International Journal of Public Policy

ISSN online: 1740-0619 - ISSN print: 1740-0600
https://www.inderscience.com/ijpp

An investigation of the delayed stabilisation hypothesis with experimental data

Adedeji Adeniran, Oluyele Akinkugbe

DOI: 10.1504/IJPP.2022.10049349

Article History:

Received:	17 February 2018
Last revised:	16 April 2022
Accepted:	12 May 2022
Published online:	08 August 2022

An investigation of the delayed stabilisation hypothesis with experimental data

Adedeji Adeniran

Research Division,
Centre for the Study of the Economies of Africa,
Abuja, Nigeria
Email: adedejimgt@gmail.com

Oluyele Akinkugbe*

School of Economics and Finance,
University of the Witwatersrand,
WITS 2050, Johannesburg, South Africa
Email: olu_akinkugbe@yahoo.com
*Corresponding author

Abstract: This paper empirically tests a version of the delayed stabilisation hypothesis, using experimental data. The hypothesis suggests the possibility of a stabilisation threshold at which a deficit bias due to common pool problem – government fragmentation is eliminated. Formulated along the line of a dynamic common pool problem earlier used in the natural resource environment literature, we extend the experimental design to a fiscal setting by including key features of the legislative bargaining game of 'divide-the-dollar'. The extent of government fragmentation is captured in the formulation by varying the size of interest groups across treatments. Our results do not support the prediction of delayed stabilisation. Moreover, deficit level tends to be highest in the period after post-stabilisation threshold predicted by the hypothesis. This finding suggests that adopting an active stabilisation policy is a more potent tool for policymakers, as against relying on budget actors to act endogenously to correct a deficit bias.

Keywords: fiscal stabilisation; delayed stabilisation hypothesis; dynamic common pool model; experimental design; panel threshold regression.

JEL codes: E63, C72, C73, C92, C24.

Reference to this paper should be made as follows: Adeniran, A. and Akinkugbe, O. (2022) 'An investigation of the delayed stabilisation hypothesis with experimental data', *Int. J. Public Policy*, Vol. 16, Nos. 2/3/4, pp.204–226.

Biographical notes: Adedeji Adeniran is the Director of Research at the Centre for the Study of the Economies of Africa (CSEA). He holds a PhD from the University of Witwatersrand, South Africa. His research interests cut across macroeconomics, development finance, public economics and policy analysis and experimental economics. His recent publication include: *The Role of Social Influence in Enforcing Tax Compliance: Experimental Evidence from Nigeria* (Institute of Development Studies, International Centre for Tax and Development Working Papers).

Oluyele Akinkugbe is currently with the Department of economics at the Mount Saint Vincent University in Halifax, Canada. Between 2016 and 2018, he served as Economic Policy Adviser to the Kingdom of Tonga in the South Pacific Islands. He also served as Macroeconomic Policy adviser to the Government of Barbados, Caribbean Islands, during the period 2010 to 2011.

1 Introduction

An important prediction of the dynamic common pool model in fiscal policy analysis is that there is a threshold fiscal stabilisation threshold at which government fragmentation no longer leads to poor fiscal outcome.[1] When attained, the threshold triggers fiscal stabilisation a change in fiscal policy, either through tax increase and/or expenditure cut, aimed at correcting the deficit bias (Velasco, 1997, 1998).[2] This phenomenon is described as the '*delayed stabilisation hypothesis*' (Alesina et al., 2006; Velasco, 1997). That is, at higher levels of public-sector wealth, aggressive appropriation of resources occurs that leads to debt accumulation. On the other hand, as public-sector wealth declines, the efficiency gain accruing from fiscal stabilisation becomes attractive to budget actors. With the use of experimental data, this paper provides empirical evidence on the hypothesis that has an important policy implication and offers insight into the possible timing and strategies for fiscal stabilisation. The paper 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) extensively investigated the effect of resource stock decline on cooperation (or stabilisation) in the natural resource setting. The focus of this paper, on the other hand is on fiscal policy.

