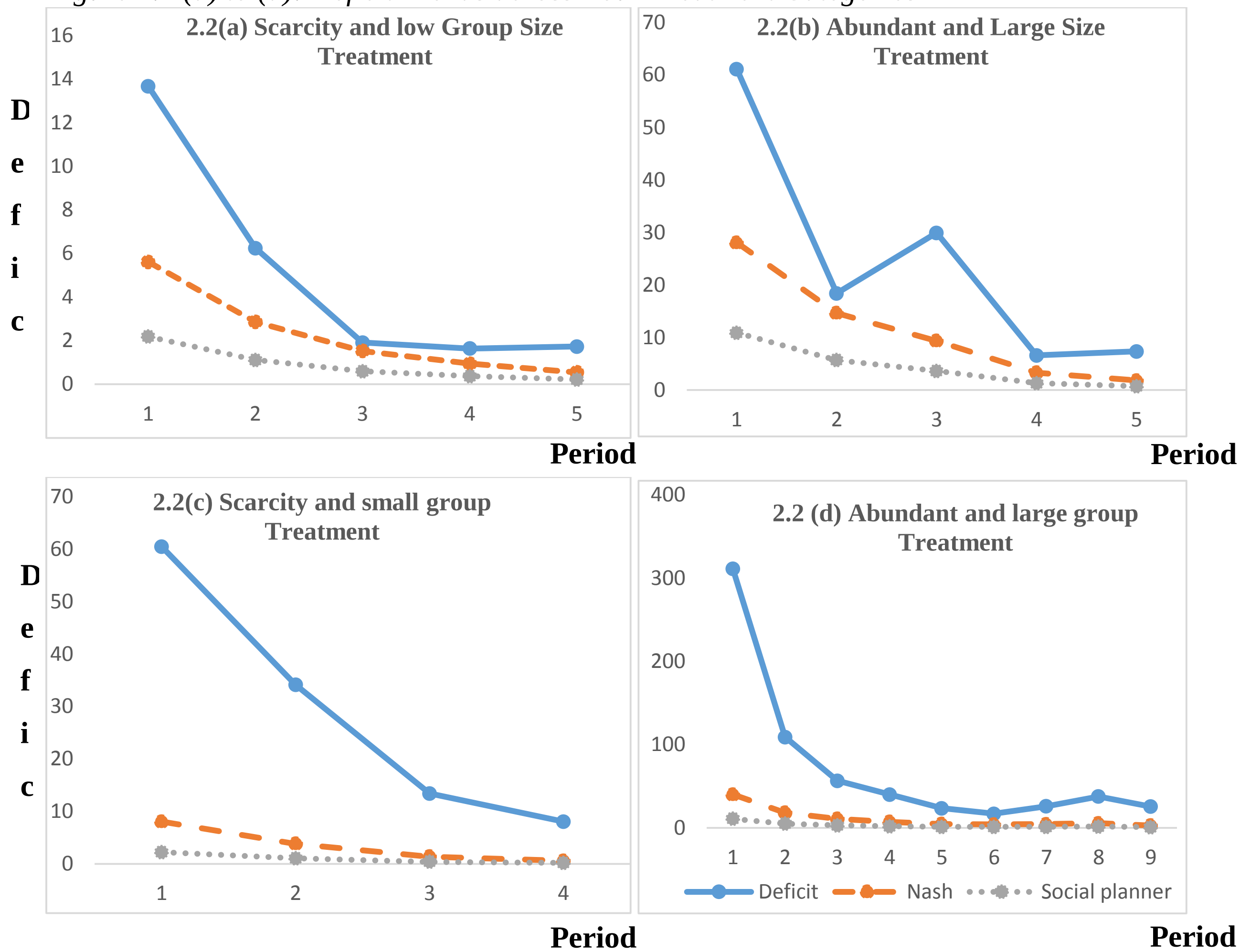


Source: Author's computation

Additional insight into the extent to which the experimental data matches the theoretical prediction is provided in figure 2.2, in which actual deficit is disaggregated across the four treatment categories. Drawn alongside actual deficit are two hypothetical lines that depict the predicted outcomes based on the Nash and social-planner strategies. Across all the treatment categories, the actual deficit lines lie above those of the social-planner. This further reinforces earlier conclusion that deficit is generated across groups or irrespective of the initial wealth level. Furthermore, the predicted deficit based on the Nash strategy—the dashed orange lines in figures 2.2a to 2.2d—closely mirrors those observed in the actual deficit, although, they differ sizably in the first period. Thus, the Nash strategy is a better predictor of the behaviour of deficit in a fragmented government compared to social-planer strategy.

Observation 2: *There is evidence of Nash equilibrium strategy in the behaviour of deficit.*

Figure 2.2 (a) to (d): Deficit Trends across Four Treatment Categories

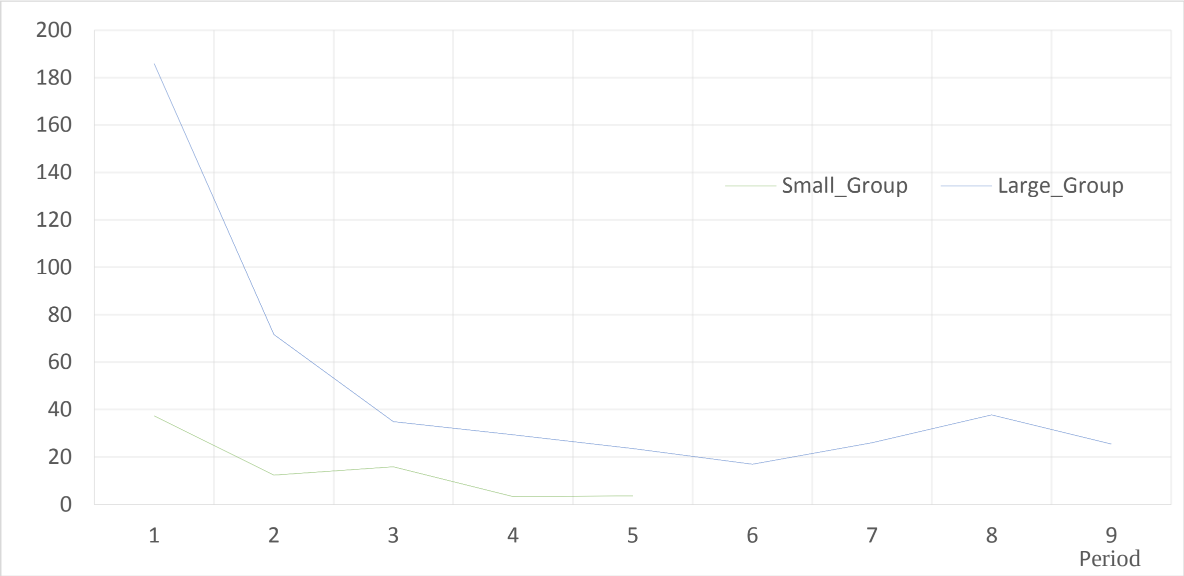


Source: Author's computation from fieldwork

Furthermore, figure 2.3 compares the average level of deficit generated between small and large groups. Clearly, the deficit size among the large groups exceed those incurred by the small groups. Again, the divergence in deficit between the two groups over time is strinking. To establish whether or not the differences are statistically significant, the Kruskal-Wallis test was undertaken. The result gave a chi-squared statistic of $\chi^2_{df=1} = 34.93$that is statistically significant at 0.0001 level. This confirms the proposition stated above that deficits rise with increases in level of fragmentation.

Observation 3: *deficit level rises proportionately with increasing government fragmentation.*

Figure 2.3: Deficit trend across group size



Source: Author's computation

2.6.2 Empirical Model

Evidence based on the non-parametric analysis, while quite revealing, remains limited as it could not help control for possible confounding factors that could bias the observed relationship between deficit and government fragmentation. Thus, following Osés-Eraso et al. (2008), we specified a model that capture the dynamic behaviour of deficit level, conditional on government fragmentation, and other control variables. Formally, the relationship between deficit and government fragmentation is described as follows:

$$D_{it} = \beta_0 D_{i,t-1} + \beta_1 n_i + \sum_{k=2}^{10} \beta_k X'_{it} + c_i + \mu_{it} \qquad (21)$$

where D_{it} is the deficit/surplus incurred by group i in period t; n_i is the group size; $D_{i,t-1}$ is the lag of deficit/surplus and X'_{it} is a vector of control variables; c_i is a group-level fixed effect; μ_{it} is the disturbance term.

2.6.3 Definition and Measurement of Variables

(i) Deficit/surplus: This is measured as the differences between appropriation level chosen by the fragmented government and social planner. Specifically, it is calculated as the differences between

equations (15) and (7). The underlying logic is that since the social planner will choose a balanced budget, any appropriation level that exceeds—or falls below—this optimal level leads to deficit (or surplus).

(ii) Government fragmentation: The two types of government fragmentation (size and procedural) are both captured by n_i. This is because the experimental setting is restricted to the case of $n_i > 1$ and where weak fiscal structure prevails in the budgeting process. Thus, the experiment, so designed in this manner, implies that n_i captures the interaction between the size and procedural fragmentations. Focusing mainly on the interaction effect between the size and procedural fragmentations is in line with Wehner (2009) argument that poor fiscal outcome only result from coincidence of both types of fragmentation. Two measures of n_i are adopted. The first measure is to introduce n_i directly as defined by the group sizes in each treatment. Alternatively, given that group size are in two categories, n_i can also be introduced as a dummy variable. Thus, we define a binary dummy variable—$Dumm\_small$, which takes a value of "1" if group size is three and "0" otherwise. In this respect, the large group is the reference category. This allows us to examine whether or not the results are robust to alternative definition of government fragmentation.

Control variables

(iii) Initial wealth: The theoretical prediction, presented in Table 2.1, shows that the deficit level is proportional to the size of the initial wealth. Therefore, to isolate the effect of initial wealth, we introduced another binary dummy variable—$Dumm\_high$ which takes the value of "1" if the group belongs to the secondary treatment and "0" otherwise.

(iv) Lag of shadow cost: Shadow cost indicates the extent to which the initial wealth is being sustained by a group. In a case where the initial wealth is declining, this could act as an incentive for members to reduce their appropriation level (see Osés-Eraso & Viladrich-Grau, 2007) or encourage a more aggressive appropriation behaviour (see De Janvry et al., 1998). To determine the type of appropriation behaviour elicited in our case, lag of shadow cost is controlled for in the analysis. We use the lag of shadow cost since players cannot observe its current values as decision are made simultaneously. Formally, it is defined as $Log(W_{i,t-1}/W_{i,t-2})$.

(v) Time index: The preliminary result from observation 1 above indicates that deficit declines over time. A period is defined as an interval between when participants make an allocation decision and when the stopping rule is invoked to determine whether or not to proceed to the next period. We account for this effect by including $1/Time$. A positive coefficient of this variable implies that deficit indeed decreases overtime.

(vi) Interaction term: We also examine all possible interactions between the control variables, government fragmentation and lag of public-sector wealth, which could affect deficit level. However, any interaction term that is not significant is dropped from the final regression results reported in section 5.3. For example, we find no effect of gender composition and participants subject area on any of the results; the variables are therefore dropped from the reported regression results.

2.6.4 Estimation Technique

The experimental data generated for this study have an unbalanced panel structure. Specifically, sixteen groups constitute the cross-sectional units, while the time dimension, defined by the total number of rounds in a game, ranges between 4 and 9 periods. The unbalanced panel structure is as a result of the random stopping and the different times at which groups exhaust their initial wealth. Estimating the model given in equation (21) by using conventional ordinary least squares (OLS) yields an inconsistent and bias estimate due to the dynamic nature of the specification and time invariant fixed effects. Alternatively, while fixed and random effects models can help address the time invariant problem, the issue around dynamic panel bias still persists Roodman (2009). Moreover, preliminary diagnostics testing shows that estimation based on fixed and random effects models are inefficient due to problem of autocorrelation and heteroscedasticity[11].

To address these plethora of econometric problems, we apply the system generalized method of moment (system GMM) estimator developed by Arellano and Bover (1995) and Blundell and Bond (1998). System GMM works by transforming the independent variables, including the lag of the dependent variable using the forward orthogonal deviation. The transformed variables are further differenced to generate additional instruments. The transformed variables and their

[11] The estimated F-statistics for Wooldridge panel autocorrelation test is 424.77, while the modified Wald test for group-wise heteroscedasticity is 188.01. Both have a p-value of 0.000, which implied the presence of autocorrelation and heteroscedasticity.

differences then constitute a system of equations, which can be estimated using GMM technique developed by L. P. Hansen (1982). The transformation process directly addresses the problem of dynamic panel bias, time invariant effect and autocorrelation. Also, by applying (Windmeijer, 2005) small-sample correction for the two-step standard error, the heteroscedasticity problem is equally dealt with.

Furthermore, the system GMM possesses several other qualities which make it most suitable for this study. For example, Roodman (2006) noted that system GMM are designed for situation with small time dimension and large cross-section units $(N > T)$. Roodman (2006) further observes that the estimator remains asymptotically efficient and consistent when the independent variables are not strictly exogenous. Also, Serlenga and Shin (2013) observed that the estimator is unbiased and robust even with measurement error in the underlying data generating process. In addition, by using forward orthogonal deviation, system GMM minimizes the gap in unbalanced panel, thereby increasing their efficiency.

2.6.5 Econometric results

Table 2.3 presents the empirical results on the relationship between deficit and government fragmentation. Four separate models are estimated. Model 1 is based on dummy variable measure of government fragmentation, while model 2 uses the direct measure. Models 3 and 4 follow models 1 and 2 respectively, in terms of the measure of government fragmentation, and further include interaction effects between lag of wealth, initial wealth level and the time index. Given that the consistency of system GMM depends on the validity of instruments used, we report, in the last two rows, the Sargan test of over-identifying restriction and the AR (2) test that detects the presence or otherwise of second-order serial correlation. The assumptions underlying Sargan and AR (2) tests were found to hold in all the models, which imply the models are correctly specified.

The result based on model 1 shows that government fragmentation has a statistically significant negative effect on deficit. Specifically, average deficit level is lower by 20.3 points in smaller groups than among the larger groups. A similar conclusion is reached based on model 2, which shows that large group size corresponds to higher deficit level. However, when the interaction term is controlled for, the marginal effect of government fragmentation falls. As the result of model 3 shows, the average deficit is lower by 15.21 points among the smaller group compared to the referenced group. In model 4, while the coefficient of government fragmentation remains positive as predicted by the dynamic common pool model, it is no longer statistically significant and the marginal effect is also considerably lower.

With regards to other explanatory variables, we found no evidence that differences in initial wealth has significant influence on deficit. Shadow cost has a consistently positive effect on deficit. This implies that when shadow cost is high in the previous period, players respond by demanding higher appropriation from the group account. Therefore, declining wealth elicits aggressive appropriation behaviour in the players in our experiment. Lastly, coefficient of time index has the predicted positive value(s) in majority of the model, but only statistically significant in model 3. This suggests that deficit indeed decreases over time, but the downward trend is mostly due to other variables in the model.

Table 2.3: Two-step System GMM Estimation of Deficit and Government Fragmentation

	Model (1)	Model (2)	Model (3)	Model (4)
Lag of deficit	0.278***	0.278***	0.208***	0.130***
	(0.0701)	(0.0701)	(0.0253)	(0.0428)
Dumm_small/Group size	-20.30**	10.15**	-15.21**	4.666
	(8.292)	(4.146)	(6.819)	(4.113)
Dumm_high	1.265	1.265	0.334	1.641
	(6.693)	(6.693)	(6.226)	(4.931)
Lag of shadow cost	7.830***	7.830***	8.097***	3.656***
	(1.612)	(1.612)	(1.256)	(1.030)
1/Time	1.730	1.730	7.872***	-3.015
	(2.222)	(2.222)	(2.708)	(3.887)
Interaction Term				
Lag of wealth*group size			-0.225	0.0577**
			(0.173)	(0.0244)
Lag of wealth *Dumm_high			-0.361*	0.258
			(0.192)	(0.206)
Constant	-28.14***	-78.89***	-39.46***	-31.03
	(6.997)	(21.21)	(9.559)	(26.03)
Wald Statistics	143.9***	143.9***	449.6***	636.7***
Specification Tests				

-Sargan test (p)	7.36	7.36	5.21	5.57
-AR(2) test (p)	-0.41	-0.41	-1.09	-0.81

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

The preceding results are in line with the prediction of the dynamic common pool theory that government fragmentation leads to poor fiscal outcomes. However, whether the impact of government fragmentation works through the strategic effect channel as suggested by the theory is not clearly evident. An approach to test the influence of strategic effect is by examining the influence of government fragmentation on a static variable, such as total group demand. Total group demand is equivalent to public expenditure in field data and should weakly relate to government fragmentation in a dynamic setting. Therefore, we re-estimated models 1 to 4, using total demand as the dependent variable.

The results, presented in Table 2.4, show that coefficients of different measures of government fragmentation has the predicted sign as those in Table 2.3. However, only in model 3 is the effect statistically significant. Furthermore, the size of the marginal effects dropped comparatively. This evidence strongly suggests that effect of government fragmentation in the dynamic setting, is more towards deficit than total demand (expenditure). This is similar to the findings of Perotti and Kontopoulos (2002) that in the static setting, government fragmentation impacts more significantly on public expenditure than deficit. This again establishes that strategic effect is the key channel through which government fragmentation affects fiscal outcomes.

Table 2.4: Two-step System GMM Estimation of Total Demand and Government Fragmentation

	Model (1)	Model (2)	Model (3)	Model (4)
Lag of total demand	0.238***	0.238***	0.208***	0.0865
	(0.0288)	(0.0288)	(0.0137)	(0.0771)
Dumm_small/Group size	-13.64	6.818	-15.91**	3.762
	(8.989)	(4.494)	(6.917)	(2.960)
Dumm_high	-5.169	-5.169	0.477	-0.792
	(6.537)	(6.537)	(5.251)	(2.876)
Lag of shadow cost	7.475***	7.475***	9.561***	4.843***
	(1.419)	(1.419)	(1.789)	(1.292)

1/Time	4.575***	4.575***	10.41***	5.376**
	(1.725)	(1.725)	(3.700)	(2.479)
Interaction term				
Lag of wealth*group size			-0.0352	0.0556**
			(0.178)	(0.0281)
Lag of wealth *Dumm_high			-0.567*	-0.230
			(0.293)	(0.187)
Constant	-30.00***	-64.09**	-50.49***	-45.29**
	(8.659)	(25.36)	(14.64)	(18.54)
Wald Statistics	138.8***	138.8***	857.2***	458.9***
Specification Tests				
-Sargan test (p)	7.85	7.85	7.6.88	7.46
-AR(2) test (p)	-0.31	-0.30	-1.24	-1.10

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

2.6.6 Sensitivity Analysis

We performed various sensitivity analyses to test the robustness and validity of our results. These include: sensitivity to alternative specification choices, estimation methods and noisy decision-making.

(i) *Sensitivity to alternative specification choices:* System GMM estimates have been observed to be very sensitive to variation in instruments (Windmeijer, 2005). Therefore, Roodman (2009) noted that testing the robustness of the estimates to changes in instrument is crucial. In the main result presented in Table 2.3, the maximum lag length used as instrument for the dependent variable is restricted at one, while those of independent variable is set at two. This gives an instrument count of 15. We tested for alternative specification of lag limit and in all cases found out that it has limited impact on the reported marginal effect. The standard errors were, however, lower, increasing the number of significant variables. Further, all the alternative specifications suffer from instrument proliferation problem, implying model overfit and inefficiency of the Sargan test. It also implies a violation of the rule of thumb technique proposed by Roodman (2006)—when the instrument count exceeds the number of cross-section units, the instrument proliferation problem is acute. We therefore conclude that our instrument selection is efficient compared to those of alternative specification.

(ii) *Sensitivity to alternative estimation methods*: Bond (2002) noted that OLS and fixed effects models provide a useful check on the validity of system GMM. This is because in a first-order autoregressive model such as equation (21), OLS will bias the estimated coefficient of the lag of deficit upward, while the fixed effect model will have a downward bias. Thus, the true estimate of the coefficient of the lag of deficit will be between OLS and fixed effect estimates and could, in essence, serve as bounds to test the validity of system GMM results (Bond, 2002). In tables 2.5 and 2.6 we present results of the re-estimated models 1 to 4 using fixed effects and OLS respectively. The results show that our initial model is correctly specified and theoretically superior, as the estimated coefficients are mostly within the appropriate bounds.

Table 2.5: Fixed Effect Model Estimation of Deficit and Government Fragmentation

	Model (1)	Model (2)	Model (3)	Model (4)
Lag of deficit	0.248***	0.248***	0.193***	0.186***
	(0.0309)	(0.0309)	(0.0418)	(0.0324)
Lag of shadow cost	7.562***	7.562***	8.129***	2.399
	(1.117)	(1.117)	(1.139)	(1.815)
1/Time	3.519*	3.519*	9.120***	-0.834
	(1.757)	(1.757)	(3.373)	(2.795)
Dumm_high	-7.083*	-7.083*	-4.982	-2.900
	(4.122)	(4.122)	(6.401)	(5.238)
Dumm_small/Group size	-18.07***	9.034***	-18.98***	4.148
	(4.077)	(2.039)	(5.170)	(2.815)
<u>Interaction effect</u>				
Lag of wealth*group size			-0.0980	0.0350***
			(0.0721)	(0.00944)
Lag of wealth *Dumm_high			-0.489*	0.178
			(0.263)	(0.262)
Constant	-26.01***	-71.18***	-37.82***	-22.18
	(7.673)	(13.10)	(9.610)	(23.01)
R^2	0.831	0.831	0.843	0.871
adj. R^2	0.815	0.815	0.822	0.854

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

Table 2.6: OLS Estimation of Deficit and Government Fragmentation

	Model (1)	Model (2)	Model (3)	Model (4)
Lag of deficit	0.286***	0.286***	0.167***	0.179***
	(0.0325)	(0.0325)	(0.0457)	(0.0343)
Lag of shadow cost	7.596***	7.596***	8.022***	1.964
	(1.173)	(1.173)	(1.208)	(1.917)
1/Time				
Dumm_high	-8.112*	-8.112*	-6.651	-5.496
	(4.365)	(4.365)	(6.955)	(5.792)
Dumm_small/Group size	-19.01***	9.504***	-18.78***	3.572
	(4.437)	(2.218)	(5.657)	(3.015)
Interaction effect				
Lag of wealth*group size			-0.0948	0.0367***
			(0.0780)	(0.00987)
Lag of wealth *Dumm_high			-0.415	0.303
			(0.300)	(0.286)
Constant	-14.12*	-61.64***	-8.672	-20.94
	(7.629)	(12.03)	(8.521)	(17.20)
R^2	0.838	0.838	0.846	0.877
adj. R^2	0.802	0.802	0.804	0.844

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

(iii) *Noisy decision making*: Economic experiment basically mimics a given market or institutional scenario in controlled setting based on a set of instructions. It has been observed in many studies that confusion regarding the game or ambiguity in instruction could create noisy decision making (Andersson, Holm, Tyran, & Wengström, 2015). In order to assess the extent of noisy decision making, most economic experiments administer questionnaire to debrief the participants' decisions (see López-Pérez & Spiegelman, 2013). Following this approach, we administered a questionnaire after each experimental session (see appendix 2B). The results, shown in Table 2.7, indicate that 40 of the participants fully understood the instruction of the pilot-test session, 19 stated they mostly understood the instruction and 5 indicated they did not understand the instruction. For the main experiment, the number of participants that indicated they fully understood the instruction remained the same, while those that did not understand fell to 4. This shows that the possibility for noisy decision making is low in our experiment. Furthermore, we found no effect on the result of excluding the groups where at least a participant indicates the instruction is not fully understood.

Table 2.7: An Assessment of Noise in Decision Making

Scale	**Practice Session Instruction**	**Main Experimental Instruction (in percent)**
Fully Understand	40 (63%)	40 (63%)
Mostly Understand	19 (30%)	20 (31%)
Didn't Understand	5 (8%)	4 (6%)
Total	**64 (100%)**	**64 (100%)**

Source: Author's computation

2.7 Conclusion

High and persistent deficit and debt experienced in many countries, in recent years, has no doubt become a major concern for policymakers and researchers. These poor fiscal outcomes have been widely attributed to the presence of static or/and dynamic common pool problems in fiscal policy making (Perotti & Kontopoulos, 2002). However, existing literature has mostly demonstrated the implication of static common pool problem for fiscal performance. In this paper, we explore both the theoretical and empirical implications of dynamic common pool problem, as measured by the level of government fragmentation, on deficit. On the theoretical side, we built on and extended the dynamic common pool model of Velasco (1998) and derived key empirically testable predictions therefrom. The empirical approach is based on experimental data, given limitations of field data in addressing the derived predictions. The experiment consists of two treatments: a primary treatment—with low initial wealth—and a secondary treatment—with high initial wealth. We also varied the level of fragmentation between three-member and five-member groups in each treatment. The generated experimental data were thereafter analysed using system GMM technique; this is further complemented with non-parametric analysis.

The empirical result confirms the theoretical prediction that dynamic common pool problem exacerbates the level of deficit bias in fiscal policy making. For example, when government fragmentation is measured using the binary dummy variable, we found that deficit level is lower by 15.21 points among the smaller groups, despite controlling for the possible confounding factors. Further sensitivity analyses, which entail comparisons between alternative instrument selection procedures and inferential analysis based on fixed-effect and OLS estimators, reveal that the system GMM technique adopted is correctly specified and theoretically superior to alternative techniques. Essentially, while deficit is incurred across the treatments, the more fragmented settings—large groups—are found to accumulate higher deficit level than smaller groups.

Additionally, we established that strategic effect is an important channel through which government fragmentation influences deficit. This means that deficit is basically incurred due to uncertainty about the future of fiscal path, resulting from collective ownership of the public-sector wealth. Overall, the results support the observation by Perotti and Kontopoulos (2002) that in the dynamic setting, the effect of government fragmentation manifests through higher deficit level, while in the static setting, the effect is reflected in higher public expenditure. Our findings also corroborate the conclusion of Sutter (2003) that strategic effect is one of the key sources of deficit bias. Furthermore, we found that players respond to declining public-sector wealth by assuming more aggregative demand positions in the subsequent periods. This however is in sharp contrast with the findings of Osés-Eraso et al. (2008) that players restrain their appropriation strategies when faced with resources scarcity.

These findings have important policy implications for budgeting institutions in the real world. For example, our finding suggests that budgeting institutions that are characterized by multiple decision makers and decentralized fiscal structure would be highly susceptible to dynamic common pool problem, which might negatively affect overall fiscal performance. Therefore, policies that limit the extent of both size and procedural fragmentations will be crucial. Size fragmentation, defined as the presence of multiple decision makers in the budgeting process, are difficult to change, as it is ingrained in the constitution (see Kontopoulos & Perotti, 1999). Procedural fragmentation, defined as the existence of weak structure governing the budget decision making processes, on the other hand is more amendable to reform in the short-run; targeting policy intervention towards improving procedural fragmentation will therefore be more productive. This could be achieved through the centralization of budgeting decision—increasing the strategic dominance of one of the budget actors, the finance minister for instance. Another policy option towards reducing the level of procedural fragmentation as proposed by Von Hagen (2002) is through limiting the amendment power of the parliament/legislative arm to the executive budget proposal.

Furthermore, since static and dynamic common pool problem work through different channels, the set of policies that address one aspect of the problem may not necessarily affect the other. Thus, understanding the nature/type of common pool problem that prevails within a given budgeting institution is crucial to designing the optimal policy response. In the case of strategic effect channel, a number of policy options have been suggested in the literature such as: improving transparency and accountability of the budget processes (Alt & Lassen, 2006); establishment of an apolitical fiscal council as is the case with central bank's autonomy in the monetary policy context (Wyplosz, 2012); use of long-term fiscal constraints (Von Hagen, 2002), among others. In essence, the appropriate policy mix for a country will depend crucially on its economic and institutional characteristics.

2A Appendix: Experimental Instructions

Instruction 1: For small group and low wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant will attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form, more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.58 payoff = 1 South African Rand

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of three (3) members and your group account has been allocated an initial endowment of 60 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- ❖ **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account (Pr).

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 6 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any non-negative real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- ❖ **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period: Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How will the random stopping rule be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

What is the new endowment for the next period? The remaining endowment from the last period attract an interest rate of 0.1 (10 percent) for the next period. For example, if there are 15 points at the end of a period, in the next period, the endowments to share become 16.5 points (15+15*0.1)

What happens if the endowment in group's account is exhausted? The game ends once the group exhaust its endowment as resources are no longer available to share during subsequent periods.

What happens to the remaining endowments if the game ends through the random stopping rule? The remaining endowments are simply discarded. This means they are not added to your payoff.

What is the endowment dynamics? It is a ratio of the endowment in the last period to endowment in the current period and indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one (>1), then the initial endowment is increasing over time if it is less than one (<1), then initial endowment is decreasing.

What happens when the game ends? Your payoff in each period is summed up using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 2: For small group and high wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant will attend to you accordingly. You will now participate in a decisionmaking experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.06 payoff =1 South African Rand

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of three (3) members and your group account has been allocated an initial endowment of 300 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 6 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any non-negative real number is allowed (e.g. 5, 0.5, 1.5 etc.)

❖ **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period: Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How will the random stopping rule be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

What is the new endowment for the next period? The remaining endowment from the last period attract an interest rate of 0.1 (10 percent) for the next period. E.g. if there are 15 points at the end of a period, in the next period, the endowments to share become 16.5 points (15+15*0.1)

What happens if the endowment in group's account is exhausted? The game ends once the group exhaust its endowment as resources are no longer available to share during subsequent periods.

What happens to the remaining endowments if the game ends through the random stopping rule? The remaining endowments are simply discarded. This means they are not added to your payoff.

What is the endowment dynamics? It is a ratio of the endowment in the last period to endowment in the current period and indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one (>1), then the initial endowment is increasing over time if it is less than one (<1), then initial endowment is decreasing.

What happens when the game ends? Your payoff in each period is summed up using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 3: For large group and low wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant will attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.35 payoff = 1 South African Rand

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of five (5) members and your group account has been allocated an initial endowment of 100 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- ❖ **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 10 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any non-negative real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- ❖ **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

<u>Completion of stage 2 marks the end of a period</u>

Whether to proceed or not to the next period: Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How will the random stopping rule be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

What is the new endowment for the next period? The remaining endowment from the last period attract an interest rate of 0.1 (10 percent) for the next period. E.g. if there are 15 points at the end of a period, in the next period, the endowments to share become 16.5 points (15+15*0.1)

What happens if the endowment in group's account is exhausted? The game ends once the group exhaust its endowment as resources are no longer available to share during subsequent periods.

What happens to the remaining endowments if the game ends through the random stopping rule? The remaining endowments are simply discarded. This means they are not added to your payoff.

What is the endowment dynamics? It is a ratio of the endowment in the last period to endowment in the current period and indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one (>1), then the initial endowment is increasing over time if it is less than one (<1), then initial endowment is decreasing.

What happens when the game ends? Your payoff in each period is summed up using the exchange rate given above. You will receive your earnings privately.

<u>Any question before we take some practice session?</u>

Instruction 4: For large group and high wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant to attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.03 payoff = 1 South African Rand

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of five (5) members and your group account has been allocated an initial endowment of 500 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 10 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any non-negative real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period? Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How the random stopping rule will be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

What is the new endowment for the next period? The remaining endowment from the last period attract an interest rate of 0.1 (10 percent) for the next period. E.g. if there are 15 points at the end of a period, in the next period, the endowments to share become 16.5 points (15+15*0.1)

What happens if the endowment in group's account is exhausted? The game ends once the group exhaust its endowment as resources are no longer available to share during subsequent periods.

What happens to the remaining endowments if the game ends through the random stopping rule? The remaining endowments are simply discarded. This means they are not added to your payoff.

What is the endowment dynamics? It is a ratio of the endowment in the last period to endowment in the current period and indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one (>1), then the initial endowment is increasing over time if it is less than one (<1), then initial endowment is decreasing.

What happens when the game ends? Your payoff in each period is summed up using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

2B Appendix: QUESTIONNAIRE

(to be filled after the experiment)

1. Terminal number (Number written on your screen)

   ____

2. Gender:

   ☐ Male ☐ Female

3. Year of study:

   ☐ 1st Year ☐ 2nd Year

4. Department:

   ____

5. Instructions for Practice Session:

   ☐ I fully understood the instructions.

   ☐ I understood most of the instructions.

   ☐ I didn't understand most of the instructions.

6. Keys for Practice Session:

   ☐ I was fully at ease with the keystrokes.

   ☐ I was at ease with most of the keystrokes.

   ☐ I was not at ease with the keystrokes.

7. Instructions for the Experiment:

   ☐ I fully understood the instructions.

   ☐ I understood most of the instructions.

   ☐ I didn't understand much of the instructions.

8. Keys for the Experiment:

   ☐ I was fully at ease with the keystrokes.

   ☐ I was at ease with most of the keystrokes.

   ☐ I was not at ease with the keystrokes.

9. Imagine you are the only person in your group, will your decision be different?
   - ☐ Yes.
   - ☐ No.
10. If your answer in Question 9 is a yes, briefly discuss what you would have done differently?

    ____

    ____

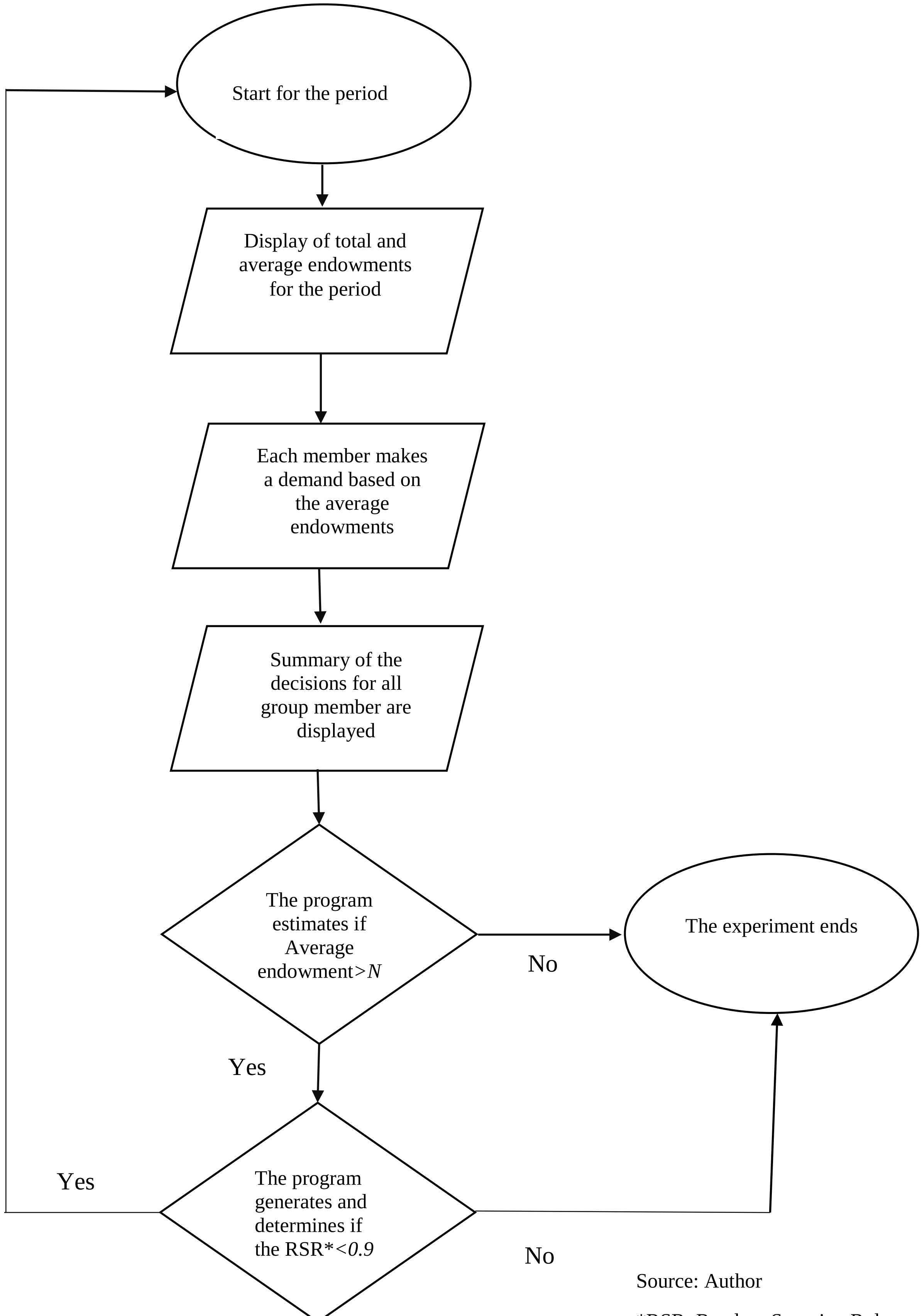
    ____

2C Appendix: Z-Tree

The Z-TREE provides a useful means of mimicking the budget institution in line with the decision setting using a network of computer systems. The z-tree comprises of two programs: z-tree and z-leaf. The z-tree program is a server program which the experimenter operates from. It is essentially made up of lines of commands that map out the desired decision setting and institution. The z-leaf program, on the other hand, translates the lines of command in z-tree into sets of interface for participants to interact and also make their budgeting decision.

Figures 0.1C illustrates the step by step process through which the z-tree program is translated into the z-leaf program, as implemented for the experiment. At the start of any period, the participants in a group are shown the total and average endowments to be appropriated upon over the entire duration of the experiment. The groups are then taken to the next stage where each member makes a decision regarding the amount of endowments to transfer from the group account to their individual account. Their demand cannot exceed the average endowments for the period, in line with the spending rule. After making their decision, groups are taken to the next stage where the summary of decisions for the period is displayed for all members. The information summarized include the individual and group demands, endowments dynamics and the remaining endowments. Subsequently, the program determined if the remaining endowment exceeds N (group size). The program proceeds to generate the random stopping rule and determines if the experiment should continue to the next period, only if the remaining endowment has not been exhausted. The loop is repeated until either the endowment is exhausted or the random stopping rule terminates the experiment.

Figure 0.1C: Diagrammatic Illustration of the Z-TREE program



Source: Author

*RSR: Random Stopping Rule

3. Chapter Three: The Delayed Stabilization Hypothesis: An Exploratory Experimental Evidence

3.1 Introduction

One of the key predictions of the dynamic common pool model is the possibility of fiscal stabilization; a change in fiscal policy, either through tax increase and/or expenditure cut, aimed at correcting deficit bias (Velasco, 1997, 1998). Essentially, fiscal stabilization implies that there is a threshold at which government fragmentation no longer leads to poor fiscal outcome. Velasco (1997) also notes that such a stabilization threshold is determined in large part by the dynamics of public-sector wealth (or public debt)[12]. Specifically, when public-sector wealth is high, aggressive appropriation of resources occur leading to debt accumulation. On the other hand, as public-sector wealth declines, the efficiency gain accruing from fiscal stabilization becomes attractive to budget actors. This phenomenon is described as the "*delayed stabilization hypothesis*" (Alesina et al., 2006; Velasco, 1997).

For several reasons, testing the delayed stabilization hypothesis using field data has been challenging. For example, assessing the effect of public-sector wealth on the stabilization process requires isolating only that aspect of deficit bias that can be ascribed to common pool problem. However, with the conceptual issues associated with gauging common pool problem (see Perotti & Kontopoulos, 2002), a precise estimate is difficult to obtain. Moreover, since dynamic common pool problem emanated from uncertainty regarding the paths of public-sector wealth, evaluating the magnitude of dynamic externalities along different wealth paths becomes crucial. Again, the magnitude of dynamic externalities cannot be easily observed using field data. This explains why existing empirical evidence are mostly limited to indirect inferences from related studies[13].

The objective of this study is to attempt an exploratory empirical verification of the delayed stabilization hypothesis. To this end, the study uses experimental data. The experimental setting follows the dynamic common pool game— an extension of the version examined in chapter two. The players are initially endowed with common resources designated as public-sector wealth,

[12] Conceptually, public-sector wealth and debt are analogous. For a given government budget constraint, debt increase implies fall in public-sector wealth, and vice versa. However, in this study we focus more on the dynamics of public-sector wealth as it is a lot easier to empirically handle using experimental data.

[13](see Eslava, 2011) for a detailed survey of the existing literature on delayed stabilization hypothesis. As the author noted, these evidences are indirect, therefore difficult to interpret either in support of or against the hypothesis.

which are to be shared over an infinite horizon. In line with the literature, infinite horizon is generated with the use of random stopping rule after a pre-specified period (see Croson, 2005). Overall, the sustainability of the resources hinges on players' appropriation behaviour over the duration of the game.

To examine the response of players to public-sector wealth decline, we choose design parameters that increase the likelihood of resources reaching an unsustainable level. In addition, the terminal period of the game is determined only by the random stopping rule[14]. Thus, in the case resources become unsustainable; a lump-sum tax is imposed on the players and transferred back into the common resources pool. If and when the game gets to this stage, the prediction of the model is that fiscal stabilization takes place as players become more constrained in their appropriation behaviour. However, an identification problem could arise as to whether the stabilization is temporal (due to pressure from taxation) or permanent (as expected if stabilization truly takes place). To separate these effects, initial wealth is restored after some rounds of play. In essence, for delayed stabilization hypothesis to be validated, we should observe an appropriation level that remains consistent with fiscal stabilization as players ought to have developed the reputation for low spending through the history of the game.

The contributions of this paper are two-fold. First and from the empirical side, the study provides an evidence on the delayed stabilization hypothesis. This evidence has an important policy implication; it offers insight into the possible timing and strategies for fiscal stabilization. Secondly, the study builds on and extends earlier experimental studies on cooperation in the commons. For example, experiments by Mason and Phillips (1997), Osés-Eraso and Viladrich-Grau (2007) and Osés-Eraso et al. (2008) have extensively investigated the effect of resource stock decline on cooperation (or stabilization) in the natural resource setting. However, the present study focuses on the fiscal setting, which is markedly different from the natural resource setting.

The reminder of this paper is organized as follows. Section two reviews literature on the effect of public-wealth dynamics on fiscal stabilization. In this discussion, debt dynamics rather than

[14] Random stopping rule is a strategy used to generate condition that mimic infinite horizon in an experimental setting. The rule uses the discount factor as the continuation probability of the experiment. Thus, when a number less than the discount factor is generated the experiment is terminated, otherwise it continues to subsequent period.

public-wealth is the focus, as both are conceptually related. Section three describes the delayed stabilization model, which motivates the design of our experiment. Section four details the experimental procedure and its implementation. Section five presents results from the non-parametric and econometric analyses. Finally, policy implications of the results are discussed in section six.

3.2 Literature Review

There is a paucity of studies that have explicitly tested the delayed stabilization hypothesis. However, two strands of existing literature on fiscal sustainability have extensively addressed related objective, from which useful inferences can be drawn. The first strand consists of studies on government's fiscal reaction to debt level, and the second relates to studies on determinants of fiscal adjustment. This section reviews the key findings from these two strands of literature in relation to delayed stabilization hypothesis.

3.1.1 Fiscal Reaction to Rising Debt

Bohn (1998) argues that government takes corrective action towards ensuring fiscal sustainability if the coefficient of debt variable is positive and significant in a fiscal reaction function (FRF). Essentially, FRF captures the marginal effect of changes in stock of government debt on primary (non-interest) surplus, conditional on cyclical components of government spending and output. Bohn (1998) tests this model with historical USA data and finds that at low debt levels government's fiscal reaction is weak, but becomes more pronounced at higher levels of debt. This evidence supports the delayed stabilization hypothesis that high debt level (or equivalently declining public-sector wealth) could influence government to initiate stabilization process.

However, efforts at testing the model outside of the USA have produced mixed results. In a cross-sectional study of seven developed countries, Lukkezen and Rojas-Romagosa (2013) find a positive response of primary surplus to high debt level only in the Netherland. The estimates for Spain and Portugal on the other hand yield negative coefficients for higher order debt-GDP ratio. This may be an indication that limited corrective actions are being undertaken. Mendoza and Ostry (2008) and Afonso and Hauptmeier (2009) also estimate the FRF based on a panel of developed countries. They find the reaction of primary surplus to be positive and significant only at low debt threshold.

A much weaker support for the hypothesis is reported for developing countries. For example, Abiad and Baig (2005) and Mendoza and Ostry (2008) find a positive and significant fiscal reaction at low debt levels, but becomes unresponsive as the debt-GDP ratio exceeds the threshold of 50 percent. Burger, Stuart, Jooste, and Cuevas (2011), using a vector error correction model (VECM), observe in the case of South Africa that every 1 percent increase in debt/GDP ratio is followed by 0.5 percent and 0.31 percent rise in primary balance/GDP over the short-run and long-run respectively. This implies that government response is stronger in the short-run than the long-run.

Several explanations have been provided in the literature for the weak fiscal reaction at high debt levels, especially in developing countries. These include: debt overhang (Daniel, Callen, Terrones, Debrun, & Allard, 2003); negative effect of debt on growth (Kaur & Mukherjee, 2012); volatile economic and financial environments of most developing economies, which is transmitted into high risk premium and interest rate on borrowing (Dell'Erba, Mattina, & Roitman, 2015), among others. Taking these factors into account imply that developing countries face lower threshold in responding to rising debt, compared to developed countries. Thus, it may be inferred that these evidences partly support the delayed stabilization hypothesis in developing countries; that government react to rising debt, but financial market and other economic forces could offset the effort at high debt thresholds.

However, it needs be highlighted that the evidence is less credible if the source of the deficit bias is not due to the fragmentation in fiscal policy making. Therefore, there is dire need to interrogate the main causes of the debt trajectory such that it could be juxtaposed with what received literature says with regards to the delayed stabilization hypothesis. Comparing Lukkezen and Rojas-Romagosa (2013) findings with previous studies on fragmentation, a much weaker support is found for the hypothesis. For example, Lukkezen and Rojas-Romagosa (2013) observed a negative and insignificant reaction of primary surplus to high debt levels in Spain and Portugal— two countries which according to De Haan et al. (1999) have highly fragmented fiscal institutions. On the other hand, the United Kingdom (UK) and Netherlands, with centralized fiscal institution recorded positive and significant fiscal reactions, irrespective of debt levels. The only exception is Belgium that had the weakest budgeting procedure in De Haan et al. (1999) sample, but with a positive and significant coefficient for higher order debt-GDP ratio in the FRF.

The main limitation of the FRF approach in understanding fiscal stabilization and sustainability is that it is backward looking; it provides rather limited information on future fiscal path. For example, Richter and Paparas (2013) examined the reliability of FRF as a predictor of fiscal sustainability using Greece as a testbed. From Greek's current situation, it is notable that her debt level is unsustainable. The question then arises as to whether *ex-ante* we can arrive at a similar conclusion from the FRF test. Based on time series data spanning 176 years before the onset of the Greek debt crisis, Richter and Paparas (2013) found that the FRF predicts a sustainable fiscal path. In essence, while the FRF indicates that government responds to rising debt, it does not reveal whether the depth of the changes is enough to ensure fiscal stabilization.

3.1.2 Determinants of Fiscal Adjustment

Another perspective from which the relation between public-sector wealth dynamics and fiscal stabilization can be assessed is through the lens of fiscal adjustment process. Fiscal adjustment is a deliberate effort by government to restore fiscal policy towards a sustainable path. While there is no consensus in the literature on how to measure fiscal adjustment, most studies follow the approach of Alesina and Perotti (1995a), which consider a decline of at least 0.5percent in the cyclically adjusted primary balance as evidence that a fiscal adjustment process is being undertaken. If sustained over adequate period of time, the fiscal adjustment process is regarded as successful (Von Hagen & Strauch, 2001). Expectedly, the broad conclusion in the literature is that initial debt level is a crucial factor in determining the need for governmental intervention in fiscal adjustment process (Dell'Erba et al., 2015; Gupta, Clements, Baldacci, & Mulas-Granados, 2004; Mierau, Jong-A-Pin, & De Haan, 2007).

However consensus of opinion seems to diverge on the impact of debt level on successful fiscal adjustment. Von Hagen and Strauch (2001) and Ardagna (2004) find that high debt-GDP ratio increases the likelihood of fiscal adjustment being successful among OECD countries. Similarly, Gupta et al. (2004) note that in the developing countries, deteriorating fiscal conditions contribute significantly to the sustainability of the adjustment process. On the contrary, Purfield (2003) finds that differences in the initial fiscal position of transition economies do not affect the probability of successful adjustment process.

A plausible explanation for these observed differences in the debt effect on successful fiscal adjustment is the prevailing state of the economy. For instance, high debt level could push the economy into recession or coincide with major economic crisis. As a consequence, economic crisis could set in "adjustment fatigue" that tends to reduce the incentive to pursue tight fiscal policy. For example, Baldacci, Gupta, and Mulas-Granados (2010) note that successful adjustment episode takes an average of 10 years and is less likely in countries facing prolong banking crisis. On the other hand, economic crisis could be beneficial by exerting pressure on government to undertake fiscal reform in earnest rather than delay further (Alesina & Drazen, 1989; Lora & Olivera, 2004).

The likelihood of successful adjustment has equally been found to be lower for countries with fragmented electoral system, such as coalition governments (Alesina & Perotti, 1995a; Alesina, Perotti, Tavares, Obstfeld, & Eichengreen, 1998; Illera & Mulas-Granados, 2002). In particular, Alesina and Perotti (1995a) report that only 8.7 percent of adjustment attempts by coalition governments in OECD countries are successful, compared to 64.3 percent success rate recorded by single governments. In addition, the study finds that adjustment targeting politically sensitive expenditure items such as transfer programs and wages are more successful than those focusing on revenue increases or cut in public investment. Mulas-Granados (2003) further explores the determinants of adjustment strategies in the EU countries between 1970 and 2000. Remarkably, coalition governments are observed to be more inclined towards revenue-based adjustment. In other words, high failure rate of adjustment in coalition regimes is partly due to limited use of expenditure-based adjustment.

A major problem with these findings is that they contradict the conventional theoretical paradigm that government fragmentation promotes distributive policy (Alesina & Perotti, 1995b; Weingast et al., 1981), that could skew expenditure towards public investment in district projects, rather than transfers and wages whose benefits are not confined to specific interest group. Overall, the key insight from this literature indicates that high debt levels motivate government to undertake fiscal adjustment, but the effect on successful stabilization is inconclusive.

3.1.3 Summary of the review

The review provides inconclusive evidence on delayed stabilization hypothesis. In line with the hypothesis, countries characterized with fragmented fiscal institutions are observed to accumulate debt (or public-sector wealth declines) and government's reaction is weak at low debt level. In contrast, evidence that stabilization will take place in the long run leading to correction of the deficit bias is not widely reported. However, given that earlier studies are not explicitly set out to test the hypothesis, there are limitations to the inferences that can be drawn from their findings. For example, the causes of deficit bias are not dealt with in these studies, whereas the dynamic common pool model links the causes and subsequent stabilization in a chain process.

Moreover, the inconclusive results could be partly due to the limitations of field data and its weak comparability across countries (see Padovano & Venturi, 2001), coupled with subjective measurement of government fragmentation (Perotti & Kontopoulos, 2002). Given this obvious gap in the literature, this study uses a laboratory experiment to examine the process through which fragmentation could generate deficit bias and responsiveness of budget actors as debt build up.

3.2 Theoretical Model

3.2.1 The Delayed Stabilization Model

The idea that declines in public-sector wealth could increase the propensity for fiscal stabilization is widely attributed in the literature to Alesina and Drazen (1989). In their "war of attrition" model, they demonstrated that when two groups are divided in opinions over how to distribute the burden of reform—the level of taxation needed to restore sustainability for instance—there arises a delay in stabilization until one of the groups attains a point where the marginal cost of delaying reform exceeds the marginal benefit of waiting for the other group to bear more of the cost of stabilization. Since delay involves debt accumulation—further declines in public-sector wealth—stabilization could therefore be attributed to public-sector wealth dynamics. Similar result was arrived at by Velasco (1997) who used the common pool model and Hsieh (2000) who on the other hand used the bargaining model. However, while the source of deficit bias is simply assumed in other models, Velasco (1997), in his analysis, explicitly demonstrates how common pool problem generates deficit bias as well as creates the inertia for stabilization. Given this theoretical advantage, the conceptualization that follows is based more heavily on Velasco's approach.

Assume a budget institution made up of $n>1$ budget actors. The budget actors have a utility function given by:

$$U_i = \sum_{s=t}^{\infty} Log(g_{is})(1+r)^{-(s-t)} \qquad (1)$$

where g_i is the level of spending/appropriation chosen by say legislator i; r is the constant real interest; and t denotes timing of the event. Also, the budgeting decision process is subjected to a spending rule given by

$$\sum_{i=1}^{n} g_{it} \leq (1+r)W_t \quad \forall t \tag{2}$$

and the conventional solvency condition:

$$\lim_{t\to\infty} W_t (1+r)^{-t} \geq 0 \tag{3}$$

where W is the public-sector wealth and n is number of legislators within the budget institution. Equations 1 to 3 describe the standard optimization problem faced by a social planner. To enable analysis of issues relating to delayed stabilization, some key features are introduced. First, we extend the budget constraint to include the cost associated with deviation from the stabilization path. As Katayama (2008) noted, in the pre-stabilization period, the government relies more on distortionary financing options which comes with significant welfare loss. Taken this cost into consideration, the government inter-temporal budget constraint is specified as:

$$W_t = (1+r)W_{t-1} - z_t - \sum_{i=1}^{n} g_{it} \tag{4}$$

where z_t denotes the deadweight loss per period of time, and captures the cost incurred in the absence of stabilization. The behaviour of z_t is described by:

$$z_t = \begin{cases} 0 & \text{if} \quad g_{it}^{*} = \dfrac{rW_{t-1} - z_t}{n} \\ z & \text{if otherwise} \end{cases} \tag{5}$$

where g_{it}^{*} is the social planner desired spending, derived by maximizing equation (1) subject to equation (4). This means that deadweight loss— z_t, is only incurred if the social planner desired spending level is not realized. In essence, this relation can be used to formally define fiscal stabilization as spending level that is equal to g_{it}^{*}. However, in case the appropriation behaviours follow the non-cooperative strategy, the spending level is derived as:

$$g_{it}^{nc} = \frac{(1+r)(rW_{t-1} - z)}{1+nr} \tag{6}$$

where g_{it}^{nc} is the spending level chosen by budget actor in a fragmented system and is greater than g_{it}^{*}. Another extension to the standard social planner problem is the introduction of tax into the economy. Tax comes into the model only when the economy exceeds its debt ceiling; that is when the initial public-sector wealth is exhausted. Once the debt ceiling is reached, the interest groups are locked perpetually into paying tax, as no borrower will lend further to government. Thus, unlike in the case of natural resource exploitation, the common pool resource is reversible in the fiscal setting. The marginal rate of transformation of private assets to public goods is given as one.