Our use of experimental data derives from challenges reported in studies that had attempted to test the delayed stabilisation hypothesis using field data. In the first instance, assessing the effect of public-sector wealth level on the stabilisation process requires isolating only that aspect of the deficit bias that can be ascribed to common pool problem. However, with the conceptual issues associated with gauging common pool problem (Perotti and Kontopoulos, 2002), a precise estimate is of such public sector wealth is difficult to obtain. Furthermore, since the dynamic common pool problem emanates from uncertainty regarding the paths of public-sector wealth, evaluating the magnitude of dynamic externalities along different wealth paths is not readily achievable using field data. Hence existing empirical evidence on the delayed stabilisation hypothesis are mostly limited to indirect inferences from related studies.[3] This is the major motivation for our use of experimental data to test the validity or otherwise of the delayed stabilisation hypothesis in the framework of a dynamic common pool problem.

The reminder of this paper is organised as follows. Section 2 reviews earlier literature on the effect of public-wealth dynamics on fiscal stabilisation; doing this, debt dynamics rather than public-wealth is the focus, as both are conceptually related. Section 3 describes the delayed stabilisation model, which motivates the design of our experiment. Section 4 details the experimental procedure and its implementation. Section 5 presents results from the non-parametric and econometric analyses. Policy implications of the results are discussed in Section 6.

2 Earlier literature

Two strands of existing literature on fiscal sustainability have extensively addressed related objectives of this paper, from which useful inferences can be drawn. The first strand consists of studies on government's fiscal reaction to debt level. In this regard 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 data from the USA 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 stabilisation hypothesis that high debt level (or equivalently declining public-sector wealth) could influence government to initiate stabilisation process.

However, efforts at further 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 Netherlands. 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. Related studies along this line also include those by Abiad and Baig (2005) and Mendoza and Ostry (2008) who find a positive and significant fiscal reaction at low debt levels, but becomes unresponsive as the debt-GDP ratio exceeds the threshold of 50%. Burger et al. (2011), using a vector error correction model (VECM), observe in the case of South Africa that every 1% increase in debt/GDP ratio is followed by 0.5% and 0.31% 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 et al., 2003); negative effect of debt on growth (Kaur and 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 et al., 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 this evidence partly support the delayed stabilisation 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.

The second strand of the literature builds on the relation between public-sector wealth dynamics and fiscal stabilisation that is assessed using the fiscal adjustment process – 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 considers a decline of at least 0.5% 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 (e.g., see Von Hagen and 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 the fiscal adjustment process (e.g., see Dell'Erba et al., 2015; Gupta et al., 2004; Mierau et al., 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) notes 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 lead to '*adjustment fatigue*' that tends to reduce the incentive to pursue tight fiscal policy. For example, Baldacci et al. (2010) note that successful adjustment episode takes an average of ten years and is less likely in countries facing prolonged 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 (e.g., see Alesina and Drazen, 1989; Lora and 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 and Perotti, 1995a; Alesina et al., 1998; Illera and Mulas-Granados, 2002). In particular, Alesina and Perotti (1995a) report that only 8.7% of adjustment attempts by coalition governments in the Organization for Economic Cooperation and Development (OECD) countries are successful, compared to 64.3% 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 European Union (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 that could skew expenditure towards public investment in district projects, rather than transfers and wages whose benefits are not confined to specific interest group (Alesina and Perotti, 1995b; Weingast et al., 1981). Overall, the key insight from this literature indicates that high debt levels motivate government to undertake fiscal adjustment, but the effect on successful stabilisation is inconclusive.

3 Theoretical conceptualisation

3.1 The delayed stabilisation model

The idea that declines in public-sector wealth could increase the propensity for fiscal stabilisation 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 over how to distribute the burden of reform – the level of taxation needed to restore sustainability for instance – there arises a delay in stabilisation 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 stabilisation. However, as delay involves debt accumulation or equivalently further declines in public-sector wealth – stabilisation 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 stabilisation. Given this theoretical advantage, the conceptualisation that follows leans more heavily on Velasco’s approach.

The economy is assumed to exhibit the same set of characteristics as those of the mechanism through which common pool problem generates deficit bias (see Chari and Kehoe, 1993; Velasco, 1997, 1998, 2000; Weingast et al., 1981; Koh, 2015). Specifically, the budget actors have a utility function given by:

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

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:

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

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. As Katayama (2008) notes, in the pre-stabilisation period, the government relies more on distortionary financing options which comes with significant welfare loss. Taking 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 stabilisation. The behaviour of z_t is described by:

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

where g_{it}^* it g is the social planner desired spending, derived by maximising equation (1) subject to equation (4). This means that deadweight loss – z_t, is only incurred if the social planner's desired spending level is not realised. This relation can be used to formally define fiscal stabilisation 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:

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

where g_{it}^{nc} is the spending level chosen by budget actor in a fragmented system and is greater than g_{it}^*. Taxation is then introduced into the model to account for situations that occur if and when the economy exceeds its debt ceiling; that is, when the initial public-sector wealth is exhausted. Once the debt ceiling is reached, 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 budget constraint, 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:

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

$$U_i^* = \frac{1+r}{r}\left[Log\frac{rW_{t-1}}{n}\right] \quad (8)$$

with

$$U_i^{nc} < U^*. \quad (9)$$

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

3.2 Fiscal stabilisation and public-sector wealth dynamics

Given the initial level of public-sector wealth, aggressive appropriation will characterise the budgeting process, thereby putting the fiscal policy on an unsustainable path.

However, with the dynamic nature of interaction, fiscal stabilisation 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 and 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:

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

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

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

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 maximising equation (10) subject to equation (11), which yields:

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

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

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

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

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

and

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

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

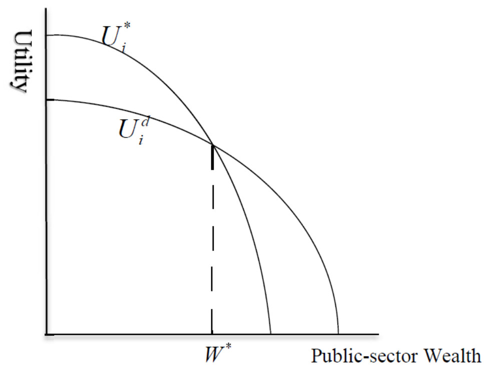
$$Log\left[\frac{rW_{t-1}}{rW_{t-1} - z}\right] \geq \frac{1}{r} Log\left(\frac{1+r}{1+nr}\right) + Log\left[n - (n-1)\left(\frac{r}{1+r}\right)\right] \tag{16}$$

Equations (15) and (16) show that the utility associated with defection path can be greater or less than that which obtains from stabilisation 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 stabilisation path exceeds that of defection path.[4] Thus, we have the situation where each group chooses spending level that restores fiscal stabilisation.

Figure 1 provides a graphical illustration of this result, with W^* denoting the threshold at which further decline in public-sector wealth leads to fiscal stabilisation.

Figure 1 Public-sector wealth dynamics and stabilisation 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 stabilisation and stock of public-sector wealth when stabilisation occurs. Specifically, if we substitute equation (3) into (1), it yields:

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

The definite solution to equation (17) gives:

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

where T is the expected time of stabilisation and W_T is the stock of public-sector wealth at the time of stabilisation. Essentially, fiscal stabilisation is feasible within a fragmented budget institution, although it comes with substantial delay. The stabilisation process is driven by public-sector wealth decline, which forces groups to fully internalise 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 stabilisation.

4 Experimental procedure

As previously highlighted, this paper seeks to empirically verify the delayed stabilisation hypothesis using experimental data. Based on the dynamic common pool game (Chari and Kehoe, 1993; Velasco, 1997, 1998, 2000), participants – players in the game – are initially endowed with common resources designated as public-sector wealth; these 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 (Croson, 2005).[5] Overall, the sustainability of the resources hinges on player's appropriation behaviour over the duration of the game. Thus, in the case that resources become unsustainable; a lump-sum tax is imposed on 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 stabilisation takes place as players become more constrained in their appropriation behaviour. However, an identification problem could arise as to whether the stabilisation is temporal (due to pressure from taxation) or permanent (as expected if stabilisation truly takes place). To separate these effects, initial wealth is restored after some rounds of play. For delayed stabilisation hypothesis to be validated, we should observe an appropriation level that remains consistent with fiscal stabilisation as players ought to have developed the reputation for low spending through the history of the game.

In what follows, we discuss 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 paper.

4.1 Decision setting

The participants/players are involved in a legislative bargaining game 'divide-the-dollar', over an infinite horizon – each player makes a sequential demand on a fixed sum of resources and a the game continues till a majority winning coalition (MWC) that satisfies the budget constraint is formed.[6] 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 ten rounds of play. The rule gives the continuation probability of the game, and it is defined by the discount factor $1 - r$. As previously mentioned, the terminal period of the game is determined by the random stopping rule.[7] This design helps prevent the possibility of the game ending prematurely, thereby providing sufficient observations to test the delayed stabilisation hypothesis. If points are exhausted – i.e., if 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 stabilisation. The implication of this is that observed stabilisation 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 stabilisation. 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 stabilising, it will simply generate a surplus.