Based on the derived spending levels, the utilities obtained along the paths when groups act according to social planner strategy (U_i^{*}) and when they adopt non-cooperative strategy (U_i^{nc}) is given as:

$$U_i^{nc} = \frac{1+r}{r}[Log\left((\frac{1+r}{1+nr})(rW_{t-1} - z) \right) + \frac{1}{r} Log(\frac{1+r}{1+nr})] \tag{7}$$

$$U_i^{*} = \frac{1+r}{r}[Log \frac{rW_{t-1}}{n}] \tag{8}$$

with

$$U_i^{nc} < U^{*}. \tag{9}$$

These results shows that deficit is incurred by the fragmented government, since its spending level (g_{it}^{nc}) exceeds the social planner desired level. However, the non-cooperative path is suboptimal as its yields lower utility, as shown in equation (9). Simply put, the non-cooperative strategy or fragmented government is characterized with inefficiencies, as well as having an inherent tendency to generate deficit bias.

3.2.2 Fiscal Stabilization and Public-Sector Wealth Dynamics

Given the initial level of public-sector wealth, aggressive appropriation will characterize the budgeting process, thereby putting the fiscal policy on an unsustainable path. However, with the dynamic nature of interaction, fiscal stabilization can be supported as an equilibrium using the trigger strategy. Trigger strategy is defined as an agreement among budget actors to follow social planner strategy as long as no group defects from this implicit arrangement and a threat of reversion to non-cooperative strategy if otherwise (Mason & Phillips, 2002). To examine the implication of this strategy, we assume that groups agree initially to follow the social planner strategy. If one of the groups decides to defect from this agreed-upon path, the valuation function, $V^d(W_t)$ is given as:

$$V^d(W_t) = [Log(g_t) + (1+r)^{-1} V^m(W_{t+1})] \tag{10}$$

The budget constraint facing the defecting group, while others continue to follow the social planner strategy is therefore:

$$W_t = (1+r)\gamma]W_{t-1} - \left[1 - \left(\frac{n-1}{n}\right)\right]z - g_{it} \tag{11}$$

where $\gamma \equiv \left[1 - \left(\frac{n-1}{n}\right)\left(\frac{r}{1+r}\right)\right]$

The optimal spending (g_{it}^d) when defecting is derived by maximizing equation (10) subject to equation (11), which yields:

$$g_{it}^d = \gamma(rW_{t-1} - z) \tag{12}$$

Also, the utility to the defecting group is derived as:

$$U_i^d = \frac{1+r}{r}[Log(rW_{t-1} - z) + \left(\frac{1+r}{r}\right)^2 Log\left(\frac{1+r}{1+nr}\right) + \frac{1+r}{r}\left[Log\gamma - Log\left(\frac{1+r}{1+nr}\right)\right] \tag{13}$$

By comparing the utility from defecting with those of non-cooperative and stabilization paths, we obtain:

$$U_i^d > U_i^{nc} \tag{14}$$

and

$$U_i^d \gtreqless U_i^* \tag{15}$$

Using equations (8) and (13), we can rewrite equation (15) as:

$$Log\left[\frac{rW_{t-1}}{rW_{t-1}-z}\right] \geq \frac{1}{r} Log\left(\frac{1+r}{1+nr}\right) + Log\left[n-(n-1)\left(\frac{r}{1+r}\right)\right] \quad (16)$$

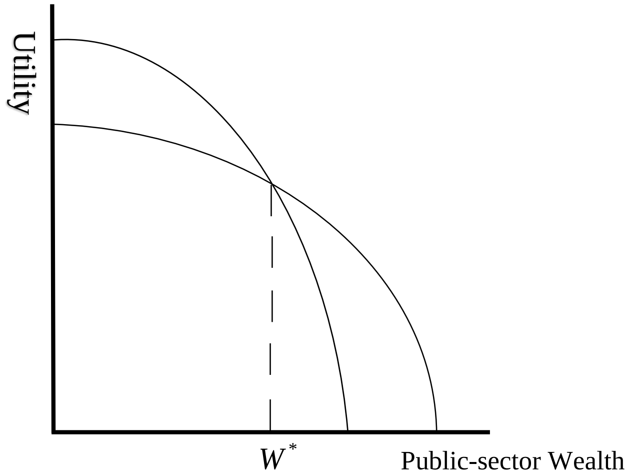
Equations (15 and (16) show that the utility associated with defection path can be greater or less than that which obtains from stabilization path depending on the level of public-sector wealth, which is the only endogenous variable.

By examining the evolution of public-sector wealth in equation (16), it can be shown that at high wealth levels, the utility from defecting exceeds that of cooperating. Thus groups are expected to adopt defection strategy, and if all groups behave accordingly, the outcome will follow non-cooperative path, U_i^{nc}. However, as public-sector wealth declines overtime, it reaches a point where utility from the stabilization path exceeds that of defection path[15]. Thus, we have the situation where each group chooses spending level that restores fiscal stabilization.

Figure 3.1 provides a graphical illustration of this result, with W^* denoting the threshold at which further decline in public-sector wealth leads to fiscal stabilization.

[15] In the case initial public-sector wealth is already low enough to ensure that utility from stabilization path exceeds that from defection path, then stabilization is immediately established.

Figure 3.1: Public-Sector Wealth Dynamics and Stabilization Path



Source: Velasco (1998)

Furthermore, if groups indeed adopt the non-cooperative strategy before reaching the wealth threshold, then it is possible to formally solve for the expected time of stabilization and stock of public-sector wealth when stabilization occurs. Specifically, if we substitute equation (3) into (1), it yields:

$$W_t = \left(\frac{1+r}{1+nr}\right)W_{t-1} + \left(\frac{n-1}{1+nr}\right)z \qquad (17)$$

The definite solution to equation (17) gives:

$$W_T = W_0\left(\frac{1+r}{1+nr}\right)^T - z\left[\left(\frac{1+r}{1+nr}\right)^T - 1\right] \qquad (18)$$

where T is the expected time of stabilization and W_T is the stock of public-sector wealth at the time of stabilization. Essentially, fiscal stabilization is feasible within a fragmented budget institution, although it comes with substantial delay. The stabilization process is driven by public-sector wealth decline, which forces groups to fully internalize the costs and benefits of their distributive policy into the decision framework.

Proposition 1: *There is a threshold at which the decline in public-sector wealth creates a coordination mechanism leading to fiscal stabilization.*

3.3 Experimental Procedure

This section discusses the experimental procedure to test proposition 1. We first describe the laboratory setting, which mimic the basic structure of the economy described in equations (1) to (5). Thereafter, the theoretical predictions based on the specified parameters are derived. These predictions serve as benchmark for comparing the experimental results from this study.

3.3.1 Decision Setting

The decision setting builds on and extends our previous experiment on dynamic common pool game. The participants/players are involved in a legislative bargaining game of "divide-the-dollar", over an infinite horizon. Each period, the players make a demand on the exogenously given public-sector wealth—denoted as experimental points—which evolve according to equation (4). Players' demands are governed by the spending rule given by equation (2), which generates payoff described by equation (1). Furthermore, the dynamic common pool problem is introduced into the game by revealing to the players at the start of each period the shadow cost—dynamic externality—associated with different spending paths. At the end of the experiment, the participants earn cash, based on the cumulative payoff at a pre-specified rate.

The infinite horizon is induced through the use of random stopping rule, after 10 rounds of play. The rule gives the continuation probability of the game and it is defined by the discount factor $1-r$. Unlike in the previous experiment, the terminal period of the game is only determined by the random stopping rule. This design helps prevent the possibility of the game ending prematurely, thereby providing sufficient observations to test the delayed stabilization hypothesis. If points are exhausted (i.e. public-sector wealth is less than group size) before the terminal period, then the economy reaches a debt ceiling point and subsequent allocation will be financed through tax. We exogenously impose a lump-sum tax (T).

The model predicts that if the game approaches this stage where the lump-sum tax has to be imposed, declines in public-sector wealth create some sort of incentives for the players to endogenously implement fiscal stabilization. The implication of this is that observed stabilization at this stage could be due to two interrelated effects. The first is the threat of taxation, which forces player to reduce their appropriation. The second is that of reputation building that emanates from repeated interaction among players; this of course may eliminate the strategic effect and corrects the source of the deficit bias. While both effects could reduce appropriation, only the second effect implies successful stabilization. To separate these two effects, we restore initial public-sector wealth back to the groups after some rounds of paying taxes, which is randomly determined by the computer. Thus, the boost in wealth represents a shock. However, if the groups are already stabilizing, it will simply generate a surplus.

3.3.2 Design Parameters and Treatments

The design parameters for the experiment are defined by W_t, r, n, z and T. For comparability purpose, we continue with parameters chosen for the previous experiment. Specifically, we implement two treatments: primary and secondary. In the two treatments, the legislature/group comprises of $n=3$ and $n=5$ members. Furthermore, the interest rate is $r=0.1$, the lump-sum tax is $T=0.2$, the deadweight loss is $z=1$ point per period and the discount factor is $1-r=0.9$. The only parameter that varies across the treatments is the initial public-sector wealth (W_t), which is normalized to 20 points and 100 points per person in the primary and secondary treatments respectively.

Based on these parameters, we derived theoretical predictions regarding expected time of stabilization and wealth threshold at which this takes place. The predictions as summarized in Table 3.1 indicate that despite the groups operating in a fragmented setting, fiscal stabilization resulted in response to public-sector wealth decline. For example, in the primary treatment with $n=3$, the initial public-sector wealth declined from 60 points to 14 points after seven periods of play, after which the group stabilizes. Another striking prediction is that the larger the group size, the shorter the expected time for stabilization to set in. Overall, the theoretical prediction validates the above proposition of a threshold at which decline in public-sector wealth creates a coordination mechanism that leads to fiscal stabilization.

Table 3.1: Theoretical Predictions

Treatment		Fragmented Government	
Initial Wealth (points)	Group size	Wealth Threshold for Stabilization (W_T, points)	Expected Time of Stabilization (T, period)
60	3	14	7
100	5	13	6
300	3	13	15
500	5	16	10

Source: Author computation

3.4 Sampling Procedure

3.4.1 Optimal Sample Selection

The experimental design for this study has varying treatment levels and continuous outcome. According to List et al. (2011), the optimal sample size (M) under these conditions is given by:

$$M = mk = 2k(t_{\alpha/2} + t_{\beta})^2 (\frac{\sigma}{\delta})^2 \quad (19)$$

where M is the total sample size; k is the number of treatment group, which is four in our design; m is the optimal size of each treatment; $t_{\alpha/2}$ is the critical t value at $\alpha/2$ level of significant, t_{β} represents the power of the test at β level of probability; σ is the variance of the treatment effect; δ is the minimum average effect. The ratio ($\frac{\sigma}{\delta}$) gives the standard deviation of the change in outcome variable and captures the minimum detectable effect size. Using the conventional level of significance of 5percent and power of the test of 0.80 give $t_{\alpha/2}$ as 1.96 and t_{β} as 0.84 respectively[16]. Thus, substituting for these parameters into equation (19) and restricting minimum detectable effect size to detect one standard deviation, the optimal sample (M) size of 64 participants is obtained.

3.4.2 Sampling Technique

The subject pool for the experiment consisted of first and second year undergraduate students of the University of the Witwatersrand, Johannesburg. The selection process involves two stages. In the first stage, emails were sent to students that meet the study's inclusion criteria, namely: (i) the participants that have attained 18 years of age; (ii) the participants are computer literate; and (iii) the participants are in the first or second year undergraduate degree programme.

A total of 446 students responded and volunteered to partake in the experiment. In the second stage, we applied the simple random sampling technique to select the required 64 participants from the initial sampling frame. The demographic characteristics of participants along sex, year of study and faculty is presented in Table 3.2. The distribution shows that the participants come from a

[16] These are the benchmark level of significance and power test reported in optimal sample selection literature (see List et al., 2011).

variety of background, which implies that our sample selection procedure is not biased towards any demographic characteristics.

Table 3.2: Demographic Characteristics of the Participants

Variable	**Frequency**	**Percentage (%)**
Sex		
Male	36	56.3
Female	28	43.8
Year of Study		
1^{st}	38	59.4
2^{nd}	26	40.6
Faculty		
Commerce, Law and Management	25	39.1
Humanities	23	35.9
Science	16	25

3.4.3 Experimental Implementation

The experiment reported in this paper was conducted at the PhD/AERC Computer Laboratory at the University of Witwatersrand in the months of September and October, 2016. Four sessions were run in all: two sessions each for primary and secondary treatments respectively. No participant is involved in more than a session. On average, participants earned approximately R68 and the sessions lasted between 60-90 minutes. The experiment was computerized using z-tree (Zurich toolbox for readymade experiment) developed by Fischbacher (2007).

On the day of the experiment, participants were first taken through the instructions. Appendix 3A shows the instruction set. The instructions detailed the decision setting and group size, but the identity of other group members is not known. Also, the initial public-sector wealth, shadow cost, interest rate and use of the random stopping rule are explained to all participants. The groups are informed of the tax rate imposed if/when the initial wealth is exhausted; they are however not made aware of the fact that their initial wealth will be restored in subsequent periods. Explicit communication is not allowed among the participants. Lastly, before the main experimental session commences, participants had a hands-on pilot session to clarify any ambiguity.

3.5 Empirical Results

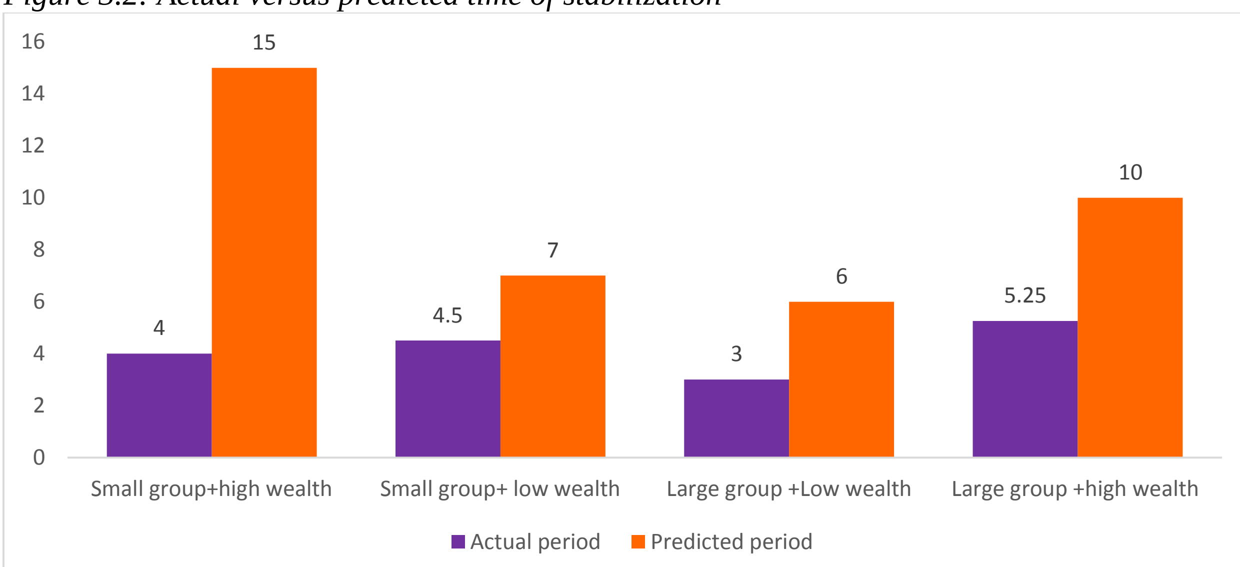
The section contains presentation and analysis of the data obtained from the experiment. The discussion proceeds as follows: we first present some non-parametric analysis regarding the theoretical prediction on delayed stabilization hypothesis. This is complemented with detailed econometric analysis.

3.5.1 Non-parametric analysis

The theoretical description of a fragmented budget institution indicates that while public-sector wealth initially declines—or deficit is generated—there is a threshold where fiscal stabilization becomes the optimal strategy for the budget actors. The expected time and public-sector wealth at this threshold are estimated and presented in Table 3.1. In figures 3.2 and 3.3, we compare these theoretical predictions with outcomes of the experiment. As shown in figure 3.2, all groups attained the expected time of stabilization much earlier than predicted. This means that the rate of appropriation across the groups exceeds the level predicted by the Nash strategy. This observation reinforces the conclusion in chapter two regarding aggressive appropriation behaviour by budget actors in a dynamic setting.

Observation 1: Expected stabilization time is shorter than what is theoretically predicted.

Figure 3.2: Actual versus predicted time of stabilization

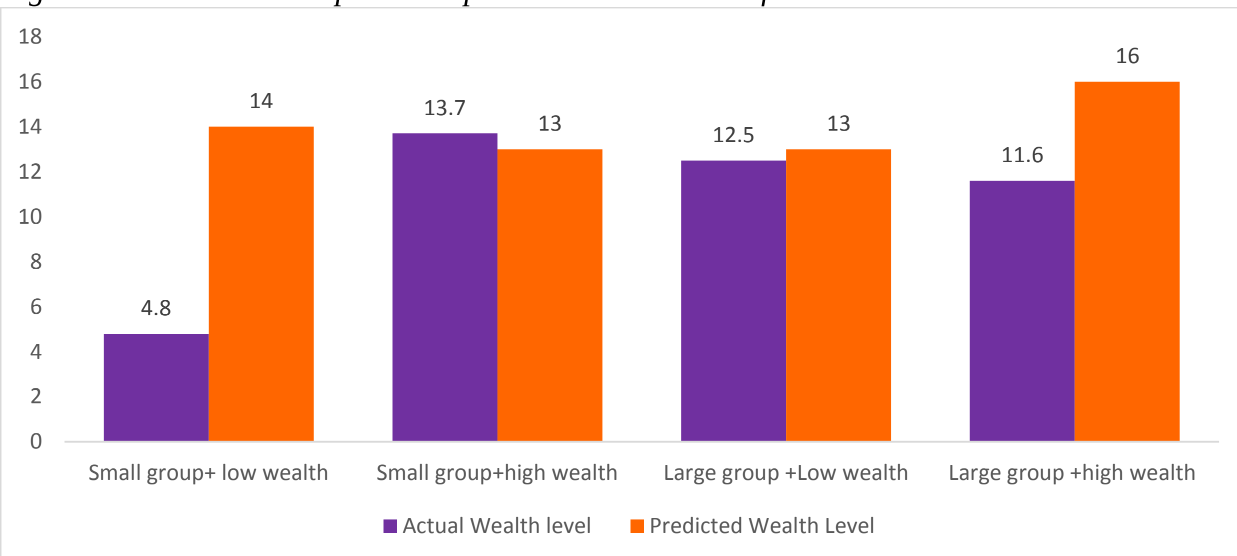


Source: Authors computation

Furthermore, figure 3.3 shows the average public-sector wealth corresponding to the expected stabilization time. Unlike the finding on the timing of stabilization, average public-sector wealth threshold is closer to its theoretically predicted value. We only observed a wide disparity in the first treatment category—small group and high initial wealth. In fact, if this treatment category is excluded, the two-sample *t*-test shows that there is no statistically significant difference between the predicted and observed public-sector wealth threshold (*t*-test= -1.196 and *p*-value=0.297).

Observation 2: The public-sector wealth threshold matches the theoretical expectation.

Figure 3.3: Actual versus predicted pubic-sector threshold for stabilization



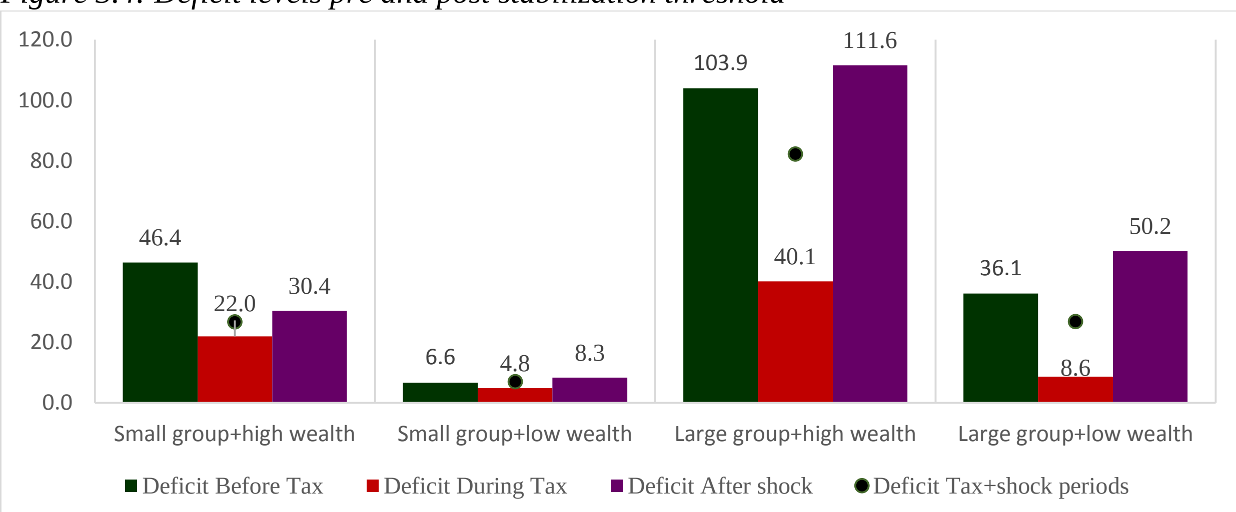
Source: Authors computation

Results of the expected time and public-sector wealth threshold for stabilization, discussed above only reveal whether or not the groups reach a stage in which fiscal stabilization is the optimal strategy. To check if the groups indeed adopted the optimal strategy and importantly, if stabilization ensues afterwards, we needed to compare the public-sector wealth dynamics or deficit before and after the groups reached the stabilization threshold. However, restoring the public-sector wealth after the stabilization threshold implies that the estimated public-sector wealth dynamics will be overstated. Therefore, we only report deficit level in figure 3.4. The deficit level after the stabilization threshold is divided into three parts: (i) when groups pay taxes only; (ii) when group initial wealth is restored only; (iii) the average deficit incurred in (i) and (ii).

The results show that deficit level is lowest when groups pay taxes. In contrast, when the initial public-sector wealth is restored, deficit spike again, even exceeding the level in the pre-stabilization threshold in three of the four treatment categories. Intriguingly, the only exception is groups in the first treatment. As observed in figures 3.2 and 3.3, groups in the first treatment exhibit more aggressive appropriation behaviour than others, prior to reaching the stabilization threshold. This may suggest a behavioural reversal, such that groups that are aggressive in their appropriation becomes more restrained after reaching the stabilization threshold and vice versa. A similar conclusion is reached by averaging the deficits in periods when groups pay taxes and after the initial wealth level is restored. Overall, the evidence did not support delayed stabilization hypothesis, as the observed deficit reduction is temporal and occur only when groups are subjected to taxation.

Observation 3: the experimental evidence contradicts the prediction of delayed stabilization hypothesis.

Figure 3.4: Deficit levels pre and post stabilization threshold



Source: Authors computation

3.5.2 Econometric Model Specification and Estimation

The key prediction of the delayed stabilization hypothesis is that the relationship between fiscal performance and government fragmentation switches regime depending on the dynamics of public-sector wealth. Several approaches have been suggested in the literature to account for this regime dependent relationship. For this study, we adopt the panel threshold regression model developed by B. Hansen (1999). This model is suitable for context in which the regime switching variable is known a priori. However, panel threshold regression model is only applicable to balanced panel data. But the experimental data for this study have an unbalanced panel structure, as the random stopping rule generates different terminal periods across groups. Specifically, the experimental data consist of 16 cross-section units and the time dimension ranges between 10 to 17 periods. Following Mason and Phillips (1997), we transform the initial data set into a balanced panel by using the game with the minimum numbers of rounds as the cut-off period. This gives a balanced panel with 10 periods and 16 cross-section units.

For a panel threshold regression, imposing a balanced panel structure is not expected to have significant effect on the efficiency of the estimator. In fact, B. Hansen (1999) argued that searching for the optimal threshold over the whole dataset is numerically intensive. He instead proposed a shortcut in which the search for threshold is restricted to given quartiles of the dataset. This shortcut will only be inefficient if the threshold value is located towards extreme ends of the distribution. Since in our case the threshold value is not achieved at the upper or lower tails of the distribution, we expect no substantial effect from imposing a balanced panel structure. Nevertheless, it is important to test the results' robustness when extended to unbalanced panel. In this regard, we complement the panel threshold regression model with piecewise linear regression model. According to Gujarati (2003), if the actual threshold point is known in advance, the piecewise linear regression can be used to test the existence or otherwise of a threshold effect. In what follows is a discussion of the panel threshold and piecewise regression models adopted for this paper.

3.5.2.1 Panel Threshold Regression

Following Wang (2015), we capture the single threshold model on the relationship between deficit and government fragmentation as:

$$D_{it} = d_i + \sum_{k=2}^{10} \alpha_k X'_{it} + \begin{cases} \theta_1 n_i + \varepsilon_{1t}, & D^*_{t-1} \leq \gamma_1 \\ \theta_2 n_i + \varepsilon_{2t}, & D^*_{t-1} > \gamma_1 \end{cases} \tag{20}$$

This can be more compactly written as:

$$D_{it} = \theta_1 n_i DM(D^*_{i,t-1} \leq \gamma_1) + \theta_2 n_i DM(D^*_{t-1} > \gamma_1) + \sum_{k=2}^{10} \alpha_k X'_{it} + d_i + \varepsilon_{it} \tag{21}$$

where D_{it} is the deficit/surplus incurred by group i in period t; n_i is the group size (measure of government fragmentation) which also represents the threshold variable. Only the direct measure of government fragmentation is used in this chapter. Also, γ_1 is the endogenously determined threshold parameter which splits the sample into two regimes; D^*_{t-1} is the lag of deficit level corresponding to public-sector wealth threshold, DM is the dummy variable which is equal to 1 if $D^*_{t-1} > \gamma_1$ or zero otherwise; d_i is the group-level fixed effect; ε_{it} is the disturbance term. Also, X'_{it} represents the control variables, which includes: shadow cost which is measured as $Log(W_{i,t-1}/W_{i,t-2})$; inverse of time index[17] and initial wealth of the group, which is captured by a binary dummy variable— $Dumm\_high$ which takes the value of "1" if the group belongs to the secondary treatment and "0" otherwise.

Furthermore, while the theoretical model suggests a single threshold in the relation between deficit and government fragmentation, the preliminary non-parametric results raise the possibility of double thresholds. For example, the deficit level varies after stabilization threshold, between periods when the groups pay tax and when the initial-wealth is restored. We therefore extend equation (21) to test the possibility of double thresholds as follows:

$$D_{it} = \theta_1 n_i DM(D^*_{i,t-1} \le \gamma_1) + \theta_2 n_i DM(\gamma_1 < D^*_{t-1} \le \gamma_2) + \theta_3 n_i DM(D^*_{t-1} > \gamma_2) + \sum_{k=2}^{4} \alpha_k X'_{it} + d_i + \varepsilon_{it} \quad (22)$$

where γ_2 is the second threshold value, with $\gamma_1 < \gamma_2$ and other variables are as previously defined.

B. Hansen (1999) proposed a fixed effect model to estimate equations (21) and (22). However, government fragmentation, which is our main variable of interest, is time invariant and applying fixed effect model will eliminate any time invariant variable. Yet, only Hansen's fixed effect model exists so far in the literature that could account for threshold effect in a panel set-up (see Wang, 2015). It therefore becomes necessary to transform equations (21) and (22) by interacting all the time invariant variables with the time dimension, represented by (*T*). Applying this transformation gives:

[17] A period is defined as an interval between when participants make an allocation decision and when the stopping rule is invoked to determine whether or not to proceed to the next period.

For single threshold:

$$D_{it} = \theta_1 n_i TDM\,(W^*_{i,t-1} \le \gamma_1) + \theta_2 n_i TDM\,(W^*_{t-1} > \gamma_1) + +\sum_{k=2}^{4} \alpha_k X'_{it} + d_i + \varepsilon_{it} \qquad (23)$$

For double thresholds:

$$D_{it} = \theta_1 n_i TDM\,(W^*_{i,t-1} \le \gamma_1) + \theta_2 n_i TDM\,(\gamma_1 < W^*_{t-1} \le \gamma_2) + \theta_3 n_i TDM\,(W^*_{t-1} > \gamma_2) + \sum_{k=2}^{4} \alpha_k X'_{it} + d_i + \varepsilon_{it} \quad (24)$$

This transformation allows us to estimate the time invariant effect, albeit with a caveat. The caveat is that, θ_1, θ_2 and θ_3 no longer measure the marginal effects of changes in government fragmentation on deficit; rather how the effect of government fragmentation on deficit varies over time. In this regard, the delayed stabilization hypothesis holds true if θ_2 is insignificant or has a negative sign in the case of single threshold, or if both θ_2 and θ_3 are insignificant or have negative signs in the case of double threshold.

3.5.2.2 Piecewise Linear Regression

The advantage of panel threshold regression model over piecewise linear regression is that it endogenously identifies the threshold point. However, since our theoretical model suggests possible threshold point, piecewise linear regression can be used to determine if indeed a threshold effect exist at the point suggested. According to our model, fiscal stabilization is expected once groups have attained the stabilization threshold. This means government fragmentation should have no effect on deficit in the period after stabilization threshold. Thus, we define a dummy variable— $Break1$, which is 1 in periods when the groups have not reached stabilization threshold and zero otherwise. Periods after the stabilization threshold are the reference category. Following, Brooks (2014), we therefore specify the piecewise linear regression relating to effect of government fragment on deficit as:

$$D_{it} = \phi_1 n_i + \phi_2 n_i * Break1 + \varphi_1 Break1 + \sum_{k=2}^{4} \alpha_k X'_{it} + d_i + \varepsilon_{it} \quad (25)$$

where all the variables are as previously defined. For the delayed stabilization hypothesis to hold, $\phi_1 + \phi_2$ should be significantly less than ϕ_1. Also, to account for the possibility of double thresholds, we include an additional dummy variable—$Break2$ which takes the value of 1 in periods when the groups pay taxes and zero otherwise. In this case, the period(s) after groups' initial wealth is restored now serve as reference category. Incorporating this into equation (25) yields;

$$D_{it} = \phi_1 n_i + \phi_2 n_i * Break1 + \phi_3 n_i * Break2 + \varphi_1 Break1 + \varphi_2 Break2 + \sum_{k=2}^{4} \alpha_k X'_{it} + d_i + \varepsilon_{it} \quad (26)$$

This implies that delayed stabilization hypothesis is validated on the condition that $\phi_1 + \phi_2$ is less than ϕ_1 and not significantly different from $\phi_1 + \phi_3$.

Equations (25) and (26) are estimated with random effect model and the standard error computed using the clustering method that accounts for possible heteroscedasticity and autocorrelation.

3.5.2.3 Diagnostics analysis

According Cheng, Liu, and Chien (2010), Hansen's fixed effect panel threshold estimator is only valid and unbiased when the variables in the model are stationary and the threshold effect is statistically significant. Against this background, the estimated model is tested for stationarity and significance of the threshold effect. Table 3.3 presents the panel unit-root results based on Levin-Lin-Chu (LLC) and Im, Pesaran and and Shin (IPS) tests (see Im, Pesaran, & Shin, 2003; Levin, Lin, & Chu, 2002). The group size is excluded from the test, since it is *ab initio* time invariant. Both tests reveal that all variables are stationary at 1% level of significance.

Table 3.3: Panel unit-root test results

Variables	**LLC**		**IPS**	
	***t*-statistics**	***p*-value**	***t*-statistics**	***p*-value**
Time	-53.99***	0.000	-8.46***	0.000
Shadow cost	-6.82***	0.000	-5.27***	0.000
Deficit	-4.88***	0.000	-5.11***	0.000

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

The results of the test for threshold effect is presented in Table 3.4. We use bootstrap method on critical values of *F*-statistics at 1%, 5% and 10% respectively to test the significance of single versus double threshold effects. The bootstrap procedure is repeated 200 times for single threshold and the result yields *F*-statistics of 32.42 and *p*-value of 0.000. However, repeating the same process for the double threshold, we obtain an *F*-statistics of 1.18 and *p*-value of 0.885. This means that single threshold model is the best fit. Also, the threshold parameter/value which is the lag of deficit that split the model into two regimes is calculated as 21.475. A further interrogation of the data shows that the estimated threshold parameter corresponds to periods when the groups are paying tax.

Table 3.4: Test of threshold effects between deficit and government fragmentation

Test	Threshold value	F	p-value	Critical value of *F* 1%	5%	10%
Single threshold effect test	21.475	34.42***	0.000	7.144	8.524	12.764
Double threshold effect test	21.475	34.42***	0.000	6.585	7.694	10.311
	5.132	1.18	0.885	6.467	8.3678	11.135

Note: *F*-statistics and p-values are from repeating the bootstrap procedures 200 times for each of the two bootstrap tests. ***, ** and * indicate significance at the 1, 5 and 10% level, respectively

3.5.3 Econometric Results

3.5.3.1 Fixed effect panel threshold regression results

Table 3.5 presents results of the single fixed effect panel threshold regression estimates using data from the experiments. For comparability purposes, we also show results of the double thresholds regression analysis. The result based on single threshold shows that θ_2 is positive and significant. This implies that in the post-stabilization threshold periods, which consist of both the periods when the groups were paying tax and when the initial wealth was restored, deficit level increases with rising government fragmentation. This evidence contradicts the prediction of delayed stabilization. Furthermore, the double threshold results, although not the best fit, help to decompose the relationship between government fragmentation and deficit in the post-stabilization threshold periods. Specifically, θ_2 that captures the effect of government fragmentation in periods when the groups are paying tax is negative and significant. This means that the prediction of delayed stabilization holds over these periods. However, θ_3, which measures the effect of government fragmentation after restoring the initial wealth is positive and significant. This indicates that the temporary decline in deficit is due to threat of taxation.

With respect to control variables included in the model, shadow cost is positive and significant; that is, an increase in rate appropriation leads to higher level of deficit. Moreover, deficit level is found to increase over time among groups with high initial wealth level than other treatment. Also, deficit has an inverse relationship with time index, which implies it declines over the periods.

Table 3.5: Estimated fixed effect panel threshold regression model

	Single	Double
1/TIME	8.192***	8.289***
	(1.478)	(1.502)
Lag of shadow cost	6.238**	6.322**
	(2.138)	(2.168)
Dumm_high**T*	9.329***	8.662***
	(1.570)	(1.343)
<u>Government fragmentation</u> (n_i)		
θ_1	-1.101**	-0.612
	(0.459)	(0.425)
θ_2	0.836*	-1.145**
	(0.447)	(0.481)
θ_3		0.893*
		(0.447)
Constant	-27.42	-28.91
	(19.64)	(20.11)
R^2	0.507	0.511
Adj. R^2	0.491	0.491
F-statistics	38.85***	40.77***

White standard errors in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.5.3.2 Piecewise linear regression results

Results from the piecewise linear regression is presented in Table 3.6. For the single threshold, ϕ_1 is estimated to be 17.39, while $\phi_1 + \phi_2$ is 24.17. This implies that the effect of government fragmentation on deficit becomes more pronounced after the stabilization threshold. However, more disaggregated estimates from double thresholds show that ϕ_1 is 29.42, $\phi_1 + \phi_2$ is 24.597, and $\phi_1 + \phi_3$ is 4.78. These results imply that effect of government fragmentation on deficit is lowest when groups pay taxes, and highest after the initial wealth is restored. The *p*-values also indicate that $\phi_1 + \phi_2$ and $\phi_1 + \phi_3$ are statistically different. This result concurs with the conclusion of the panel fixed effect threshold model; players in the experiment do not behave as predicted by the delayed stabilization model.

Table 3.6: Estimated piecewise linear regression model

	Single	Double
L.shadow cost	6.350***	5.269***
	(2.190)	(0.732)
1/TIME	11.24***	12.02**
	(3.595)	(4.767)
Dumm_high	49.51**	47.30**
	(22.83)	(21.98)
Government fragmentation (n_i)	17.39***	29.42***
	(0.183)	(0.779)
$n_i * Break1$	7.172***	-4.823***
	(1.287)	(0.830)
$n_i * Break2$		-24.64***
		(2.457)
$Break1$	-66.03***	-26.69*
	(12.21)	(14.65)
$Break2$		83.27***
		(7.503)
Constant	-99.78***	-135.6***

	(25.20)	(19.05)
R^2	0.822	0.835

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

3.5.4 Sensitivity Analysis

The results discussed in section 3.5.3 are robust to alternative estimation techniques as already showed. Therefore, the discussion that follows focuses only on concerns with noisy decision making.

(i) *Noisy decision making*: The response of participants to the administered questionnaire (see appendix 3B) is presented in Table 3.7. Majority of the participants indicate they fully or mostly understood instructions of the pilot session, main experiment as well as keystroke for the z-tree program. Only four participants indicated that instructions of the pilot session are not understood; the number dropped to one for the main experiment. This suggests that decisions of the participants are not affected by lack of clarity of instructions.

The participants were also debriefed about their reactions to tax imposition and subsequent reinstatement of initial wealth level. In line with findings from non-parametric and econometric analyses, majority of the participants indicated that they reduced their appropriation demand in response to tax being imposed, but reacted in a diametrically opposite manner when the initial wealth is restored. This indicates that observed behaviour of participants in the experiment is not due to noisy decision making.

Table 3.7: An Assessment of Noise in Decision Making

Response to questions relating to the instruction			
Scale	**Pilot Session Instruction**	**Main Experimental Instruction**	**Keystroke**
Fully Understand	34 (53.1%)	39 (60.9)	55 (85.9%)
Mostly Understand	26 (40.6%)	24 (37.5%)	9 (14.1%)
Didn't Understand	4 (6.3%)	1 (1.6%)	0
Response to questions relating to the decision making			
Scale	**Reaction to tax imposition**	**Reaction to initial wealth being restored**	
Reduced demand	31 (48%)	8 (13%)	

Maintain demand	23 (36%)	11(23%)
Increase demand	10 (16%)	41(64%)

Source: Author's computation

3.6 Conclusion

This paper empirically tests the prediction of delayed stabilization hypothesis as formulated by Velasco (1997). Specifically, we examined the possibility of a public-sector wealth threshold at which government fragmentation exert no further effect on deficit. Given the difficulties in using field data to directly test the hypothesis, we adopted the experimental approach. The experimental data are analysed using a battery of econometric techniques—non-parametric test; panel fixed effect threshold regression and piecewise linear regression.

In general, the findings do not support the prediction of delayed stabilization. Regardless of treatment categories, groups attained the pubic-sector wealth threshold for stabilization as suggested in the model. However, we did not observe the convergence in their appropriation levels towards the social planner level. Instead, players displayed aggressive appropriation behaviour, leading to higher deficit that even exceeded the level in periods before pre-stabilization threshold is reached. In essence, the negative effect of government fragmentation on deficit becomes even more pronounced in post-stabilization threshold phase, contrary to the prediction of delayed stabilization.

The results corroborate Lukkezen and Rojas-Romagosa (2013) and Purfield (2003) studies that also find weak support for delayed stabilization hypothesis but differ from the optimistic conclusion of Bohn (1998) and Ardagna (2004). Moreover, this paper highlights a probable explanation for the contrasting findings in the literature. Specifically, we observed a temporary stabilization during the periods when groups are being taxed. This suggests that designing means of augmenting public-sector wealth could moderate the relationship between government fragmentation and deficit. Previous studies such as Botelho, Dinar, and Costa (2013) and Thaler (2008) noted that effect of common pool problem is less severe when the main source of resources is taxation. Thus, our findings reaffirm this conclusion. Similarly, our results seem to identify with the "*crisis hypothesis*" as suggested by Lora and Olivera (2004) and Tommasi and Velasco

(1996)—the crisis hypothesis postulates that the probability of successful stabilization is higher during periods of economic crisis, than otherwise.

These findings have important implications for fiscal policy making, especially in developing economies. For example, it has been widely observed that an increase in public-sector wealth in these economies leads to higher spending and debt accumulation. This observation matches the experimental findings in this study, where subjects respond to increased wealth with aggressive spending behaviour. This suggests the fundamentals driving spending and debt trajectories could be due to common pool problem, which are rampant in budget institutions in developing countries. The implication of the findings on delayed stabilization hypothesis in this context is that budget actors, on their own, should not be relied upon to correct the deficit bias. In essence, it becomes crucial for policy makers to adopt proactive active stabilization policy such as fiscal rules or centralization of the budget institution. More importantly, this paper suggests that periods of economic crisis are the optimal time to introduce holistic stabilization strategy. In addition, it is imperative for developing countries to diversify government revenue base away from natural resources—analogous to a given initial public-sector wealth in our experiments—and instead develop a viable tax capacity. As our results clearly show, taxation tends to be effective in ameliorating the negative effect of government fragmentation on fiscal performance.

3A Appendix: Experimental Instructions

Instruction 1: For small group and low wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant to attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.27 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of three (3) members and your group account has been allocated an initial endowment of 60 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- ❖ **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 6 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any positive real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- ❖ **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period? Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How the random stopping rule will be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

How is the endowments for subsequent periods calculated? The remaining endowment in the previous period is adjusted for endowment dynamics and interest rate to derive the endowments for the subsequent periods. The adjustment involves:

- A deduction of a point out of the remaining endowment whenever endowment dynamics is greater than interest rate. E.g. given the interest rate is 0.1 (10 percent), while the remaining endowment is 21 and the endowments dynamics for the period is 0.11 (11 percent). The endowments for the next period therefore reduce to 20. This deduction does not occur if the endowment dynamics is less than or equal to interest rate.
- The total endowment left after the first adjustment attract an interest rate of 0.1 (10 percent) for the next period. E.g. in the case above, since there are 20 points left after adjusting for endowment dynamics, hence in the subsequent periods, the total endowments to share become 22 points (20+20*0.1). Note that the computer managing the experiment will automatically implement these adjustments.

What happened when the initial endowment is exhausted? In the case the group's initial endowment is exhausted and the random stopping has indicated that the game should proceed, *a 20 percent tax is then levelled on the amount of points in each individual account and transfer back to the group account*. It is from this new endowment that subsequent transfer decision will be made. If the game reaches this stage, you will see message "Taxation now in progress" displayed on your monitor.

What is the endowment dynamics? It is ratio of the endowment in the last period to endowment in the current period and it indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one, then the initial endowments is increasing overtime, but if it is less than one, then initial endowments is decreasing.

What happens when the game ends and there are still some endowments? The remaining endowments are simply discarded. This means they are not added to your payoff.

What happens when the game ends? Your payoff in each period is summed up, and using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 2: For small group and high wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant to attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. _**All handouts must be turned in at the end of the experiment.**_

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.03 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of three (3) members and your group account has been allocated an initial endowment of 300 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 6 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any positive real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period? Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How the random stopping rule will be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

How is the endowments for subsequent periods calculated? The remaining endowment in the previous period is adjusted for endowment dynamics and interest rate to derive the endowments for the subsequent periods. The adjustment involves:

- ❖ A deduction of a point out of the remaining endowment whenever endowment dynamics is greater than interest rate. E.g. given the interest rate is 0.1 (10 percent), while the remaining endowment is 21 and the endowments dynamics for the period is 0.11 (11 percent). The endowments for the next period therefore reduce to 20. This deduction does not occur if the endowment dynamics is less than or equal to interest rate.
- ❖ The total endowment left after the first adjustment attract an interest rate of 0.1 (10 percent) for the next period. E.g. in the case above, since there are 20 points left after adjusting for endowment dynamics, hence in the subsequent periods, the total endowments to share become 22 points (20+20*0.1). Note that the computer managing the experiment will automatically implement these adjustments.

What happened when the initial endowment is exhausted? In the case the group's initial endowment is exhausted and the random stopping has indicated that the game should proceed, *a 20 percent tax is then levelled on the amount of points in each individual account and transfer back to the group account*. It is from this new endowment that subsequent transfer decision will be made. If the game reaches this stage, you will see message "Taxation now in progress" displayed on your monitor.

What is the endowment dynamics? It is ratio of the endowment in the last period to endowment in the current period and it indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one, then the initial endowments is increasing overtime, but if it is less than one, then initial endowments is decreasing.

What happens when the game ends and there are still some endowments? The remaining endowments are simply discarded. This means they are not added to your payoff.

What happens when the game ends? Your payoff in each period is summed up, and using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 3: For large group and low wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant to attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.18 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of five (5) members and your group account has been allocated an initial endowment of 100 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 10 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any positive real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period? Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How the random stopping rule will be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

How is the endowments for subsequent periods calculated? The remaining endowment in the previous period is adjusted for endowment dynamics and interest rate to derive the endowments for the subsequent periods. The adjustment involves:

- ❖ A deduction of a point out of the remaining endowment whenever endowment dynamics is greater than interest rate. E.g. given the interest rate is 0.1 (10 percent), while the remaining endowment is 21 and the endowments dynamics for the period is 0.11 (11 percent). The endowments for the next period therefore reduce to 20. This deduction does not occur if the endowment dynamics is less than or equal to interest rate.
- ❖ The total endowment left after the first adjustment attract an interest rate of 0.1 (10 percent) for the next period. E.g. in the case above, since there are 20 points left after adjusting for endowment dynamics, hence in the subsequent periods, the total endowments to share become 22 points (20+20*0.1). Note that the computer managing the experiment will automatically implement these adjustments.

What happened when the initial endowment is exhausted? In the case the group's initial endowment is exhausted and the random stopping has indicated that the game should proceed, *a 20 percent tax is then levelled on the amount of points in each individual account and transfer back to the group account*. It is from this new endowment that subsequent transfer decision will be made. If the game reaches this stage, you will see message "Taxation now in progress" displayed on your monitor.

What is the endowment dynamics? It is ratio of the endowment in the last period to endowment in the current period and it indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one, then the initial endowments is increasing overtime, but if it is less than one, then initial endowments is decreasing.

What happens when the game ends and there are still some endowments? The remaining endowments are simply discarded. This means they are not added to your payoff.

What happens when the game ends? Your payoff in each period is summed up, and using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 4: For large group and high wealth

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, you can only do so by raising your hand and the laboratory assistant to attend to you accordingly. You will now participate in a decision making experiment.

What material do you have for the experiment? You are to make use of three handouts and the computer monitor on your desk in the course of the experiment. The computer monitor is where you input all your decisions in each period. The first handout is the instruction that you are free to read any time during the course of the experiment. The second handout is the personal history form; more of a scratch sheet that you can for note taking during the experiment. Please note that the use of the personal history form is optional but must be turned in at the end of the experiment(s). The last handout is a questionnaire to be completed and handed in after the experiment. Do not complete the questionnaire until you are told to do so. *All handouts must be turned in at the end of the experiment.*

What determines how much is earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding how to divide some currency denominated in **points**. The amount you earn depends on your decisions and decisions of other members of your group. Specifically, in each period you will make a transfer from the group's account to your individual account; the total point transferred is used to calculate your payoff.

The payoff is calculated as:

$$payoff = Logarithm\ (Your\ Transfer\ \ for\ the\ Period)$$

[THE COMPUTER WILL ESTIMATE THIS]

The amount you earn is given by the sum of payoffs in each of the individual period. The exchange rate is:

0.016 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about whom he or she is paired with.

How does the experiment work? You have been randomly assigned to a group consisting of five (5) members and your group account has been allocated an initial endowment of 500 points. You have only this initial endowment to decide upon over an undefined periods. Each period has two stages described as follows:

- **Stage 1**: Each member has to enter the number of points to demand from the group account (W) and transfer into his/her individual account $(\Pr)$.

Information

- ✓ Your transfer must not be greater than the average endowment which is available at any period. E.g. if the endowment for the current period is 10 points, then your transfer can be less but must not exceed 2 points. The average endowment available for each period will always be indicated on your computer monitor.
- ✓ Each member makes his/her decision privately (no consultation with other members)
- ✓ Any positive real number is allowed (e.g. 5, 0.5, 1.5 etc.)

- **Stage 2:** Your pay-off for the period is then calculated by the computer. The computer will also display the summary of the last period (e.g. total demand by all members of your group in the period ending, endowment dynamics & the remaining endowments in the group's account) and information for the next period (e.g. endowment for the period).

Completion of stage 2 marks the end of a period

Whether to proceed or not to the next period? Whether or not the game is terminated or continued is determined by a random stopping rule. This rule states that the game is terminated with probability of 0.1 (10 percent), otherwise the game continues.

How the random stopping rule will be implemented? The computer managing the experiment will generate random number from 0 to 1. If any number between 0 and 0.9 comes up, the game continues, but if the number is greater than 0.9, game terminates. You will be notified when this occurs.

What happens if the random stopping indicates the game's continuation? You will follow the procedures as described in Stages 1 and 2, based on new endowments.

How is the endowments for subsequent periods calculated? The remaining endowment in the previous period is adjusted for endowment dynamics and interest rate to derive the endowments for the subsequent periods. The adjustment involves:

- ❖ A deduction of a point out of the remaining endowment whenever endowment dynamics is greater than interest rate. E.g. given the interest rate is 0.1 (10 percent), while the remaining endowment is 21 and the endowments dynamics for the period is 0.11 (11 percent). The endowments for the next period therefore reduce to 20. This deduction does not occur if the endowment dynamics is less than or equal to interest rate.
- ❖ The total endowment left after the first adjustment attract an interest rate of 0.1 (10 percent) for the next period. E.g. in the case above, since there are 20 points left after adjusting for endowment dynamics, hence in the subsequent periods, the total endowments to share become 22 points (20+20*0.1). Note that the computer managing the experiment will automatically implement these adjustments.

What happened when the initial endowment is exhausted? In the case the group's initial endowment is exhausted and the random stopping has indicated that the game should proceed, *<u>a 20 percent tax is then levelled on the amount of points in each individual account and transfer back to the group account</u>*. It is from this new endowment that subsequent transfer decision will be made. If the game reaches this stage, you will see message "Taxation now in progress" displayed on your monitor.

What is the endowment dynamics? It is ratio of the endowment in the last period to endowment in the current period and it indicates the rate at which initial endowment evolved over time. If the endowment dynamics is greater than one, then the initial endowments is increasing overtime, but if it is less than one, then initial endowments is decreasing.

What happens when the game ends and there are still some endowments? The remaining endowments are simply discarded. This means they are not added to your payoff.

What happens when the game ends? Your payoff in each period is summed up, and using the exchange rate given above. You will receive your earnings privately.

<u>Any question before we take some practice session?</u>

3B Appendix: QUESTIONNAIRE

(To be filled after the experiment)

1. Terminal number (Number written on your screen)

____

2. Gender:

☐ Male ☐ Female

3. Year of study:

☐ 1st Year ☐ 2nd Year

4. Department:

____

5. Instructions for Practice:

☐ I fully understood the instructions.

☐ I understood most of the instructions.

☐ I didn't understand much of the instructions.

6. Instructions for the Experiment:

☐ I fully understood the instructions.

☐ I understood most of the instructions.

☐ I didn't understand much of the instructions.

7. Keystrokes for the Experiment:

☐ I was fully at ease with the keystrokes.

☐ I was at ease with most of the keystrokes.

☐ I was not at ease with the keystrokes.

8. How did you react when the taxes were imposed on your group?

☐ Reduce the total transfer demanded than previously.

☐ Maintain the total transfer demanded.

☐ Increase the total transfer demanded than previously.

9. How did you react when the initial endowment was restored?

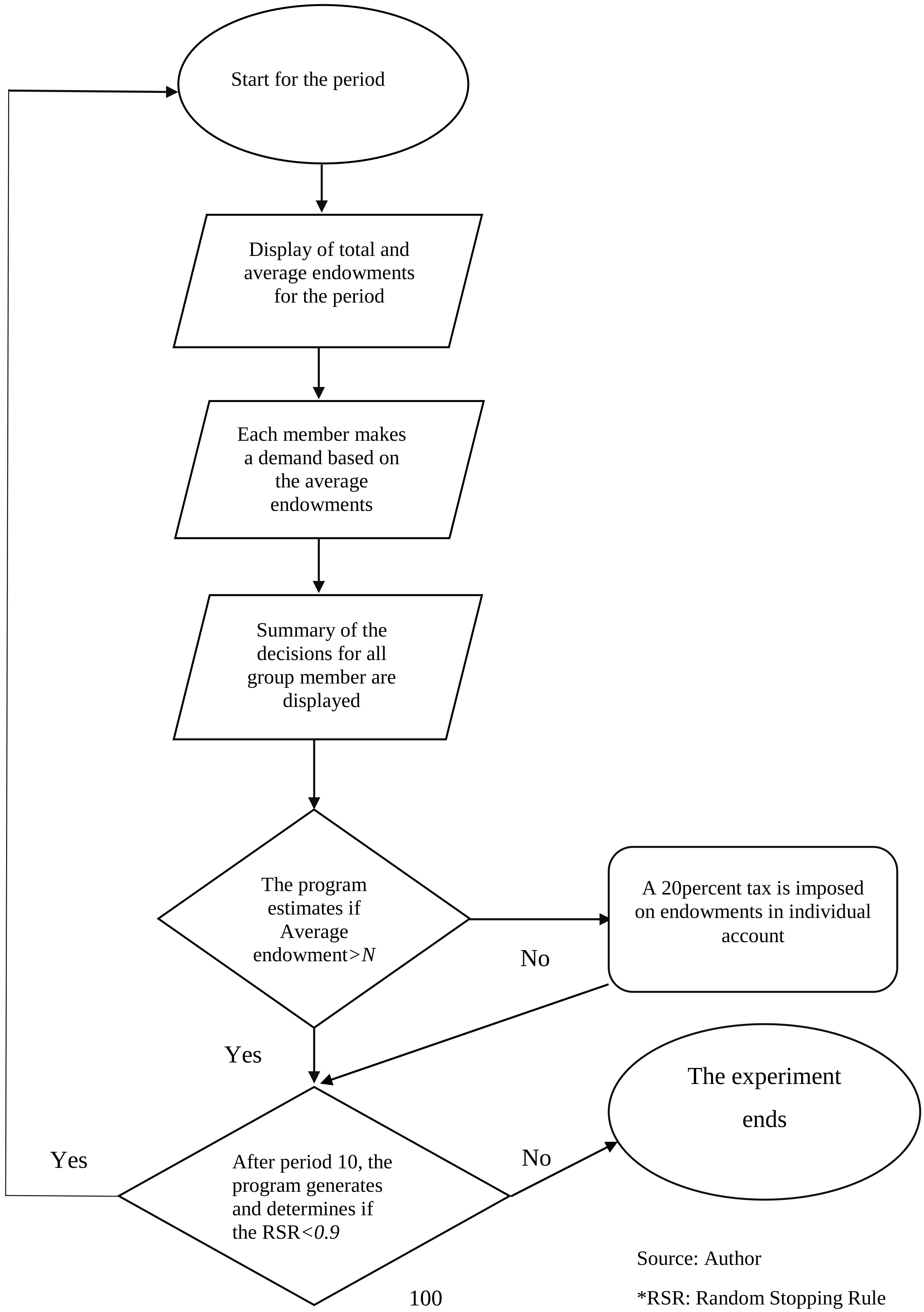
- ☐ Reduce total transfer demanded than previously.
- ☐ Maintain the level of total transfer demanded.
- ☐ Increase the total transfer demanded than previously.