4.2 Design parameters and treatments

The design parameters for the experiment are defined by W_t, r, n, z and T; the parameters being as previously defined. 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 normalised 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 stabilisation and wealth threshold at which this takes place. The predictions as summarised in Table 1 indicate that despite the groups operating in a fragmented setting, fiscal stabilisation 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 stabilises. Another striking prediction is that the larger the group size, the shorter the expected time for stabilisation 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 stabilisation.

Table 1 Theoretical predictions

Treatment		*Fragmented government*	
Initial wealth (points)	*Group size*	*Wealth threshold for stabilisation (W_T, points)*	*Expected time of stabilisation (T, period)*
60	3	14	7
100	5	13	6
300	3	13	15
500	5	16	10

Source: Author computation

4.3 Sampling procedure

4.3.1 Optimal sample selection

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

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

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 $\left(\frac{\sigma}{\delta}\right)$ gives the standard deviation of the change in outcome variable and captures the minimum detectable effect size. Using the conventional level of significance of 5% and power of the test of 0.80 give $t_{\alpha/2}$ as 1.96 and t_β as 0.84 respectively.[8] 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 was obtained.

4.3.2 Sampling technique

The subject pool for the experiment consisted of first and second year undergraduate students at 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:

1. the participants have attained 18 years of age
2. the participants are computer literate
3. the participants are in the first or second year undergraduate degree program.

A total number 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 2. The distribution shows that the participants come from a variety of background, which implies that our sample selection procedure is not biased towards any demographic characteristics.

4.3.3 Experimental implementation

The experiment reported in this paper was conducted at the Computer Laboratory of the School of Economics and Business of the University of Witwatersrand in September and October 2016. Four sessions were run in all: two sessions each for primary and secondary treatments, respectively. No participant was involved in more than a session. On average, participants earned ZAR68, and the sessions lasted between 60–90 minutes.

Table 2 Demographic characteristics of the participants

Variable	*Frequency*	*Percentage (%)*
Sex		
Male	36	56.3
Female	28	43.8
Year of study		
1st	38	59.4
2nd	26	40.6
Faculty		
Commerce, law and management	25	39.1
Humanities	23	35.9
Science	16	25

Source: Authors' computation

On the day of the experiment, participants were first taken through the instructions.[9] The instructions detailed the decision setting and group size; identities other group members are not revealed. 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 about the tax rate imposed if/when the initial wealth is exhausted; they were, however, not intimated with the fact that their initial wealth would be restored in subsequent periods. Explicit communication is not allowed among the participants. Lastly, before the main experimental session, participants had a hands-on pilot session to clarify any ambiguity. The experiment was computerised using the z-tree program (Zurich toolbox for readymade experiment) developed by Fischbacher (2007).

4.4 Empirical results

4.4.1 Non-parametric analysis

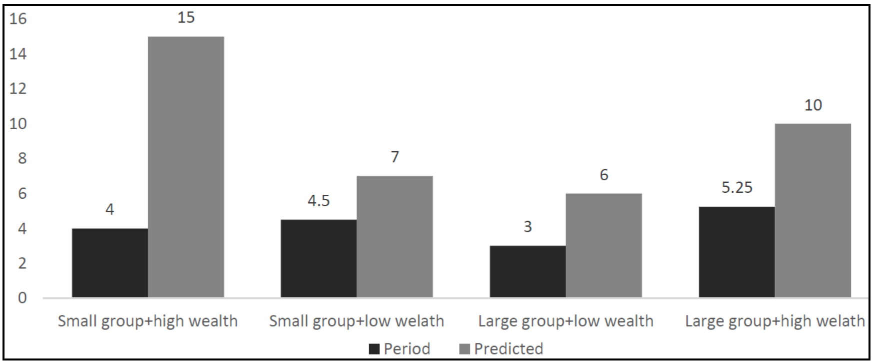
The theoretical description of a fragmented budget institution indicates that while public-sector wealth initially declines – or deficit is generated – there is a threshold where fiscal stabilisation becomes the optimal strategy for the budget actors. The expected time and public-sector wealth at this threshold are estimated and presented in Table 1. In Figures 2 and 3, we compare these theoretical predictions with outcomes of the experiment. As shown in Figure 2, all groups attained the expected time of stabilisation 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 stabilisation time is shorter than what is theoretically predicted.