3C Appendix: Z-Tree

The z-tree provides a convenient means of mimicking the budget institution and decision setting using a networks of computer systems. The z-tree consists of two programs: z-tree and z-leaf. The z-tree program is the server program which the experimenter works with. It is basically made up of lines of command that map out the desired decision setting and institution. On the other hand, z-leaf program translates the lines of command in z-tree into sets of interface for participants to interact and also make their budgeting decision.

Figures 0.2C illustrates the step by step process through which the z-tree program is translated into the z-leaf program, as implemented for the experiment. At the start of any period, the participants in a group are shown the total and average endowments to be appropriated upon over the entire duration of the experiment. The groups are then taken to the next stage where each member makes a decision regarding the amount of endowments to transfer from the group account to their individual account. This demand cannot exceed the average endowments for the period, in line with the spending rule. After making their decision, groups are taken to the next stage where the summary of decisions for the period is displayed for all members, which include: the individual and group demands, endowments dynamics and the remaining endowments. Any group that exhaust their initial endowment are taxed at the rate of 20percent. However, after period 10, the random stopping rule is invoked and this determines when the experiment will be terminated. The loop is repeated until the random stopping rule terminate the experiment.

Figure 0.2C: A Diagrammatic Illustration of the Z-TREE program



Source: Author

*RSR: Random Stopping Rule

4. Chapter Four: Multiyear Budgeting and Fiscal Performance: Experimental Evidence

4.1 Introduction

In recent years, several countries have introduced multiyear budgeting—one of the most important phases of the Medium Term Expenditure Framework (MTEF)—in order to address observed common pool problem in fiscal policy making. World Bank (2013) estimated that as at 2008, about 80 percent of countries globally have adopted multiyear budgeting in one form or another. It is posited that multiyear budgeting helps to improve fiscal performance through the establishment of a sound framework for fiscal discipline, enhancement of transparency and accountability in budgeting process, and correction of perceived short-sightedness embedded in annual budgeting (Boex, Martinez-Vazquez, & McNab, 2000; Vlaicu, Verhoeven, Grigoli, & Mills, 2014; Wyplosz, 2012).

However, empirical evidence on the effect of multiyear budgeting on fiscal performance have produced rather divergent results. One explanation for this lack of consensus is provided in the voting equilibrium model developed by Ferejohn and Krehbiel (1987); they opined that a rational and forward-looking budget actor, deciding on alternative budget proposals will vote to maximize his/her spending preference. The implication is that budget processes, such as top-down or multiyear budgeting, do not necessarily improve fiscal performance; their effect depends largely on the configuration of voters' preferences. Existing literature have examined the implication of this model in relation to top-down budgeting, even though portends similar consequence for multiyear budgeting.

Against this background, this paper investigates implication(s) of the voting equilibrium model on the relationship between multiyear budgeting and fiscal performance in a laboratory experiment. Given that voters' preferences are difficult to measure using field data, experimental approach becomes more relevant in addressing the key objective of this study. The experimental setting consists of n legislators, who are to decide the aggregate budget size over a defined planning horizon. We implement a factorial design with two configurations of voters' preferences and budget processes—multiyear and annual budgeting. Also, fiscal performance is measured by aggregate budget size, such that a smaller-sized budget corresponds to improved fiscal performance. Based on this setting, comparison of budgetary outcomes under various treatments helps to determine the extent to which voters' preferences affect the relationship between multiyear budgeting and fiscal performance.

This paper contributes to existing literature on the subject in two significant ways. First, we extend the voting equilibrium model by applying it to a setting in which budgetary decision involves intertemporal allocation. This departs from earlier studies (cf Ehrhart et al., 2007; Gardner & von Hagen, 2012) which focus on sectoral allocation of budget for a single period. Second, the paper contributes to the empirical literature on the efficacy of multiyear budgeting. Specifically, we examine the likelihood of voters' preference acting as a confounding factor on the relationship between multiyear budgeting and fiscal performance. This understanding is crucial for policymakers in designing a more effective fiscal rule to mitigate the common pool problem.

The reminder of the paper is organised as follows. Section 2 discusses the conceptual issues around multiyear budgeting and medium-term expenditure framework. Section 3 briefly reviews the literature regarding effect of multiyear budgeting on fiscal performance. Section 4 presents the theoretical model and the basic predictions of the voting equilibrium model. Section 5 describes the experimental procedure adopted for the study. Section 6 presents the experimental results. Finally, policy implications of the results are discussed in section 7.

4.2 Conceptual Issues

At the outset, it is crucial to highlight the distinction between multiyear budgeting and the closely related concept of medium-term expenditure framework (MTEF). Multiyear budgeting is an institutional tool aimed at extending planning horizon of public-sector budget beyond a fiscal period; usually to include estimates for two to three additional years (Boex et al., 2000; Filc & Scartascini, 2011). The objective of multiyear budgeting is to address the short-sightedness and conservatism entrenched in annual budgeting (Wildavsky, 2003). Multiyear budgeting also eliminates the uncertainty regarding paths of fiscal policy, thereby reducing the extent of strategic effect(s) among budget actors. If correctly implemented, according to received literature (*Ibid*), multiyear budgeting may eliminate deficit bias and result in improved fiscal performance.

On the other hand, MTEF focuses on translating macro-fiscal objectives and constraints into broad budget aggregates and sectoral expenditure plans, while retaining the defining attribute of multiyear budgeting (World Bank, 2013). In practice, MTEF involves three phases, as shown in

Figure 4.1. The first phase of MTEF—medium term fiscal framework (MTFF)—details on key fiscal targets over the defined planning horizon. The main difference between the MTFF and multiyear budgeting is the inclusion of aggregate fiscal projections to the MTFF. However, given the weak fiscal projections in many developing countries, MTFF and multiyear budgeting can be taken as the same in this context. Other phases of MTEF introduce elements of program logic and result-oriented budgeting along with multiyear budgeting; the aim being to improve the allocative and technical efficiencies of government finance (Grigoli et al., 2012).

Figure 4.1: Stages in Medium Term Expenditure Framework

<table>
<tr><th colspan="3">Phase</th><th>Projections</th></tr>
<tr><td rowspan="6">Medium-Term Performance Framework (MTPF)</td><td rowspan="4">Medium-Term Budgetary Framework (MTBF)</td><td>Medium Term Fiscal Framework (MTFF)</td><td>Aggregated fiscal projections (e.g. expenditure, income, deficit)</td></tr>
<tr><td rowspan="3"></td><td>Expenditure projections by Administrative Unit</td></tr>
<tr><td>Expenditure projections by Functions</td></tr>
<tr><td>Disaggregated Income Projections</td></tr>
<tr><td colspan="2" rowspan="2"></td><td>Expenditure Projections by Program</td></tr>
<tr><td>Results Projections</td></tr>
</table>

Source: Filc and Scartascini (2011)

For this study, we focus on multiyear budgeting or the medium term fiscal framework. Clearly, given the interest in fiscal performance, multiyear budgeting is the aspect of MTEF that is most crucial. In addition, medium term fiscal framework remains the most operationalized aspect of MTEF, especially among developing countries amongst which 67 percent are still at the first phase of MTEF (World Bank, 2013). This trend, coupled with concerns regarding the effectiveness of multiyear budgeting further accentuates the focus of this study.

4.3 Literature Review

Over the years, there has been no convergence of views, even as results of empirical studies on the effect of multiyear budgeting on fiscal performance remain largely inconclusive. For example, Von Hagen (1992), using non-parametric regression analysis, finds that multiyear budgeting has no significant effect on debt level among European countries. Volkerink and De Haan (2001) note that Von Hagen sample is limited to the 1980s, the onset of MTEF in most countries. Furthermore, Volkerink and De Haan (*ibid*) observed that budgeting processes have changed significantly with the introduction of Maastricht Treaty in 1992 that emphasized multiyear budgeting as part of fiscal reform in the Euro zone. Hence in their 2001 study, Volkerink and De Haan re-examined the effect of multi-year term budgeting on fiscal performanc but found only a modest impact in post-Maastricht period; a unit increase in the index measuring extent of multiyear budgeting adoption only reduced primary deficits by 0.001 percent.

A number of earlier studies have also observed weak influence of multiyear budgeting on fiscal performance in developing countries. For instance, Le Houerou and Taliercio (2002) in a pre and post-MTEF comparison of the four most developed MTEF countries in Africa noted that primary deficit did not show significant improvement in the post-MTEF period for data that spanned 1985 to 2000. For Latin America, Filc and Scartascini (2011) found mixed results; some countries witnessed improvement in fiscal outcome, while stagnation in performance were observed in others.

However, Bevan and Palomba (2000), observed that in Uganda, one of the very first countries to adopt the MTEF in Africa, introduction of multiyear budgeting resulted in significant positive effect on fiscal performance. They also highlighted improved macroeconomic stability and increased allocative efficiency of the budget in Uganda. Vlaicu et al. (2014) undertook what appears to be the most extensive study of this nature in the literature; covering 120 countries over the period 1990 and 2008. Their conclusions were similar to those of Bevan and Palomba (2000); multiyear budgeting has a significant and positive impact on fiscal performance as well as on allocative efficiency. The World Bank, in a 2013 study, reported the same results as those of Bevan and Palomba (2000) and Vlaicu et al. (2014).

From the foregoing, the seemingly divergence of views and contrasting findings are attributed in the literature to several factors such as: initial conditions of the public expenditure management and reform sequencing (Holmes & Evans, 2003); disparity in technical and institutional capacities across countries (Schiavo-Campo, 2009); types of government (Hallerberg & von Hagen, 1999); among others. Another plausible factor, largely ignored in the literature until recent times, is the spending preference of budget actors. According to the voting equilibrium theory developed by Ferejohn and Krehbiel (1987), a rational and forward looking legislator will always vote to maximize his/her spending preference. Therefore, fiscal outcome hinges on the ideal point of the legislator's spending preference rather than on the budget process. The implication of this theory has been examined for bottom-up and top-down budget processes by Ferejohn and Krehbiel (1987) and Ehrhart et al. (2007). Using various experimental techniques[18], the studies validate the prediction of voting equilibrium theory that budgeting process has an ambiguous effect on fiscal performance.

While the voting equilibrium theory is yet to be widely tested in relation to multiyear budgeting process, a glean through the existing literature suggests it deserves more recognition and validation; this constitutes the motivation for this chapter. Von Hagen (2002) and Von Hagen and Wolff (2006), the few studies that are readily available in this regard find that governments, when subjected to fiscal rule tend to develop ways to circumvent the restriction. When fiscal rules are for instance imposed on recurrent budget, a shift to capital budget is induced, leaving the total expenditure the same. Alternatively, the preferred expenditure level can be attained by basing the budget estimate on overly optimistic revenue projections. This evidence strongly suggests that spending preference of the budget actors is crucial to the effectiveness of multiyear budgeting. Given the obvious gap in the literature, we explore in what follows, the validity or otherwise of voting equilibrium theory and the implication on the relationship between multiyear budgeting and fiscal performance using laboratory experiment.

4.4 Theoretical Model

4.4.1 A Voting Equilibrium Model of Budgeting

The voting equilibrium model (VEM) adopted for this study derives in large part from Ferejohn and Krehbiel (1987). The VEM was proposed in an attempt to explain the persistence of budget deficit in the US, despite the change from a bottom-up to top-down budget process in the mid-1970s. Under the bottom-up budget process, the legislators first determine the appropriation across spending units, which are summed up to derive the aggregated budget. In contrast, the top-down budget process requires the legislators to initially set a binding budget aggregate, which are subsequently divided among the spending units. The main premise is that top-down budget process will result in smaller budget size, as it integrates sustainability concern into budgetary decision

[18] Ferejohn and Krehbiel (1987) explored the natural experimental condition generated with the promulgation of the 1974 Budget and Impoundment Control Act (PL 93-344) in the US, while Ehrhart et al. (2007) used laboratory experiment to examine the implication of voting equilibrium model for top-down vis-à-vis bottom-up budgeting processes.

(Ljungman, 2009). However, Ferejohn and Krehbiel (1987) show that if the budget actors are rational and forward looking, budget size does not necessarily depends on budget process but rather on the spending preference of budget actors.

We extend this basic model to the case of multiyear vis-à-vis annual budget processes. Similar to the top-down budget process, multiyear budgeting entails initially setting a binding budget size for a defined planning-horizon, out of which the yearly estimates are derived. Conversely, the annual budgeting mirrors the bottom-up budget process, with estimate for each year determined separately. Thus, with a change in nomenclature from spending units to planning horizon, the underlining logic remains the same in both top-down/bottom-up and multiyear/annual budget processes.

4.4.2 Model Structure

The budget institution is made up of n (odd) legislators indexed $i = 3,5...,n$. The key assumption of the model is that legislators are sophisticated. This means the legislators are rational and forward-looking, such that when faced with alternative budget proposals, they consider each other's preferences and the consequences of their present choice on future choices (Yuval & Herne, 2005). For simplicity, we assume that revenue and allocation across spending units are exogenously determined. Hence, the legislators only decide, through voting, the aggregate budget size, B_t, and its division over a planning-horizon, T. In this instance, any budgetary decision can be represented as a point on the non-negative orthant of the T-dimensional Euclidean space. Taking $T = 2$ in the interim and using a Cartesian plane, the x and y axes represent appropriations in periods t and $t+1$ respectively, and any point on the plane implicitly define the aggregate budget size over the planning horizon.

Furthermore, each legislator is assumed to have an 'ideal budget', B_{it}^{*}, which maximizes his/her utility. Essentially, a legislator's ideal budget corresponds to the appropriation level that minimizes the net tax to his/her constituency (see Ehrhart et al., 2007). Therefore, the closer the realized budget is to the ideal budget, the higher the utility level. Formally, this implies that, for a given budget outcome, legislator i utility, $u_i(B)$, can be described by the Euclidean distance function as follows:

$$u_i(B_t) = V_i(B_{it}^*) - \sqrt{\sum_{t=1}^{T}[B_t - B_{it}^*]^2} \quad (1)$$

where $V_i(B_{it}^*)$ is the utility derived from the ideal budget of legislator i at time t. The Euclidean preference is particularly useful in describing allocation across periods. Specifically, the resulting voting equilibrium is order-independent and depends only on the sophistication of the outcome (see proof in Kramer, 1972)[19]. Thus, Euclidean preference allows us to model and test an institutional setting in which the budget actors are assumed to be sophisticated. Finally, the legislators use the majority voting rule in deciding among alternative budget proposals.

Based on this setting, it is well-established in the literature that a unique voting equilibrium cannot be guaranteed—a consequence of Condorcet's paradox (Roberts, 2007). However, by restricting budgetary decision to a single dimension, a unique voting equilibrium can be established. For this study, we examine two of such restrictions on budgetary decision: annual and multiyear budgeting procedures.

4.4.3 Budgetary Outcomes under Annual versus Multiyear Budgeting

We begin by considering a legislature that consists of five members who are to decide the aggregate budget size over two periods. In the case of annual budgeting, members make a proposal first for period t and subsequently for period $t+1$. Expectedly, each legislator will propose his/her ideal budget and the voting equilibrium will be determined by the ideal budget of the median voter. Figure 4.2 shows one of the possible configurations of legislators' ideal budgets over the two periods. In period t, the median voter is the third member and the budgetary outcome is depicted by *line a*. Following the same process for period $t+1$, the ideal budget of the first member wins over other proposals, as represented by *line b*. The voting equilibrium under the annual budgeting, denoted by B^a, is obtained at the point of intersection of *lines a* and *b*.

For comparability purpose, Figure 4.2 is again used to illustrate the budgetary outcome in the case of multiyear budgeting. First, a vote is taken to set a binding aggregate budget size over the planning horizon. Given the preferred mix of allocation across periods, members propose the sum of their ideal budgets for periods t and $t+1$ as the aggregate budget size. The median voter is the

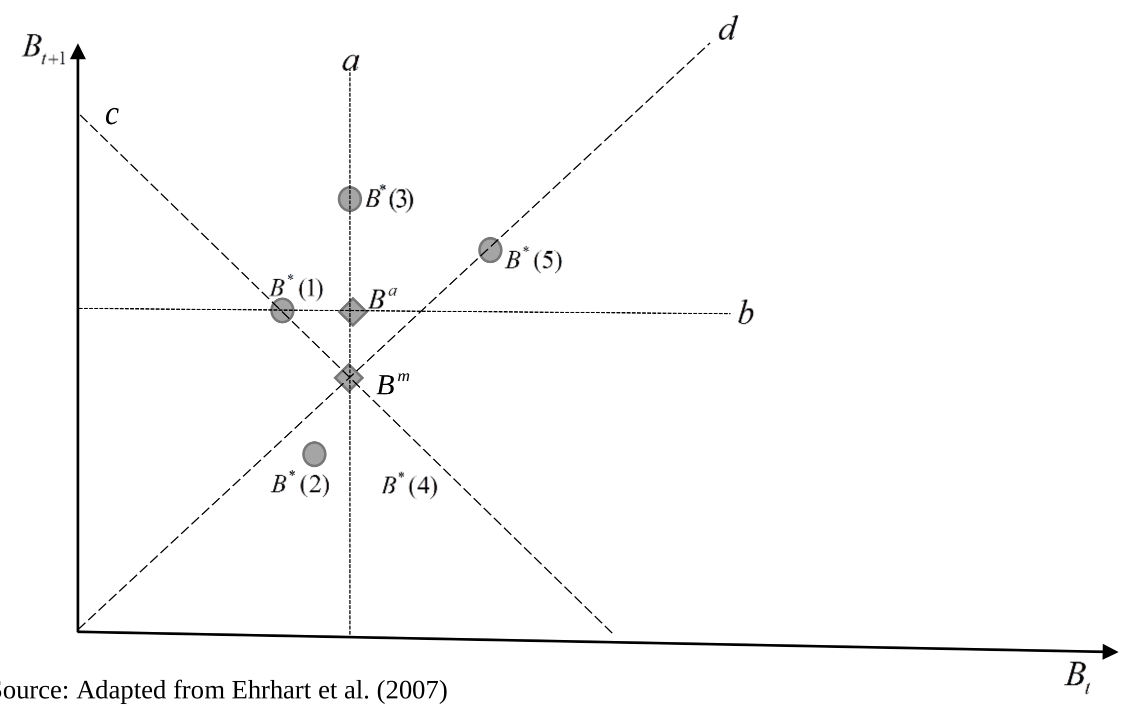
[19] For an extensive proof of this theory, see Kramer (1972: pages 165 to 180)

first member and the equilibrium aggregate budget size derived is represented by *line c*. In the second stage, members determine the allocation mix across periods, based on the pre-determined budget cap. The median voter preferred mix is depicted by the *line d*. The voting equilibrium under the multiyear budgeting is given by B^m, at the point of intersection of *lines c* and *d*.

With the configuration of the ideal budgets in Figure 4.2, the multiyear budgeting leads to a smaller budget size compared to annual budgeting. However, when the configuration of ideal budgets is altered as depicted in Figure 4.3, a diametrically opposite conclusion is reached. Specifically, the configuration of ideal budgets yields a voting equilibrium in which annual budgeting produces a smaller budget size. The implication of these results is that multiyear budgeting does not necessarily produce a lower budget size, rather the spending preference of legislators moderate the final budgetary outcome.

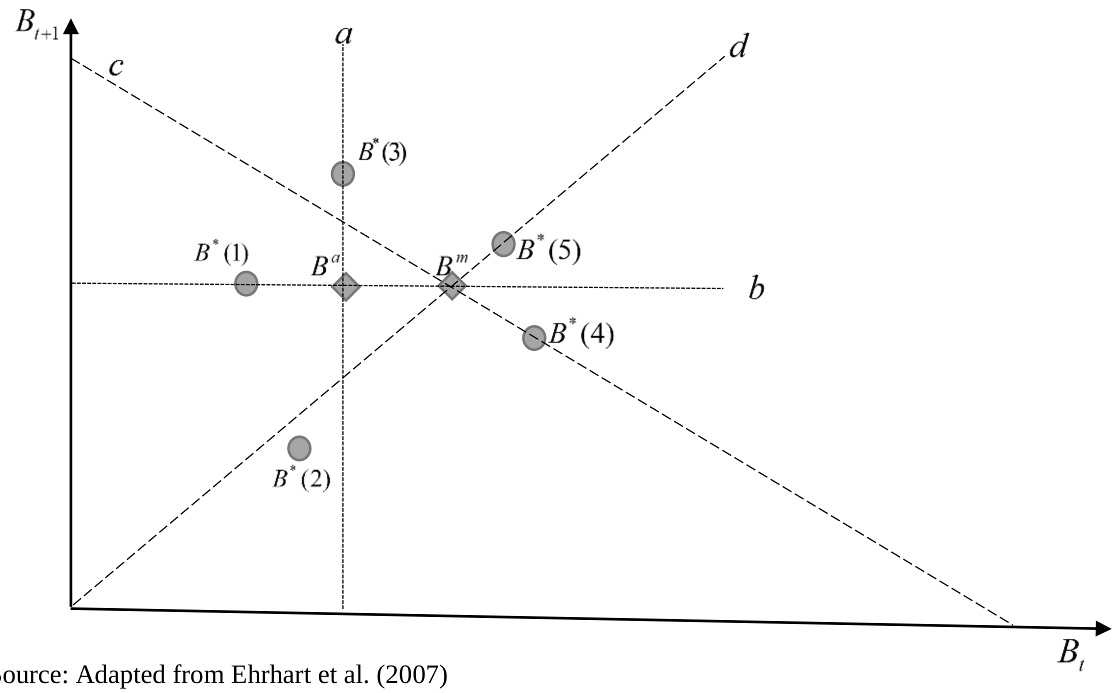
Proposition 1 *Multiyear budgeting has an ambiguous effect on fiscal performance.*

Figure 4.2: Voting Equilibrium Outcome with $B^m < B^a$



Source: Adapted from Ehrhart et al. (2007)

Figure 4.3: Voting Equilibrium Outcome with $B^m > B^a$



Source: Adapted from Ehrhart et al. (2007)

4.5 Experimental Procedure

This section details the experimental procedure for testing the proposition derived above. It begins by describing the laboratory setting in line with basic model developed above. The voting equilibrium outcomes based on the specified parameters of the model are then derived. These outcomes provide a benchmark to compare the experimental findings from this study.

4.5.1 Decision Setting

The decision setting draws from previous experiments on budget processes by Ehrhart et al. (2007) and classical work of Plott and Krehbiel (1979). The participants are to decide the aggregate budget size (represented by the experimental points) and its division over a defined planning-horizon using either annual or multiyear budgeting process. Participants have dispersed preferences, which is common knowledge within the group, and is given by the configuration of the ideal budget. Based on the voting outcome, participants receive payoff determined by the extent of deviation between the actual and ideal budgets as defined in equation (1). At the end of the experiment, the cumulative payoff is exchanged for cash at a predetermined conversion rate.

In the setting modelled along annual budgeting, the parliamentary procedure starts with a baseline proposal which contains zero allocation for the first period. Any member can propose an amendment to the baseline proposal and the legislature vote on the two proposals. The proposal with most votes becomes the status quo. Another member can either propose a further amendment or move a motion to accept the status quo as final. The first period appropriation is completed when a motion for acceptance of status quo is moved and it receives the majority vote. The same procedure is followed in subsequent periods until the endpoint, T . The aggregate budget size is derived by summing up appropriation over the planning-horizon.

The same parliamentary procedure is employed in the setting that adopts multiyear budgeting. However, the initial decision is to agree on a binding aggregate budget for the planning-horizon. Thereafter, the legislature decides on the allocation for the first period up till period $T-1$. Finally, appropriation for the last period can be trivially estimated based on past decisions.

4.5.2 Design Parameters and Treatments

Our experimental technique utilizes a 2×2 factorial design defined by dimension (length) of the planning-horizon (T) and the configuration of ideal budget. Table 4.1 details the design parameters adopted in this study. The legislative size is fixed at $n=5$ across designs (legislatures are labelled A to E). For the dimension of the planning-horizon, we implement two treatments; $T=2$ (two-dimensional) and $T=4$ (four-dimensional). As Le Houerou and Taliercio (2002) noted, a longer planning-horizon could undermine the credibility of multiyear budgeting by complicating the decision space. Thus, our design allows for an explicit test of this conjecture. Also, two contrasting configuration of ideal budgets are considered, depending on whether annual or multiyear budgeting leads to smaller aggregate budget. The two configurations are differentiated by the fourth legislator ideal budget.

Table 4.1: Configuration of ideal budget and utility function

Legislature	Two-dimensional $(T=2)$				Four-dimensional $(T=4)$							
	Design I		Design II		Design III				Design IV			
	$B_1^*(i)$	$B_2^*(i)$	$B_1^*(i)$	$B_2^*(i)$	$B_1^*(i)$	$B_2^*(i)$	$B_3^*(i)$	$B_4^*(i)$	$B_1^*(i)$	$B_2^*(i)$	$B_3^*(i)$	$B_4^*(i)$
A	7	14	7	14	7	14	15	22	7	14	15	22
B	8	10	8	10	8	10	16	18	8	10	16	18
C	9	17	9	17	9	17	17	25	9	17	17	25

D	12	13	10	10	12	13	20	21	10	10	18	18
E	12	15	12	15	12	15	20	23	12	15	20	23
	Two-dimensional				Four-dimensional							
Utility function, $u_i(B)$, of voter i	$16-\sqrt{\sum_{t=1}^{2}[B_t - B_{it}^*]^2}$ ……………………………….(2)				$40-\sqrt{\sum_{t=1}^{4}[B_t - B_{it}^*]^2}$ ………..(3)							

Source: Author's computation from experimental design.

Table 4.2 summarizes the theoretical prediction derived from the voting equilibria model. In designs I and III, which consider the case of $T = 2$ and $T = 4$ respectively, the resulting voting equilibria predict that annual budgeting leads to smaller aggregate budget than multiyear budgeting. On the other hand, design II and IV show a voting equilibria outcome in which multiyear budgeting produces a smaller aggregate budget. Essentially, the theoretical prediction based on the VEM indicates that budgetary outcome varies under different budgeting processes. More importantly, the predictions confirm earlier derived proposition that configuration of the legislators' ideal budgets is a key determinant of fiscal performance.

Table 4.2: Theoretical prediction of the VEM

Process	Two-dimensional				Four-dimensional							
	Design I		Design II		Design III				Design IV			
	$B_1^*(i)$	$B_2^*(i)$	$B_1^*(i)$	$B_2^*(i)$	$B_1^*(i)$	$B_2^*(i)$	$B_3^*(i)$	$B_4^*(i)$	$B_1^*(i)$	$B_2^*(i)$	$B_3^*(i)$	$B_4^*(i)$
Annual budgeting	9	14	9	14	9	14	17	22	9	14	17	22
Aggregate	**23**		**23**		**62**				**62**			
Multiyear budgeting	9	16	9	12	9	16	17	24	9	12	17	20
Aggregate	**25**		**21**		**66**				**58**			

Source: Author's computation from experimental design.

4.6 Sampling Procedure

4.6.1 Optimal Sample Selection

According to List et al. (2011), a factorial design with $m \times k$ treatments requires at least $2mk$ trials. This implies that each treatment must be replicated at least once over the experiment. However, if there are large treatments and in the absence of any interaction effect among them, List et al. (2011) suggests $mk+1$ trials as sufficient. In line with the literature and given that there are only 4 treatments in this study, we conduct $2mk$ (8) trials. This means each treatment is replicated twice. Since the legislative size in each treatment is 5, the total sample required for the study is 40 participants.

4.6.2 Sampling Technique

Participants for the study are selected from first and second year undergraduate students of University of the Witwatersrand, South Africa. Additional selection criteria includes (i) participant must have attained 18 years of age; (ii) participant must be computer literate.

Emails were sent to the targeted students that meet our inclusion criteria, to which an initial 446 students volunteered. The simple random sampling technique is thereafter applied in selecting the 40 participants required for the experiment. Table 4.3 shows the distribution of the participants along key demographic characteristics: sex, year of study, race and faculty of study. The result shows that each demographic group is well-represented in the final sample, except for the distribution by racial group, which is skewed towards black participants. This is as a result of low turn-out of volunteers from the other races. However, the sample distribution is relatively proportional to size of each race in the total student population. The proportionality of the sampling distribution has two advantages that reduce sampling error. First, the standard deviation is lower, as actual student composition is better reflected. Second, it yields improved statistical precision at lower sample size. Essentially, this indicates that sample selection bias is less of a concern for this study.

Table 4.3: Demographic Characteristics of the Participants

Variable	**Frequency**	**Percentage (%)**
Sex		
Male	22	55.0
Female	18	45.0
Year of Study		
1^{st}	17	43.3
2^{nd}	23	56.7
Race		

Blacks	36	90.0
White	4	10.0
Faculty		
Commerce, Law and Management	15	36.7
Health Sciences	7	16.7
Humanities	8	20.0
Science	11	26.7

Source: Author's computation based on questionnaire response (see Appendix 4B)

4.6.3 Experimental Implementation

The experiment reported in this paper took place at the PhD/AERC Computer Laboratory at the University of Witwatersrand in the month of October, 2016. Four sessions were conducted in total: two sessions each for annual and multi-year budgeting respectively. In each session, the participants make budget decision using designs I and II or designs III and IV. This implies that a between-subject approach is used, because subjects were exposed to more than one designs. In this regard, there is possibility of history and learning effects on the final outcome. To address these problems, a cross-over design is used. The cross-over design involves varying the order of treatments across the sessions. For example, half of the groups were first exposed to design I before design II, while the other half were exposed to the later before former. A similar strategy was applied for sessions based on designs III and IV. Overall, participants average earning is approximately R65 and no session lasted more than 55 minutes.

Following the tradition of Plott and Krehbiel (1979), we run a semi-computerized experiment. Specifically, the participants made their decision manually using the decision card provided. The votes are counted in the backroom by laboratory assistants and final tallies displayed on a computer monitor visible to all participants. The decision card used is shown in Figure 4.4. Each participant is provided with a hundred decision cards, and additional cards are provided if the initial once are exhausted in the course of the experiment. Essentially, the decision card allows the participants to vote on a budget motion as well as make an alternative proposal.

Furthermore, the decision space is simplified by converting the utility function in Table 4.1 into a payoff table. Each design has a separate payoff table, which are generated based on the ideal budgets and by extending the possible proposals to some reasonable range. Given that budget size that resulted into negative or very low utility will reasonably not be proposed, we therefore vary the budget sizes between 0 to 30 for design I and II, and 0 to 40 for designs III and IV. The payoff tables are presented in appendix 4C. The instruction set used for each session is also presented in the appendix. The instruction details the decision setting, the group size, budgeting process to be adopted, the dimension of the planning-horizon, and how the majority voting system works. After going over the instruction, participants had a hands-on training/pilot session before proceeding to the main experimental session.

Figure 4.4: Decision Card

Decision Card

Please tick (√) appropriate box that matches your decision.

Should the motion on the floor be implemented?

- ❖ Yes:
- ❖ No:

If you have choose "No", please propose an alternative budget size in the box below:

New Proposal:

4.7 Empirical Results

This section details the analysis of results from the experiment. The analysis is conducted in three parts. First, we compare the fiscal outcomes under the annual and multiyear budgeting processes. Second, we analyse the extent to which observed decisions reflect the sophistication of voters. The last part explores the efficacy of voting equilibrium model (VEM) as a good predictor of the experimental observations.

4.7.1 Comparative analysis of annual and multi-year budgeting processes

The voting equilibrium prediction regarding budget size under alternative budgeting processes is examined in Table 4.4. We essentially compare the aggregate budget size as predicted with those observed based on the experimental data. As shown in Table 4.4, the aggregate budget sizes differ between the annual and multiyear budgeting processes. This implies that budget processes matter and they significantly influence the fiscal outcomes. More importantly, the voting equilibrium prediction on the effect of voters' preference on budget outcome is mostly confirmed. Specifically, multiyear budgeting produces lower average budget size than annual budgeting in designs II (18 versus 22) and IV (57 versus 62), while, in design III, annual budgeting yields a lower budget size (62 versus 66). The only exception is in design I, in which the multiyear budgeting also leads to lower budget size (16 versus 10), contrary to voting equilibrium prediction.

A further disaggregation of the result over the dimension of planning-horizon shows that while the magnitude does exactly match, the trends follow the pattern suggested by the voting equilibrium. This can be more clearly seen by restacking the data in Table 4.4 such that all the observed budget sizes are pooled together as one variable and all the predicted budget sizes as another variable. Thereafter a pairwise correlation is performed on the two variables. The estimated correlation coefficient is 0.96 and it is significant at 5 percent. This high and positive correlation estimate implies that the trends between the observed and predicted budget sizes are closely matched.

Table 4.4: Average Budget Size by Dimension and Budgeting Process

Budgeting Process	**Dimension**	**Design I**		**Design II**		**Design III**		**Design IV**	
		Observed	**Predicted**	**Observed**	**Predicted**	**Observed**	**Predicted**	**Observed**	**Predicted**
Annual	1	4	9	12	9	10	9	10	9
	2	12	14	10	14	12	14	14	14
	3					15	17	19	17
	4					20	22	23	22
	Aggregate	**16**	**23**	**22**	**23**	**57**	**62**	**66**	**62**
Multiyear	1	2	9	8	9	10	16	8	9
	2	8	16	10	12	12	17	14	12
	3					18	24	20	17
	4					22	9	20	20
	Aggregate	**10**	**25**	**18**	**21**	**62**	**66**	**62**	**58**

Source: Author's computation based on the experimental data

4.7.2 Mann-Whitney Test

To determine the statistical significance of the result above, we apply the Mann-Whitney test. Mann-Whitney test is appropriate for assessing whether two independent samples, such as alternative budgeting processes—annual versus multiyear budgeting—come from the same distribution (have the same median). Previous studies that have examined the implication of the VEM also adopt Mann-Whitney test (see Ehrhart et al., 2007; Gardner & von Hagen, 2012). The key advantage of the Mann-Whitney test is that it is distribution free. This means it remains valid for data that are not normally distributed. It is equally useful for semi-qualitative data or relatively small dataset (Nachar, 2008). Also, the test is robust even in the presence of outliers or when the sample size differs in the two groups being compared.

Formally, the Mann-Whitney test is defined as:

$$Z = \frac{U - N_1 N_2 / 2}{\sqrt{N_1 N_2 (N_1 + N_2 + 1)/12}} \qquad (4)$$

where N_1 and N_2 are the number of observations in the first and second samples respectively, and U is smallest of the Mann-Whitney distributions between the two samples. Basically, the Mann-Whitney test combines the two samples and thereafter assigns ranks to all the observations. The sum of ranks for each sample is compared to establish if there is significant difference. The logic is that if there is any systematic difference between the two sums of rank, this will be reflected by the ranking being disproportionately skewed towards the sample with higher values. Otherwise, lack of systematic difference in the sums of rank implies the samples are drawn from the same distribution.

Taking the annual and multiyear budgeting as the first and second samples respectively, we apply Mann-Whitney test as defined in equation (4). The result is reported in Table 4.5 for each design. Evidently, the *p*-values are well above the statistical threshold of 1%, 5% and 10%. This implies that the budget size produced by both annual and multi-year budgeting processes are not statistically different across the designs.

Table 4.5: Mann-Whitney Test

	Design I	Design II	Design III	Design IV
Mann-Whitney *U*	29	31	99	111
χ^2	2.619	1.725	0.21	0.093
p-value	0.106	0.189	0.647	0.76

Source: Author's computation based on the experimental data

Observation 1: *Multi-year budgeting do not necessarily lead to smaller budget size than annual budgeting, rather voting preference moderate the eventual budgetary outcomes.*

4.7.3 Budgeting processes and delay in decision making

The effectiveness of a budgeting process is not only measured by its impact on fiscal outcome, but also in terms of extent of delay it causes in decision making (Palfrey, 2013). Indisputably, a budget process with long delay will be less preferred to those with shorter delays. Therefore, we compare the extent of delay between the annual and multiyear budgeting processes. For this purpose, the extent of delay is measured by the number of moves it takes to arrive at the final budget decision. Also, for ease of comparison across treatments, the data reported is restacked such that designs I and II are pooled together and accordingly labelled 'two-dimension'. Designs III and IV are similarly combined and labelled, 'four-dimension'. The new classifications of data are referred to henceforth, except otherwise stated.

The result, presented in Table 4.6, shows that under the two-dimension, multiyear budgeting takes slightly more moves than annual budgeting; the opposite conclusion is reached in the case of four-dimension. In addition, the differences are not statistically significant based on the chi-squared test (χ^2 =0.12 and *p*-value=0.34). Similarly, reaching the final budget decision requires about twice more moves in the four-dimension than in the two-dimension, regardless of the budgeting process. However, it must be noted that procedural-wise, the number of budgeting decision involved under the four-dimension is twice that of the two-dimension. Therefore, we can conclude that the number of moves to reach the final budget decision is the same across dimensions. A related experiment by Ehrhart et al. (2007) reached a similar conclusion regarding the absence of differences in decision delay among alternative budgeting processes.

Observation 2: *The number of moves required to make final budget decision is the same across budgeting processes and dimensions.*

Table 4.6: Average number of moves to reach the budget decision

Budgeting Process	Two-dimension	Four-dimension
Annual	5	9
Multiyear	4	9.5
Average	**4.5**	**9.25**

Source: Author's computation based on the experimental data

4.7.4 Sophistication of budget actors

Next we examine the extent to which individuals' decisions and voting are sophisticated. The assumption of budget actors being sophisticated is central to the VEM and therefore testing for sophistication in decision making is crucial. Table 4.7 presents the result on the frequency at which the participants' proposals match or are close to the ideal budget. Following Ehrhart et al. (2007) approach, a proposal is considered as close if it is within two points neighbourhood of the ideal budget. The rationale is that since utility is maximized when ideal budget is realized, a sophisticated—rational and forwarding—budget actor will often time propose his/her ideal budget or budget size close to it. The result shows that, for two-dimension, the ideal budget or budget size close to it is proposed at least 54 percent of the times under annual budgeting, but the frequency reduces to 52.5 percent for multiyear-term budgeting. In the case of four-dimension, the frequency reduces to 47.8 percent under annual budgeting and 42.1 percent under multiyear term budgeting.

Table 4.7: Percentage of proposal that are or close to ideal budget

Budgeting Process	Two-dimension	Four-dimension	Average
Annual	54	47.8	**50.9**
Multiyear	52.5	42.1	**42.1**
Average	**53.25**	**44.95**	**49.1**

Source: Author's computation based on the experimental data

Another means of assessing the extent of voting sophistication is by examining the direction of transition in proposals made. A sophisticated voter more frequently adjust the preceding motion towards his/her ideal budget or stay put if it is already the motion on the floor. Thus, we divide transition of proposals into three categories: those approaching the ideal budget, those that stay put on the ideal budget and those drifting away from the ideal budget. The result, as presented in Table 4.8, shows that majority of the proposal are either moving towards the ideal budget or equal to it. Overall, evidence based on Tables 4.7 and 4.8 indicates that, on average, quality proposals were being made by the participants, which suggest sophisticated voting.

Table 4.8: Transition in proposals

Budgeting Process	Dimension	Approaching Ideal Budget	Equal to Ideal Budget	Drifting from Ideal Budget
Annual	Two	74	20	6
Multiyear	Two	50	30	20
Annual	Four	67.37	27.37	5.26
Multiyear	Four	60	27.78	12.22

Source: Author’s computation based on the experimental data

Observation 3: *Participants act sophisticatedly—rational and forward-looking—in making their proposals.*

4.7.5 Efficacy of VEM

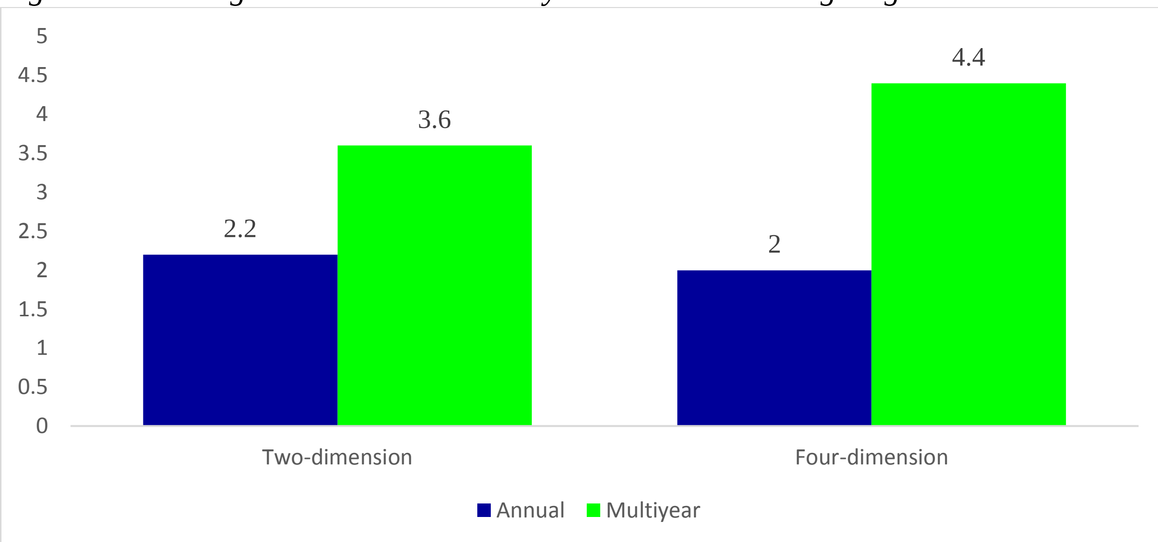
Finally, we examine the performance of the voting equilibrium model vis-à-vis the generated experimental data. Two measures of performance are used: Euclidean distance and Selten’s predictive success. Euclidean distance measures the relative gap between observed and predicted budget size, taking into account the peculiarities of the circular utility function. The Euclidean distance formula is given by:

$$\text{Euclidean distance formula} = \sqrt{\sum_{i}^{n}(B_i - B_p)^2} \qquad (5)$$

where B_i is the observed budget size and B_p is the predicted budget size. Both B_i and B_p are already shown in Table 4.4. Thus, the estimated Euclidean distance is presented in Figure 4.5. The result shows that on average, the Euclidean distance is lower for annual budgeting and two-dimension than for multiyear budgeting and four-dimension. This implies that the observed budget sizes are much closer to the predicted budget sizes under the annual budgeting and two-dimension.

This result could in part be attributed to the obvious fact that the decision space is more complex in the case of multiyear budgeting and over four-dimension. Nevertheless, the estimated distance of 3.6 for multiyear budgeting and 4.4 for four-dimension are still lower compared to the predicted total sum of between 21 and 25 for multiyear budgeting and 58 and 66 for four-dimensional treatment. Evidently, the VEM on average is a good predictor of budget size.

Figure 4.5: Average Euclidean Distance by Dimension and Budgeting Process



Source: Author's computation based on the experimental data

An alternative measure of performance of VEM is given by the Selten's predictive success measure (SPSM). The SPSM is a benchmark instrument for comparison of area theories for characteristic function experiments (Selten, 1991). The VEM is an example of area theory—defined as a theory whose prediction covers a subset of all possible outcomes. Formally, SPSM is given by:

$$SPSM = r - a \quad (6)$$

where r is the hit rate and is measured as the frequency of outcomes that are close to the voting equilibrium; a is the area rate and is defined as the relative size of the predicted subset compared with the set of all possible outcomes. Following Ehrhart et al. (2007), the hit rate is measured by the percentage of budget proposals that are within 2 points neighbourhood of the ideal budget. In essence, the information provided in Table 4.7 already captures the hit rate. For the area rate, the predicted subset is defined as all final budgets that exactly match the ideal budget, while the set of all possible outcomes is captured by all budget outcomes that are located within the minimum and maximum values of the subjects' ideal budgets. Base on this definition, the area rate is estimated and result is shown in Table 4.9.

Table 4.9: Area Rate

Budgeting Process	Two-dimensional	Four-dimensional	Average
Annual	25.00	20.83	22.92
Multiyear	16.67	12.00	14.33
Average	20.83	16.42	18.63

Source: Author's computation based on the experimental data

Lastly, Table 4.10 presents the estimated SPSM, based on equation (6). The predictive success is higher under the multiyear budgeting than annual budgeting and in the two dimension than four dimension. On average, the estimated predictive success is higher relative to the level reported in the literature. For example, Keser and Gardner (1999) found that predictive success for experiments on Nash equilibrium is around 5 percent. Thus, we can again conclude on the basis of predictive success measure that the VEM performs reasonably well in explaining the experimental observations.

Table 4.10: Selten's Predictive Success Measure

Budgeting Process	Two-dimension	Four-dimension	Average
Annual	29.00	26.97	27.98
Multiyear	35.83	30.10	27.77
Average	32.42	28.53	30.48

Source: Author's computation based on the experimental data

Observation 4: *Voting equilibrium model (VEM) performs well in predicting the magnitude of budget size across treatments.*

4.7.6 Sensitivity to Noisy Decision making

The experimental data generated is only valid if the participants' behaviours are not influenced by noisy decision setting. Noisy decisions mostly arise through confusion regarding the instruction set or other relevant materials used for the experiment. Therefore, it is crucial to determine extent to which the experimental results is impacted by noisy decision making. To achieve this, questionnaires were administered on the participants immediately after the experiment. The questionnaire used is shown in appendix 4B.

The results based on participants' responses are presented in Table 4.11. Overall, the results indicate that participants mostly understood the instruction set in both practice and the main experimental sessions. Moreover, none of the participants indicate any difficulties in reading the payoff table or in making use of the decision card. In fact, no experimental session is allowed to proceed without participants demonstrating adequate understanding of use of the payoff table and the decision card.

Table 4.11: An Assessment of Noise in Decision Making

Scale	**Practice Session Instruction**	**Main Experimental Instruction**	**Reading of the payoff table**	**Use of the decision card**
Fully Understand	30 (75%)	34 (85%)	31 (77.5%)	34 (85%)
Mostly Understand	7 (17.5%)	5 (12.5%)	9 (22.5%)	6 (15%)
Didn't Understand	3 (7.5)	1(2.5%)	0 (0%)	0 (0%)
Total	**40 (100%)**	**40 (100%)**	**40 (100%)**	**40 (100%)**

Source: Author's computation based on the questionnaire response

4.8 Conclusion

In the past two decades, governments around the globe have been implementing various forms of fiscal reforms in response to the fiscal problems arising from common pool problem. Multiyear budgeting is no doubt a key component of this reform effort. However, empirical evidence remains inconclusive despite a priori expectation of a positive impact of multi-year budgeting on fiscal performance. In this paper, we explore a possible confounding factor that could account for this inconclusive evidence as suggested by the voting equilibrium model. Specifically, we investigate the effect of voting preference on the relationship between multi-year budgeting and fiscal performance. The theoretical discussion was motivated by the classical work of Ferejohn and Krehbiel (1987) on voting equilibrium model. However, in line with focus of the present study, we extend the generic model to an intertemporal allocation problem. On the empirical side, experimental approach is adopted, in which multiyear budgeting is used as the baseline treatment and fiscal outcome under it is compared with annual budgeting. Also, dimension of planning horizon is varied between two and four-dimensions.

The key findings of the study are in three-fold. First, the fiscal outcome is not significantly different between multi-year and annual budgeting processes. For example, we observed larger aggregate budget size under the multi-year budgeting than the annual budgeting when the voters' preferences cluster around higher budget levels and vice versa. In addition, there is no difference in the number of moves it takes to make budget decision under the alternative budgeting processes. This implies that in considering alternative budget processes the issue of decision delay might be taken as trivial. Secondly, the quality of proposals being made suggests that participants are sophisticated in their decision making. Lastly, we found that performance of the VEM is higher in the two dimensional treatment than in the four dimension. This is attributed to complexity in higher dimension. Thus, increasing the dimension of planning horizon has the tendency to increase uncertainty in the budgeting decision.

In the light of these findings, a number of relevant policy implications emerge. In designing appropriate fiscal reform, consideration of preference of budget actor is crucial. In essence, introducing multiyear budgeting alone does not guarantee good fiscal performance, especially if it has a negligible effect on voting preference. Therefore, understanding the preference of budget actors and taking it into account is fundamental. Also, given the possible uncertainty that could arise in higher dimension of planning horizon, needlessly long planning horizon should be avoided in future fiscal reform. However, what constitute the optimal dimension of planning horizon is shaped by budget institution specific and economic-wide idiosyncrasies.

In concluding this chapter, it is important to emphasize the main point of this paper. The study has merely examined possible confounding factor in the relationship between multiyear budgeting and fiscal performance. This does not imply that multiyear budgeting is irrelevant for budget institutions or fiscal performance more broadly; rather the main point is that neglecting the voting preference of budget actors could limit multi-year budgeting effectiveness as fiscal reform tool. In fact, by extension, this consideration could also hold for other fiscal reform measures.

4A Appendix: Experimental Instructions

Instruction 1: annual budgeting and two-dimension

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, raise your hand and the laboratory assistant will attend to you accordingly. You will now take part in a decision making experiment.

What material do you have for the experiment? You have four types of handouts. (1) First you have a copy of this instruction which you can look at any time during the experiment. (2) The second item you have is a set of payoff table, which shows each participant's payoff for any given budget size. Note that there is a separate payoff table for each period. (3) The third item is a set of decision cards. The decision card will be used to make your proposals and also vote on alternative proposals. (4) The fourth item you have been given is the personal history form. The personal history form is like a scratch sheet that you can use for taking note of proceedings during the experiment. Please note that the use of the personal history form is optional but it must be turned in at the end of the experiment. The label on your personal history form indicates your subject ID. *All handouts must be turned in at the end of the experiment.*

What determines how much earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding the aggregate budget size over two periods. The amount you earn depends on the final budget size jointly voted for by all members of the group. In the Payoff Table, you will see the payoff associated with all possible budget sizes. At the end of the experiment, you will be able to exchange this payoff for Rands. The exchange rate is:

2.19 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about the subject ID of other members of their group.

How does the experiment work? You have been assigned to a group consisting of five (5) members and each of you have been allocated an ID (A to E), as indicated by the subject ID on your personal history form. The group's task is to decide the aggregate budget size to implement over two periods. The group will first decide on the final budget size for period 1 and thereafter decide the final budget for the period 2. This marks the end of the game.

How will members decide on the final budget in each period? The group decides the final budget size for any period following the procedure below:

- *Step 1*: Every period starts with an initial motion on the floor which indicates that zero budgets should be implemented for the period. Each member is then asked to vote on the proposal by ticking "yes" or "no" on the decision card. Ticking "yes" indicate zero budget should be implemented, while "no" indicates you wish to propose an alternative proposal. Any member that indicates "no" must proceed to propose an alternative budget size for the period. You have only a minute to make your decision, after which the laboratory assistant will collect your decision card.
- *Step 2*: The laboratory assistant will announce if majority votes for or against the initial motion. If majority has ticked "yes", then the final budget for period 1 will be zero. But if the majority ticks "no", the laboratory assistance assistant will randomly pick any of the decision card and budget size contained in the selected card becomes the new motion on the floor.
- *Step 3*: A new round of voting takes place to decide whether to implement or not the new motion. If majority ticks "yes", then the final budget for period 1 will be the new motion on the floor. If the majority ticks "no", the laboratory assistant will randomly pick among the decision cards ticked "no" in step 2 and announce to all members the budget size proposed on the decision card. This becomes the new motion on the floor.
- *Step 4*: A new round of voting takes place to decide whether or not to implement the new motion. If majority ticks "yes", then the final budget for period 1 will be new motion on the floor. If majority ticks "no", step 3 is repeated again until a motion with majority support emerges.

What does it means for a proposal/motion to have a majority vote/support? It means that at least 3 members of your group tick "yes" on their decision card; this indicates that the proposal should be implemented.

Can a member proposed a budget size outside range given in the payoff table? No.

What happens when the game ends? Your payoff in each period is summed up using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 2: annual budgeting and four-dimension

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, raise your hand and the laboratory assistant will attend to you accordingly. You will now take part in a decision making experiment.

What material do you have for the experiment? You have four types of handouts. (1) First you have a copy of this instruction which you can look at any time during the experiment. (2) The second item you have is a set of payoff table, which shows each participant's payoff for any given budget size. Note that there is a separate payoff table for each period. (3) The third item is a set of decision cards. The decision card will be used to make your proposals and also vote on alternative proposals. (4) The fourth item you have been given is the personal history form. The personal history form is like a scratch sheet that you can use for taking note of proceedings during the experiment. Please note that the use of the personal history form is optional but it must be turned in at the end of the experiment. The label on your personal history form indicates your subject ID. *All handouts must be turned in at the end of the experiment.*

What determines how much earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding the aggregate budget size over four periods. The amount you earn depends on the final budget size jointly voted for by all members of the group. In the Payoff Table, you will see the payoff associated with all possible budget sizes. At the end of the experiment, you will be able to exchange this payoff for Rands. The exchange rate is:

2.28 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment nor afterwards will anybody be informed about the subject ID of other members of their group.

How does the experiment work? You have been assigned to a group consisting of five (5) members and each of you have been allocated an ID (A to E), as indicated by the subject ID on your personal history form. The group's task is to decide the aggregate budget size to implement over four periods. The group will first decide on the final budget size for period 1 and thereafter decide the final budget for the periods 2, 3 and 4. This marks the end of the game.