Furthermore, Figure 3 shows the average public-sector wealth corresponding to the expected stabilisation time. Unlike the finding on the timing of stabilisation, 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).

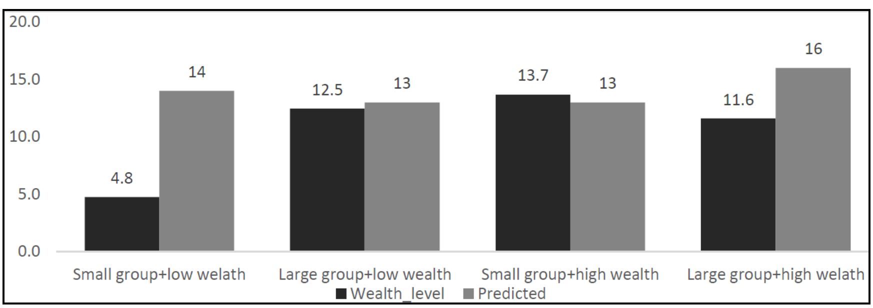
Observation 2: The public-sector wealth threshold matches the theoretical expectation.

Figure 2 Actual versus predicted time of stabilisation



Source: Authors computation

Figure 3 Actual versus predicted pubic-sector threshold for stabilisation



Source: Authors computation

Results of the expected time and public-sector wealth threshold for stabilisation, discussed above only reveal whether or not the groups reach a stage in which fiscal stabilisation is the optimal strategy. To check if the groups indeed adopted the optimal strategy and importantly, if stabilisation ensues afterwards, we need to compare the public-sector wealth dynamics or deficit before and after the groups reached the stabilisation threshold. However, restoring the public-sector wealth after the stabilisation threshold implies that the estimated public-sector wealth dynamics will be overstated. Therefore, we only report deficit level in Figure 4. The deficit level after the stabilisation threshold is divided into three parts:

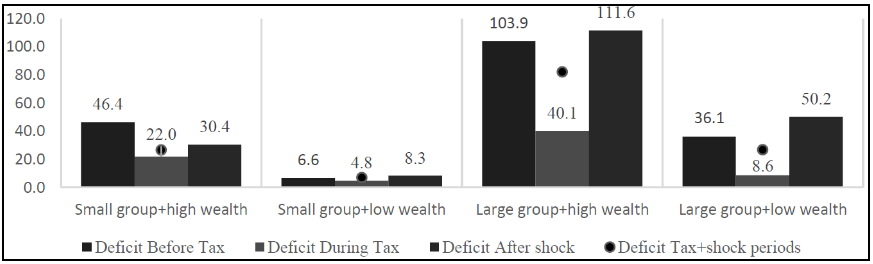
1 when groups pay taxes only

2 when group initial wealth is restored only

3 the average deficit incurred in 1 and 2.

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-stabilisation threshold in three of the four treatment categories. Intriguingly, the only exception is groups in the first treatment. As observed in Figures 2 and 3, groups in the first treatment exhibit more aggressive appropriation behaviour than others, prior to reaching the stabilisation threshold.

Figure 4 Deficit levels pre and post stabilisation threshold



Source: Authors computation

This may suggest a behavioural reversal, such that groups that are aggressive in their appropriation become more restrained after reaching the stabilisation 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 stabilisation 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 stabilisation hypothesis.

4.4.2 Econometric model specification and estimation

The key prediction of the delayed stabilisation 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 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, Hansen (1999) argues 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.

4.4.2.1 Panel threshold regression

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

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

This can be more compactly written as:

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

where D_{it} is the deficit/surplus incurred by group i in period t; n_i is the group size – measure of government fragmentation – that also represents the threshold variable. 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} or zero otherwise; d_i is the group-level fixed effect; ε_{it} is the disturbance term. Also, it X'_{it} represents the control variables, which includes: shadow cost which is measured as $Log\left(W_{i,t-1}/W_{i,t-2}\right)$; inverse of time index 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 stabilisation 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:

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

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

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 setup (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:

- For single threshold:

$$D_{it} = \theta_1 n_i TDM\left(W_{i,t-1}^{*} \le \gamma_1\right) + \theta_2 n_i TDM\left(W_{t-1}^{*} > \gamma_1\right) + \sum_{k=2}^{4} \alpha_k X_{it}' + d_i + \varepsilon_{it} \quad (23)$$