How will members decide on the final budget in each period? The group decides the final budget size for any period following the procedure below:

- *Step 1*: Every period starts with an initial motion on the floor which indicates that zero budgets should be implemented for the period. Each member is then asked to vote on the proposal by ticking "yes" or "no" on the decision card. Ticking "yes" indicate zero budget should be implemented, while "no" indicates you wish to propose an alternative proposal. Any member that indicates "no" must proceed to propose an alternative budget size for the period. You have only a minute to make your decision, after which the laboratory assistant will collect your decision card.
- *Step 2*: The laboratory assistant will announce if majority votes for or against the initial motion. If majority has ticked "yes", then the final budget for period 1 will be zero. But if the majority ticks "no", the laboratory assistance assistant will randomly pick any of the decision card and budget size contained in the selected card becomes the new motion on the floor.
- *Step 3*: A new round of voting takes place to decide whether to implement or not the new motion. If majority ticks "yes", then the final budget for period 1 will be the new motion on the floor. If the majority ticks "no", the laboratory assistant will randomly pick among the decision cards ticked "no" in step 2 and announce to all members the budget size proposed on the decision card. This becomes the new motion on the floor.
- *Step 4*: A new round of voting takes place to decide whether or not to implement the new motion. If majority ticks "yes", then the final budget for period 1 will be new motion on the floor. If majority ticks "no", step 3 is repeated again until a motion with majority support emerges.

What does it means for a proposal/motion to have a majority vote/support? It means that at least 3 members of your group tick "yes" on their decision card; this indicates that the proposal should be implemented.

Can a member proposed a budget size outside range given in the payoff table? No.

What happens when the game ends? Your payoff in each period is summed up using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session?

Instruction 3: multiyear budgeting and two-dimension

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, raise your hand and the laboratory assistant will attend to you accordingly. You will now take part in a decision making experiment.

What material do you have for the experiment? You have four types of handouts. (1) First you have a copy of this instruction which you can look at any time during the experiment. (2) The second item you have is a set of payoff table, which shows each participant payoff for any given budget size. Note that there is a separate payoff table for each period. (3) The third item is a set of decision cards. The decision card will be used to make your proposals and also vote on alternative proposals. (4) The fourth items you have been given is the personal history form. The personal history form is like a scratch sheet that you can use for taking not of proceedings during the experiment. Please note that the use of the personal history form is optional but it must be turned in at the end of the experiment. The label on your personal history form indicate your subject ID. *All handouts must be turned in at the end of the experiment.*

What determines how much earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding the aggregate budget size over two periods. The amount you earn depends on the final budget size jointly voted for by all members of the group. In the Payoff Table, you will see the payoff associated with all possible budget sizes. At the end of the experiment, you will be able to exchange this payoff for Rand. The exchange rate is:

16 payoff = 35 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment or afterwards will anybody be informed the subject ID of other members of their group.

How does the experiment work? You have been assigned to a group consisting of five (5) members and each of you have been allocated an ID (A to E), as indicated by the subject ID on your personal history form. The group's task is to decide the aggregate budget size to implement for two periods. The group will first decide on the aggregate budget size for both periods 1 and 2. Thereafter, you will proceed to decide the final budget specifically for period 1. This marks the end of the game.

How will members decide on the final budget in each period? The group decides the budget size for any period following the procedure below:

- *Step 1*: Every period starts with an initial motion on the floor which indicates that zero budget should be implement for the period. Each member are then ask to vote on the proposal by ticking "yes" or "no" on their decision card. Ticking "yes" indicates zero budget should be implemented, while "no" indicates you wish to propose an alternative proposal. Any member that indicates "no" must proceed to propose an alternative budget size for the period. You have only a minute to make your decision, after which the laboratory assistant will collect your decision card.
- *Step 2*: The laboratory assistant will announce if majority votes for or against the initial motion. If majority ticks "yes", then the final budget for period 1 will be zero. But if the majority ticks "no", the laboratory assistant will randomly picked any of the decision card and budget size contained in the selected card becomes the new motion on the floor.

- *Step 3*: A new round of voting takes place to decide whether to implement or not the new motion. If majority ticks "yes", then the final budget for period 1 will be the new motion on the floor. If the majority ticks "no", the laboratory assistant will randomly pick among the decision cards that ticked "no" in step 2 and announce to all members the budget size proposed on the decision card. This becomes the new motion on the floor.

- *Step 4*: A new round of voting takes place to decide whether or not to implement the new motion. If majority ticks "yes", then the final budget for period 1 will be new motion on the floor. If majority ticks "no", step 3 is repeated again until a motion with majority support emerges.

What does it means for a proposal/motion to have a majority vote/support? It means that at least 3 members of your group tick "yes" on their decision card, which indicates that the proposal should be implemented.

Can a member proposed a budget size outside range given in the payoff table? No.

What happens when the game ends? Your payoff in each period is summed up and using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session

Instructions 4: multiyear budgeting and four-dimension

Welcome to the experiment. Please remain silent and seated until the end of the experiment. If you have questions at any stage of the experiment, raise your hand and the laboratory assistant will attend to you accordingly. You will now take part in a decision making experiment.

What material do you have for the experiment? You have four types of handouts. (1) First you have a copy of this instruction which you can look at any time during the experiment. (2) The second item you have is a set of payoff table, which shows each participant payoff for any given budget size. Note that there is a separate payoff table for each period. (3) The third item is a set of decision cards. The decision card will be used to make your proposals and also vote on alternative proposals. (4) The fourth items you have been given is the personal history form. The personal history form is like a scratch sheet that you can use for taking not of proceedings during the experiment. Please note that the use of the personal history form is optional but it must be turned in at the end of the experiment. The label on your personal history form indicate your subject ID. *All handouts must be turned in at the end of the experiment.*

What determines how much earned? The experiment consists of a series of group decision-making periods in which you will participate with others in deciding the aggregate budget size over four periods. The amount you earn depends on the final budget size jointly voted for by all members of the group. In the Payoff Table, you will see the payoff associated with all possible budget sizes. At the end of the experiment, you will be able to exchange this payoff for Rand. The exchange rate is:

2.28 payoff = 1 South African Rand (scaled)

[THIS WILL BE CALCULATED FOR YOU]

You will be privately paid at the end of the experiment to ensure that your decisions remain anonymous. Neither during the experiment or afterwards will anybody be informed the subject ID of other members of their group.

How does the experiment work? You have been assigned to a group consisting of five (5) members and each of you have been allocated an ID (A to E), as indicated by the subject ID on your personal history form. The group's task is to decide the aggregate budget size to implement for four periods. The group will first decide on the aggregate budget size for all the periods from 1 to 4. Thereafter, you will proceed to decide the final budget specifically for period 1, and thereafter periods 2 and 3. This marks the end of the game.

How will members decide on the final budget in each period? The group decides the budget size for any period following the procedure below:

- *Step 1*: Every period starts with an initial motion on the floor which indicates that zero budget should be implement for the period. Each member are then ask to vote on the proposal by ticking "yes" or "no" on their decision card. Ticking "yes" indicates zero budget should be implemented, while “no” indicates you wish to propose an alternative proposal. Any member that indicates “no” must proceed to propose an alternative budget size for the period. You have only a minute to make your decision, after which the laboratory assistant will collect your decision card.

- *Step 2*: The laboratory assistant will announce if majority votes for or against the initial motion. If majority ticks “yes”, then the final budget for period 1 will be zero. But if the majority ticks “no”, the laboratory assistant will randomly picked any of the decision card and budget size contained in the selected card becomes the new motion on the floor.

- *Step 3*: A new round of voting takes place to decide whether to implement or not the new motion. If majority ticks “yes”, then the final budget for period 1 will be the new motion on the floor. If the majority ticks “no”, the laboratory assistant will randomly pick among the decision cards that ticked “no” in step 2 and announce to all members the budget size proposed on the decision card. This becomes the new motion on the floor.

- *Step 4*: A new round of voting takes place to decide whether or not to implement the new motion. If majority ticks “yes”, then the final budget for period 1 will be new motion on the floor. If majority ticks “no”, step 3 is repeated again until a motion with majority support emerges.

What does it means for a proposal/motion to have a majority vote/support? It means that at least 3 members of your group tick “yes” on their decision card, which indicates that the proposal should be implemented.

Can a member proposed a budget size outside range given in the payoff table? No.

What happens when the game ends? Your payoff in each period is summed up and using the exchange rate given above. You will receive your earnings privately.

Any question before we take some practice session

4B Appendix: QUESTIONNAIRE

(to be filled after the experiment)

1. Terminal number (Number written on personal history form)

   ____

2. Gender:

   ☐ Male ☐ Female

3. Year of study:

   ☐ 1st Year ☐ 2nd Year

4. Department:

   ____

5. Instructions for Practice Session:

   ☐ I fully understood the instructions.

   ☐ I understood most of the instructions.

   ☐ I didn't understand much of the instructions.

6. Instructions for the Experiment:

   ☐ I fully understood the instructions.

   ☐ I understood most of the instructions.

   ☐ I didn't understand much of the instructions.

7. The readings of the payoff table:

   ☐ I fully understood the readings of the payoff table.

   ☐ I moderately understood the readings of the payoff table.

   ☐ I didn't understand at all the reading of the payoff table.

8. Use of the decision cards:

   ☐ I fully understood how to use the decision card.

   ☐ I moderately understood how to use the decision card.

   ☐ I didn't understand at all how to use the decision card.

4C Appendix: Pay-off Tables

Payoff Table: Design I

*Table 0.1Ä: Period 1**

Payoff Table					
Budget Size	Players				
	A	**B**	**C**	**D**	**E**
0	9	8	7	4	4
1	10	9	8	5	5
2	11	10	9	6	6
3	12	11	10	7	7
4	13	12	11	8	8
5	14	13	12	9	9
6	15	14	13	10	10
7	16	15	14	11	11
8	15	16	15	12	12
9	14	15	16	13	13
10	13	14	15	14	14
11	12	13	14	15	15
12	11	12	13	16	16
13	10	11	12	15	15
14	9	10	11	14	14
15	8	9	10	13	13
16	7	8	9	12	12
17	6	7	8	11	11
18	5	6	7	10	10
19	4	5	6	9	9
20	3	4	5	8	8
21	2	3	4	7	7
22	1	2	3	6	6
23	0	1	2	5	5
24	-1	0	1	4	4
25	-2	-1	0	3	3
26	-3	-2	-1	2	2
27	-4	-3	-2	1	1
28	-5	-4	-3	0	0
29	-6	-5	-4	-1	-1
30	-7	-6	-5	-2	-2

* The planning horizon for design I is two-dimensional; hence the payoff tables for the design I consists of two periods presented as Tables 0.1A and 0.2A. Each of the column in the payoff table is derived by substituting into

equation (2) the ideal budgets for subjects A to E as presented in columns (2) and (3) of Table 2.1. Also, given that subjects can make proposals other than their ideal budget, the possible budget size is allowed to range between 0 and 30.

Table 0.2A: Period 2[*]

Payoff Table					
Budget Size	Players				
	A	**B**	**C**	**D**	**E**
0	2	6	-1	3	1
1	3	7	0	4	2
2	4	8	1	5	3
3	5	9	2	6	4
4	6	10	3	7	5
5	7	11	4	8	6
6	8	12	5	9	7
7	9	13	6	10	8
8	10	14	7	11	9
9	11	15	8	12	10
10	12	16	9	13	11
11	13	15	10	14	12
12	14	14	11	15	13
13	15	13	12	16	14
14	16	12	13	15	15
15	15	11	14	14	16
16	14	10	15	13	15
17	13	9	16	12	14
18	12	8	15	11	13
19	11	7	14	10	12
20	10	6	13	9	11
21	9	5	12	8	10
22	8	4	11	7	9
23	7	3	10	6	8
24	6	2	9	5	7
25	5	1	8	4	6
26	4	0	7	3	5
27	3	-1	6	2	4
28	2	-2	5	1	3
29	1	-3	4	0	2
30	0	-4	3	-1	1

31	-1	-5	2	-2	0
32	-2	-6	1	-3	-1

* *refer to the note below table 0.1A*

Pay-off Table-Design II

*Table 0.3A: Period 1**

Payoff Table					
Budget Size	Players				
	A	B	C	E	F
	7	8	9	10	12
1	10	9	8	7	5
2	11	10	9	8	6
3	12	11	10	9	7
4	13	12	11	10	8
5	14	13	12	11	9
6	15	14	13	12	10
7	16	15	14	13	11
8	15	16	15	14	12
9	14	15	16	15	13
10	13	14	15	16	14
11	12	13	14	15	15
12	11	12	13	14	16
13	10	11	12	13	15
14	9	10	11	12	14
15	8	9	10	11	13
16	7	8	9	10	12
17	6	7	8	9	11
18	5	6	7	8	10
19	4	5	6	7	9
20	3	4	5	6	8
21	2	3	4	5	7
22	1	2	3	4	6
23	0	1	2	3	5
24	-1	0	1	2	4
25	-2	-1	0	1	3
26	-3	-2	-1	0	2
27	-4	-3	-2	-1	1
28	-5	-4	-3	-2	0
29	-6	-5	-4	-3	-1
30	-7	-6	-5	-4	-2

31	-8	-7	-6	-5	-3
32	-9	-8	-7	-6	-4

* Tables 0.3A to 0.4A are based on design II. Each of the column in the payoff tables is computed by substituting into equation (2) the ideal budgets for subjects A to E as presented in columns (4) and (5) of Table 2.1. Also, given that subjects can make proposals other than their ideal budget, the possible budget size is allowed to range between 0 and 30.

*Table 0.4A: Period 2**

Payoff Table					
Budget Size	Players				
	1	2	3	4	5
	14	10	17	10	15
1	3	7	0	7	2
2	4	8	1	8	3
3	5	9	2	9	4
4	6	10	3	10	5
5	7	11	4	11	6
6	8	12	5	12	7
7	9	13	6	13	8
8	10	14	7	14	9
9	11	15	8	15	10
10	12	16	9	16	11
11	13	15	10	15	12
12	14	14	11	14	13
13	15	13	12	13	14
14	16	12	13	12	15
15	15	11	14	11	16
16	14	10	15	10	15
17	13	9	16	9	14
18	12	8	15	8	13
19	11	7	14	7	12
20	10	6	13	6	11
21	9	5	12	5	10
22	8	4	11	4	9
23	7	3	10	3	8
24	6	2	9	2	7
25	5	1	8	1	6
26	4	0	7	0	5
27	3	-1	6	-1	4
28	2	-2	5	-2	3
29	1	-3	4	-3	2

30	0	-4	3	-4	1

* *refer to the note below table 0.3A*

Pay-off Table-Design III

*Table 0.5A: Period 1**

Budget Size	Players					Budget Size	Players				
	A	B	C	D	E		A	B	C	D	E
0	33	32	31	30	28	**21**	26	27	28	29	31
1	34	33	32	31	29	**22**	25	26	27	28	30
2	35	34	33	32	30	**23**	24	25	26	27	29
3	36	35	34	33	31	**24**	23	24	25	26	28
4	37	36	35	34	32	**25**	22	23	24	25	27
5	38	37	36	35	33	**26**	21	22	23	24	26
6	39	38	37	36	34	**27**	20	21	22	23	25
7	40	39	38	37	35	**28**	19	20	21	22	24
8	39	40	39	38	36	**29**	18	19	20	21	23
9	38	39	40	39	37	**30**	17	18	19	20	22
10	37	38	39	40	38	**31**	16	17	18	19	21
11	36	37	38	39	39	**32**	15	16	17	18	20
12	35	36	37	38	40	**33**	14	15	16	17	19
13	34	35	36	37	39	**34**	13	14	15	16	18
14	33	34	35	36	38	**35**	12	13	14	15	17
15	32	33	34	35	37	**36**	11	12	13	14	16
16	31	32	33	34	36	**37**	10	11	12	13	15
17	30	31	32	33	35	**38**	9	10	11	12	14
18	29	30	31	32	34	**39**	8	9	10	11	13
19	28	29	30	31	33	**40**	7	8	9	10	12
20	27	28	29	30	32						

***** Tables 0.5A to 0.8A are based on design III. Each of the column in the payoff tables is computed by substituting into equation (3) the ideal budgets for subjects A to E as presented in columns (6) to (9) of Table 2.1. Also, given that subjects can make proposals other than their ideal budget, the possible budget size is allowed to range between 0 and 40.

Table 0.6A: Period 2[*]

Budget Size	Players					Budget Size	Players				
	A	B	C	D	E		A	B	C	D	E
0	26	30	23	30	25	**21**	33	29	36	29	34
1	27	31	24	31	26	**22**	32	28	35	28	33
2	28	32	25	32	27	**23**	31	27	34	27	32
3	29	33	26	33	28	**24**	30	26	33	26	31
4	30	34	27	34	29	**25**	29	25	32	25	30
5	31	35	28	35	30	**26**	28	24	31	24	29
6	32	36	29	36	31	**27**	27	23	30	23	28
7	33	37	30	37	32	**28**	26	22	29	22	27
8	34	38	31	38	33	**29**	25	21	28	21	26
9	35	39	32	39	34	**30**	24	20	27	20	25
10	36	40	33	40	35	**31**	23	19	26	19	24
11	37	39	34	39	36	**32**	22	18	25	18	23
12	38	38	35	38	37	**33**	21	17	24	17	22
13	39	37	36	37	38	**34**	20	16	23	16	21
14	40	36	37	36	39	**35**	19	15	22	15	20
15	39	35	38	35	40	**36**	18	14	21	14	19
16	38	34	39	34	39	**37**	17	13	20	13	18
17	37	33	40	33	38	**38**	16	12	19	12	17
18	36	32	39	32	37	**39**	15	11	18	11	16
19	35	31	38	31	36	**40**	14	10	17	10	15
20	34	30	37	30	35						

* *refer to the note below table 0.5A*

*Table 0.7A: Period 3**

Budget Size	Players					Budget Size	Players				
	A	B	C	D	E		A	B	C	D	E
0	25	24	23	22	20	**21**	34	35	36	37	39
1	26	25	24	23	21	**22**	33	34	35	36	38
2	27	26	25	24	22	**23**	32	33	34	35	37
3	28	27	26	25	23	**24**	31	32	33	34	36
4	29	28	27	26	24	**25**	30	31	32	33	35
5	30	29	28	27	25	**26**	29	30	31	32	34
6	31	30	29	28	26	**27**	28	29	30	31	33
7	32	31	30	29	27	**28**	27	28	29	30	32
8	33	32	31	30	28	**29**	26	27	28	29	31
9	34	33	32	31	29	**30**	25	26	27	28	30
10	35	34	33	32	30	**31**	24	25	26	27	29
11	36	35	34	33	31	**32**	23	24	25	26	28
12	37	36	35	34	32	**33**	22	23	24	25	27
13	38	37	36	35	33	**34**	21	22	23	24	26
14	39	38	37	36	34	**35**	20	21	22	23	25
15	40	39	38	37	35	**36**	19	20	21	22	24
16	39	40	39	38	36	**37**	18	19	20	21	23
17	38	39	40	39	37	**38**	17	18	19	20	22
18	37	38	39	40	38	**39**	16	17	18	19	21
19	36	37	38	39	39	**40**	15	16	17	18	20
20	35	36	37	38	40						

* *refer to the note below table 0.5A*

Table 0.8A: Period 4[*]

Budget Size	Players					Budget Size	Players				
	A	B	C	D	E		A	B	C	D	E
0	18	22	15	22	17	**21**	39	37	36	37	38
1	19	23	16	23	18	**22**	40	36	37	36	39
2	20	24	17	24	19	**23**	39	35	38	35	40
3	21	25	18	25	20	**24**	38	34	39	34	39
4	22	26	19	26	21	**25**	37	33	40	33	38
5	23	27	20	27	22	**26**	36	32	39	32	37
6	24	28	21	28	23	**27**	35	31	38	31	36
7	25	29	22	29	24	**28**	34	30	37	30	35
8	26	30	23	30	25	**29**	33	29	36	29	34
9	27	31	24	31	26	**30**	32	28	35	28	33
10	28	32	25	32	27	**31**	31	27	34	27	32
11	29	33	26	33	28	**32**	30	26	33	26	31
12	30	34	27	34	29	**33**	29	25	32	25	30
13	31	35	28	35	30	**34**	28	24	31	24	29
14	32	36	29	36	31	**35**	27	23	30	23	28
15	33	37	30	37	32	**36**	26	22	29	22	27
16	34	38	31	38	33	**37**	25	21	28	21	26
17	35	39	32	39	34	**38**	24	20	27	20	25
18	36	40	33	40	35	**39**	23	19	26	19	24
19	37	39	34	39	36	**40**	22	18	25	18	23
20	38	38	35	38	37						

* *refer to the note below table 0.5A*

Pay-off Table-Design IV

Table 0.9A: Period 1[*]

Budget Size	Players					Budget Size	Players				
	A	B	C	D	E		A	B	C	D	E
0	33	32	31	28	28	**21**	26	27	28	31	31
1	34	33	32	29	29	**22**	25	26	27	30	30
2	35	34	33	30	30	**23**	24	25	26	29	29
3	36	35	34	31	31	**24**	23	24	25	28	28
4	37	36	35	32	32	**25**	22	23	24	27	27
5	38	37	36	33	33	**26**	21	22	23	26	26
6	39	38	37	34	34	**27**	20	21	22	25	25
7	40	39	38	35	35	**28**	19	20	21	24	24
8	39	40	39	36	36	**29**	18	19	20	23	23
9	38	39	40	37	37	**30**	17	18	19	22	22
10	37	38	39	38	38	**31**	16	17	18	21	21
11	36	37	38	39	39	**32**	15	16	17	20	20
12	35	36	37	40	40	**33**	14	15	16	19	19
13	34	35	36	39	39	**34**	13	14	15	18	18
14	33	34	35	38	38	**35**	12	13	14	17	17
15	32	33	34	37	37	**36**	11	12	13	16	16
16	31	32	33	36	36	**37**	10	11	12	15	15
17	30	31	32	35	35	**38**	9	10	11	14	14
18	29	30	31	34	34	**39**	8	9	10	13	13
19	28	29	30	33	33	**40**	7	8	9	12	12
20	27	28	29	32	32						

*Tables 0.9A to 0.12A are based on design IV. Each of the column in the payoff tables is computed by substituting into equation (3) the ideal budgets for subjects A to E as presented in columns (11) to (13) of Table 2.1. Also, given that subjects can make proposals other than their ideal budget, the possible budget size is allowed to range between 0 and 40.

*Table 0.10A: Period 2**

Budget Size	Players					Budget Size	Players				
	A	**B**	**C**	**D**	**E**		**A**	**B**	**C**	**D**	**E**
0	26	30	23	27	25	**21**	33	29	36	32	34
1	27	31	24	28	26	**22**	32	28	35	31	33
2	28	32	25	29	27	**23**	31	27	34	30	32
3	29	33	26	30	28	**24**	30	26	33	29	31
4	30	34	27	31	29	**25**	29	25	32	28	30
5	31	35	28	32	30	**26**	28	24	31	27	29
6	32	36	29	33	31	**27**	27	23	30	26	28
7	33	37	30	34	32	**28**	26	22	29	25	27
8	34	38	31	35	33	**29**	25	21	28	24	26
9	35	39	32	36	34	**30**	24	20	27	23	25
10	36	40	33	37	35	**31**	23	19	26	22	24
11	37	39	34	38	36	**32**	22	18	25	21	23
12	38	38	35	39	37	**33**	21	17	24	20	22
13	39	37	36	40	38	**34**	20	16	23	19	21
14	40	36	37	39	39	**35**	19	15	22	18	20
15	39	35	38	38	40	**36**	18	14	21	17	19
16	38	34	39	37	39	**37**	17	13	20	16	18
17	37	33	40	36	38	**38**	16	12	19	15	17
18	36	32	39	35	37	**39**	15	11	18	14	16
19	35	31	38	34	36	**40**	14	10	17	13	15
20	34	30	37	33	35						

* *refer to the note below table 0.9A*

*Table 0.11A: Period 3**

Budget Size	Players					Budget Size	Players				
	A	B	C	D	E		A	B	C	D	E
0	25	24	23	20	20	**21**	34	35	36	39	39
1	26	25	24	21	21	**22**	33	34	35	38	38
2	27	26	25	22	22	**23**	32	33	34	37	37
3	28	27	26	23	23	**24**	31	32	33	36	36
4	29	28	27	24	24	**25**	30	31	32	35	35
5	30	29	28	25	25	**26**	29	30	31	34	34
6	31	30	29	26	26	**27**	28	29	30	33	33
7	32	31	30	27	27	**28**	27	28	29	32	32
8	33	32	31	28	28	**29**	26	27	28	31	31
9	34	33	32	29	29	**30**	25	26	27	30	30
10	35	34	33	30	30	**31**	24	25	26	29	29
11	36	35	34	31	31	**32**	23	24	25	28	28
12	37	36	35	32	32	**33**	22	23	24	27	27
13	38	37	36	33	33	**34**	21	22	23	26	26
14	39	38	37	34	34	**35**	20	21	22	25	25
15	40	39	38	35	35	**36**	19	20	21	24	24
16	39	40	39	36	36	**37**	18	19	20	23	23
17	38	39	40	37	37	**38**	17	18	19	22	22
18	37	38	39	38	38	**39**	16	17	18	21	21
19	36	37	38	39	39	**40**	15	16	17	20	20
20	35	36	37	40	40						

* *refer to the note below table 0.9A*

*Table 0.12A: Period 4**

Budget Size	**Players**					**Budget Size**	**Players**				
	A	**B**	**C**	**D**	**E**		**A**	**B**	**C**	**D**	**E**
0	18	22	15	19	17	**21**	39	37	36	40	38
1	19	23	16	20	18	**22**	40	36	37	39	39
2	20	24	17	21	19	**23**	39	35	38	38	40
3	21	25	18	22	20	**24**	38	34	39	37	39
4	22	26	19	23	21	**25**	37	33	40	36	38
5	23	27	20	24	22	**26**	36	32	39	35	37
6	24	28	21	25	23	**27**	35	31	38	34	36
7	25	29	22	26	24	**28**	34	30	37	33	35
8	26	30	23	27	25	**29**	33	29	36	32	34
9	27	31	24	28	26	**30**	32	28	35	31	33
10	28	32	25	29	27	**31**	31	27	34	30	32
11	29	33	26	30	28	**32**	30	26	33	29	31
12	30	34	27	31	29	**33**	29	25	32	28	30
13	31	35	28	32	30	**34**	28	24	31	27	29
14	32	36	29	33	31	**35**	27	23	30	26	28
15	33	37	30	34	32	**36**	26	22	29	25	27
16	34	38	31	35	33	**37**	25	21	28	24	26
17	35	39	32	36	34	**38**	24	20	27	23	25
18	36	40	33	37	35	**39**	23	19	26	22	24
19	37	39	34	38	36	**40**	22	18	25	21	23
20	38	38	35	39	37						

* *refer to the note below table 0.9A*

5. Chapter Five: Conclusion, Policy Recommendations and Suggestions for Further Research

5.1 Summary of Findings

This thesis explores, using experimental methods, the implications of common pool problem in fiscal policy making on fiscal performance. We set out to achieve three main objectives (i) examine the effect of dynamic common pool problem on fiscal performance; (ii) test the prediction of dynamic common pool model regarding the possibility for delayed stabilization; and (iii) investigate the effect of multiyear budgeting on fiscal performance. A concise summary of major findings are set out in what follows.

5.1.1 Dynamic Common Pool Problem and Fiscal Performance

Chapter 2 examines effects of the dynamic common pool within budget institution on fiscal performance. Dynamic common pool problem occurs when fragmentation within budgeting institution creates uncertainty in future fiscal path, thereby increasing the proclivity for budget actors to strategically incur deficit and draw down on public-sector wealth. Given the limitation of field data in exploring a dynamic and strategic setting, the experimental method is adopted. Following previous experiments by Fréchette et al. (2005) and Mason and Phillips (1997), we implement a legislative bargaining game of "divide-the-dollar" to capture the dynamic common pool problem in a fiscal setting. The experimental data generated are estimated using non-parametric techniques and system GMM.

We find that dynamic common pool problem indeed results in poor fiscal performance, measured by the deficit level. In addition, we observe that budget actors react to declining public-sector wealth with more aggressive appropriation behaviour, which further exacerbates the poor fiscal outcomes. Furthermore, the effect of dynamic common problem is found to be propagated through the strategic channel, as predicted in the model. In essence, despite the absence of conventional fundamentals for incurring fiscal deficit—the divergence between discount rate and interest rate—poor fiscal performance persists due to the dynamic common pool problem. This evidence therefore underscores the need for policymakers to be more proactive in addressing the problem of fragmentation within the budget institution, especially with regards to strategic effect in order to improve fiscal performance.

5.1.2 A Test of Delayed Stabilization Hypothesis

Chapter three empirically tests the delayed stabilization hypothesis, as predicted by the dynamic common pool model. Specifically, the hypothesis suggests that stabilization implies that there is public-sector wealth threshold at which presence of dynamic common pool problem no longer exerts a negative effect on fiscal outcome. In testing this hypothesis, we adopt an experimental design which builds on and extends the legislative bargaining game of "divide-the-dollar". This is implemented by applying the random stopping rule after a pre-specified period. Also, tax is imposed on a group that exhausts its initial public-sector wealth. However, to eliminate possible identification problem, the initial public-sector wealth is restored after some periods of paying tax, although subjects are not aware of this replacement *ab initio*. Thus, it becomes possible to evaluate if the subjects appropriation behaviours follow the pattern predicted by delayed stabilization hypothesis. Furthermore, the experimental data generated is analysed using both fixed-effect panel threshold regression model and piecewise linear regression model.

Our results do not support the prediction of delayed stabilization. In fact, the deficit level was highest in the periods after which a group's initial public-sector wealth is restored, a clear contradiction of the delayed stabilization hypothesis. However, it is observed that a temporary stabilization is achieved during the periods that players are paying tax to replenish the public-sector wealth, when it is exhausted. This evidence in part supports the crisis hypothesis, which posits that the probability of successful stabilization is enhanced when it coincides with economic crisis. In this case, a group's resort to taxation implies fiscal crisis—since this only arises because the initial public-wealth is exhausted—which invariably helps constrain their aggressive appropriation behaviour. Thus, this suggests a possible role for the components of public-sector wealth (whether taxation or resource revenue) and economic crisis in designing and achieving a successful fiscal stabilization strategy.

5.1.3 Multiyear Budgeting and Fiscal Performance

Chapter three examines the implication of the voting equilibrium model for relationship(s) between multiyear budgeting and fiscal performance. Voting equilibrium model recognizes that budget actors are rational individuals, who seek to maximize their utility in relation to the budget proposal. Thus, voting equilibrium model predicts that voting preference of the budget actor will influence the eventual budget outcome, thereby also influencing any possible effect of multiyear budgeting on fiscal performance. Building on the work of Ehrhart et al. (2007), this implication of voting equilibruim model is tested using a laboratory experiment. We design a configuration of voting preference for a five-member budgeting institution. In the baseline treatment, subjects make their budgeting decision using multiyear budgeting procedure; annual budgeting procedure is on the other hand adopted in the secondary treamnent. The budget outcomes are thereafter compared under the two treatments to determine which produces the smallest budget size.

We find no signifcant differences in budget sizes under the multiyear and annual budgeting processes. As predicted by voting equilibruim, the multiyear budgeting generates lower budget size than annual budgeting and vice versa, according to the configuration of the voters' preference. Results also show that the predictive success of the voting equilibruim model reduces as dimension of the planning horizon increases. In essence, increasing the dimension of planning horizon has the tendency to increase uncertainty in the budgeting outcomes.

5.2 Policy Recommendations

These findings have important policy implications for budgeting institutions in the real world, especially in developing economies. For example, our findings suggest that budgeting institutions that are characterized by multiple decision makers (size fragmentation) and decentralized fiscal structure (procedural fragmentation) would be highly susceptible to dynamic common pool problem that may negatively affect overall fiscal performance. Therefore, policies that limit the extent of both size and procedural fragmentations will be crucial. In reality, procedural fragmentation is more amendable to reform in the short-run than size fragmentation which is ingrained in the constitution. Thus, targeting policy intervention towards reducing procedural fragmentation will be more productive. This could be achieved through centralization of budgeting decision—increasing the strategic dominance of one of the budget actors, the finance minister for instance. Another policy option that could reduce the level of procedural fragmentation is through limiting the amendment power of the parliament/legislative arm to the executive budget proposal.

Furthermore, since static and dynamic common pool problem work through different channels, the set of policies that address one aspect of the problem may not necessarily affect the other. Thus, understanding the nature/type of common pool problem that prevails within a given budgeting institution is crucial to designing the optimal policy response. In the case of strategic effect channel, a number of policy options have been suggested in the literature. Such options include improving transparency and accountability of the budget processes (Alt & Lassen, 2006); establishing an apolitical fiscal council as is the case with central bank's autonomy in the monetary policy context (Wyplosz, 2012), using long-term fiscal constraints (von Hagen, 2002) and making government revenue tilt more towards taxation (Botelho et al., 2013). One of the experiments conducted for this study particularly give credence to the potential of using taxation to mitigate common pool problem. This evidence could also help explain poor fiscal and economic performance of resources abundant countries, which are beleaguered by problem of resource curse and weak institution (Sachs & Warner, 2001), compared to tax-reliant economies that have been observed to have better fiscal performance, given that the system promotes transparency and accountability.

Also, in view of the lack of evidence regarding delayed stabilization, it becomes crucial for policy makers to adopt active stabilization policy such as fiscal rules or centralization of the budget institution, as against relying on budget actors to act endogenously to correct deficit bias. More importantly, the result suggests a possible benefit from economic crisis, as it enhances the probability of success of stabilization efforts. Thus, countries undergoing fiscal crisis, such as those with high and unsustainable debt level, could exploit this benefit through appropriate timing of their fiscal reform initiatives. In addition, given that resource dependency amplifies the effect of fragmentation within budget institution, it is imperative for developing countries to diversify government revenue base away from natural resources and instead develop a viable tax capacity.

Finally, given that introduction of multiyear budgeting alone does not guarantee good fiscal performance, it becomes crucial for policymakers to take into account the likely effect of voting preference when designing fiscal reform and rules. The cases of Ghana and Malawi, which implemented a similar form of multiyear budgeting framework but experienced contrasting performance afterwards, underscore the important role of preference of budget actors. In the case of Ghana, the multi-year budgeting is formally linked to budget making process and the implementation is centrally coordinated by the Ministry of Finance. On the other hand, the multiyear budgeting and the actual budget making process are detached in the case of Malawi, while the coordination among line ministries is minimal. The implication of this separate arrangement implies that preference of budget actors will be affected differently by the multiyear budgeting in Ghana than in Malawi. This difference is partly evident in fiscal performance between the two countries after decades of implementing multiyear budgeting, with Debt/GDP ratio higher by more than 10 percentage points in Malawi compared to Ghana (World Development Indicators, 2015). There is also the need for more profound analysis of concentration of developing countries on MTFF or multiyear budgeting stage and lack of progression to higher level of MTEF which improves fiscal discipline. This will shed more light on other aspect of budget actors' preferences that affect fiscal performance.

Also, given the possible uncertainty that could arise in higher dimension of planning horizon, needlessly long planning horizon should be avoided in future fiscal reform initiatives. This corroborates with Le Houerou and Taliercio (2002) observation that credibility of multiyear budgeting might actually be undermined in countries like Mozambique with six years planning horizon. In general, what constitutes the optimal dimension of planning horizon will be shaped by individual country's budget institution specific and economic-wide idiosyncrasies.

5.3 Suggestion for Further Research

The present study has explored, using experimental approach, the implication of the common pool problem in fiscal policy making for fiscal performance. Going forward, there are several areas that the present study can be deeply enriched and extended. As it is the tradition in the field of experimental economics, a single experiment is not definitive; as such more experiments are needed to ensure the robustness of the results under alternative designs and institutional settings. For instance, to simplify the decision space and interpretation of result, only the parameters of interest are varied across treatments in our experiment. Further research can explore the implication of varying these parameters that were fixed in our case. For example, it will be revealing to examine the comparative static outcomes when the discount and interest rates are higher compared to the specification adopted in chapter two, or when a tax regime different from the lump-sum tax used in chapter three is implemented.

In the same vein, discussion in chapter four can be extended by experimenting with alternative legislative size and configuration of ideal budgets. Future studies can also explore the aspect of non-distributive policies. The present study focuses on the distributive policy, which the presence of common pool problem inherently promotes. However, some aspects of non-distributive policies have been observed to be influenced by the persistence of common pool problem within the budget institution. Thus, examining the implication of common pool problem in this domain is a fertile area for further research.

Finally, the empirical evidences provided by the present study are derived from laboratory experiments. Hence, exploring alternative data generating techniques will be crucial. While there is doubt regarding suitability of field data, there are some innovative ways that it could be explored to address some relevant research questions. For example, Ferejohn and Krehbiel (1987) tested their voting equilibrium model using the natural experimental condition generated with the promulgation of the 1974 Budget and Impoundment Control Act (PL 93-344) in the US. Thus, innovative use of alternative data generating technique to explore the research questions posed in the present study is another promising area for further research.

References

Abiad, A. G., & Baig, T. (2005). Underlying factors driving fiscal effort in emerging market economies. *IMF Working Paper 05*(106).

Afonso, A., & Hauptmeier, S. (2009). Fiscal behaviour in the European Union: rules, fiscal decentralization and government indebtedness. *European Central Bank Working Paper*(1054).

African Economic Outlook. (2015). *Database on African Fiscal Performance*. Retrieved from Ethiopia:

Alesina, A., Ardagna, S., & Trebbi, F. (2006). Who adjusts and when? On the political economy of reforms. *National Bureau of Economic Research working paper*

Alesina, A., & Drazen, A. (1989). Why are stabilizations delayed? *National Bureau of Economic Research working paper*.

Alesina, A., Hausmann, R., Hommes, R., & Stein, E. (1999). Budget institutions and fiscal performance in Latin America. *Journal of Development Economics, 59*(2), 253-273.

Alesina, A., & Perotti, R. (1995a). Fiscal expansions and fiscal adjustments in OECD countries. *National Bureau of Economic Research working paper*.

Alesina, A., & Perotti, R. (1995b). The Political Economy of Budget Deficits. *Staff Papers-International Monetary Fund*, 1-31.

Alesina, A., & Perotti, R. (1999). Budget deficits and budget institutions *Fiscal institutions and fiscal performance* (pp. 13-36): University of Chicago Press.

Alesina, A., Perotti, R., Tavares, J., Obstfeld, M., & Eichengreen, B. (1998). The political economy of fiscal adjustments. *Brookings Papers on Economic Activity*, 197-266.

Alt, J. E., & Lassen, D. D. (2006). Fiscal transparency, political parties, and debt in OECD countries. *European Economic Review, 50*(6), 1403-1439.

Andersson, O., Holm, H. J., Tyran, J. R., & Wengström, E. (2015). Risk aversion relates to cognitive ability: Preferences or Noise? *Journal of the European Economic Association*.

Ardagna, S. (2004). Fiscal stabilizations: When do they work and why. *European Economic Review, 48*(5), 1047-1074.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29-51.

Baldacci, E., Gupta, S., & Mulas-Granados, C. (2010). Getting Debt under Control. *Finance and Development, 47*(4), 18-21.

Barro, R. J. (1979). On the determination of the public debt. *The Journal of Political Economy*, 940-971.

Battaglini, M., Nunnari, S., & Palfrey, T. R. (2012a). The dynamic free rider problem: A laboratory study. *Economic Theory Center Working Paper*(48-2012).

Battaglini, M., Nunnari, S., & Palfrey, T. R. (2012b). Legislative bargaining and the dynamics of public investment. *American Political Science Review, 106*(02), 407-429.

Bevan, D., & Palomba, G. (2000). Uganda: The Budget and medium term expenditure framework set in a wider context. *Background paper for Poverty Reduction Support Credit with DFID finance, Oct*.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*(1), 115-143.

Boex, L., Martinez-Vazquez, J., & McNab, R. M. (2000). Multi-Year Budgeting: A Review of International Practices and Lessons for Developing and Transitional Economies. *Public Budgeting & Finance, 20*(2), 91-112.

Bohn, H. (1998). The behavior of US public debt and deficits. *Quarterly Journal of Economics*, 949-963.

Bond, S. (2002). Dynamic panel data models: a guide to micro data methods and practice. *Portuguese economic journal, 2*(2), 141-162.

Borensztein, E., & Panizza, U. (2009). The Costs of Sovereign Default. *IMF Staff Papers, 56*(4).

Botelho, A., Dinar, A., & Costa, L. M. (2013). Linking Appropriation and Provision of Public Goods Decisions Decreases Rate of Destruction of the Commons. *Working paper, Water Science and Policy Center Department of Environmental Sciences, University of California, Riverside*

Brooks, C. (2014). Introductory econometrics for finance. *Cambridge university press*.

Buiter, W. H. (1985). A guide to public sector debt and deficits. *Economic policy, 1*(1), 13-61.

Burger, P., Stuart, I., Jooste, C., & Cuevas, A. (2011). Fiscal sustainability and the fiscal reaction function for South Africa. *IMF Working Papers*, 1-27.

Cassar, A., & Friedman, D. (2004). *Economics lab: an intensive course in experimental economics*: Routledge.

Chari, V. V., & Kehoe, P. J. (1993). Sustainable plans and debt. *Journal of Economic Theory, 61*(2), 230-261.

Cheng, Y. S., Liu, Y. P., & Chien, C. Y. (2010). Capital structure and firm value in China: A panel threshold regression analysis. *African Journal of Business Management, 4*(12), 2500.

Corsetti, G., Feld, L. P., Lane, P. R., Reichlin, L., Rey, H., & Vayanos, D. (2015). A New Start for the Eurozone: Dealing with Debt. *Centre for Economic Policy Research Publication*.

Croson, R. (2005). The method of experimental economics. *International Negotiation, 10*(1), 131-148.

Cullis, J., Jones, P., & Jones, P. R. (2009). *Public finance and public choice: analytical perspectives*: Oxford University Press.

Daniel, J., Callen, T., Terrones, M., Debrun, X., & Allard, C. (2003). Public debt in emerging markets: is it too high? *World Economic Outlook*, p.113.

De Haan, J., Moessen, W., & Volkerink, B. (1999). Budgetary procedures-aspects and changes: New evidence for some European countries *Fiscal institutions and fiscal performance* (pp. 265-300): University of Chicago Press.

De Janvry, A., McCarthy, N., & Sadoulet, E. (1998). Endogenous provision and appropriation in the commons. *American Journal of Agricultural Economics, 80*(3), 658-664.

Dell'Erba, S., Mattina, T., & Roitman, A. (2015). Pressure or prudence? Tales of market pressure and fiscal adjustment. *Journal of International Money and Finance, 51*, 196-213.

Ehrhart, K.-M., Gardner, R., von Hagen, J., & Keser, C. (2007). Budget processes: Theory and experimental evidence. *Games and Economic Behavior, 59*(2), 279-295.

Elgie, R., & McMenamin, I. (2008). Political fragmentation, fiscal deficits and political institutionalisation. *Public Choice, 136*(3-4), 255-267.

Eslava, M. (2011). The political economy of fiscal deficits: a survey. *Journal of Economic Surveys, 25*(4), 645-673.

Ferejohn, J., & Krehbiel, K. (1987). The budget process and the size of the budget. *American Journal of Political Science*, 296-320.

Filc, G., & Scartascini, C. G. (2011). Is Latin America on the Right Track? An Analysis of Medium-Term Frameworks and the Budget Process. *SSRN Working Paper Series*.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*(2), 171-178.

Fischer, M.-E., Irlenbusch, B., & Sadrieh, A. (2004). An intergenerational common pool resource experiment. *Journal of Environmental Economics and Management, 48*(2), 811-836.

Frechette, G. R. (2011). Laboratory Experiments: Professionals Versus Students. *SSRN Working Paper Series*.

Fréchette, G. R., Kagel, J. H., & Morelli, M. (2005). Behavioral identification in coalitional bargaining: An experimental analysis of demand bargaining and alternating offers. *Econometrica, 73*(6), 1893-1937.

Fréchette, G. R., Kagel, J. H., & Morelli, M. (2012). Pork versus public goods: an experimental study of public good provision within a legislative bargaining framework. *Economic Theory, 49*(3), 779-800.

Gardner, R. J., & von Hagen, J. (2012). Sequencing and the Size of the Budget Experimental Evidence. *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten*, 464.

Gollwitzer, S. (2011). Budget institutions and fiscal performance in Africa. *Journal of African Economies, 20*(1), 111-152.

Grigoli, F., Mills, Z., Verhoeven, M., & Vlaicu, R. (2012). MTEFs and fiscal performance: panel data evidence. *World Bank Policy Research Working Paper*(6186).

Gujarati, D. N. (2003). Basic Econometrics. 4th. *New York: McGraw-Hill*.

Gupta, S., Clements, B., Baldacci, E., & Mulas-Granados, C. (2004). The persistence of fiscal adjustments in developing countries. *Applied Economics Letters, 11*(4), 209-212.

Hallerberg, M., & von Hagen, J. (1999). Electoral Institutions, Cabinet Negotiations, and Budget Deficits in the European Union. *NBER Chapters*, 209-232.

Hansen, B. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics, 93*(2), 345-368.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029-1054.

Herr, A., Gardner, R., & Walker, J. M. (1997). An experimental study of time-independent and time-dependent externalities in the commons. *Games and Economic Behavior, 19*(1), 77-96.

Holmes, M., & Evans, A. (2003). A review of experience in implementing Medium Term Expenditure Frameworks in a PRSP context: a synthesis of eight country studies. *London: Overseas Development Institute*.

Horton, M., & El-Ganainy, A. (2006). What Is Fiscal Policy. *Finance & Development, 43*.

Hsieh, C.-T. (2000). Bargaining over reform. *European Economic Review, 44*(9), 1659-1676.

Illera, R. M., & Mulas-Granados, C. (2002). *Duration of fiscal budgetary consolidations in the European Union*. Retrieved from

Im, K. S., Pesaran, M. H., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics, 115*(1), 53-74.

Katayama, K. (2008). Delay in fiscal reform. *East Asian Bureau of Economic Research, No. 23075*.

Kaur, B., & Mukherjee, A. (2012). Threshold Level of Debt and Public Debt Sustainability: The Indian Experience. *Reserve Bank of India, 1*.

Keser, C., & Gardner, R. (1999). Strategic behavior of experienced subjects in a common pool resource game. *International Journal of Game Theory, 28*(2), 241-252.

Koh, Y. (2015). Universalism and Inefficiency in Distributive Politics. *Job Market paper*.

Kontopoulos, Y., & Perotti, R. (1999). Government fragmentation and fiscal policy outcomes: Evidence from OECD countries *Fiscal institutions and fiscal performance* (pp. 81-102): University of Chicago Press.

Kramer, G. H. (1972). Sophisticated voting over multidimensional choice spaces†. *Journal of Mathematical Sociology, 2*(2), 165-180.

Krogstrup, S., & Wyplosz, C. (2010). A common pool theory of supranational deficit ceilings. *European Economic Review, 54*(2), 269-278.

Le Houerou, P., & Taliercio, R. (2002). *Medium term expenditure frameworks: from concept to practice-preliminary lessons from Africa*: The World Bank.

Levin, A., Lin, C. F., & Chu, C. S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics, 108*(1), 1-24.

List, J. A., Sadoff, S., & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics, 14*(4), 439-457.

Ljungman, G. (2009). *Top-Down Budgeting: An Instrument to Strengthen Budget Management*: International Monetary Fund.

Lopes da Veiga, J., Ferreira-Lopes, A., & Sequeira, T. (2014). Public Debt, Economic Growth, and Inflation in African Economies.

López-Pérez, R., & Spiegelman, E. (2013). Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics, 16*(3), 233-247.

Lora, E., & Olivera, M. (2004). What makes reforms likely: Political economy determinants of reforms in Latin America. *Journal of Applied Economics, 7*(1), 99-135.

Lukkezen, J., & Rojas-Romagosa, H. (2013). An early-warning indicator for sovereign debt sustainability. *CPB background information: Updated version of CPB discussion paper, 212*(7).

Mason, C. F., & Phillips, O. R. (1997). Mitigating the tragedy of the commons through cooperation: an experimental evaluation. *Journal of Environmental Economics and Management, 34*(2), 148-172.

Mason, C. F., & Phillips, O. R. (2002). In Support of Trigger Strategies: Experimental Evidence from Two-Person Noncooperative Games. *Journal of Economics & Management Strategy, 11*(4), 685-716.

Mendoza, E. G., & Ostry, J. D. (2008). International evidence on fiscal solvency: Is fiscal policy “responsible”? *Journal of Monetary Economics, 55*(6), 1081-1093.

Mierau, J. O., Jong-A-Pin, R., & De Haan, J. (2007). Do political variables affect fiscal policy adjustment decisions? New empirical evidence. *Public Choice, 133*(3-4), 297-319.

Momeni, F. (2015). Voluntary and Mandatory Provision of Common Pool Resources with Heterogeneous Appropriators. *SPI Working Paper*(131-SPI).

Mulas-Granados, C. (2003). The political and economic determinants of budgetary consolidation in Europe. *European Political Economy Review, 1*(1), 015-039.

Nachar, N. (2008). The Mann-Whitney U: A test for assessing whether two independent samples come from the same distribution. *Tutorials in Quantitative Methods for Psychology, 4*(1), 13-20.

Osés-Eraso, N., Udina, F., & Viladrich-Grau, M. (2008). Environmental versus human-induced scarcity in the commons: do they trigger the same response? *Environmental and Resource Economics, 40*(4), 529-550.

Osés-Eraso, N., & Viladrich-Grau, M. (2007). Appropriation and concern for resource scarcity in the commons: An experimental study. *Ecological Economics, 63*(2), 435-445.

Ostrom, E., Gardner, R., & Walker, J. (1994). *Rules, games, and common-pool resources*: University of Michigan Press.

Padovano, F., & Venturi, L. (2001). Wars of attrition in Italian government coalitions and fiscal performance: 1948–1994. *Public Choice, 109*(1-2), 15-54.

Palfrey, T. R. (2013). Experiments in political economy. *Handbook of Experimental Economics,, 2*.

Perotti, R., & Kontopoulos, Y. (2002). Fragmented fiscal policy. *Journal of Public Economics, 86*(2), 191-222.

Plott, C., & Krehbiel, K. (1979). Voting in a Committee: an Experimental Analysis. *American Political Science Review*.

Purfield, C. (2003). Fiscal Adjustment in Transition Countries: Evidence From the 1990's. *IMF Working Papers, No. 03/36*.

Richter, C., & Paparas, D. (2013). How reliable are budget sustainability tests? A case study for Greece. *International Journal of Public Policy 14, 9*(1-2), 23-43.

Roberts, K. (2007). Condorcet cycles? A model of intertemporal voting. *Social Choice and Welfare, 29*(3), 383-404.

Romer, D. (2004). Advanced macroeconomics. *McGraw-Hill Education*(4).

Roodman, D. (2006). How to do xtabond2: An introduction to difference and system GMM in Stata. *Center for Global Development working paper, 103*.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. *Stata Journal, 9*(1), 86-136.

Rose, A. K. (2005). One reason countries pay their debts: renegotiation and international trade. *Journal of Development Economics, 77*(1), 189-206.

Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European Economic Review, 45*(4), 827-838.

Schiavo-Campo, S. (2009). Potemkin Villages:"The" Medium-Term Expenditure Framework in Developing Countries. *Public Budgeting & Finance, 29*(2), 1-26.

Selten, R. (1991). Properties of a measure of predictive success. *Mathematical Social Sciences, 21*(2), 153-167.

Serlenga, L., & Shin, Y. (2013). The Euro Effect on Intra-EU Trade: Evidence from the Cross Sectionally Dependent Panel Gravity Models. *Unpublished mimeo*.

Shepsle, K. A., & Weingast, B. R. (1981). Political preferences for the pork barrel: A generalization. *American Journal of Political Science*, 96-111.

Smith, V. L. (1976). Experimental economics: Induced value theory. *The American Economic Review*, 274-279.

Smith, V. L. (1982). Microeconomic systems as an experimental science. *The American Economic Review*, 923-955.

Sutter, M. (2003). The political economy of fiscal policy: an experimental study on the strategic use of deficits. *Public Choice, 116*(3-4), 313-332.

Tabellini, G., & Alesina, A. (1990). Voting on the budget deficit. *American Economic Review*

80, 37–49.

Thaler, R. H. (2008). Mental accounting and consumer choice. *Mental accounting and consumer choice, 27*(1), 15-25.

Tommasi, M., & Velasco, A. (1996). Where are we in the political economy of reform?. . *The Journal of Policy Reform, 1*(2), 187-238.

Velasco, A. (1997). *A model of endogenous fiscal deficits and delayed fiscal reforms*. Retrieved from

Velasco, A. (1998). The common property approach to the political economy of fiscal policy. *The Political Economy of Reform*, 165-184.

Velasco, A. (2000). Debts and deficits with fragmented fiscal policymaking. *Journal of Public Economics, 76*(1), 105-125.

Vlaicu, R., Verhoeven, M., Grigoli, F., & Mills, Z. (2014). Multiyear budgets and fiscal performance: Panel data evidence. *Journal of Public Economics, 111*, 79-95.

Volkerink, B., & De Haan, J. (2001). Fragmented government effects on fiscal policy: New evidence. *Public Choice, 109*(3-4), 221-242.

Von Hagen, J. (1992). Budgeting procedures and fiscal performance in the European Communities. *Economic papers*(96), 1-79.

Von Hagen, J. (2002). Fiscal rules, fiscal institutions, and fiscal performance. *Economic and Social Review, 33*(3), 263-284.

Von Hagen, J., & Strauch, R. R. (2001). Fiscal Consolidations: Quality, economic conditions, and success. *Public Choice, 109*(3-4), 327-346.

Von Hagen, J., & Wolff, G. B. (2006). What do deficits tell us about debt? Empirical evidence on creative accounting with fiscal rules in the EU. *Journal of Banking & Finance, 30*(12), 3259-3279.

Wang, Q. (2015). Fixed-effect panel threshold model using Stata. *Stata Journal, 15*(1), 121-134.

Wehner, J. (2009). Institutional constraints on profligate politicians: The conditional effect of partisan fragmentation on budget deficits. *Comparative Political Studies*.

Weingast, B., Shepsle, K., & Johansen, C. (1981). The political economy of benefits and costs: A neoclassical approach to distributive politics. . *Journal of Political Economy, 89*(4), 642-664.

Wildavsky, A. (2003). Controlling public expenditure. *OECD Journal on Budgeting, 2*(4), 27-47.

Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics, 126*(1), 25-51.

World Bank. (2013). Beyond the Annual Budget: Global Experince with Medium Term Expenditure Frameworks. *Washington, DC: World Bank*(doi:10.1596/978-0-8213-9625-4).

World Development Indicators. (2015). Washington DC: World Bank. http://data.worldbank.org/data-catalog/world-development-indicators. *Accessed: December, 2016*.

Wyplosz, C. (2012). Fiscal rules: Theoretical issues and historical experiences *Fiscal Policy after the Financial Crisis* (pp. 495-525): University of Chicago Press.

Yuval, F., & Herne, K. (2005). Sophisticated behavior under majoritarian and non-majoritarian voting procedures. *Political Behavior*, *27*(3), 217-237.