- For double thresholds:

$$\begin{aligned} D_{it} &= \theta_1 n_i TDM\left(W_{i,t-1}^{*} \le \gamma_1\right) + \theta_2 n_i TDM\left(\gamma_1 < W_{t-1}^{*} \le \gamma_2\right) \\ &\quad + \theta_3 n_i TDM\left(W_{t-1}^{*} > \gamma_2\right) + \sum_{k=2}^{4} \alpha_k X_{it}' + d_i + \varepsilon_{it} \end{aligned} \quad (24)$$

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 stabilisation 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.

4.4.2.2 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 stabilisation is expected once groups have attained the stabilisation threshold. This means government fragmentation should have no effect on deficit in the period after stabilisation threshold. Thus, we define a dummy variable – *Break*1, which is 1 in periods when the groups have not reached stabilisation threshold and zero otherwise. Periods after the stabilisation 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 stabilisation 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 – *Break*2 which takes the value of 1 in period 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 stabilisation hypothesis is validated on the condition that $\phi_1 + \phi_2$ is less than 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.

4.4.2.3 Diagnostics analysis

According to Cheng et al. (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 presents the panel unit-root results based on Levin-Lin-Chu (LLC) and Im, Pesaran and Shin (IPS) tests (see Im et al., 2003; Levin et al., 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 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

Notes: Standard errors in parentheses; *$p < 0.1$, **$p < 0.05$ and ***$p < 0.01$.

Table 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

Notes: 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.

The results of the test for threshold effect are presented in Table 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.

4.4.3 Econometric results

4.4.3.1 Fixed effect panel threshold regression results

Table 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-stabilisation 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 stabilisation. Furthermore, the double threshold results, although not the best fit, help to decompose the relationship between government fragmentation and deficit in the post-stabilisation 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 stabilisation 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 5 Estimated fixed effect panel threshold regression model

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

Notes: White standard errors in parentheses; $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

4.4.3.2 Piecewise linear regression results

Results from the piecewise linear regression is presented in Table 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 stabilisation 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 stabilisation model.

Table 6 Estimated piecewise linear regression model

	Single	*Double*
L.shadow cost	6.350*** (2.190)	5.269*** (0.732)
1/TIME	11.24*** (3.595)	12.02** (4.767)
Dumm_high	49.51** (22.83)	47.30** (21.98)
Government fragmentation (n_i)	17.39*** (0.183)	29.42*** (0.779)
n_i * *Break*1	7.172*** (1.287)	–4.823*** (0.830)
n_i * *Break*2		–24.64*** (2.457)
*Break*1	–66.03*** (12.21)	–26.69* (14.65)
*Break*2		83.27*** (7.503)
Constant	–99.78*** (25.20)	–135.6*** (19.05)
R^2	0.822	0.835

Notes: Standard errors in parentheses; $*p < 0.1$, $**p < 0.05$ and $***p < 0.01$.

5 Conclusions

This paper empirically tests the prediction of delayed stabilisation 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 stabilisation. Regardless of treatment categories, groups attained the pubic-sector wealth threshold for stabilisation 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-stabilisation threshold is reached. The negative effect of government fragmentation on deficit becomes even more pronounced in post-stabilisation threshold phase, contrary to the prediction of delayed stabilisation.

The results corroborate Lukkezen and Rojas-Romagosa (2013) and Purfield (2003) studies that also find weak support for delayed stabilisation 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 stabilisation 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 et al. (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 stabilisation is higher during periods of economic crisis, than otherwise.

These findings have important implications for fiscal policy making, especially in developing economies. In view of lack of evidence regarding delayed stabilisation, it becomes crucial for policy makers to adopt proactive active stabilisation policy such as fiscal rules or centralisation of the budget institution, as against relying on budget actors to act endogenously to correct deficit bias. More importantly, this paper suggests that periods of economic crisis are the optimal time to introduce holistic stabilisation 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.

Appendices/Supplementary materials are available on request by emailing the corresponding author.

References

Abiad, A.G. and Baig, T. (2005) *Underlying Factors Driving Fiscal Effort in Emerging Market Economies*, IMF Working Paper, Vol. 5, No. 106.

Afonso, A. and Hauptmeier, S. (2009) *Fiscal Behaviour in the European Union: Rules, Fiscal Decentralization and Government Indebtedness*, European Central Bank Working Paper, No. 1054.

Alesina, A. and Drazen, A. (1989) *Why are Stabilizations Delayed?*, National Bureau of Economic Research Working Paper.

Alesina, A. and Perotti, R. (1995a) *Fiscal Expansions and Fiscal Adjustments in OECD Countries*, National Bureau of Economic Research Working Paper.

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

Alesina, A., Ardagna, S. and Trebbi, F. (2006) *Who Adjusts and When? On the Political Economy of Reforms*, National Bureau of Economic Research Working Paper.

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

Ardagna, S. (2004) 'Fiscal stabilizations: when do they work and why', *European Economic Review*, Vol. 48, No. 5, pp.1047–1074.

Baldacci, E., Gupta, S. and Mulas-Granados, C. (2010) 'Getting debt under control', *Finance and Development*, Vol. 47, No. 4, pp.18–21.

Battaglini, M., Nunnari, S. and Palfrey, T.R. (2012) *The Dynamic Free Rider Problem: A Laboratory Study*, Economic Theory Center Working Paper, No. 48-2012.

Bohn, H. (1998) 'The behavior of US public debt and deficits', *Quarterly Journal of Economics*, Vol. 113, No. 3, pp.949–963.

Botelho, A., Dinar, A. and 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*, 3rd ed., Cambridge University Press, Cambridge.

Burger, P., Stuart, I., Jooste, C. and Cuevas, A. (2011) *Fiscal Sustainability and the Fiscal Reaction Function for South Africa*, IMF Working Papers, pp.1–27.

Chari, V.V. and Kehoe, P.J. (1993) 'Sustainable plans and debt', *Journal of Economic Theory*, Vol. 61, No. 2, pp.230–261.

Cheng, Y.S., Liu, Y.P. and Chien, C.Y. (2010) 'Capital structure and firm value in China: a panel threshold regression analysis', *African Journal of Business Management*, Vol. 4, No. 12, p.2500.

Croson, R. (2005) 'The method of experimental economics', *International Negotiation*, Vol. 10, No. 1, p.131148.

Daniel, J., Callen, T., Terrones, M., Debrun, X. and Allard, C. (2003) 'Public debt in emerging markets: is it too high?', *World Economic Outlook*, p.113.

Dell'Erba, S., Mattina, T. and Roitman, A. (2015) 'Pressure or prudence? Tales of market pressure and fiscal adjustment', *Journal of International Money and Finance*, Vol. 51, No. C, pp.196–213.

Eslava, M. (2011) 'The political economy of fiscal deficits: a survey', *Journal of Economic Surveys*, Vol. 25, No. 4, pp.645–673.

Fischbacher, U. (2007) 'z-tree: Zurich toolbox for ready-made economic experiments', *Experimental Economics*, Vol. 10, No. 2, pp.171–178.

Fréchette, G.R., Kagel, J.H. and Morelli, M. (2005) 'Behavioral identification in coalitional bargaining: an experimental analysis of demand bargaining and alternating offers', *Econometrica*, Vol. 73, No. 6, pp.1893–1937.

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

Gupta, S., Clements, B., Baldacci, E. and Mulas-Granados, C. (2004) 'The persistence of fiscal adjustments in developing countries', *Applied Economics Letters*, Vol. 11, No. 4, pp.209–212.

Hansen, B. (1999) 'Threshold effects in non-dynamic panels: estimation, testing, and inference', *Journal of Econometrics*, Vol. 93, No. 2, pp.345–368.

Hsieh, C-T. (2000) 'Bargaining over reform', *European Economic Review*, Vol. 44, No. 9, pp.1659–1676.

Illera, R.M. and Mulas-Granados, C. (2002) *Duration of Fiscal Budgetary Consolidations in the European Union*, European Economy Group.

Im, K.S., Pesaran, M.H. and Shin, Y. (2003) 'Testing for unit roots in heterogeneous panels', *Journal of Econometrics*, Vol. 115, No. 1, pp.53–74.

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

Kaur, B. and Mukherjee, A. (2012) *Threshold Level of Debt and Public Debt Sustainability: The Indian Experience*, Reserve Bank of India, No. 1.

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

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

List, J.A., Sadoff, S. and Wagner, M. (2011) 'So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design', *Experimental Economics*, Vol. 14, No. 4, pp.439–457.

Lora, E. and Olivera, M. (2004) 'What makes reforms likely: political economy determinants of reforms in Latin America', *Journal of Applied Economics*, Vol. 7, No. 1, pp.99–135.

Lukkezen, J. and Rojas-Romagosa, H. (2013) *An Early-Warning Indicator for Sovereign Debt Sustainability*, CPB Background Information: Updated Version of CPB Discussion Paper, Vol. 212, No. 7.

Mason, C.F. and Phillips, O.R. (1997) 'Mitigating the tragedy of the commons through cooperation: an experimental evaluation', *Journal of Environmental Economics and Management*, Vol. 34, No. 2, pp.148–172.

Mason, C.F. and Phillips, O.R. (2002) 'In support of trigger strategies: experimental evidence from two person noncooperative games', *Journal of Economics & Management Strategy*, Vol. 11, No. 4, pp.685–716.

Mendoza, E.G. and Ostry, J.D. (2008) 'International evidence on fiscal solvency: is fiscal policy', *Journal of Monetary Economics*, Vol. 55, No. 6, pp.1081–1093.

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

Mulas-Granados, C. (2003) 'The political and economic determinants of budgetary consolidation in Europe', *European Political Economy Review*, Vol. 1, No. 1, pp.15–39.

Osés-Eraso, N. and Viladrich-Grau, M. (2007) 'Appropriation and concern for resource scarcity in the commons: an experimental study', *Ecological Economics*, Vol. 63, No. 2, pp.435–445.

Osés-Eraso, N., Udina, F. and Viladrich-Grau, M. (2008) 'Environmental versus human-induced scarcity in the commons: do they trigger the same response?', *Environmental and Resource Economics*, Vol. 40, No. 4, pp.529–550.

Perotti, R. and Kontopoulos, Y. (2002) 'Fragmented fiscal policy', *Journal of Public Economics*, Vol. 86, No. 2, pp.191–222.

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

Thaler, R.H. (2008) 'Mental accounting and consumer choice', Vol. 27, No. 1, pp.15–25.

Tommasi, M. and Velasco, A. (1996) 'Where are we in the political economy of reform?', *The Journal of Policy Reform*, Vol. 1, No. 2, pp.187–238.

Velasco, A. (1997) *A Model of Endogenous Fiscal Deficits and Delayed Fiscal Reforms*, National Bureau of Economic Research.

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

Velasco, A. (2000) 'Debts and deficits with fragmented fiscal policymaking', *Journal of Public Economics*, Vol. 76, No. 1, pp.105–125.

Von Hagen, J. and Strauch, R.R. (2001) 'Fiscal consolidations: quality, economic conditions, and success', *Public Choice*, Vol. 109, Nos. 3–4, pp.327–346.

Wang, Q. (2015) 'Fixed-effect panel threshold model using Stata', *Stata Journal*, Vol. 15, No. 1, pp.121–134.

Weingast, B., Shepsle, K. and Johansen, C. (1981) 'The political economy of benefits and costs: a neoclassical approach to distributive politics', *Journal of Political Economy*, Vol. 89, No. 4, p.642664.

Weingast, B.R., Shepsle, K.A. and Johnsen, C. (1981) 'The political economy of benefits and costs: a neoclassical approach to distributive politics', *The Journal of Political Economy*, August, Vol. 89, No. 4, pp.642–664.

Notes

1 Dynamic common pool problem occurs when fragmentation within budgeting institution size and/or procedural fragmentation – 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 (Perotti and Kontopoulos, 2002; Chari and Kehoe, 1993; Velasco, 1997, 1998, 2000). Also, common pool problem and government fragmentation will be used interchangeably in this paper.

2 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 paper we focus more on the dynamics of public-sector wealth as it is easier to empirically handle using experimental data.

3 Eslava (2011) undertakes a detailed survey of existing literature on delayed stabilisation hypothesis. As the author notes, this evidence are indirect, therefore difficult to interpret either in support of or against the hypothesis.

4 In the case in which the initial public-sector wealth is already low enough to ensure that utility from stabilisation path exceeds that from defection path, then stabilisation is immediately established.

5 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.

6 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.

7 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. (2012). 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 $1 - r$.

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

9 In line with the treatment levels, four different instruction set were used. The instruction sets are available upon request from the authors.